中部地区英語教育学会　第40回石川大会
英語教育研究法セミナー

2010年6月27日

# 量的研究デザインの方法


**本田勝久（k-honda@faculty.chiba-u.jp）**
**千葉大学（教育学部）**


## 0. 要旨


英語教育の研究を行うとき、もし研究者が、実証主義的な視点 (positivist) に立てば、客観性、予測、反復可能性を重視し、科学的一般化や現象を説明する法則を明らかにすることが目的となる。そのため、実験研究や質問紙調査などの方法論 (methodology) を選択し、測定や選択式質問紙などの量的手法 (method) でデータを収集して、統計的な分析を行うことになる。

本セミナーでは、具体的な研究テーマを例として、量的手法で研究を行う場合の研究デザインについて提示する。具体的には、(1)リサーチ・クエスチョンの設定、(2)データ収集の方法、(3)データ分析の方法、(4)研究の評価に関して、それぞれ提示していく。

また本年度は、統計的検定における検出力 (power) に基づくサンプルサイズの設計方法について触れる。検出力やサンプルサイズの設計は、統計的検定を行うにあたって基本的な概念であり、重要である。英語教育における量的手法では、自由度(n−1)の t 分布の両側 5%点より大きければ「有意差あり」と判定し、「帰無仮説を棄却して対立仮説が成り立っている」と判断することがある。しかし、検出力とサンプルサイズの関係を検討すれば、実質的な意味のある差について考察することができる。

・対立仮説が成り立っている→サンプルサイズ（大）、検出力（高）
・帰無仮説が成り立っている→サンプルサイズ（小）、検出力（低）

本発表では、正規分布、t 分布、$\chi^2$分布、*F* 分布の分布関数を計算できる（例えば、表計算ソフト Excel などを使って）ことを想定している。統計的検定におけるサンプルサイズの設計では、検出力の考え方と計算方法が密接に関連している。サンプルサイズや検出力を検討することで、量的研究デザインの方法について議論したい。


## 1. 理論的視点

研究デザインを行う際に、研究者が立つ理論的視点・枠組み(paradigm)は大きく 3 つに分類される。

① 実証主義的視点（positivist/scientific paradigm）

人々の行為には普遍的な法則・規則がある。客観性、予測、反復可能性を重視し、因果関係を説明するための調査を目的とする。科学的で実験主義的な調査を通して得られた知識は客観的で定量的であり、現実世界(reality)は静的で観察・測定可能なものである。

② 解釈主義的視点（interpretive paradigm）

人々の経験は文脈に縛られており、時、場所、人間という行為者の心と切り離すことはできない。多元的な現実世界(reality)は人々により社会的に構築されている。参加者の行為や経験、教育のプロセスの意味を探究し理解することを目的とする。

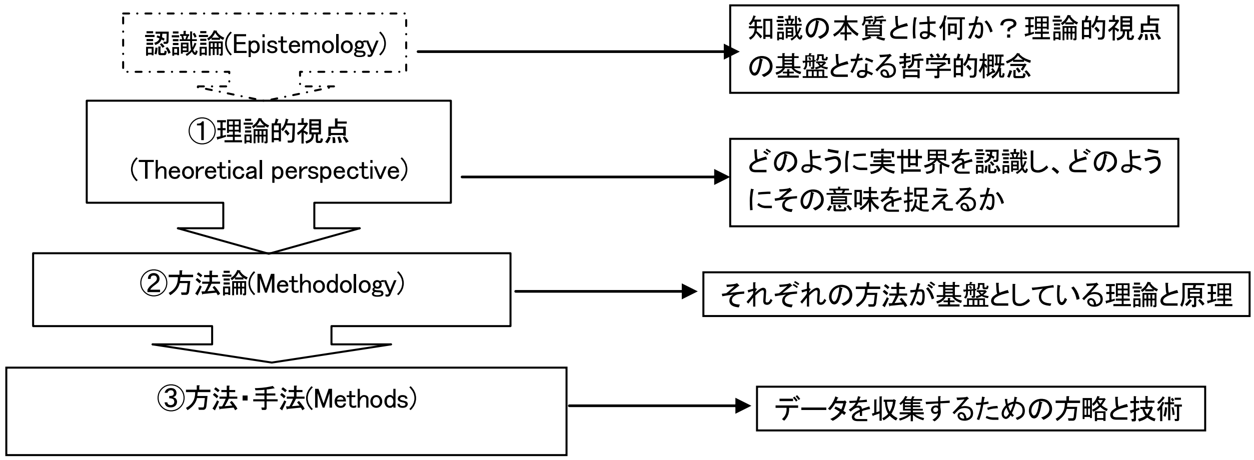
③ 批判的視点（critical paradigm）

客観的や中立的な知識は存在せず、知識は常に社会的利害に影響されている。人間の行動を解釈し、理解するだけでなく、社会的批判をすることで、社会的あるいは組織的な変化を起こすことを目的とする。

