予測・説明関係を検討する統計的検定法の分類

従属変数			独立変数（質的）			
量的・質的	尺度水準	統計量	群数	対応の有無：対応あり	対応の有無：対応なし	
量的	比 間隔	平均値	2群	対応のあるt検定	対応のないt検定	パラメトリックな方法
			2群以上	被験者内要因分散分析	被験者間要因分散分析	
質的	順序	分布位置	2群	ウィルコクソンの符号つき順位検定	マン・ホイットニの検定 ＝ウィルコクソンの順位和検定	ノンパラメトリックな方法
			2群以上	フリードマンの検定	クラスカル・ウォリスの検定	
	名義	度数 比率	2群	マクネマーの検定	カイ2乗検定	
			2群以上	コクランのQ検定	カイ2乗検定	

分析目的による多変量データ解析法の分類

従属変数に対する独立変数の影響の強さを検討する分析

重回帰分析	数量化Ⅰ類
因子分析	数量化Ⅱ類
ロジスティック回帰分析	ロジット対数線形モデル
判別分析	など

変数や個体のまとまりを構成する分析法

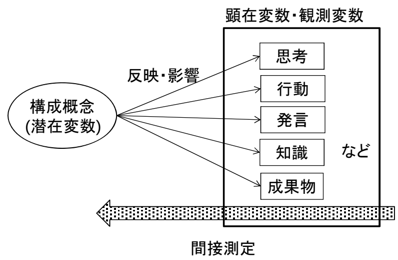
クラスター分析	対応分析
主成分分析	数量化Ⅲ類
因子分析	数量化Ⅳ類
潜在クラス分析	多次元尺度構成法　など

変数間の関連を説明するモデルを構成する分析法

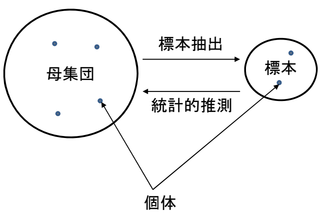
パス解析	対数線形モデル
共分散構造分析(SEM)	など

心理的な量

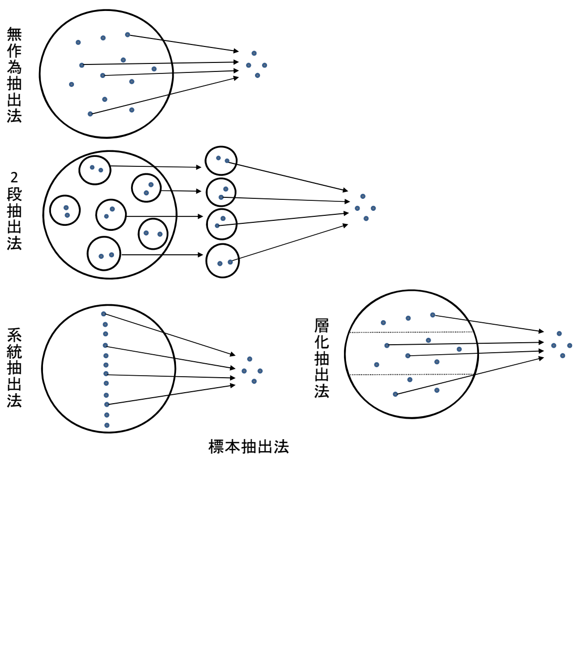
- 心理的な量：　性格，感情，気持ち，知能，能力，学力など，心理的なものの程度を表す
- 性格，感情，気持ち，知能，能力，学力などは，頭の中で考え出した心理的なもの．　構成概念
- 心理的なもの（構成概念）は，物理的に存在しない
- それを用いると，何らかの現象を都合良く（上手に）説明することができる抽象的なもの



構成概念の測定



母集団，標本，個体の関係



標本抽出法

■データ収集法
実験法…条件(処理，処遇）を研究者がコントロールして，データを収集
調査法…質問紙やテストの項目に対する応答として，データを収集
面接法…自由発話や質問項目に対する応答などにより，データを収集
観察法…特定の観点に従って対象を観察することにより，データを収集

■倫理的配慮
・研究者の氏名，所属，連絡(問合わせ）先等を明示する
・研究参加者にとっての，十分な説明と同意．「インフォームド・コンセント」
・苦痛を与える可能性がある場合は，そのこと，その程度，予後などを説明
・調査に協力せずとも，また，途中でやめても，何ら不利益を受けない
・任意参加．研究参加者が決められる
・同意後でも，自由に研究への参加を取りやめることができる
　社会心理学実験などで，デブリーフィング後，同意を撤回してもよい
　デブリーフィング：　真の研究目的を伝えること，<u>及び，それを口外しないと約束させること</u>
・同意がない限り，個人が分かってしまうような分析や結果の公表はしない　など

■質問紙作成にあたっての注意点
倫理的配慮，著作権を意識して作成する
　既存の尺度を使う場合，著作権を侵害しない

必ず一度自分で回答してみる
　自分でもやる気も起きないような質問紙は，他人ならなおさらやりたくない

必ず他の人に回答してもらう
　不明な点，わかりにくいところなどを指摘してもらう

必要かつ最小限の項目数にする
　回答者の負担を考える．あまり長いとやる気を喪失する
　成人でも，時間にして 15 分程度まで．項目数にして 100 項目くらいまで
　小さい子や年長者を対象とする場合は，もっと注意を．

依頼文ではリサーチクエスチョンを前面に出し過ぎない
　回答にバイアスがかかる危険性がある
　ex）　親のしつけが子どもの成長に及ぼす効果　→　親子関係

集団の特徴を聞くフェイスシート項目を入れておく
　性別，年齢など，標本の基本的属性を把握
　不必要なことは聞かない(学校名，学年など．分析に使わないのであれば)

潜在変数の測定は複数の項目を用いて聞く
　その特性が高い人だったらこう答えるだろう，低い人だったらこう答えるだろうという区別がなるべくつくような項目を提示する．
　複数の項目を使って，全体として潜在変数を捉えるようにする
　1 つ潜在変数を測定する項目数は経験的に 10 項目前後．多くても 20 項目

