

# アンケートの作成と分析の基本 PART2

## —もう悩まないアンケート調査—

帝京平成大学ヒューマンケア学部看護学科

北川 明

### 本日も話せる内容

- ◆ 昨年度セミナーの振り返り
  - ✓ アンケート調査の基本
  - ✓ アンケート作成上の注意
- ◆ データ処理
  - ✓ 単純集計とクロス集計
- ◆ 分析方法
  - ✓ 各種検定について
  - ✓ パラメトリック検定とノンパラメトリック検定
  - ✓ 有意でないとは？

2

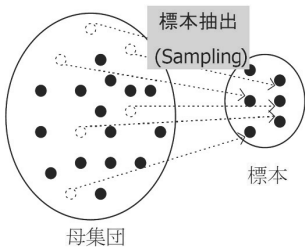
### まず、はじめに

## なぜ、アンケート調査を行うのでしょうか？

3

### それは・・・

#### 母集団のことを知りたいから



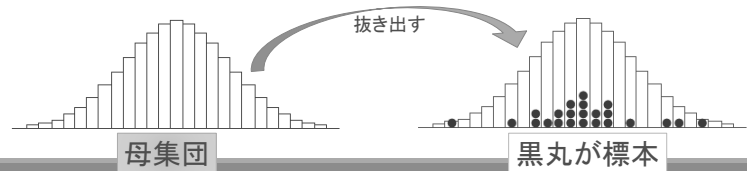
**母集団とは・・・**  
われわれが知識・情報を得たいと考えている対象の全体のこと

**標本とは・・・**  
母集団から抽出された一部分のこと

4

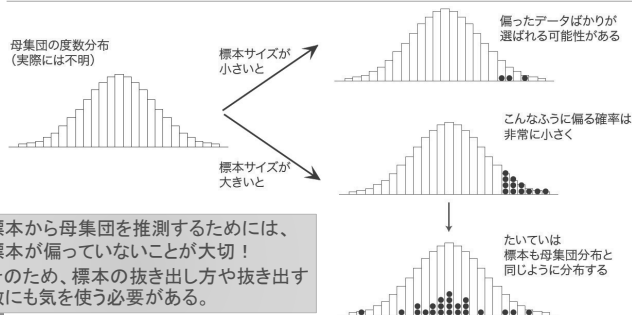
### 標本を調べると母集団が推測できる

ランダムに抜き出すと  
当然、度数が多いところほど、選ばれる可能性が高い



5

### 抜き出し方(標本抽出方法)や抜き出す数(標本サイズ)は大切



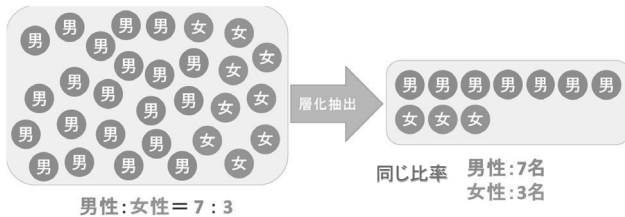
- ✓ 標本から母集団を推測するためには、標本が偏っていないことが大切！
- ✓ そのため、標本の抜き出し方や抜き出す数にも気をを使う必要がある。

### 様々な標本抽出法①

<b>単純無作為抽出法</b>	母集団に通し番号をつけ、乱数表などを用いて必要な標本の数だけ抽出する方法 ランダム性が保証される最も基本的な抽出法 ※母集団全員の名簿が必要なため実際に実施するのは難しい。
<b>系統抽出法</b>	 母集団に通し番号をつけ、はじめの1人だけを無作為に選び、2人目以降は必要な標本数を勘案し一定の間隔で抽出する方法 ※並び方に法則性があるとランダムではなくなる。