## 2. 研究計画

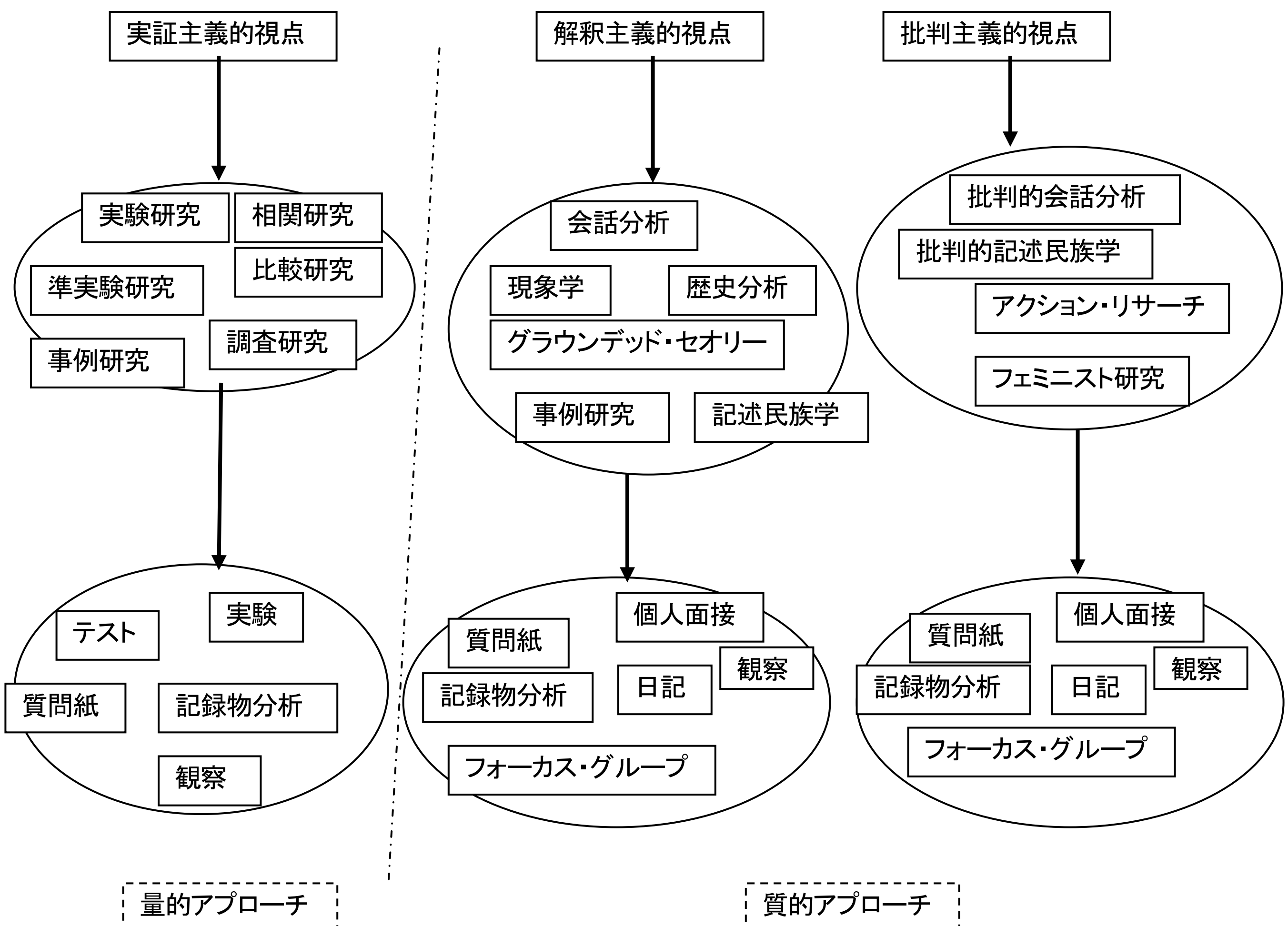
### 2.1 研究計画に関わる 4 要素

研究計画を立てるときに以下の 4 つの要素について認識しておく必要がある。




Crotty (1998, p.4)より

### 2.2. 量的アプローチと質的アプローチ



## 3. 研究課題の設定

### 3.1 一般から特定へ

教師としての経験や関心(あるいは勘)⇒日々の教育実践の中で抱く疑問や課題から
↓
測定される特定の変数の用語によって具体的に言い表されたものへ

- 研究論文は、研究課題が興味を引くに足ることを立証しなければならない。
- 研究課題は、その研究が適切かつ重要であり、新たな寄与をおこなうものであることを示さなければならない。