先行する質問が後の質問への回答に影響しないようにする
　抑うつ尺度など，ネガティブな動機づけが高まってしまうようなら後ろへ

複数の場面設定をするような質問紙では，場面の順序を入れ替えた複数の冊子を作る
　順序効果が結果に影響を及ぼす危険性がある
　多数の場面（条件）を設定したい場合，全部の条件を 1 つの冊子に入れ込むのではなく，いくつかの条件だけを入れた何種類かの冊子を作ると，各回答者の負担は軽くなる．
　なおこの場合，対応のない要因として分析する

値が大きい方が，その特性が高いような段階評定にするとわかりやすい
　○　あてはまらない　1　2　3　4　5　　あてはまる
　×　あてはまる　　　1　2　3　4　5　　あてはらない

評定段階の数は4，5くらいが適当．7が限界
　小学校低学年や年長者の場合には，2，3 くらいにすることもある

小学校低学年が回答者の場合，選択枝は設問ごとに縦にならべるとよい
　3　あてはまる
　2　よくわからない
　1　あてはまらない

中央カテゴリ(どちらとも言えない)を入れるかどうかの決まりはない
　行政が行う調査では入れないことが多い
　「どちらとも言えない」と「(経験がないから)わからない」を区別したければ，段階評定とは別に「わからない」というカテゴリを作ることもある．分析時，このカテゴリは除外する

逆転項目を利用する
　不適切なデータ（全項目に 111 …など）を除外することが可能
　合計点を求めるときは，逆転項目のデータは，他の項目のデータと方向性そろえる

　変換後データ　＝　最小カテゴリ値　＋　最大カテゴリ値　－　データ値

■予備調査・本調査
予備調査
　本調査の前に行う調査で，その結果を踏まえて，質問紙を修正したり，場合によっては研究仮説を修正したりする．必要があれば予備調査を繰り返す．
　場合によっては予備予備調査を行う．

予備調査で確認すること
　教示文や調査項目に不備や分かりにくいところはないか
　機能していない項目はないか(全員の回答が同じなど)
　ただし，項目レベルで天井効果や床効果を議論するのはナンセンス
　条件の統制はきちんとできているか
　調査実姉の手続きに問題はないか
　研究者が手続きに十分慣れているか(とくに実験研究)　など

本調査

研究における結果を導出するために実施する調査
対象としたい集団に属する人に実施

研究の種類と経験的な研究参加者数（標本サイズ）

	事例研究 面接・観察	実験研究 実験	調査研究 質問調査・尺度作成	大規模試験 尺度の標準化
予備予備調査			数名	数十名
予備調査	若干名	若干名～数名	数十名	数百名
本調査	数名～数十名	数十名	数百名（～数千名）	数千名～数万名

■尺度水準

名義尺度

数値が特性の違いを表すが，数量的意味を持たない
分類符号として用いられる
1 対 1 対応になっていれば任意に数値を変換可能
四則計算不可
ex）　性別，学部

順序尺度

数値が特性の順序関係を表す
順序性のある分類符号として用いられる
大小関係が保存されれば，任意に変換可能
四則計算不可
ex）　順位（PISA など），薬の副作用の程度，（本来なら，テスト得点，段階評定）

間隔尺度

数値の差（間隔）が特性の大きさの差（違いの程度）を反映する
一般に 0（ゼロ）が「無い」状態を表さない
数値の比が意味を持たない
線形変換可能（$x \rightarrow ax+b,\ a \neq 0$）
加減算可，乗除算不可
ex）　偏差値，知能指数，原点を移動した比尺度（摂氏（華氏）温度），（テスト得点，段階評定）

比尺度

数値の比が特性の大きさの比を反映する
0（ゼロ）が「無い」状態を表す
単位を変換可能（$x \rightarrow ax,\ a \neq 0$）
四則演算可
ex）　物理的な量，面積，体積，力，密度，濃度，人数，割合（比率）

データの構造

多変量データ

番号	入学年度	学科	卒業後の進路
1	20Y1	看護	医療
2	20Y1	教育	教員
3	20Y2	化学	進学
4	20Y2	文	公務員
5	20Y3	医	医療
:	:	:	:

対応のないデータ

学科	学習意欲 専門科目
看護	16
看護	18
:	:
心理	16
心理	15
:	:

反復測定データ

学習意欲 入学時	学習意欲 1年前期末	学習意欲 1年後期末
13	15	18
17	18	17
19	15	14
14	17	17
13	8	2
:	:	:

対応のあるデータ

学習意欲 2年教養	学習意欲 2年専門	学習意欲 2年その他
13	18	12
11	16	17
18	15	18
10	19	8
11	11	11
:	:	:

統計分析における，要因，水準，群

対応のあるデータ

要因(条件の違い)

学習意欲 教養科目	学習意欲 専門科目	学習意欲 その他科目
13	18	12
11	16	17
18	15	18
10	19	8
11	11	11
:	:	:

水準（各条件）

群

対応のないデータ

要因(群分け変数)

学科	学習意欲 専門科目
看護	16
看護	18
:	:
心理	16
心理	15
:	:

水準（各カテゴリ）

群

■質的データの要約

度数分布表

変数の各カテゴリに属する個体の度数や割合(%)をまとめた表

クロス集計表

2 つ以上の質的変数があるとき，それぞれの水準の組合せに属する個体の度数や割合(%)をまとめた表

行パーセント，列パーセント，セルパーセント

円グラフ

変数の各カテゴリに属する個体の割合を，扇形の中心角の大きさで示した図

帯グラフ

変数の各カテゴリに属する個体の割合を，帯の長さで示した図

質的変数の分布の時系列的な変化を概観するときに威力を発揮

棒グラフ

各群における何らかの特性の程度を，棒の長さや高さで表した図

横軸は名義尺度なので，順番を看護学，心理学，医学のように変更しても構わない

■量的データの要約

度数分布表

階級：分割された得点範囲　　　　　　階級幅：階級の得点幅

境界値：階級の上限と下限の値　　　　階級値：階級の中央の値

相対度数：全個体数に対する度数の割合

累積度数：値の小さい階級から積み上げた度数

累積相対度数：全個体数に対する累積度数の割合

ヒストグラム

量的データにおいて，各階級に属する個体の度数や割合を，棒の面積で表した図

箱ひげ図

量的データの分布をよりシンプルに表した図

折れ線グラフ

各階級(群)における何らかの特性の程度を点の位置で示し，階級間の点を直線で結んだ図

散布図

2 つの量的変数の関係を，座標平面上の点で表した図

度数分布表（質的変数）

表4.2　過去3年間の各学科の卒業者数

	医学	看護学	心理学	計
人数	76	100	94	270
パーセント	28%	37%	35%	100%

クロス集計表

表4.3　入学年度と卒業時進路のクロス集計

入学年度	就職	進学	不明	行計
20Y1	77	4	6	87
	0.885	0.046	0.069	0.322
	0.352	0.105	0.462	
	0.285	0.015	0.022	
20Y2	80	11	7	98
	0.816	0.112	0.071	0.363
	0.365	0.289	0.538	
	0.296	0.041	0.026	
20Y3	62	23	0	85
	0.729	0.271	0.000	0.315
	0.283	0.605	0.000	
	0.230	0.085	0.000	
列計	219	38	13	270
	0.811	0.141	0.048	

n
n ／ 行計
n ／ 列計
n ／ 全体

度数分布表（量的変数）

表4.4　自己効力感尺度得点の度数分布表

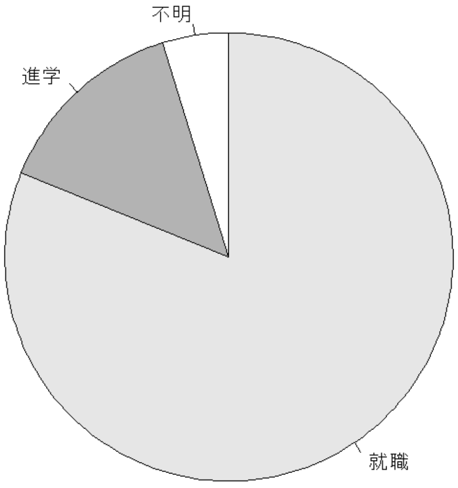
階級(階級幅=5) 下の境界値	上の境界値	階級値	度数	相対度数(%)	累積度数	累積相対度数(%)
25	30	27.5	4	1.5	4	1.5
30	35	32.5	8	3.0	12	4.4
35	40	37.5	29	10.7	41	15.2
40	45	42.5	42	15.6	83	30.7
45	50	47.5	54	20.0	137	50.7
50	55	52.5	50	18.5	187	69.3
55	60	57.5	40	14.8	227	84.1
60	65	62.5	28	10.4	255	94.4
65	70	67.5	9	3.3	264	97.8
70	75	72.5	5	1.9	269	99.6
75	80	77.5	1	0.4	270	100.0

