

現代日本学演習Ⅰ 2019年度(5セメスタ) 東北大学文学部

質問紙調査の基礎

田中重人 (東北大学文学部准教授)

3年生対象：2019年度5セメスタ<金4>605演習室(文学部棟6F)

※ 教科書は第3版(2017)を指定していますが、第1版(2010)あるいは第2版(2013)でもかまいません。

1 授業の概要

講義日程は別紙を参照

講義と並行して、受講生各自の興味に基づいてそれぞれが調査をおこなう。

- 5月中旬に調査企画書を提出し、企画内容について検討
- 作成した調査票について、授業中に検討する機会を持つ(6月上旬の予定)
- 6月中旬から7月にかけて各自が調査をおこなう
- 調査結果を口頭報告する(7月上旬の予定)
- レポートにまとめて提出(8/14締切)
- レポート講評と返却(9月5日以降)

2 次回までの宿題

教科書の第1~4章の「基礎」の部分を読んでおくこと

3 連絡先

田中重人 (東北大学文学部現代日本学研究室)

〒: 980-8576 仙台市青葉区川内27-1 文学部棟6F

E-mail: tanakas2013@tsigeto.info

Homepage: <http://tsigeto.info/welcomej.html>

Blog: <http://b.tsigeto.info/school>

オフィス・アワーは定めていない。質問等がある場合は、あらかじめ適当な時間に予約をとること。

現代日本学演習 I 「質問紙調査の基礎」

第1講 調査課題の設定

田中重人 (東北大学文学部准教授)

[テーマ] 理論とデータの対応

1 課題

配布した調査票を見て、理論仮説と作業 (操作) 仮説との対応を考える。

- 用紙の左半分に理論仮説、右半分に作業仮説を書く。複数取り上げた場合は、対応関係がわかるよう、番号をつけるか線で結ぶなどすること。
- 作業仮説については、どの項目についてどのような分析結果が出ればよいかかわかるように書くこと
- 教科書のほか、何を調べてもよい。何を調べたかは、用紙下部に記入する。
- 周囲の人に見せて意見をもらい、用紙下部の該当欄に記入すること

2 理論とデータ

頭の中で考えたことを理論 (theory) という。実証研究においては、理論的なレベルで検証すべき「仮説」をはっきりさせたうえで、それに対応したデータを集めるよう調査方法等を工夫する。

仮説 (hypothesis): 研究によって真偽を決定すべき命題

理論仮説 (theoretical —): 理論的なレベルで考えた仮説

作業仮説 (working —): データによって検証可能な仮説 (操作仮説 operational — ともいう)

理論仮説を作業仮説に翻訳する作業を「操作化」(operationalization) という。

- 理論は直接検証できない
- 検証できるのは作業仮説のみ → 「もしその理論が正しいとしたら、●●の現象が観察されるはずである」
- 作業仮説が棄却された場合 → 理論仮説か操作化のどちらかがまちがっている
- 作業仮説が支持された場合、理論仮説や操作化が正しいということが主張できるわけではない
- 科学的研究における「通説」(accepted theory) は、たくさんの研究の結果、その理論仮説を否定することができず、かつ、対立する理論仮説がすべて棄却されていることをもって成立する

理論仮説と作業仮説は必ずしも2つに区分されるものではなく、抽象的なレベルの仮説から具体的なレベルの仮説にいたる連続的・多段階のものと考えたほうがよい。

また、理論仮説と作業仮説は1対1で対応しているものではない。おなじ理論仮説を検証するために複数の作業仮説がありうる。逆に、ひとつの作業仮説に複数の理論仮説が対応することもある。

いずれにせよ、理論仮説 \leftrightarrow 作業仮説の双方を往復する想像力が重要。

3 次回予定

来週 (4/26) は休講。その代り、各自の調査課題について面談をおこなう (日時は今日決定)。自分が興味のある分野について、どのようなことを調査したいか、またそのためには誰を対象にどのような質問項目が必要かを考えてくる。

第2講 既存調査と先行研究の探索

田中重人 (東北大学文学部准教授)