「目的」はその研究の目標についての一般的な表現
「仮説」はその実験で起こると思われることについてのより特定化した予想(量的)

(一般)学生が好んで選択する心理学のコースは、教える講師のタイプと関係があるか
(特定)個人的・社会的なスキルを強調する講師は、臨床のコースでより好まれるか
分析的・客観的なスキルを強調する講師は、実験的なコースでより好まれるか

(Findlay 1993, p. 17改訂)

### 3.2. 先行研究から

・関連文献を探す(雑誌論文, 電子ジャーナル, 人と人とのコンタクト etc.)
学習者要因に関する多くの研究論文は、*The Modern Language Journal*や*Language Learning*に代表されるSLAの主なジャーナルにも掲載されている。

・関連文献をまとめる(文献目録, コンピュータのデータベース etc.)
「対象者の属性」「何がなされたか」「何がわかったか」「どのような結論になったか」などの関連文献ごとのリストを作成する。

・用語を整理する(検索エンジン, 用語辞典 etc.)
questionnaires, inventories, forms, tests, batteries, checklists, scales, surveys, schedules, studies, profiles, index/indicators, sheets, etc.

先行研究の追試や先行研究への疑問から研究課題が見つかることも...

### 3.3 自分なりの研究課題の設定へ

(1)位置づけができること→先行研究から自分の研究はどのような位置づけになるのか？
(2)オリジナリティがあること→先行研究との違いは何か？単なる追試になっていないか？
(3)検証可能であること→無理な課題設定はしていないか？変数の統制や条件は整備できるか？(量的)

| 先行研究では | | 研究課題では | | 研究方法の決定へ |
|---|---|---|---|---|
| (1)何が主張されていて.. | ⇒ | (1)何を知らなければならないのか.. | ⇒ | では、どのようにしたら、 |
| (2)何がわかっていて.. | | (2)何を知ることができるのか.. | | 知ることができるか？ |

## 4. 量的アプローチによるリサーチ・デザイン：調査研究

### 4.1 調査研究とは？

**実験研究（experimental study）**
特定の事象・現象を分析する目的のために、人為的に整えられた条件（例：実験(experimental group)・統制群（control group)）のもとで、特定の道具や装置を用いた実験を行い、そこで得られる実験結果を分析するもの。
**調査研究（survey）**
観察、検査、アンケート、面接などを行いそれによって得られた資料を分析するもので、人為的な条件は設定しない。
白畑知彦・冨田祐一・村野井仁・若林茂則（1999, p.261）

①観察によるもの⇒「見ること」により学習者を理解しようとするもの
・観察法：学習者の行動を観察・記録・分析し、行動の質的・量的特徴や行動の法則性を解明すること
②言語を媒介とするもの⇒「聞くこと」により学習者を理解しようとするもの
・質問紙法＆面接法：行動そのものよりも学習者の感情や価値観、動機など、心の内面を理解すること

### 4.2 研究手法の決定

**<u>リサーチ・デザインを決定する際、選択する研究手法によりプロセスが異なることに留意する。</u>**

例えば…質問紙法では、

①測定対象を明確にする
↓
②項目の候補を収集する
＊自分で考える
＊人に尋ねる
＊関連文献にあたる
↓＊項目の分類，絞込み
③予備データを収集する
↓
④項目を決定する
＊反応分布の検討
＊G-P分析, I-T相関分析
↓＊因子分析
⑤本調査を行う
↓
⑥信頼性の検討
＊再検査法
＊折半法
↓＊α係数
⑦妥当性の検討
＊基準関連妥当性
＊構成概念妥当性

例えば…観察法では、

①対象行動の適切性の検討
＊予備観察
↓＊継続時間，観察単位
②観察対象者の決定
↓
③行動の下位カテゴリーとその操作的定義の決定
↓
④適切な時間間隔と観察回数の決定
↓
⑤記録用紙の設計
＊自由記述法
＊サンプリング法（1/0，ポイント）
↓＊サインシステム，カテゴリーシステム
⑥観察の実行
↓
⑦一致度の算出
＊信頼性の検討
↓＊カッパー係数
⑧データ処理