質的変数：　名義尺度，順序尺度の変数

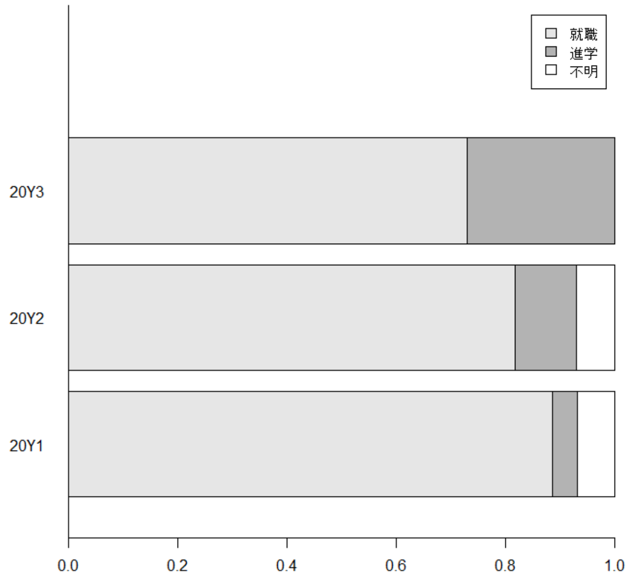
量的変数：　間隔尺度，比尺度の変数

質的変数のデータを要約する図

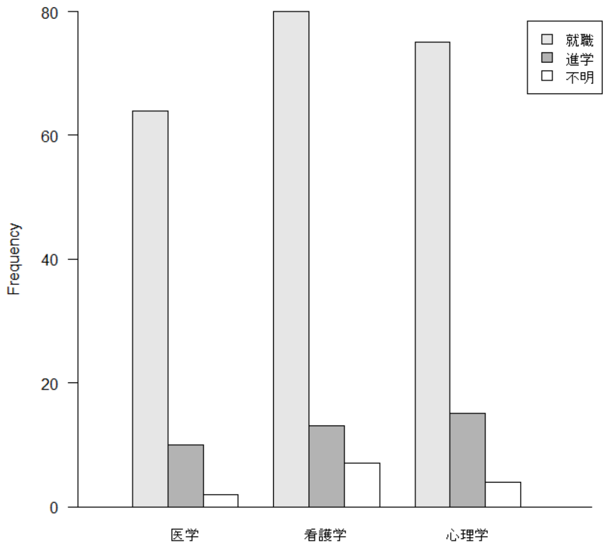
円グラフ



帯グラフ

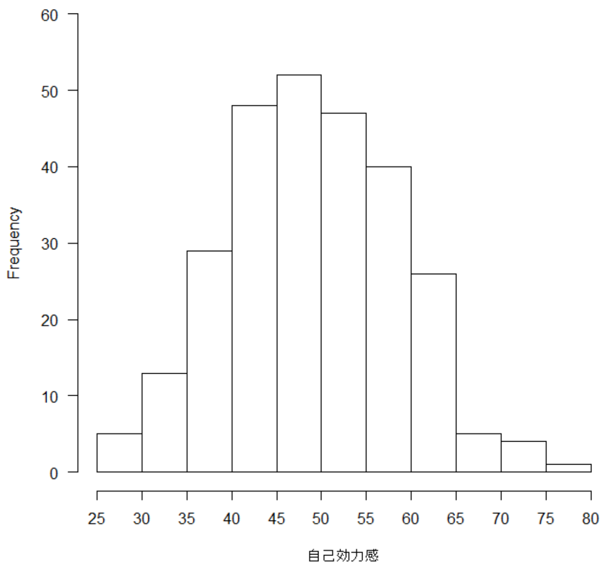


棒グラフ

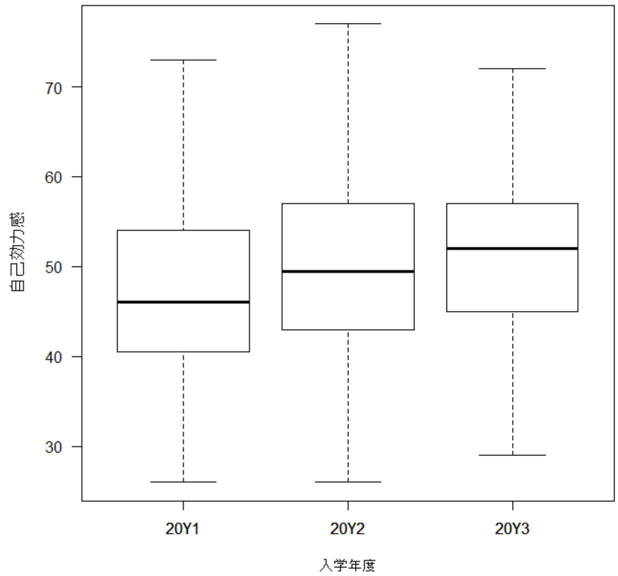


量的変数のデータを要約する図

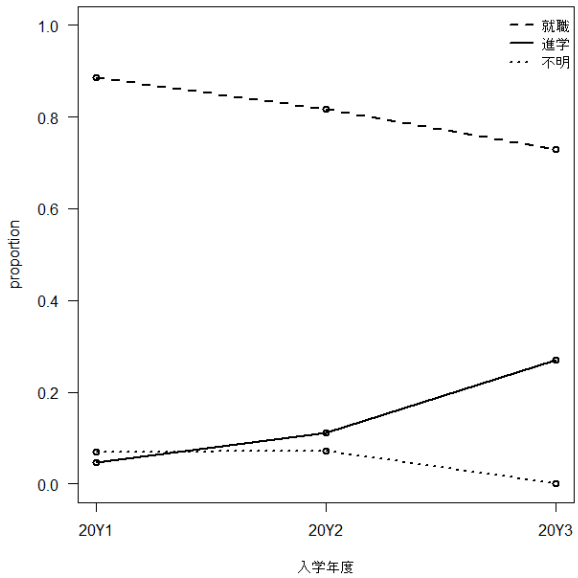
ヒストグラム



箱ひげ図



折れ線グラフ



散布図

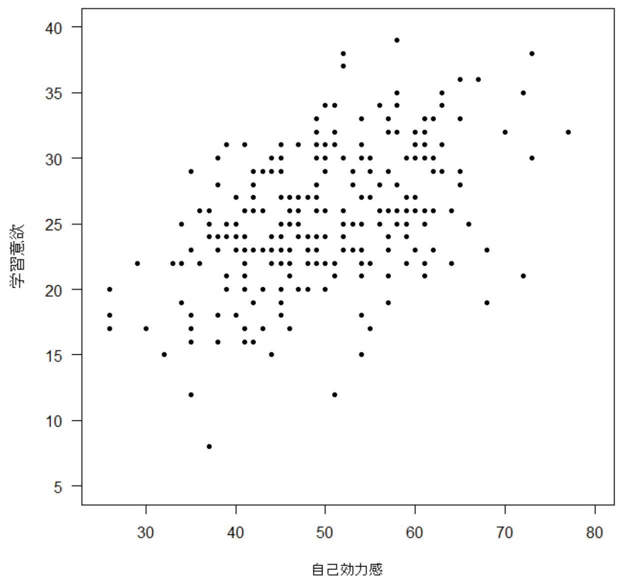
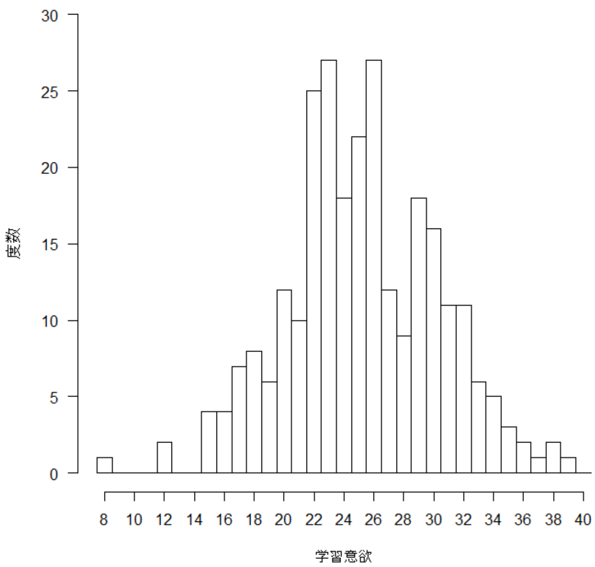


表5.2　学習意欲データの度数分布表

階級値	度数	相対度数(%)	累積度数	累積相対度数(%)
8	1	0.4	1	0.4
9	0	0.0	1	0.4
10	0	0.0	1	0.4
11	0	0.0	1	0.4
12	2	0.7	3	1.1
13	0	0.0	3	1.1
14	0	0.0	3	1.1
15	4	1.5	7	2.6
16	4	1.5	11	4.1
17	7	2.6	18	6.7
18	8	3.0	26	9.6
19	6	2.2	32	11.9
20	12	4.4	44	16.3
21	10	3.7	54	20.0
22	25	9.3	79	29.3
23	27	10.0	106	39.3
24	18	6.7	124	45.9
25	22	8.1	146	54.1
26	27	10.0	173	64.1
27	12	4.4	185	68.5
28	9	3.3	194	71.9
29	18	6.7	212	78.5
30	16	5.9	228	84.4
31	11	4.1	239	88.5
32	11	4.1	250	92.6
33	6	2.2	256	94.8
34	5	1.9	261	96.7
35	3	1.1	264	97.8
36	2	0.7	266	98.5
37	1	0.4	267	98.9
38	2	0.7	269	99.6
39	1	0.4	270	100.0
40	0	0.0	270	100.0