[テーマ] 既存調査を探す

1 課題

各自の関心にしたがって、参考になりそうな既存調査をさがす。結果をまとめて印刷して提出。

- 分量は A4 用紙 1 枚
- 用紙上部に学生番号と氏名を書く

つぎのデータベースをひとつおとり使ってみる。各データベースの性格や収録されているデータを把握すること。必要に応じて、他のサービスも使ってよい。

- SSJ データアーカイブ (東京大学) <<http://csrda.iss.u-tokyo.ac.jp/access/choice/>>
- 科学研究費補助金データベース (国立情報学研究所) <<http://kaken.nii.ac.jp>>
- 人間文化研究機構 統合検索システム <<http://int.nihu.jp>>
- Google Scholar <<http://scholar.google.co.jp>>
- Google <<http://www.google.co.jp>>

見つかった調査について、つぎのような項目を書く。不明の場合は「不明」などと書いておく。

- 調査主体
- 実施期間
- 名称
- 調査法
- 調査対象
- 報告書などの書誌情報
- どうやって探したか
- 自分の研究にとってどのような点で参考になりそうか

2 既存調査を特定するのに必要な情報

- 調査主体・連絡先
- 実施期間
- 名称 (対象者向け／研究者向け)
- 調査法
- 対象 (母集団・標本)
- 報告書・論文・Web サイトなど
- データ・アーカイブなどの登録情報

3 既存調査の探索

3.1 探す対象

- 先行研究
- 調査
- 質問文案・回答選択肢

3.2 探しかた

- 人に聞く
- 入門書・概説書・展望論文
- 芋づる式
- 白書、データブック
- 文献データベース
- 調査データベース、データ・アーカイブ
- 尺度集

いずれの場合も、情報の収集範囲と方法、収録基準を理解して利用すること。

調査に関する資料(質問紙など)は、非売品の報告書に載せるのがふつう。論文や本を出版する段階では省略されることが多い。このため、自分の興味にあった質問文案を探すことは非常にむずかしい。

自分の研究分野については、代表的な調査をおさえておくこと。たとえば、

あたらしい言語表現の使用実態: 文化庁「国語に関する世論調査」

日本人の社会意識: 統計数理研究所「日本人の国民性調査」

また、一度の探索で網羅的に情報が集められるわけではないので、ふだんからアンテナを立てておくことが大切である。

3.3 先行研究(論文・書籍)のデータベース

調査を利用した研究成果は、最終的には論文や書籍になるので、それらのデータベースから情報をたどれることが多い

- 日本語教育学研究室のデータベースリスト <<http://www2.sal.tohoku.ac.jp/nik/student/litsurv.html>>
- 国立国会図書館サーチ <<http://iss.ndl.go.jp>>
- CiNii Article <<http://ci.nii.ac.jp>>
- CiNii Books <<http://ci.nii.ac.jp/books/>>

報告書は図書館では書籍とおなじあつかいになっている。また、調査データを利用した論文には、調査の概要の説明があるのがふつう。

「報告書」「調査報告」などをキーワードにふくめて探すとよい。

最近では各出版社の電子ジャーナルや大学などの「機関レポジトリ」(repository)の整備が急速に進み、全文をオンラインで読んだり検索したりできる文献が増えてきている。

3.4 質問文・尺度のコレクション

特定のテーマで既存の質問文をリストした本やデータベースもある。

- 『社会調査ハンドブック』(安田・原, 1982) など
- 『心理測定尺度集』(吉田編, 2001) など
- PSDB_Mie: 三重大学 心理尺度(質問項目) データベース <http://www.minamis.net/scale_search/mpsbmain.html>

3.5 調査のデータベース

調査そのものについてのデータベースはほとんど整備されていない。ただ、公開されている調査データを集めて2次利用のための便宜をはかる「データ・アーカイブ」(data archive)が最近発達しており、事実上、調査データベースとしてつかえる。

- SSJ データアーカイブ (東京大学) <<http://csrda.iss.u-tokyo.ac.jp>>
- 質問紙法に基づく社会調査データベース (大阪大学) <<http://srdq.hus.osaka-u.ac.jp>>

3.6 研究課題・研究者データベース

多くの調査研究は科学研究費補助金(文部科学省または日本学術振興会)の助成を受けておこなわれているので、その研究課題のデータベース中に調査の情報がかなりある。

- 科学研究費補助金データベース(国立情報学研究所) <<http://kaken.nii.ac.jp>>

また、大学などでは所属する研究者(教員・研究員・博士課程学生などをふくむ)の研究成果の情報を収集している。これを集積したデータベースが公開されており、そこから各研究者がおこなった調査の情報を得ることができる。