*これらのプロセスを視野に入れて*
*リサーチ・デザインを決定するのが望ましい…*

(1)調査対象を定義する→調査目的や方法に応じて調査対象の範囲を想定し、直接の調査対象者（サンプリング）と母集団との関係を決定すること

(2)調査内容を吟味する→調査内容の候補を収集・精選し、予備調査や分析（信頼性＆妥当性）などを通して実際の調査項目を作成すること

(3)調査方法の種類を知る→いろいろな調査方法の短所と長所を知り、研究課題や時間・労力・資金などに合わせて調査方法を選択すること（面接調査，集団調査，委託調査，郵送調査 etc.）

(4)分析手法やデータ処理を予測する→調査対象者の数やその構成（性別・年齢）、調査項目への回答の分布、欠損値の有無などのデータ内容の記述統計的な概略から、データの集計と統計的分析手法を予測すること（質的・量的データ，ノンパラメトリック検定 etc.）

(5)倫理の問題を考慮する→調査対象者とのあいだに信頼関係を築き、人間尊重の精神で調査を実施すること（説明と同意，プライバシーの保護，結果報告，データの保存または破棄 etc.）

「はじめに研究手法ありき」「ただやっているだけ」ではなく、まずは自分の研究課題をじっくり明らかにすることから、「よい良い研究手法」を決定して欲しい…

## 4.3 データ処理

### 4.3.1 統計を用いたデータ処理の5つの条件（なぜデータ処理をするのか？）

Systematic Research: A statistical study has a clear structure with definite procedural rules that must be followed.

Logical Research: The rules and procedures underlying these studies from a straightforward, logical pattern－a step-by-step progression of building blocks, each of which is necessary for the logic to succeed.

Tangible Research: Statistical research is tangible in that it is based on the collection and manipulation of data from the real world.

Replicable Research: Statistical research should be also replicable. The researcher's proper presentation and explanation of the system, logical data collection, and data manipulation in a study should make it possible for the reader to replicate the study (do it again under the same conditions).

Reductive Research: Statistical research can reduce the confusion of facts that language and language teaching frequently present, sometimes on a daily basis.

(Brown 1988, pp. 4-5 参照)

### 4.3.2 記述統計と推測統計（統計における基礎的な述語とは？）

変数，構成概念，独立変数，従属変数，母集団，標本，検定，統計的有意，危険率，水準 etc.

→記述統計(descriptive statistics)：標本のcentral tendencyとdispersionを調べること
平均，分散，標準偏差，尖度，歪度，z得点/偏差値，中央値，最頻値，パーセンタイル etc.

→推測統計(inferential statistics)：標本のnormal distributionから母集団を推測すること
正規分布，等分散検定，5%水準，片側＆両側検定，対応ありと対応なし，繰り返し要因 etc.

(1)パラメトリック検定：母集団の数理的特性を推定し、母集団の分布型についての仮定を含める検定（母集団が正規であることを前提とする）→t検定，分散分析，ピアソンの相関係数 etc.
(2)ノンパラメトリック検定：標本の統計的特性から、母集団の分布が特定の形であることを前提としない検定（分布にかかわらない統計法）→$\chi^2$検定，マン-ホイットニーの検定 etc.

#### 4.3.3 相関研究とカテゴリカルデータ分析

カテゴリカルデータとは、相互に排他的なカテゴリーに分類されたデータで、数量としての意味をもたない数字や、記号によって示される名義尺度上のデータをいう。

①数量として平均値や標準偏差を求めない　②母集団について特別な仮定をもたない

↓でも…

そのカテゴリーが量的に変化に沿って設定されていて、量的な意味を完全に失っていない場合には、量的なデータとしてそれに対応する分析を行うこともある。

↓重要なのは…

<u>**尺度（scale）が決まれば、検定（test）が決まる**</u>

4つの尺度の使い分け、差の有無、関係の有無

例えば…関係の検討（分析typeⅡ）では、
名義尺度→$\chi$2検定，順序尺度→順位相関係数，間隔＆比尺度→ピアソンの積率相関係数etc.

| | | 大小 | 加算 |
|---|---|---|---|
| 名義尺度（nominal scale） | 性別，職種，所属クラブ，好きな食べ物 etc. | × | × |
| 順序尺度（ordinal scale） | 満足度，階級，コンテスト順位（1<2<3）etc. | ○ | △ |
| 間隔尺度（interval scale）<br>比尺度（ratio scale） | 身長，体重，年齢，金額，試験成績，等間隔に点数化された回答 etc. | ○<br>（順序） | ○<br>（平均） |