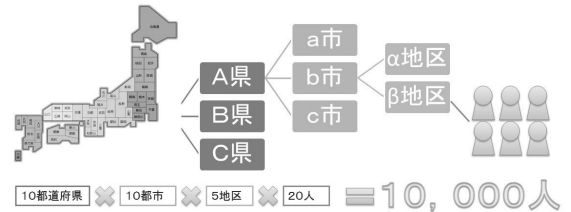
## 様々な標本抽出法②

**層化抽出法** 母集団をあらかじめいくつかの層(グループ)に分けておき、各層の中から必要な数の調査対象を無作為に抽出する方法  
※母集団がどのような層の構成になっているか事前に必要



## 様々な標本抽出法③

**多段抽出法** 母集団をいくつかのグループに分け、そこから無作為抽出でいくつかグループを選び、さらにその中から無作為抽出する...という操作を繰り返して、最終的に選ばれたグループの中から調査対象を無作為抽出する方法



## 標本サイズの決め方

参考

アンケート調査などで、母集団を推測するためには、一定数の標本がなければ、推測が大雑把になってしまう。

標本サイズがあまりに少ないと、推測の幅が広くなり、あまり役に立たない場合がある。

しかし、標本サイズを膨大な数にすると、調査コストが多くなってしまふ。

どのような分析をやりたいかによって、標本サイズを計算する必要がある。

10

## アンケートの基本

11

## アンケート調査とは

1. 調査対象(例えば、一型糖尿病の患者、1年目の看護師など)の意識や行動などを把握するために(目的設定)
2. 一定のルールで調査の対象を選び(対象者の定義)
3. 様々な調査方法で、多数の人に回答を求め(調査方法)
4. 特定の期間内で(調査期間)
5. 様式化した質問への回答をもとに(アンケート作成)
6. 統計的処理を行うもの(分析)

12

## 質問票の項目と構成①

### 1. タイトル

- ・内容を簡潔に表すものにする。(例「大学生のアルバイトの実態に関する調査」)
- ・言葉づかいに気をつける。

### 2. 企画者・実施者名

- 例:「帝京平成大学 北川 明」
- ・企画した人(機関)と調査を実施する人(機関)が異なる場合、併記。
  - ・相手が信用するための重要な要素のひとつ

### 3. あいさつ文

- 例:「私は帝京平成大学で精神看護学を専門とする...」
- ・この書き方で、回収率が大きく異なるほど重要なもの
  - ・誰が、どんな目的で、どの抽出方法で、結果はどう扱うのか、など

13

## 質問票の項目と構成②

### 4.記入上の注意事項

- 例:「回答は、該当する番号に○をつける...」
- ・わかりやすく、目立つように工夫する。

### 5.回収方法・回収日時・回収者

- 例:「この調査票は7月10日に、北川明が回収に伺います。」
- 例:「記入の終わった調査票は、同封の封筒に入れ(切手は不要です)、7月10日までに投函してください」

### 6.連絡先

- 例:「このアンケートについてのお問い合わせは...」
- 必ず連絡先を明記することが必要

14

## 質問票の項目と構成③

### 7.倫理的配慮

- ・どのようなアンケート調査も倫理的配慮がないと実施してはいけません！
- 例:「アンケートの参加は自由であり、断っても不利益になることはありません...」

### 8.アンケート本文

- ・質問番号は、問1、問2...Q1、Q2、などとする。
- ・質問が分岐する場合は、問1-1などとする。

### 9.フェイス・属性(後述)

- ・答えづらい質問は避ける。(例:年収など)

### 10.お礼

- 例:「以上で終わりです。長い間ご協力ありがとうございました」

15

## 倫理的配慮は必ず記載！

看護研究では多くの場合、人間を対象としているばかりでなく、立場の弱いケアの受け手からデータを多く得ることが多い

- 研究への参加は強制になっていないか
- 研究への参加同意はどのようにして得ているか
- 対象者は研究に参加拒否しても不利益はないか
- 研究への参加により、対象者にどのような問題が発生する可能性があるか
- 対象者はその問題について理解して参加しているか
- 問題を回避もしくは解決するためにはどのようにすればよいか
- その問題は倫理的に承認される範囲のものであるか

16

## アンケートの質問文は 回答者の負担を考える

17

## アンケート作成上の注意①

### 誰が読んでも分かる言葉を使う！

- ✓ 調査対象者のすべてが理解・回答できる言葉を使う  
調査対象者のすべてが言葉の意味を理解できないと調査は成立しない。特に、以下の言葉は注意が必要
  - ・ 専門用語(例えば、コンピュータ関係など)
  - ・ 流行語(特にファッション関係など)
  - ・ カタカナ言葉(システムエンジニア、イノベーション、MS-Wordなど)
- ✓ 多様に解釈できる言葉は避ける。あいまいな言葉は使わない  
ひとつの言葉でいくつもの解釈が出来る言葉は避けるか限定する
  - ・ お店はどつでしたか？(良かった やや良かった...) 何について聴いているのか不明