- 科学技術総合リンクセンター J-Global (科学技術振興機構) <<http://jglobal.jst.go.jp>>
- Researchmap (国立情報学研究所) <<http://researchmap.jp>>

3.7 分野別研究資源情報

- 国内言語学関連研究機関 WWW ページリスト (言語学研究室・後藤斉教授) <<http://www2.sal.tohoku.ac.jp/~gothit/kanren.html>>
- 国立国語研究所 データベース・リスト <<http://www.ninjal.ac.jp/database/>>

3.8 インターネット検索エンジン

インターネットで公開されている情報は、サーチ・エンジンでヒットするはずである。「報告書」「調査報告」「調査票」「質問紙」などをキーワードにふくめて探すといふ。ノイズが非常に多いので必要な情報をふるいわけするにはかなりの熟練を要する。

Google Scholar <<http://scholar.google.co.jp>> は学術情報に特化した検索サービスである。非学術的な情報がカットされているのでその分ノイズがすくない。

4 本人への問い合わせ

調査主体はわかったが詳細な情報が公開されていない場合は、本人に問い合わせるとよい。調査報告書などには連絡先が通常書いてあるし、雑誌論文にも著者所属やメールアドレスなどが書いてあることが多い。また上記の J-Global などでも連絡先を調べることができる。ただし、問い合わせの前に、公開されている情報をできる限り集めてから。

5 宿題

各自の関心にしたがって先行研究あるいは調査を探す。なるべく報告書や調査票など、その調査自体の内容について詳しいことがわかる資料を探すこと。

集めた情報のうち少なくともひとつについて、つぎのことをまとめて提出

- 調査を特定するための情報
- その情報源についての書誌情報
- どういう点が自分の研究に役立ちそうか
- 調べたプロセスと、苦労した点

また、次回の授業で各自の調査企画について考えるので、どのような調査をしたいか考えておくこと。

文献

川端亮 (編)(2010)『データアーカイブ SRDQ で学ぶ社会調査の計量分析』ミネルヴァ書房.

佐藤博樹・石田浩・池田謙一 (編)(2000)『社会調査の公開データ: 2次分析への招待』東京大学出版会.

安田三郎・原純輔 (1982)『社会調査ハンドブック [第3版]』有斐閣.

吉田富二雄 (編)(2001)『心理測定尺度集II』サイエンス社.

第3講 調査対象者と調査方法

田中重人 (東北大学文学部准教授)

[テーマ] 対象者の選定と調査の方法

1 レポートについて

1.1 課題

各自の関心にしがって調査を企画、質問紙を作成し、調査を実施して報告書を作成

- 有効回収数 20 以上
- 原則として、印刷・製本した質問紙を使った調査とするが、特殊な調査法を使いたい場合には、田中まで相談

1.2 スケジュール

- 調査企画について面談 (4月下旬)
- 企画書提出 (5/24)
- 調査企画と質問紙について討論 (5月下旬)
- 各自で調査実施
- 各自の調査 (結果) について報告 (7/12?)
- レポート提出 (8/14 締切)
- 講評と返却 (9月5日以降)

1.3 企画書について

各自の調査について「企画書」を作成して次回提出

- A4用紙 1 枚程度
- 調査の目的、対象者、配布・回収等の方法、参考にした既存調査と先行研究についてまとめること

2 今回の課題

自分の関心にしがって、対象とする人々の範囲、検証すべき仮説、収集すべき情報を考える。用紙にまとめて提出。

なお、自分の関心が絞り込めない場合は、たとえばつぎのような課題についてとりあえず考えてみるとよい。

- いわゆる「バイト敬語」についての人々の意識
- 自分の出身地の方言の使用状況
- 大学生の進路選択と両親とのコミュニケーションとの関係

3 調査の企画にあたって注意すべき事項

- 誰を対象とするか (母集団: population)
- 検証すべき仮説
- どのような情報を収集するか

これを具体化して、次のようなことを決めていかなければならない

- 調査で接触する対象者 (標本: sample) の抽出方法と人数 (sample size)
- 仮説検証のために必要な項目と測定・分析の精度
- 調査方法、個々の質問項目と回答選択肢、分析の方法