### 4.4 例題

これまでのセミナーでは、$\chi^2$検定、ウイルコクスンの順位和検定、重回帰分析、スピアマンの順位相関係数、対応ありのt検定、1元配置の分散分析と異なった検定を取り上げてきた。本セミナーでは、昨年度の分析手法を取り上げ、分析方法の違いについて解説する。これまでの分析結果および本年度の発表資料などは、浦野 研（北海学園大学）先生のHPからダウンロードできるので、ご参照頂きたい。

http://www.urano-ken.com/research/seminar/

#### 4.4.1 対応ありのt検定（dependent t-test）

mouth movement の提示は学習者の音素の識別に有効であると考えている N 先生は、「音声だけを用いた学習よりも音声と映像を併用した学習の方が音素の識別には有効である。」ということを実証するために、以下のような実験を試みた。まず、被験者である大学生 40 名をTOEIC（listening section）のスコアによって均等になるような2つのグループ（音声だけを用いた学習をするグループ：G1）と、音声と映像を併用した学習をするグループ：G2）に分類した。さらに、弁別が困難であるといわれる/æ/,/ʌ/,/ɑː/,/ə/の4母音を取り上げ、それぞれの音素を識別する pre-test を実施した。pre-test は音声テープにより 1 問 30 秒（15 秒×2 回）で計 40 問であった。両グループの違いが「音声」と「音声＋映像（ビデオテープ）」になるように、それ以外の変数はできるだけ統制するような授業を4週間にわたって行った。最後に、post-test として pre-test と同じテストを実施し、以下のようなデータが得られた。

デ ー タ 表

（G1）

| No | Listening | Pre-test | Post-test |
|---|---|---|---|
| 1 | 155 | 25 | 22 |
| 2 | 185 | 28 | 38 |
| 3 | 125 | 18 | 10 |
| 4 | 150 | 28 | 29 |
| 5 | 180 | 37 | 33 |
| 6 | 115 | 9 | 19 |
| 7 | 95 | 23 | 20 |
| 8 | 165 | 37 | 44 |
| 9 | 120 | 16 | 18 |
| 10 | 185 | 41 | 40 |
| 11 | 95 | 7 | 10 |
| 12 | 145 | 10 | 8 |
| 13 | 175 | 12 | 15 |
| 14 | 85 | 17 | 13 |
| 15 | 180 | 33 | 35 |
| 16 | 95 | 18 | 19 |
| 17 | 155 | 41 | 38 |
| 18 | 115 | 10 | 12 |
| 19 | 135 | 16 | 14 |
| 20 | 205 | 29 | 21 |

（G2）

| No | Listening | Pre-test | Post-test |
|---|---|---|---|
| 1 | 80 | 16 | 19 |
| 2 | 155 | 14 | 17 |
| 3 | 240 | 45 | 43 |
| 4 | 125 | 18 | 25 |
| 5 | 145 | 29 | 36 |
| 6 | 140 | 19 | 16 |
| 7 | 195 | 29 | 33 |
| 8 | 165 | 25 | 31 |
| 9 | 25 | 12 | 15 |
| 10 | 125 | 27 | 29 |
| 11 | 205 | 21 | 33 |
| 12 | 150 | 16 | 13 |
| 13 | 55 | 10 | 15 |
| 14 | 145 | 15 | 26 |
| 15 | 145 | 21 | 39 |
| 16 | 250 | 27 | 44 |
| 17 | 120 | 33 | 45 |
| 18 | 120 | 18 | 24 |
| 19 | 145 | 27 | 37 |
| 20 | 110 | 19 | 30 |

どのような統計処理をしたらよいのでしょうか？また、N先生の考えは支持されたのでしょうか？

### 4.4.2 1元配置の分散分析（one-way ANOVA）

「リーディングテキストにおける未知語の推測は、学習者の語彙のリテンションに効果的である」と考えているH先生は、4週間にわたって次のような調査を行った。まず、Nation (1990) のVocabulary Levels Test (2,000-3,000 word level) を30人の大学生に実施し、それぞれ10人からなる語彙力において等質な3グループを作成した。それら3つのグループに異なるインストラクションの授業（Group1: vocabulary guessing, Group2: vocabulary glossing, Group3: accessing dictionaries）を行い、vocabulary testを4週間後に実施した。このvocabulary testは、Flesch grade level 14のリーディングテキストから、パイロットスタディとして学習者が未知語として選出した29語の中から16語をターゲットとして作成されたものである（correct=1, incorrect=-0.33, I don't know=0)。それぞれのグループにおけるvocabulary testの結果は以下の表のとおりであった。はたして、H先生の考えは支持されたのであろうか。