平均　25.21

中央値　25

最頻値　23 と 26

■量的データの分布の記述

分布のどのような様子を表すか

おもに，位置，散らばり(広がり)，左右対称性からの歪み，とんがり度

■代表値

量的データの分布の位置を簡潔に記述する値

平均値(算術平均，mean)

標本平均 $\overline{X}$ ＝(X1 + X2 + …+Xn)／ n

データの重心

外れ値に弱いが便利

統計学上，極めて重要なもの．分散や SD の計算で使われている

中央値(median)

データを大きさ順にならべて中央にくる値．外れ値に強い(頑健)

順番しか見ないので，どの程度外れているか考慮しない

最頻値(mode)

出現頻度が最も多いデータ値．外れ値に強い(頑健)

連続分布における点推定値としては有効

単峰で，

分布が左右対称なら，3 つの値は一致する → 平均値 ＝ 中央値 ＝ 最頻値

左に裾が長い= 値の小さい方のデータが多い → 平均値 ＜ 中央値 ＜ 最頻値

右に裾が長い= 値の大きい方のデータが多い → 平均値 ＞ 中央値 ＞ 最頻値

■データの中心化

中心化データ

各データからその変数の平均値を引いた値

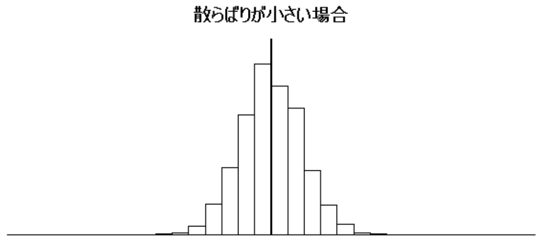
平均からの位置を表す

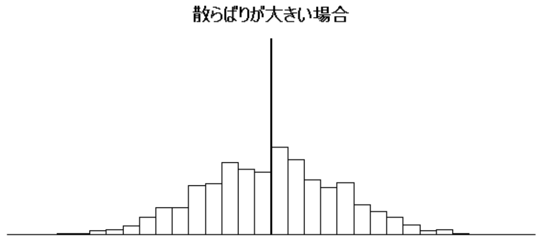
平均より大きいデータの中心化データは＋（平均より小さいデータは？）

中心化データの平均は？

表5.1　学習意欲データ

番号	データ	データ － 平均	(データ － 平均)^2
1	23	23 － 25.21	$(23 - 25.21)^2$
2	29	29 － 25.21	$(29 - 25.21)^2$
3	23	23 － 25.21	$(23 - 25.21)^2$
:	:	:	:
270	29	29 － 25.21	$(29 - 25.21)^2$





■散らばり（散布度）

量的データの分布の広がり（散らばり）具合を表す値

散らばり小さい　→　中心化データがゼロ近辺に集中

散らばり大きい　→　中心化データが正負に大きく散らばる

分散　$= \{(X1 - \overline{X})^2 + (X2 - \overline{X})^2 + \cdots + (Xn - \overline{X})^2\} / (n - 1)$

標準偏差（standard deviation, SD）　$= \sqrt{\text{分散}}$

統計ソフトで計算される分散や標準偏差は，これらの値

■その他の散布度の指標

範囲（range）

R = 最大値- 最小値

外れ値の影響を受けやすい

四分位範囲

データを大きさ順に並べたとき，データの真ん中半数が収まっている幅の大きさ

＝ 第 3 四分位数（75 パーセンタイル，Q3）－　　第 1 四分位数（25 パーセンタイル，Q1）

上位 25% のところに位置するデータの値　　　下位 25% のところに位置するデータの値

■分布の歪み

歪度（skewness）

分布の対称性が正規分布に比べどれくらい離れているかを表す値

負に大きい：左に裾が長い．平均値が中央値よりも左にくる傾向がある．

0：対称性が正規分布と同等．左右対称

正に大きい：右に裾が長い．平均値が中央値よりも右にくる傾向がある．

■分布のとんがり度・裾の重さ

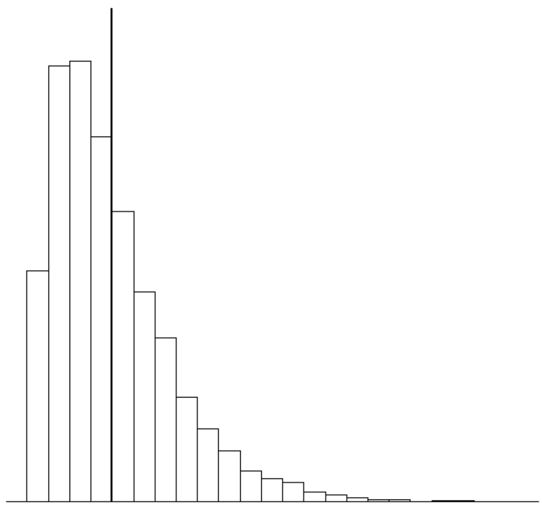
尖度（kurtosis）

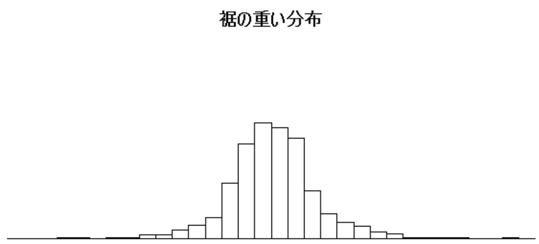
分布の裾の重さが正規分布に比べどれくらい離れているかを表す値

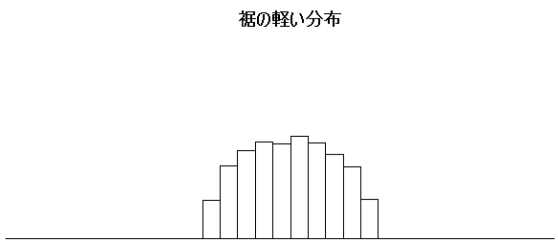
負に大きい：正規分布より裾が軽い．中央付近は正規分布より平べったい．

0：裾の重さが正規分布と同等．中央付近は正規分布と同じ尖んがり．

正に大きい：正規分布よりも裾が重い．中央付近は正規分布より尖っている．







■正規分布

平均値を挟んで分布が左右対称

平均に近いほど発生確率が高く，平均から遠いほど発生確率が小さい　→　単峰

平均± 1SD のところに変曲点(上に凸と下に凸の境目）がある　→　釣り鐘型(ベル型)