制約条件として日程、人手、費用などを考慮して、実現可能な企画を立てること。

4 母集団と標本

「母集団」(population) の3つの意味 (教科書7章)

- 興味の対象となる人や事物の全体 (理論母集団)。通常、非常に巨大で、境界のあいまいな、漠然とした集団
- 調査の対象としたい具体的な範囲 (調査母集団)。教科書ではこの意味
- 標本を抽出するための名簿 (標本抽出枠) に載っている全個体

「標本」(sample) とは

- 母集団から抽出した対象者のことをいう
- 調査の結果、標本の一部からしかデータが集まらないのが普通である。この点を明確にするため、本来の意味での標本を「計画標本」、有効なデータのみを「有効標本」と呼びわけることがある。

無作為抽出 (random sampling) とは、すべての個体をおなじ確率で抽出するものである。これは推測統計手法を使う際の前提となるものであるため、非常に重要。

ただし、無作為抽出がおこなわれるのは、標本抽出枠から計画標本の抽出のところだけである。また、完全な無作為抽出をおこなうのはむずかしい。

→ しかし、できるかぎり無作為抽出にちかい標本を得るように考えることが大切

5 調査の方法

調査の実施 (「実査」と呼ぶことがある) にあたって重要なポイント (教科書1-3章、5章、8章)

- 配布の方法
- 回答の方法
- 回収の方法
- 調査員の関与の程度

現代日本学演習 I 「質問紙調査の基礎」

第4講 調査の企画

田中重人 (東北大学文学部准教授)

[テーマ] 調査の企画と質問項目のリストアップ

1 今回の課題

各自の調査企画に基づいて、具体的な質問項目を考える。

- ペアを組んで議論すること
- 「質問」と「回答」の作りかたについては、教科書第6章を参照
- 用紙にまとめて、宿題の調査企画とともに提出

2 今後の予定と予習事項

- 来週以降、調査票の作成に入ります。教科書第6章を読んでおくこと。

現代日本学演習 I 「質問紙調査の基礎」

第5講 質問文と回答欄

田中重人 (東北大学文学部准教授)

[テーマ] 質問文作成にあたって気を付けること

1 課題

つぎの各項目についてたずねる質問文 (および回答欄) を考える。調査対象としては、東北大学の日本人学生を想定すること。

- 出身地 (言語使用についての分析に使う目的で)
- 授業スタイルに関する好み
- 外国語会話に関する苦手意識
- 自分の家族とのコミュニケーションの頻度
- 1日の時間の使いかた

2 質問文作成上の注意

2.1 用語について

- (1) 明確な質問
- (2) 日常的なことばで
- (3) ステレオタイプ的な表現を避ける

2.2 文章について

- (1) Double-barreled question を避ける → ひとつの質問ではひとつの事柄だけを聞く
- (2) 短く、意味のつかみやすい文章に
- (3) 誘導尋問を避ける
- (4) 黙従傾向に注意

2.3 質問のタイプ

- (1) 個人的質問 vs. 一般的質問
- (2) 意識 vs. 実態
- (3) actual status vs. usual status
- (4) 単一質問 vs. 複数質問群

3 いろいろな回答形式

3.1 自由回答とアフターコード

回答欄に自由に答えを書いてもらい、あとで値をあたえる (after-coding)。回答の予想がつかない場合、選択肢が多すぎる場合 (職業など)、数値をそのまま記入してもらう場合 (年齢など) に有効

自由回答の利点：

- 多面的な把握が可能
- 詳細な情報が得られる
- 誘導尋問になりにくい
- 回答者が不満を感じにくい

自由回答の欠点：

- 回答者の負担が大きい
- さまざまな次元の回答が混入する可能性
- 回答が標準化できない

3.2 回答選択肢 (プリコード)