Group 1

| Nation's Vocabulary Test | Vocabulary Test（4週間後） |
|---|---|
| 25 | 43.6 |
| 28 | 56.8 |
| 31 | 27.3 |
| 19 | 35.0 |
| 40 | 48.4 |
| 32 | 42.4 |
| 23 | 25.3 |
| 19 | 51.7 |
| 20 | 31.4 |
| 22 | 29.1 |

Group 2

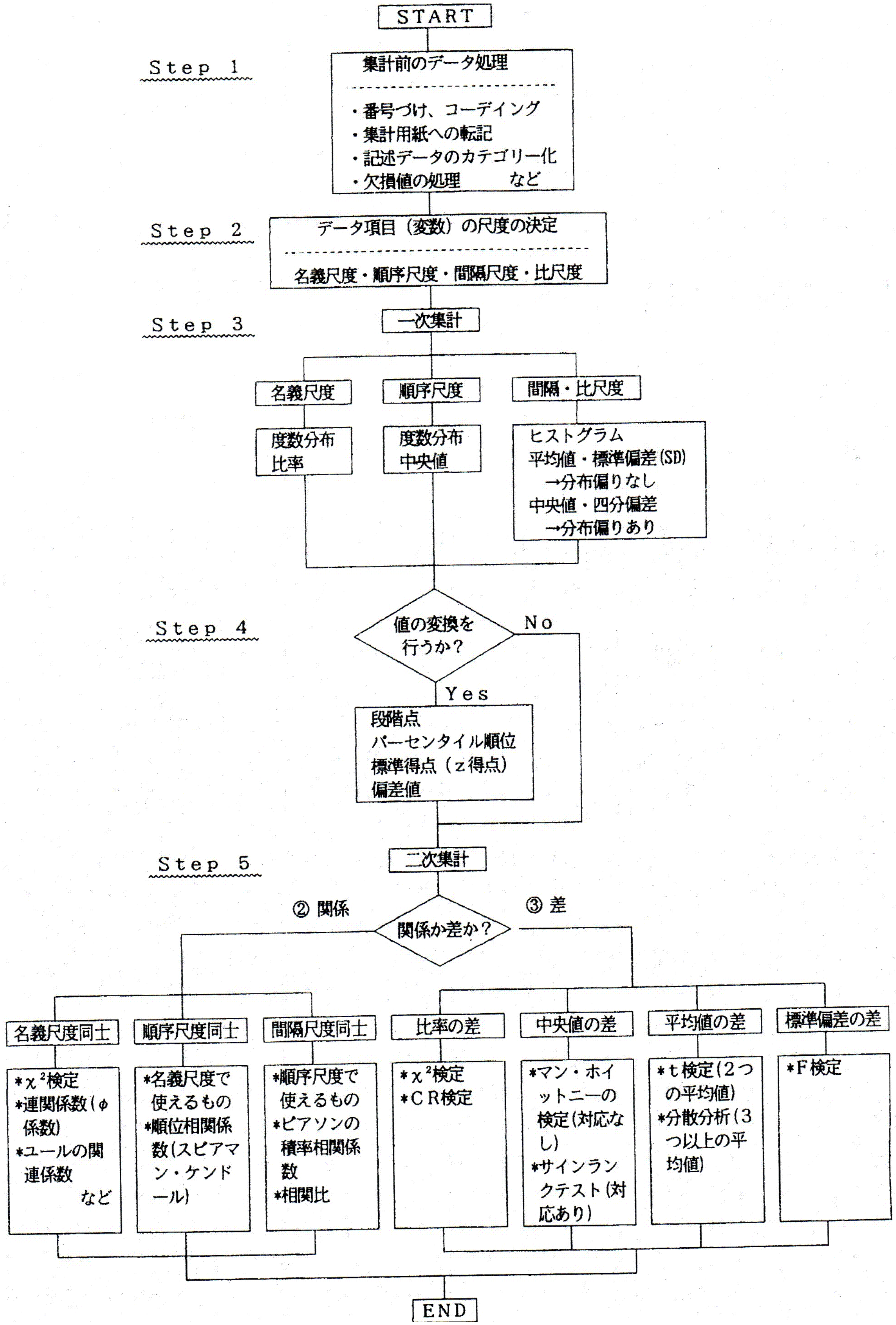
| Nation's Vocabulary Test | Vocabulary Test（4週間後） |
|---|---|
| 24 | 27.4 |
| 23 | 38.9 |
| 31 | 59.4 |
| 42 | 43.2 |
| 18 | 15.9 |
| 15 | 22.2 |
| 24 | 52.4 |
| 27 | 28.9 |
| 21 | 30.6 |
| 28 | 25.7 |