18

## アンケート作成上の注意②

### 選択肢のバランスと強い修飾語

「1.とても満足 2.やや満足 3.やや不満 4.とても不満」

- 選択肢間の間隔は一定とみなし、間隔尺度として、平均値等を分析に用いることがある。
- そのような場合、間隔が一定と思えるような選択肢にする必要がある。

○1.とても満足 2.やや満足 3.やや不満 4.とても不満 バランス尺度

×1.とても満足 2.やや満足 3.どちらかといえば満足 4.不満 アンバランス尺度

強い修飾語である、「非常に」「全く」などの語を使うと、その選択肢は選ばれなくなる傾向がある。

19

### アンケート作成上の注意③

#### 個人的なことをききすぎる

- かなり個人的なことを具体的に訊く

年収	万円
----	----

- ✓ まずは、本当にその質問が必要かを検討することが大切  
どうしても訊く必要がある場合は以下のようにすると回答者に親切

Q8. あなたの年収をお教えてください。(〇は1つだけ)

- |                    |                     |
|--------------------|---------------------|
| 1. 400万円未満         | 2. 400万円以上 600万円未満  |
| 3. 600万円以上 800万円未満 | 4. 800万円以上 1000万円未満 |
| 5. 1000万円以上        |                     |

20

### アンケート作成上の注意④

#### 質問文に二つの意味が含まれている

- 2つ以上の意味を有している

Q 看護師の身だしなみや対応はどうでしたか？

- |       |       |           |
|-------|-------|-----------|
| とても悪い | やや悪い  | どちらともいえない |
| やや良い  | とても良い |           |

- ✓ この訊き方の場合、対応は良いが身だしなみが悪いと感じた回答者がいたときに不都合が生じるため、質問は分けたほうが良い。

21

### アンケート作成上の注意⑤

#### 回答者に負担がかかる

- 順位をつけさせる

Q 以下の4種類の紅茶について、好きなものの順に「1位」から「4位」まで順位をつけてください。

- |            |   |        |   |
|------------|---|--------|---|
| 1. ストレートティ | → | 【    】 | 位 |
| 2. レモンティ   | → | 【    】 | 位 |
| 3. ミルクティ   | → | 【    】 | 位 |
| 4. アップルティ  | → | 【    】 | 位 |

個人の傾向が知りたいのではなく、全体の傾向が知りたいのであれば、好みのものを単一選択や複数選択してもらうだけで良い。

- ✓ 順位をつけさせる質問も分析手法によっては用いることがあるが、同率1位が存在する場合などは、回答者が困る。

22

### アンケート作成上の注意⑥

#### 回答の誘導

- 回答を誘導している。

Q. 日本は資源が少ないため、科学技術に関する教育は21世紀においてますます重要だといわれています。さて、あなたは、中学校の今後の理科教育をどう考えますか。(〇は1つだけ)

- |                  |             |
|------------------|-------------|
| 1. もっと充実させたほうがよい | 2. 従来どおりでよい |
|------------------|-------------|

- ✓ 多くの人が「1.もっと充実させたほうがよい」と回答せざるを得ない心境になる

23

### アンケート作成上の注意⑦

#### 無駄な段階

- 段階が多すぎる

Q. あなたは、一人暮らしのための住所を選ぶ際に、以下のことをどれくらい重視しますか。(〇は行ごとに1つだけ)

	極めて重視しない	非常に重視しない	重視しない	やや重視しない	どちらともいえない	やや重視する	重視する	非常に重視する	極めて重視する
a. 最寄りの駅からの距離	1	2	3	4	5	6	7	8	9
b. 日当たり	1	2	3	4	5	6	7	8	9
c. 収納スペース	1	2	3	4	5	6	7	8	9
d. 家賃	1	2	3	4	5	6	7	8	9

- ✓ 厳密に何段階との決まりはないが、最大でも7段階くらいが実用的であろう。

24

### アンケート作成上の注意⑧

#### 自由回答が多すぎる

- 自由回答を募る質問

Q7. 当院へのご意見ご要望などありましたら、ご自由にご記入ください。

- ✓ 自由回答は、調査者からすると、予想しなかった興味深い回答を得られる可能性がある。しかしながら、そのデータを入力し、分析することを考えたとき、多種多様な回答が得られるほど負担が増す。量的調査は、母集団の特性を知ることが目的であるため、自由回答は本当に必要か再考してみるとよい。