あらかじめ回答の種類を制限して、値をあたえておく (pre-coding)。

- ひとつの準拠枠に沿って標準化した反応が引き出せる
- 回答者の負担がすくない
- 想定外の反応を抑制してしまう

「その他」などの選択肢を設けて、一部自由回答を併用することもできる。

プリコードの場合の注意事項

- ひとつだけ選択 (single answer) か複数選択 (multiple answer) か
- ありうる回答を網羅しているか?
- 選択肢は相互に排他的か?
- 間隔尺度としてあつかうには、最低限4つの選択肢が必要
- 選択肢が多すぎると回答者の負担が大きい (7つ以下がいちおうの目安)
- 不明・回答拒否を選択肢に入れるか?
- 選択肢のレイアウト
- 一対比較

4 宿題

自分の関心に沿って、質問項目を列挙してくる。完全な質問文になっていなくても、おおよそどのようなことを聞きたいかが書いてあればよい。2部作成して次回持ってくる (1部は提出、もうひとつは授業で使用)。

第6講 調査票の構成

田中重人 (東北大学文学部准教授)

[テーマ] 調査票全体の構成と予備調査

1 今日の課題

各自の質問文案について討論

2 調査票全体の様式

2.1 分量

調査対象者の負担や回答時間を考えて上限を決める。多くても16ページくらい?

2.2 印刷の形式

両面刷りで、製本して冊子体にするのがふつう

- 余白をじゅうぶんにとること
- 重要な部分が目立つように、レイアウトや字体などを工夫する
- 回答者への指示、質問番号、質問文、回答選択肢などの役割によって、字体などを変えてわかりやすくする
- ひとまとまりの質問項目を1ページにおさめること
- 1ページ内のレイアウトが複雑になる場合は、DTPソフトを使うか、印刷所に組版を依頼するとよい
- 各ページにページ番号をふる (番号がないと、落丁を摘発しにくい)
- 色紙や高級紙を使ったり、カラー印刷すると、通常書類と差別化できる
- 印刷するときは、原版の汚れに注意すること。トナー節約モードOFFで、解像度を高くして出力するとよい。コピー機ガラス面の汚れにも注意。

3 1ページ目の様式

3.1 タイトル

- 調査の内容を的確に表すタイトルをつける。
- ただし、難解な語句や調査対象者の反発を招く表現は避ける
- 研究者向けのタイトルと調査対象者向けのタイトルを別にするのも多い

3.2 調査主体・連絡先

- 調査をおこなう主体と責任者名、所属、学年
- 住所・電話番号・メールアドレス (授業用の代表用アドレスを用意しておくので、それを使ってもよい)

3.3 あいさつ文

- 自己紹介
- 調査の目的や意義 (詳しく書くことで、回答者の動機づけの高まることが期待できる)
- 個人情報やプライバシーの保護について
- 結果の報告先や公表媒体
- 対象者の選定方法

調査票とは別の紙に「あいさつ」を用意する場合もある。

3.4 記入上の注意

選択肢のえらびかた、数字の記入法、「わからない」ばあいはどうするかなど。回収方法まで指示することもある。

4 調査票の最後

- 終わりのあいさつを入れる
- 調査への感想や調査主体への意見を書く欄を設けることも多い

5 質問のならばかた

5.1 全体的な配列

答えやすいものを最初に、私生活に関するものを最後に。

- Face sheet (性別・年齢などの基本的な属性): ただしプライバシーに注意 (教科書 88 ページ)
- 調査の中心的な項目 (意識など)
- 私生活に関する項目 (家族構成、収入など)

たがいに関連する項目や、回答対象者を限定する項目はならべて配置する。

- 前後にどういう項目がくるかによって、おなじ質問文でも回答が変わることがある (carry-over effect; 同化効果; 対比効果) → 教科書 88-89 ページ

5.2 回答上の指示

- 対象者を限定する場合: 太字にするなどして目立たせる (できればページの最上部に来るように工夫する)
- 非該当の対象者のジャンプ先 (通常は質問番号で指示するが、ページ数を付加する場合もある。矢印などで視覚的に誘導できるとわかりやすい) → 教科書 91 ページ

5.3 質問番号と見出し

各項目には番号をつけ (問 1, 問 2, ... など)、その後に質問文をつづける。
質問番号の前に話題転換のための文を入れることがある

6 予備調査 (pre-test)