Group 3

| Nation's Vocabulary Test | Vocabulary Test（4週間後） |
|---|---|
| 24 | 18.3 |
| 36 | 21.7 |
| 30 | 29.5 |
| 26 | 15.6 |
| 18 | 9.7 |
| 14 | 16.0 |
| 20 | 7.5 |
| 42 | 31.7 |
| 37 | 13.4 |
| 31 | 22.9 |

※上記のリサーチ・デザインやデータはあくまで架空のものです。

## 4.5 データ処理の手順


START
Step 1
集計前のデータ処理
・番号づけ、コーディング
・集計用紙への転記
・記述データのカテゴリー化
・欠損値の処理　　など
Step 2
データ項目（変数）の尺度の決定
名義尺度・順序尺度・間隔尺度・比尺度
Step 3
一次集計
名義尺度
順序尺度
間隔・比尺度
度数分布
比率
度数分布
中央値
ヒストグラム
平均値・標準偏差(SD)
→分布偏りなし
中央値・四分偏差
→分布偏りあり
Step 4
値の変換を行うか？
No
Yes
段階点
パーセンタイル順位
標準得点（z得点）
偏差値
Step 5
二次集計
② 関係
③ 差
関係か差か？
名義尺度同士
順序尺度同士
間隔尺度同士
比率の差
中央値の差
平均値の差
標準偏差の差
*χ²検定
*連関係数(φ係数)
*ユールの関連係数
など
*名義尺度で使えるもの
*順位相関係数(スピアマン・ケンドール)
*順序尺度で使えるもの
*ピアソンの積率相関係数
*相関比
*χ²検定
*CR検定
*マン・ホイットニーの検定(対応なし)
*サインランクテスト(対応あり)
*t検定(2つの平均値)
*分散分析(3つ以上の平均値)
*F検定
END

## 5. 研究の評価

研究の評価を行う際には、量的アプローチに適した判断基準で、研究の評価を行う必要がある。

研究の判断基準

・ 内的妥当性（internal validity）

測定したい対象が性格に測定できているか。データから結果が導けるか。

→独立変数（原因となる変数）と従属変数（結果となる変数）の間の因果関係について、その因果関係が存在するという記述（因果検証）

・ 外的妥当性（external validity）

結果がどれだけ一般化できるか。条件を変えて研究をしても同じ結果が得られるか。

→研究結果の一般化可能性（研究の結果がどの程度一般的なものであるか）を確認。異なる被験者や違う状況設定での追試実験（一度きりの結果を一般化していないか）

・ 信頼性（reliability）

結果に一貫性があるか。同じ調査方法で、同じような結果が得られるか。

→同じ測定を同じ被験者や同じ状況で実施した時、何度やっても同じ結果になるかという検査法の一貫性。信頼性の測定法として再検査法（ある検査を日時をおいて２度施行した場合の相関）。信頼性の指標として信頼性係数（coefficient of reliability: 例 Cronbach α）。内的整合性の検証としての α 係数。

・ 客観性（objectivity）

中立的で偏りがないか。

→あらゆる被験者や場面に適応する測定。採点者の個人的判断が採点に影響していないこと。客観的検査として、真偽法、多肢選択法、組み合わせ法、完成法。

## 6. 推薦文献

英語教師のための研究法の必読書

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston: Heinle & Heinle.

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.

________ (2001). Using surveys in language programs. Cambridge: Cambridge University Press.

Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. London: Lawrence Erlbaum Associates.

Bachman, L. F. (2004). *Statistical Analyses for Language Assessment*. Cambridge: Cambridge University Press.

英語教師のための研究法の必読書として上記の5冊を推薦したい。リサーチ・クエスチョンの設定から変数の扱い、名義尺度やノンパラメトリック検定を含む統計手法の意義などが解説されている。言語テストや動機づけ研究の第一人者からのメッセージは、傾聴に値する。

最近出版された研究法書籍

Mackey, A., & Gass, S. (2005). *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ellis, R., & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.