25

## アンケート作成のまとめ

1. 個人情報や順位付けなど、対象者の負担を考えて検討する。
2. 意味がとりにくい文章にしない(解釈の幅が広がるので、回答が絞れない)
  - 否定疑問文にしない(「...ませんか?」など)
3. ひとつの質問では、必ずひとつのことを聞く(2つ入っていること:ダブルバーレル)
4. 誘導的な質問にしない。
5. 自由回答は分析に負担がかかることを覚えておく。

26

# データ処理方法

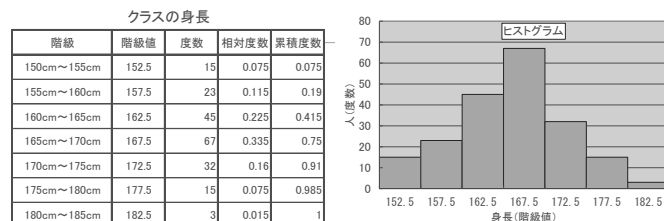
27

## 回収したらチェックする

- 調査票に通し番号をふり、記入ミス点を点検する。
  - ✓ 誤記入、順位の重複などをチェックし、ルールを決めて修正する。
  - ✓ 回答もれや無回答が多すぎる場合は無効
  - ✓ 同じ選択肢しか選んでいない、賛否の回答パターンがすべて同じものは無効
- 度数分布表の作成
  - ✓ 「1」、「2」、「3」など、取り得る値が決まっているカテゴリデータの場合は、度数分布表を作成し、本来存在しない値が入力されていないかどうか確認する。
- クロス集計表の作成
  - ✓ 変数と変数との間の整合性を確認する場合はクロス集計表を作成する。
  - ✓ 例:女性(未婚)なのに、専業主婦 等

28

## 度数分布表とヒストグラム



- ある集団のデータ(観測値)をいくつかの階級に分け、それぞれの階級でデータがいくつあったか(度数)を数え、表にしたものを度数分布表と言う。
- また、度数分布表を棒グラフに表したものをヒストグラムという。
- 相対度数は、全度数に対するある階級の度数の割合を表わし、累積度数はその階級以下の度数の合計を示している。

## 単純集計とクロス集計例

あなたは朝食を食べますか？	
食べる	42
食べない	58
計	100
単純集計	
あなたのBMIは25以上ですか？	
25以上	16
25未満	84
計	100

クロス集計		BMI		総計
		25以上	25未満	
朝食	食べない	15	43	58
	食べる	1	41	42
総計		16	84	100

属性やグループ別で集計することがクロス集計

29

## クロス集計のポイント

- クロス集計表の左側を表側、上側を表頭と呼び、表側には、結果に影響を及ぼす要因を配置し、表頭に結果を配置する。
  - ✓ クロス集計の際は、1グループ最低30サンプルは必要。
  - ✓ 性別と年齢を10歳代~60歳以上の6区分で分析する場合、最低360人は必要であり、年齢構成比が少ない層を考慮すると500人程度は必要と考えられる。
- クロス集計をもとに検定( $\chi^2$ 乗検定)を行う場合は、1グループ50サンプル程度あると良い。

31

## 天井効果とフロア効果

### ◆ アンケートの回答が、どちらかに偏ること

- ✓ 天井効果は、平均値+標準偏差
- ✓ フロア効果は、平均値-標準偏差

得点分布が高い方(低い方に)歪んでいないかを見る。

5件法で、値5以上だと天井効果が、値1以下だとフロア効果があるといえる。

質問文が悪いか、選択肢が少なすぎるのが考えられる。

32

# データ分析方法

33

## 研究によって得られるデータと その分析方法は予め想定しておく！

研究タイプ	分析によって知りたいこと	主な統計手法
因子探索研究	現象や特徴	統計処理は使わない
実態調査研究	実態はどうか	平均値と標準偏差 度数 など(記述統計)
関係探索研究	何が関係していたか	平均値と標準偏差(記述統計) 相関係数の検定、カイ二乗検定 など(推計統計)
比較研究	どこが違うか	比較する群それぞれの平均値と 標準偏差(記述統計)
実験研究	効果があったか	差の検定(推計統計)

34

## 記述統計と推計統計

- ◆ 得られたデータからその特徴を抜き出すものを記述統計
  - 記述統計とは、収集されたデータをうまく整理して、集団の特徴を適切な統計量(平均や標準偏差など)で表現しようというもの。
- ◆ 得られたデータから全体や未来を推測するものを推計統計
  - 記述統計学の手法と確率理論をミックスして、「全体を把握しきれないほど大きな対象」や「まだ起きておらず未来に起きること」に関する推測を行うもの(検定など)。