平均と標準偏差(または分散）の値で，分布のすべての性質が決まる

標準正規分布：平均 0，標準偏差 1 の正規分布

いろいろな統計分析の基礎となる

■正規分布における平均からの位置と確率

平均± 1SD の範囲に，68.26% のデータが存在する．　片側 34.13%

平均± 2SD の範囲に，95.44% のデータが存在する．　片側 47.72%

平均± 3SD の範囲に，99.72% のデータが存在する．　片側 49.86%

2.5SD 以上離れるデータがあったら，入力ミスや測定ミスをまず疑う

■正規分布を用いた「標準偏差」の理解

T 得点(正規偏差値)：平均 50，標準偏差 10 の正規分布における偏差値

正規偏差値 60 は，平均から 1SD 上で，上位 16% くらい

正規偏差値 40 は，平均から 1SD 下で，下位 16% くらい

IQ(ウェクスラー系知能検査)：　平均 100，標準偏差 15 の分布

IQ(ビネー系知能検査)：　平均 100，標準偏差 16 の分布

IQ 120 は，上位 10% くらい

IQ　80 は，下位 10% くらい

IQ　70 は，平均から 2SD くらい下で，下位 2.5% くらい

これらは，IQ　の高い人は，偏差値が高く，学業成績も良い，ということを意味するもではない

■データの標準化

心理的構成概念の測定値は単位が定まらない

　項目数

　評定段階数

同じ心理的構成概念を捉える尺度を用いれば，項目数や評定段階数によらず，同じ測定になってほしい

データを標準化しておけば，元の尺度の項目数や評定段階数などによらなくなり，測定値は，各個体の位置を表す値になる

■合成変数の平均・分散

変数X，変数Yの値を合計したときの平均・分散・SD

X＋Yの平均　＝　Xの平均　＋　Yの平均

X＋Yの分散　＝　Xの分散　＋　Yの分散　＋　2（XとYの共分散）

X＋YのSD　＝　√(X＋Yの分散)

変数Xの値から，変数Yの値を引いたときの平均・分散・SD

X－Yの平均　＝　Xの平均　－　Yの平均

X－Yの分散　＝　Xの分散　＋　Yの分散　－　2（XとYの共分散）

X－YのSD　＝　√(X－Yの分散)

■共分散

$$\text{XとYの共分散} = \frac{(X - \overline{X})(Y - \overline{Y})\ \text{を全個体について合計したもの}}{\text{個体数} - 1}$$

＝　各変数の平均からの偏差の積の平均（のようなもの）

共分散が＋に大きい

一方が平均よりも大きければ，他方も平均より大きいデータが多い

散布図は右上がり

共分散が－に大きい

一方が平均より大きく，他方が平均より小さいデータ多い

散布図は右下がり

共分散が0に近い

散布図は右上がりでも右下がりでもない

■得点を合計することの合理性

テストや性格検査では，特性の個人差を捉えようとしている

合計点は，個人差をより良く捉えたものでなければならない

共分散が＋なら，個体差はより増強される

項目得点間の関係は，共分散が＋になるような関係でなければならない

逆転項目の得点を逆転して，他の項目と方向を揃えるのはそのため

■相関（統計学における）

2 つの量的変数間の直線的な関係

正の相関：　一方の変数の値が高いほど他方の変数の値も高い傾向．散布図は右上がり

負の相関：　一方の変数の値が高いほど他方の変数の値は低い傾向．散布図は右下がり

無相関：　上記のどちらの傾向でもない場合

■相関係数（ピアソンの積率相関係数）

変数 X と変数 Y の相関係数（２つとも量的変数）

$$r = \frac{\text{X と Y の共分散}}{\text{X の SD} \times \text{Y の SD}} = \frac{sxy}{sx\ \ sy}$$

■相関係数の大きさの目安（心理・教育など社会科学系での）

0.0 ～ 0.2　相関無し（0：無相関）

0.2 ～ 0.3　ほとんど相関無し

0.3 ～ 0.4　弱い相関　…社会科学において相関傾向を認めるのはこの辺から

0.4 ～ 0.6　中程度の相関　…散布図を見て相関ありと思えるのはこの辺から

0.6 ～ 0.8　強い相関

0.8 ～ 1.0　非常に強い相関

完全な相関（± 1）：　散布図が右上（下）がりの直線

相関係数は，相関関係を－ 1 ～＋ 1 の値になる

■合成変数の分散

X＋Y の分散　＝　X の分散　＋　Y の分散　＋　2（X と Y の共分散）

　　　　　　　＝　X の分散　＋　Y の分散　＋　2（X と Y の相関係数）（X の SD）（Y の SD）

2 つの変数の相関が高いほど，和得点の分散は大きくなる

相関の高いデータほど，和得点は個人差が増強される

X－Y の分散　＝　X の分散　＋　Y の分散　－　2（X と Y の共分散）

　　　　　　　＝　X の分散　＋　Y の分散　－　2（X と Y の相関係数）（X の SD）（Y の SD）

2 つの変数の相関が高いほど，差得点の分散は小さくなる

相関の高いデータほど，差得点は個人差を見いだせなくなる

■相関係数に関する注意点等

外れ値に弱い

　極端な値のデータの混入により，相関係数の値が大きく変化してしまう

　大多数の人が「よくあてはまる」と答えるいくつかの項目に対して，ごく一部の人が「まったくあてはまらない」と評定したら(いいかげんな回答も含む)，その一部のデータの影響で，その項目間に大きな相関関係が生じ，解釈を誤る可能性がある．

　スピアマンの順位相関係数　=　データの値を順位に変換して求めたピアソンの積率相関係数

見かけの相関(疑似相関)

　対象としている 2 つの変数以外の変数の影響で相関関係が歪んで見えること

　→　不自然な相関(または無相関) が観察された場合，大発見をしたと思う前に疑似相関を疑う．

ex)　小学生において，足が大きいほど算数がよくできるというのは，発達の影響(共変数)