本調査の前に、かならず予備調査をおこない、調査票を修正する。印刷・製本の状態をふくめ、できるかぎり**最終版の調査票と同一**のものを使う。

- 質問文が適切に理解されるか
- 質問の流れに支障はないか
- 回答にかかる時間
- 回答者の負担感

対象者は少数でよい (数人から 20 人程度) が、**本調査と同様の対象者** を選ぶこと。回答に迷ったところや答えにくいところ、意味のつかみにくい質問文などについて、回答者の意見を集めることが目的である。このため、実際に回答してもらったあとで、かかった時間、気になったところ、疲労感などについて聞いておくこと。

7 調査票の校正

調査票は、調査開始前に入念にチェックする。小さなミスがのこっていると、それだけで調査が台無しになることがある。細心の注意を払って校正すること。

- ざっとながめる、版面の配置や印刷の美しさを確認する、ページをぱらぱらめくって見る
- 小さい単位に区切って、文字列を確認する
- 数字、記号、固有名詞などのチェック
- チェック済みのところに鉛筆や蛍光ペンなどでマークをつけながら、作業を進めるとよい。

校正作業に入る前に、しばらく調査票から離れて、内容についての記憶をできるだけうすれさせておくと効果的。

8 今後の予定

各自で調査を実施すること

- 質問紙を完成させ、予備調査を経て、本調査を実施
- 適宜、田中の指導を受けること。調査実施前に、必ず OK をとらなければならない。
- 7月の授業で調査結果の口頭報告

第7講 エディティングとコーディング

田中重人 (東北大学文学部准教授)

[テーマ] 調査票の回収からデータ作成までの手順を理解する

1 回収直後の作業

1.1 ID 番号の付与

調査票上部などに通し番号を振る。以降は、この ID 番号によって回答者を識別する。

1.2 エディティング (editing)

記入漏れ、不完全な回答、誤りなどを探して訂正する。可能であれば、再調査をおこなうこともある。訂正する際には、回答とは区別できるように色ペンなどを使う。

論理的にありえない回答の組み合わせや、珍しい回答についても注意する。

記入漏れの非常に多いもの、内容の信頼性が低いものは無効とする。

この時点で有効回収率を計算する。

$$\text{有効回収率} = \text{有効票数} / \text{配布票数}$$

2 コード表

2.1 コード表の作成

コンピュータでの分析のためには、データはすべて数値で表すのが便利である。回答を数値に変換するための対応表を「コード表」という。

コード表は、調査票作成と同時に作っておくとよい。ただし、調査後に自由回答のアフターコードなどを追加する場合や、データの内容をみて変数の配列を変えざるをえない場合もありえる。そのときには、コード表のほうも改訂する必要がある。

2.2 自由回答のコード

文章での回答を求めた場合や、選択肢に「その他」が入っていて具体的な内容を書いてもらった場合は、適切な数値 (code) を与える必要がある (アフターコード)。

既存のコードがある場合は、それを使えばよい。たとえば、都道府県・市町村コード、学校名コード、職業分類など。各省庁のサイトや日本工業規格 (JIS) などを探してみるとよい。

具体的な内容を見ないとコード自体が決められない場合は、つぎのようにする

- すべての回答をみてから、コードのあたえかたを決めるのがのぞましい。回答の量が多い場合には、一部だけをみて方針を決めることもありうるが、その場合、実際にコーディング作業に入ってから、対応できないケースが出てくるおそれがある。
- 「7. その他」などの場合、もとのコード「7」はそのまま残しておき、それよりも大きい数値 (たとえば 11 以降) をアフターコード用に使うとよい
- 分類の基準が微妙である場合には、3 人以上でコーディングをおこなって、2 人以上が一致したコードを採用する、などの方法をとることがある。

- 文章をいったんコンピュータ・ファイルに入力してからコーディングをおこなう方法もある。また、そのような作業がある程度自動的におこなうプログラムも存在する。
- どのような基準でコードを付与したかは、コード表に書いておくこと。
- コードはできる限り細かくしておくのがよい。あとで分析する時には「変数値の再割り当て」(SPSSの場合)などで統合することができる。

2.3 欠損値のコード

通常の調査では、欠損値としては、無回答 (NA = No Answer) によるものと非該当 (対象外の質問) によるものがある。これ以外に、「わからない」(DK = Don't Know) やクリーニング段階で欠損値を与えたものなどについても別扱いにする場合がある。