アンケートの分析は、まずは記述統計で、データの特徴を記述することから始める。

35

## アンケートの分析

- アンケートを分析する時は、まずは単純集計を行う。
- その後、代表値やばらつきといった分布を確認していく。
- 次に、様々な変数間の関係をクロス集計などで見ていくことになる。
- 2つの群の関連を推定したいのであれば、相関係数や $\chi^2$ 乗検定により関連があるかを確認する。
- 群間の差を見たいのであれば、t検定や一元配置分散分析を行う。
- どのようなことを見たいかにより、行う分析は決まっている。

36

## 検定早見表(平均値と比率の差)

検定目的	検定方法
男女間で、検査値に差があるかどうかなど、平均値の差を検定する。	t検定
対応のない2標本の代表値の差	U検定
指導を受ける前と後で、検査値に差があるかどうかなど、対応があるものの差を検定する。	マニウエル検定
対応のある2標本の代表値の差	付順位検定
若年者、壮年者、高齢者で、血圧の平均値に差があるかどうかなど、3群以上の差を検定する。	ANOVA検定
対応のない3標本以上の代表値の差	検定
内閣支持率が男女間で差があるかどうかなど、比率の差を検定する。	検定
対応のない比率の差	検定
指導を受ける前と後で、合格率に差があるかどうかなど、対応があるものの比率の差を検定する。	検定
対応のある比率の差	検定
体重と心拍出量など、要因間に関連があるかどうかを検定する。	検定
独立性	フィッシャーの正確確率検定

### 対応のない2標本の代表値の差 実例 (t検定)

例: 40歳代男性における飲酒習慣がある者とないない者でアンケート尺度得点に差があるかを知りたい。

例: アンケートに回答した男女間で年齢に差があるか知りたい。

例: 尺度得点の高い群と低い群で、平均血圧に差があるかどうか知りたい。

2つのデータが正規分布に従うならば、t検定を使用する。

飲酒者	非飲酒者
124	97
154	130
119	74
146	69
155	137
93	112
133	78
101	87
144	126
150	107
96	63
111	94
80	70
129	102
138	78

### 対応のない2標本の代表値の差 実例 (マン・ホイットニーのU検定)

例: 20歳代と30歳代の男性の飲酒傾向を調べた。年代よって飲酒傾向に差があるかを知りたい。

	飲まない	時々飲む	よく飲む	毎日飲む
20歳代	12	20	12	5
30歳代	6	10	21	10

例: 男女間で、5件法や4件法の回答結果の差を知りたい。

2つのデータが正規分布に従わない(順序データも)場合は、マン・ホイットニーのU検定を使用する。

### 対応のある2標本の代表値の差 実例 (t検定 ウィルコクソンの符号付順位検定)

例: 糖尿病患者に1年間外来受診時に看護師の面談支援を実施した。介入前後でHbA1c値に差があるかを知りたい。

例: 学生の1年後の尺度得点の差を知りたい。

2つのデータが、同一個体からのデータであり、正規分布に従うならば、t検定を使用する。

同一個体であるが、正規分布に従わないのであれば、ウィルコクソンの符号付順位検定を使用する。

介入前	介入後
9.1	8.3
8.4	8.2
8.8	7.8
9.7	9.2
8.7	8.6
9.6	8.8
9.6	8.2
8.9	8.6
8.9	8.6
9.0	7.6
8.3	8.8
8.7	9.1
8.0	8.0
8.8	8.2
9.4	7.4
9.5	8.4
7.9	7.2

### 対応のない3標本以上の代表値の差 実例 (一元配置分散分析 クラスカル・ウォリス検定)

例: 看護師経験の差によって、職場満足度尺度得点に差があるかどうかを知りたい。

例: 内科、外科、精神科の病棟間で、1ヶ月の看護師の残業時間に差があるか知りたい。

3つのデータが正規分布に従い、等分散と見なせるならば、一元配置分散分析を使用する。

これらの条件が満たせない場合は、クラスカル・ウォリス検定を使用する。

新人	中堅	熟練
74	59	103
40	54	92
62	83	96
113	95	85
85	81	76
81	116	94
91	111	88
81	73	92
72	77	87
106	87	93
107	95	84
93	87	75
52	108	85
93	125	122
113	92	99
90		103
68		