曲線相関

　相関係数としてはゼロ(無相関) に近い値になるが，変数間に関連がないわけではない

ex) 自己開示度と好感度，開示しないのも，しすぎるのも悪く，ほどよい開示がよい．

　→　区間を区切って分析．曲線のあてはめ．

選抜効果

　入試の成績と入学後の成績の相関が弱いから入試は役に立たないという議論は誤り．

　　不合格者に関する入学後の成績が無く散布図が切断されている．

　大学生に，高校時代のことを振り返って調査しても高校に関する調査にならない．

　　進学率は 5 割程度．進学しなかった 5 割のデータに関する情報がない．

相関関係は因果関係ではない

　因果関係があれば相関関係が観察されるが，相関関係が観察されるからといって因果関係があることにはならない．

　「一方の変数の値が高い人は他方の変数の値も高い」と，

　「一方の変数の値が高まれば，他方の変数の値も高まる」は，同じではない

ex)　本を良く読む親の子どもほど成績がよい，という相関関係があったとしても，

　　親が本をたくさん読んだら子どもの成績が上がるわけではない．

■尺度（テスト）を評価する必要性

心理的構成概念の測定は，通常，間接測定であるから，測定の適切さを確認するために，心理特性を測る尺度(テスト）を評価することが必要

■測定の妥当性

尺度が実際に測定している特性が，実施者が対象としたい特性をどの程度適切に捉えているか

測りたいと思っている特性を，本当に測っていると言えるか？

テスト得点または各項目の得点が，測定したい特性を正しく反映している程度の高さで評価

■妥当性の確認

内容関連妥当性

どの程度適切に作られたか，適切に仕上がったかなどを評価

・論理的妥当性…専門家による評価

・表面的妥当性…一般的なみたくれの良さ

収集したデータ(測定）についての妥当性は示されていない

基準関連妥当性

外的な客観的基準との関連の強さを評価

・予測的妥当性…外的基準が後になって観測される．職業適性検査→営業成績

・併存的妥当性…外的基準が同時に観測される．健康意識度検査→コレステロール値

その基準が妥当なものかという堂々巡りになるという議論も出てくる

構成概念妥当性

対象としたい構成概念に関する解釈を支持する根拠の強さを評価

・収束的妥当性…似たような概念と正の相関がみられる．

・弁別的妥当性…関連のない概念とは相関が見られない．

・因子的妥当性…因子分析をすると，似たような概念どうしの群ができる．

便宜上，歴史上，いろいろな妥当性のラベルがあるが，妥当性の確認は構成概念の確認という作業に終始する

■測定の信頼性

尺度(テスト）が実際に測定している特性をどの程度精度良く測定しているか

誤差の少ない(安定した）測定値が得られるか？

テストの得点または各項目の得点が各受検者において一貫している程度の高さで評価される

信頼性は，尺度が実際に測定している特性をどの程度精度良く測定しているかを考えるものであり，実施者が対象としたい特性を捉えているかという議論は含まない

■信頼性係数

古典的テストモデル

観測得点（データ）　＝　真の得点（潜在変数）　＋　誤差

信頼性係数の定義

誤差が入る分，観測得点（データ）の分散は，真の得点の分散より大きくなるので，観測得点の分散の何割を真の得点の分散が占めているかで，測定の精度（誤差の小ささ）を評価する．

$$\text{信頼性係数} = \frac{\text{真の得点の分散}}{\text{観測得点の分散}}$$

■信頼性係数の推定

真の得点は潜在変数だから，データとして得ることはできず，真の得点の分散は分からない．

そこで，なんとかして，信頼性係数の値を推定する．

再検査法

同じ検査を行ったら，同じような結果になるはず，という意味での信頼性

信頼性係数　＝　同じ集団に，同じ検査を２回行ったときの，相関係数

内適整合法（内部一貫法）

同じ概念を測定している項目には，同じように回答するはず，という意味での信頼性

$$\alpha\text{（アルファ）係数} = \frac{\text{項目数}}{\text{項目数} - 1} \times \left(1 - \frac{\text{各項目の分散の合計}}{\text{合計得点の分散}} \right)$$

■信頼性係数の大きさの評価の経験的な目安

理想は信頼性係数 ＝1

しかし，学力のような特性を考えると，例えば小学校 6 年生の算数でも，計算とか図形とか測定したい特性に，ある程度の広がりがある

計算ができても図形が苦手とか，その逆とか，各受験者において各項目の得点の一貫性を貫く（全部に正答できる，逆に，全部正答できない）ことが，もともと難しくなる．

実際には，信頼性係数の値は 1 より小さくなるのが普通

ある程度の広がりをもった構成概念を測定する場合には，信頼性係数は少し下げざるを得ない

経験的に言われているライン

個人の処遇を決めるテスト 0．9 以上

学力・能力テスト 0．8 以上

性格検査など 0．7 以上

■相関係数の希薄化

信頼性が低いと，観測変数間の相関係数は，本来の相関係数よりも，小さい値になってしまう

尺度（テスト）の信頼性が不揃いだと，本来の相関関係が，歪んで観察されてしまう

ex） 国語，社会，物理 の例

信頼性係数

国語＝ 0.7 社会＝ 0.5 物理＝ 0.9

本来の相関係数

	国語	社会	物理
国語	1	0.8	0.6
社会	0.8	1	0.4
物理	0.6	0.4	1

観測される相関係数

	国語	社会	物理
国語	1	0.47	0.48
社会	0.48	1	0.27
物理	0.47	0.27	1

■妥当性と信頼性の関係

妥当性高い　→　信頼性高い

　測定したいものを適切に測定するためには，対象を精度良く測定していることが必要

　ex)　　優秀なピッチャーは，コントロールが良い．

信頼性が低い　→　妥当性は低い

　精度の低い尺度で測定しても，適切な測定は行えない．

　ex)　　コントロールの悪いピッチャーは，優秀なピッチャーとは言えない．

信頼性が高くても，妥当性は低い場合がある

　ex)　　コントロールの良いピッチャーでも，優秀なサッカー選手とはならない．

信頼性は妥当性の必要条件(一部)