いずれの場合も、「9999」「9998」などの通常は出てこない値を入れておいて、それをSPSSで処理する際に「ユーザー欠損値」として指定するとよい。

2.4 複数回答のあつかい

「いくつでも ○ をつけてください」のような質問の回答は、ひとつの回答についてひとつ変数をつくり、つぎのような数値をあたえるのが通例である。

- 0: 丸がついていない
- 1: 丸がついている

0と1にしておくと、そのまま合計 (= その項目を選択した人数) や平均 (= その項目を選択した人の比率) を使えるという利点がある。

3 コーディング

コード表にしたがって回答を数値化していく作業。つぎのどちらかの方式による。

- 調査票の回答選択肢の数値をそのまま使う。ただし、判断を要する場合や自由回答のコードなどは、調査票に直接書き込む (色ペンで回答と区別する)
- 専用の用紙に転記する

このコード結果をみながらコンピュータに打ち込むことになる。

4 課題

配布した調査票について、つぎの手順で作業し、コード表にしたがってシートに転記する

- 調査票とコード表の内容を確認
- IDは付与済み
- エディティング (色ペンを使う)
 - 全体をみて、おかしいところをなんとかする →ある程度は原則をたてて考える
- コーディング
 - 住所コードの付与
 - 非該当の処理 → コード表をみる
 - 無回答については、この段階で処理する必要はない
 - シートに転記

5 次回予定

次回はコンピュータ実習室 (7F) で「データ入力」の授業

第8講 データの入力と点検

田中重人 (東北大学文学部准教授)

[テーマ] データ作成と基本的クリーニング

1 コンピュータファイルの作成

変数名を決めておく (コード表をつくる時に決めておくとよい)

- 英字と数字のみ (A-Z 0-9 _)。ただし最初の1文字は数字を使えない
- 英字は大文字/小文字を区別しない
- 8文字以内
- 問番号にしたがって、**Q10_4A** などとつけると便利
- 転記用紙を使うときは、最初から変数名を書き込んでおくとよい

下線 (_) 以外の記号のほとんどは変数名には使えないので、注意すること。

※ 最近の SPSS では、変数名に日本語文字が使えて、長くても OK ではあるが、トラブルが起こりやすい

コンピュータに打ち込むには、最初から SPSS 等で作成してもよいが、Excel 等の表計算ソフトが使いやすい。その場合、つぎのように進める。

- 最初の行に変数名を書き込んでおく
- 「ウインドウ枠を固定」しておくと便利
- 別の人に打ち込んでもらって、ファイルの相違をチェックするのがよい。Excel では、おなじファイルの複数のシートを使った計算ができる。対応するセル同士の差をとって、すべてゼロになっているかを確認する。

SPSS で分析するときは、ファイルの種類で「xls」などを指定すると、Excel のファイルをそのまま読み込める。オプションで、1行目に変数名であることを指定すること。この方式では変数のラベルを読み込めないで、ラベルが必要ならあらかじめ SPSS の「変数」ビュー画面で入力する。SPSS で 9999 などを欠損値にするには、「変数ビュー」で、ひとつの変数について欠損値を指定しておいて、それをほかの変数のところにコピーすると、一気に処理できる。

2 クリーニング

データにみられるまちがいを摘発・修正する

- 度数分布表やクロス表を出力して、ありえない値をチェックする。コード表作成の段階で、変数の値の範囲や、変数間の依存関係をあらかじめリストアップして、コード表に書き込んでおくとよい。
- データをすべて調査票とつきあわせて、一致しているかを確認することもある。この方法の場合、データを打ち込んだのはちがう人がおこなうのがよい。
- 調査票の回答そのものがおかしい場合にどのように修正するかについては、さまざまな立場がある。教科書を参照のこと。

あとで分析を進めていくうちにまちがいがみつかることもある。そのときは適宜データを修正して、データセットのバージョンを更新していく。一定の規則にしたがってファイル名をつけていくとよい。

3 課題

- (1) 前回作成した転記シートデータを、Excel に入力
- (2) 二人分の入力結果を1ファイルにまとめ、シート間の対応するセル同士の差をとって、すべてゼロになっているかを確認する
- (3) SPSS に読み込む
- (4) 9999, 9998などを「ユーザー欠損値」に指定