### 対応のある比率の差 実例 (マクネマー検定)

例: 看護師の接遇教育の前後で、患者に対して接遇の満足度を調査したところ、以下の表ようになった。接遇により満足度は変化したと言えるか知りたい。

同一個体の比率の差を比較したい時にマクネマー検定を使用する。

		教育後		合計
		満足	不満足	
教育前	満足	16	5	21
	不満足	18	15	33
合計		34	20	54

### 独立性 実例 (相関係数の検定)

例: A病棟の20歳代の女性のデータで、「体重」と「総コレステロール」との間に相関関係があるかを知りたい。

2つのデータが正規分布に従うと見なせるならば、ピアソンの積率相関係数の算出とその検定を行う。

2つのデータが正規分布に従わない(順序尺度も)ときは、スピアマンの順位相関係数の算出とその検定を行う。

体重	TC
51.9	188
47.7	169
45.3	177
57.1	181
57.9	189
46.3	178
49.9	199
41.4	172
47.3	167
42.7	176
60.0	160
48.8	186
51.9	180
51.2	200
53.3	176
48.2	166

## 独立性 実例 ( $\chi^2$ 検定)

例: 胃腸炎異常発生の喫食調査を行った。ケーキと食中毒症状とで以下の関係があった。ケーキと食中毒との間に関連があるか知りたい。

2群の関連性や比率の差を比較したいときにカイ2乗検定を使用する。

		胃腸炎		合計
		発症あり	発症なし	
ケーキ	食べた	27	13	40
	食べなかった	19	16	35
合計		46	29	75

44

## パラメトリック検定と ノンパラメトリック検定の使い分け

- サンプルサイズが30を超えている時かつ母集団が連続データであれば、パラメトリック検定が使用できると考えて良い。
- サンプルサイズが30以下の場合は正規性の検定を行う。
- 母集団が順序尺度(尿検査など)、名義尺度(性別など)の場合はノンパラメトリック検定を使用する。
- 比較をする場合は、比較集団が等分散(後で説明)でないときはノンパラメトリック検定を使用する。

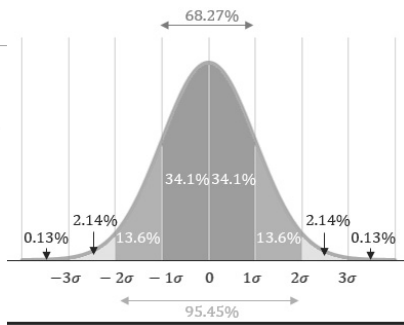
**ノンパラメトリック検定は有意になりにくい傾向あり!**

## 正規性って?

名前の所以は、自然界や人間の行動・性質など様々な現象に対して、よく当てはまるところから来ている。

左右対称で、平均が最も高い分布である。

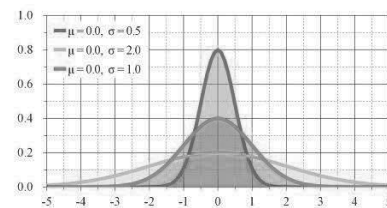
データがこの分布に従うことを正規性という。



46

## 等分散性って?

等分散性とは、読んで字のごとく、分散が等しいとみなせることを意味している。



- ✓ 分布にはばらつきというものがある。
- ✓ データの広がり方を表す指標を分散と呼ぶ。
- ✓ 青と緑の分布では緑の方が分散は大きい。

## 有意差なし=“差がない”ではない!

平均値の差の検定を行った結果、有意ではなかったとする。

▶ 有意であるとは、“今回の調査で出た平均値の差は偶然起こるには確率が低すぎるから、差があると考えよう”ということを意味している。

▶ よって、有意ではないとは、“今回の調査で出た平均値の差は偶然出た可能性が考えられるから差があるとは言えない”という意味になるので、差がないことを意味しない。

## アンケート調査をするときにあると良い文献

1. 酒井隆. 図解 アンケート調査と統計解析がわかる本. 2012. 日本能率協会マネジメントセンター. ISBN978-4820747529.
2. 浅井隆. いまさら誰にも聞けない医学統計の基礎のキソ1~3. 2010. アトムス.
3. 富田真佐子. パソコンで進める やさしい看護研究(第2版). 2014. オーム社. ISBN978-4274215186.
4. 対馬栄輝, 石田水里. 医療系データのとり方・まとめ方—SPSSで学ぶ実験計画と分散分析. 2013. 東京図書. ISBN978-4489021435.
5. 小塩 真司編, 西口 利文編. 質問紙調査の手順(心理学基礎演習). 2007. ナカニシヤ出版. ISBN978-4779502002

48