統計分析力尺度

以下の各項目は統計分析力に関する項目です．各項目の内容を読んで，あてはまる程度を，1～5の数値から選んで○をつけて下さい．

	まったくあてはまらない	あまりあてはまらない	どちらともいえない	まああてはまる	よくあてはまる
1） 分析結果全体を説明できるような解釈を導くことができる	1 －	2 －	3 －	4 －	5
2） 分析結果を見ていて思わぬ発見をすることがある	1 －	2 －	3 －	4 －	5
3） 分析結果でよく分からない出力があったら，何を意味するものか調べる	1 －	2 －	3 －	4 －	5
4） 結果から自分の仮説が支持されなかったとき，仮説は誤りだったと素直に認めることができる	1 －	2 －	3 －	4 －	5
5） 研究目的を見失わずに分析を進めることができる	1 －	2 －	3 －	4 －	5
6） 1つの分析法がうまく適用できないとき，他の方法を考えることができる	1 －	2 －	3 －	4 －	5
7） いろいろな分析法を用いたことがある	1 －	2 －	3 －	4 －	5
8） いろいろな分析法を知っている	1 －	2 －	3 －	4 －	5
9） パソコンの扱いは得意なほうだ	1 －	2 －	3 －	4 －	5
10） 数値で考えるのは得意なほうだ	1 －	2 －	3 －	4 －	5

妥当性の検討

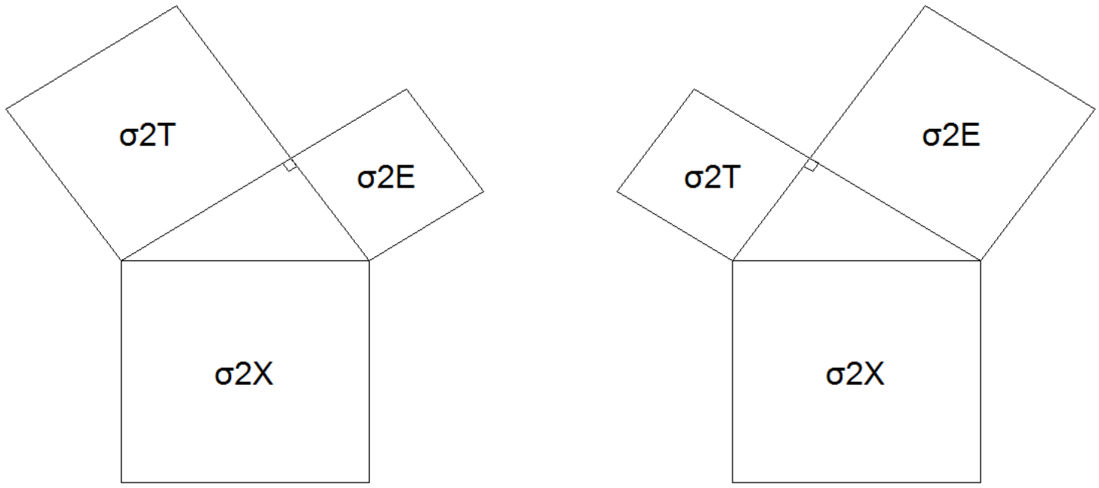
	n	M	SD	min	max	統計分析力との相関係数
統計能力テスト	365	64.83	12.47	25	99	.51
数学	365	54.17	16.07	0	98	.33
批判的思考力	365	30.11	6.14	13	47	.39
国語(現代文)	365	65.09	9.28	42	89	.08
自己効力感	365	51.99	9.76	26	78	.13

信頼性係数の推定

	n	M	SD	min	max	相関係数(n=156)		α係数
1回目	365	30.68	6.08	13	46	1	.82	.84
2回目	156	29.79	5.90	15	42	.82	1	.83

観測得点の分散（σ_X^2），真の得点の分散（σ_T^2），誤差分散（σ_E^2）の関係

誤差分散が小さい場合　　　　誤差分散が大きい場合



信頼性と妥当性の関係

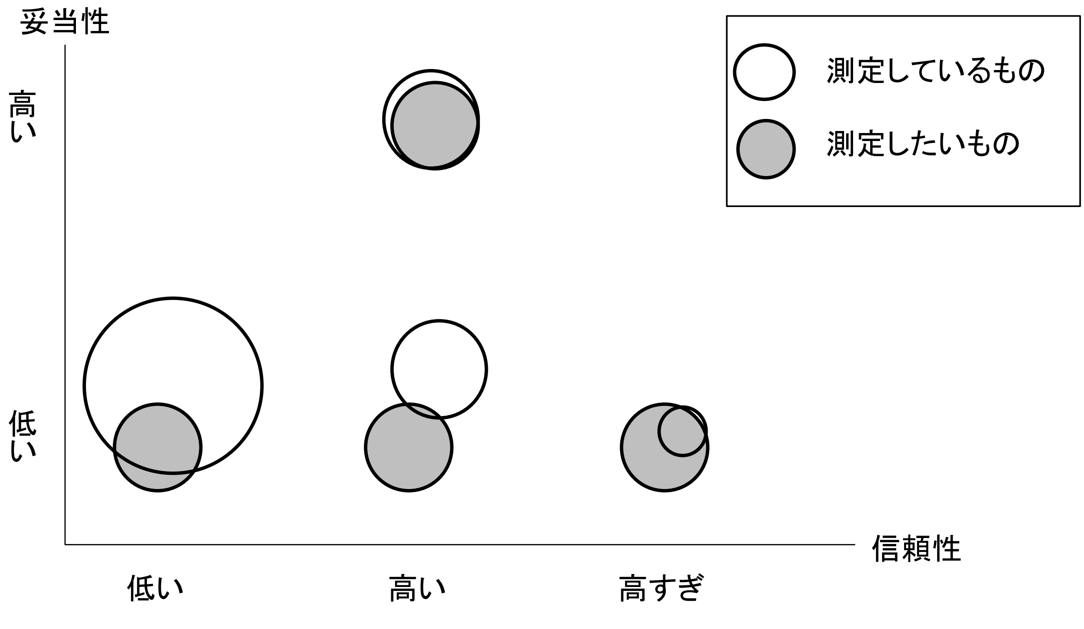


図7.3　信頼性と妥当性の関係
信頼性：○の大きさ
妥当性：○と●の重なり