

現代日本論演習「質問紙法の基礎」

<金2>授業観察室（文学部・法学部合同棟 2F）

3年生対象：2012年度 5セメスタ

『講義概要』 記載内容

- ◆ 講義題目：質問紙法の基礎
- ◆ 到達目標：(1) 質問紙調査の長所と短所を把握する；(2) 質問紙調査の実際のプロセスについて、体験を通して習得する
- ◆ 授業内容・目的・方法：質問紙を使った調査の方法についての講義と実習をおこないます。講義では、質問紙調査の基本的な概念と方法、仮説設定からレポート作成までの一連のプロセスについて解説します。実習では、受講者が各自の選んだ研究テーマに沿って文献収集をおこない、テーマへの理論的アプローチを検討し、質問紙を作成し、調査を実施し、その結果をレポートとして提出します。
- ◇ 成績評価の方法：授業中の課題と宿題（40%）、学期末に提出する質問紙（30%）、調査結果に基づく口頭発表とレポート（30%）を合計して評価する。
- ◇ 教科書：轟亮・杉野勇（編）(2010)『入門・社会調査法』法律文化社。
- ◇ その他：受講者は、現代日本論演習「統計分析の基礎」（火4）をあわせて受講することが望ましい。

授業の概要

目次は別紙を参照

※ 日付は、学期前のおおよその計画をあらわしているが、実際の授業の進行状況によって前後にずれることがある。

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。

- 5/25 調査企画書を提出
- 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
- 調査の進捗について、個別面談の機会を7月下旬に持つ
- 調査結果を口頭報告する（7/6, 13 の予定）
- レポートにまとめて提出（8/14 締切）→ 9/11 以降に返却

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

講師連絡先

田中重人（東北大学文学部日本語教育学研究室）
〒980-8576 仙台市青葉区川内 27-1 文学部・法学部合同研究棟 2F

E-mail: tanakas2009@sal.tohoku.ac.jp

オフィス・アワーは定めていません。適当な時間に予約をとってください。

第2講 仮説の設定 (4/20)

田中重人 (東北大学文学部准教授)

[今回のテーマ] 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、どの変数についてどのような分析をす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 (教科書の対応箇所は pp. 51-54)。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」をはっきりさせたいうえで、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 次回準備

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、またそのためにはどのような質問項目が必要かを考えてくる。

4 文献

小林淳一・木村邦博 (編) (1991) 『考える社会学』 ミネルヴァ書房.

レイブ, C. A. ・マーチ, J. G. (1991) 『社会科学のためのモデル入門』 (佐藤嘉倫ほか訳) ハーベスト社.

第3講 既存調査と先行研究の探索 (4/27)

田中重人 (東北大学文学部准教授)

[今回のテーマ] 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

とりあえず、つぎのデータベースをひとつお試してみよう。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://ssjda.iss.u-tokyo.ac.jp/access/flow/>>
- GeNii (国立情報学研究所) <<http://ge.nii.ac.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

- 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

不明の場合は「不明」などと書いておく。

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

<http://www.sal.tohoku.ac.jp/nik/student/litsurv.html>

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://ssjda.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>>
- 社会・意識調査データベース (札幌学院大学) <<http://www.sgu.ac.jp/soc/sordhp/main.htm>>

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/>> (GeNii から一括検索できる)

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 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よい。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 ReaD 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。

6 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4講 調査の企画 (5/11)

田中重人 (東北大学文学部准教授)

[今回のテーマ] 調査の企画の際に考慮すべき事項を理解する

1 レポートについて

1.1 課題

各自の関心にしたがって調査を企画、質問紙を作成し、調査を実施して報告書を作成

- 対象者は20人以上
- 原則として、印刷・製本した質問紙を使った調査とする。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

- 企画書提出 (5/25)
- 調査企画と質問紙について報告 (6/1, 8)
- 各自で調査実施
- 調査結果について報告 (7/6, 13)
- 個別面談 (7月下旬)
- レポート提出 (8/14 締切)

1.3 企画書について

各自の調査について「企画書」を作成して次々回提出

- 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

自分の関心にしたがって、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
- 携帯電話の「絵文字」の印象
- 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。

第5講 対象者の選定と調査の方法 (5/18)

田中重人 (東北大学文学部准教授)

[今回のテーマ] 対象者選定と実査の方法について基礎知識を学ぶ

1 母集団と標本

「母集団」(population) の3つの意味

- 興味の対象となる人や事物の全体 (理論母集団、universe)。通常、非常に巨大で、境界のあいまいな、漠然とした集団
- 調査の対象とする具体的な範囲 (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

2 実査の方法

- 配布の方法
- 回答の方法
- 回収の方法
- 調査員の関与の程度

3 問題

- (1) 内閣支持率の調査を RDD (教科書参照) で行う。この場合の問題について考えよう。
- (2) 日本語学習者の長音の習得について調査したい場合、母集団と標本はどのように設定すればよいか。
- (3) 東北大学の知り合いの学生を通じて配布する調査方法は、どのように分類できるか。

第6講 調査票の作成・1 (5/25)

田中重人 (東北大学文学部准教授)

[テーマ] 質問文を作成する

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)
- ら抜きことばに対する意識
- お酒をどれくらい飲んでいるか
- 結婚に関する意識

3-4人程度のチームを組んで案を作る

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。次回、人数分のコピーを用意してくる。

つぎのことを明記：

- どのような調査対象を想定しているか
- 調査方法
- 調査の目的

第7講 調査票の作成・2 (6/29)

田中重人 (東北大学文学部准教授)

1 調査票全体の様式

1.1 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?

1.2 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
- 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
- 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

2 1ページ目の様式

2.1 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

2.2 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あるので、それを使ってもよい)

2.3 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。

2.4 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。

3 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

4 質問のならばかた

4.1 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 87 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 ページ

4.2 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) → 教科書 92 ページ

4.3 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

5 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。

- 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

6 調査票の校正

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。

7 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。

現代日本論演習「質問紙法の基礎」(東北大学文学部) 2012年度

第8講 調査票回収からデータセット作成まで (7/6)

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、一部だけを見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、そのような作業を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にしておくこと、そのまま合計や平均を計算して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 コンピュータファイルの作成

変数名を決めておく（コード表をつくるときに決めておくとうい）

- 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うい

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。

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」しておくとうい
- 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、ファイルの種類で「xls」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行目が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ためて 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。

5 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と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りうる。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。

6 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データセットを作成する

- 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
 - 問14-1: 自分が入っているか?
- 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

【以下は宿題】

- Excel でデータ入力
 - コード表にしたがって、1行目に変数名を入力
 - 調査票をみながらデータ入力
 - ファイルを次回までにメールで提出