現代日本学演習 I 「質問紙調査の基礎」

第9講 報告書の執筆

田中重人 (東北大学文学部准教授)

[テーマ] 報告書の役割と書きかたを理解する

1 分析から報告まで

1.1 公表する媒体

- 報告書
- 雑誌論文
- 書籍
- プレス発表

学位論文や通常の授業レポートは、報告書と雑誌論文の中間くらいか。

1.2 報告書の役割

報告書はつぎのような役割を持つ (用途別にちがう版をつくることもある)。印刷した物を関係者に配布するほか、図書館に寄贈したり、インターネットで公開したりする。

- 協力者への結果報告
- 他の研究者との情報の共有
- 分析結果を利用して論文を書く場合の参照先
- データ・チェックの基礎資料

調査の企画段階から、報告書作成のための予算・時間を見越しておくこと

2 報告書の様式

2.1 調査の概要

- 背景
- 目的
- 調査の内容
- 調査主体 (場合によっては費用負担)
- 企画から調査票作成、予備調査、実施、分析までのスケジュールと内容

- 調査方法 (配布・記入・回収の過程をくわしく)
- 調査母集団と計画標本の抽出法
- 回収数、回収率、非回収の原因、基本的属性の分布
- 調査協力者

文章で詳しく書いてもよいし、表の形式で簡潔にまとめてもよい。

2.2 分析結果

- 分析結果の提示 (表または図)
- 分析結果の解釈 (文章)
- 提言など (報告書の目的による)

2.3 資料

- 調査票
- 単純集計 (度数分布表)
- あいさつ状など (あれば)
- 調査票をつくるうえで参考にした既存調査等の一覧
- コーディングやエディティングのルール
- コード表 (データを公開・共有する場合; データ分析用マニュアルとして、別に用意する場合もある)

「資料」には、通常、すべての変数 (ただし ID などのケース別に固有の値のものは除く) について度数分布を表示する (巻末に「資料」などとして載せる)。カテゴリ数の多いものについては、適当な階級に分けてよい (たとえば、年齢を5歳刻みにするなど)。

度数分布表に載せる相対度数 (パーセンテージ) には、欠損値を含めて計算したものと、除いて計算したものがある。どちらを使ってもよい。ただし、どのくらいの欠損値が出ているかという情報は重要なので、欠損値の種類と数は必ず表に載せること。

3 報告書本文の分析

報告書の中心部分には、データ分析をおこなった結果を載せる。調査をおこなうにあたっては、問題関心 (または仮説) があるはずなので、それにしたがって分析し、結論を出すことが望ましい。

分析結果をどのようなかたちで報告書に記述するかは、想定読者によって変わる

- (1) 分析結果だけを載せる (解釈しない)
- (2) 分析結果 + 解釈
- (3) 実証研究論文としての体裁をとったもの
- (4) 政策提言や企画立案をおこなうもの

また、分析結果そのものの表示の仕方もうろいろである。

- 一般の読者相手であれば、表よりもグラフを多用して視覚的に訴えるのが望ましいことが多い。また、細かい技術的な情報は割愛して、結論がストレートにわかるようにする。
- 学術的なところで発表する場合には、表を利用して細かい数値をきちんと書く。技術的な情報を十分に書いておくこと（たとえばカテゴリーへの数値の与え方、欠損値の処理、統計的検定の方法など）

報告書は繰り返しコピーされることを念頭において作成する。網掛けやカラー印刷はなるべく使わないのがよい。グラフを使用する場合、細部まではっきり見えるかどうかに注意すること。3次元(3-D)グラフはわかりにくくなるので、使わないほうがよい。

4 この授業のレポート

締切: 8/14 (水) 17:00

提出先: 電子メールで提出 (tanakas2013 @ tsigeto. info: かならず返信を確認すること)

内容: 調査報告書 (上記の (2) または (3) のかたちで分析結果を含めたもの。調査票や度数分布表などの資料もつける)

レポートは、9/5 以降に返却しますので、取りに来てください。

5 発表会

- 持ち時間は、それぞれ10分(説明3分+質疑7分)。
- 配布資料を9人分作成してくること。
- 上記の「調査の概要」にあたる情報を必ず含めること
- 調査票そのものは配布資料に含めなくてよいが、コピーを2部用意しておくこと(発表中に回覧)