Wall, D. (2006). *Studies in Language Testing 22: The impact of high-stakes testing on classroom*. Cambridge: Cambridge University Press.

比較的新しい書籍としては上記の3冊を推薦したい。Mackey & Gassは、リサーチ・クエスチョンの設定からデータ収集、さらにはコーディングからレポート作成におけるチェックリストまで、質的・量的手法ともにかなりを網羅した内容になっている。Ellis & Barkhuizenは、SLAのアプローチにもとづく分析を紹介している（Computer-based analysesはMichael Barlow）。Language Testingシリーズ の Appendixの豊富さには驚かされる（今回はスリランカのプロジェクト、他）。

<u>リサーチデザインの設定</u>

三浦省吾（監修）前田啓朗・山森光陽（編）磯田貴道・廣森友人（2004）『英語教師のための教育データ分析入門』東京：大修館書店

清川英男・濱岡美郎・鈴木純子（2003）『英語教師のためのExcel活用法』東京：大修館書店

リサーチ・デザインの設定に役立つものとして上記の2冊を推薦したい。英語教育の事例が紹介され、実践に即したデータを使って解説されている。Q＆A集や使用するソフトの情報源が掲載され、実証研究に取り組む人にとってはイメージがつかみやすい構成になっている。

<u>心理学研究法</u>

中澤　潤・大野木裕明・南　博文（編）(1997)『心理学マニュアル　観察法』京都：北大路書房

鎌倉雅彦・宮下一博・大野木裕明・中澤　潤（編）（1998）『心理学マニュアル　質問紙法』京都：北大路書房

大野木裕明・中澤　潤・保坂　亨（編）(2000)『心理学マニュアル　面接法』京都：北大路書房

C.ウィリッグ（著）上淵　寿・大塚まゆみ・小松孝至（訳）（2003）『心理学のための質的研究法入門[創造的な探求に向けて]』東京：培風館

心理学研究法の入門書としては上記の4冊を推薦したい。それぞれの手法における基本的な問題だけでなく、集計や統計的解析、尺度構成や信頼性の検討などの問題も含んでいる。『心理学マニュアル』シリーズとして『要因計画法』や『事例研究法』に関するものも刊行されている。

<u>心理統計法の解説書</u>

山内光哉（1998）『心理・教育のための統計法<第2版>』東京：サイエンス社

森　敏昭・吉田寿夫（1990）『心理学のためのデータ解析テクニカルブック』京都：北大路書房

山田剛史・村井潤一郎（2004）『よくわかる心理統計』京都：ミネルヴァ書房

心理統計法の解説書としては上記の3冊を推薦したい。分布や平均の意味からノンパラメトリック検定まで詳細に解説されている。多変量解析は少ないが、特に分散分析については大変詳しい。さらなる分析手法を望む場合には、分野別の解説書が必要となるので注意して欲しい。

<u>統計手法解説書</u>

市川伸一・岸本淳司・大橋靖雄・浜田知久馬（1993）『SASによるデータ解析入門<第2版>』東京：東京大学出版会

石村貞夫（2001）『SPSSによるカテゴリカルデータ分析の手順』東京：東京書籍

柳井久江（2004）『4 Stepsエクセル統計<第2版>』東京：オーエムエス出版

統計パッケージ関係としては上記の3冊を推薦したい。使用するソフトによって使い分けをして欲しいが、いずれもデータの読み方を知る上では有用な本である。統計の知識がないと理解しにくい部分もありが、マニュアルやプログラムの掲載本として活用して欲しい。

<u>英語教育・統計用語辞典</u>

石村貞夫・デズモンドアレン（1997）『すぐわかる統計用語』東京：東京書籍

白畑知彦・冨田祐一・村野井仁・若林茂則（1999）『英語教育用語辞典』東京：大修館書店

Longman (2005). *Longman dictionary of English language and culture: Gets to the heart of the language (3rd ed.)*. Edinburgh : Pearson Education Limited.

用語辞典としては上記の3冊を推薦したい。『すぐわかる統計用語』には英語で論文を書く時の基礎知識が付記してある。Longman Dictionary (3rd ed.) は40,000語の語義と15,000の文化的レファレンスボックスを最新化し、アメリカ文化の項目を増やしている。

## 7. **他の例題**もやってみましょう！

### 7.1 $\chi^2$ 検定 (Chi-square test)

I 先生の研究室に所属していた 4 人のグループは、ある日の NHK 総合テレビで放送された番組を 24 時間録画し、発声された外来語を全て採取・記録した。それら外来語の中で、出現回数が 6 回以上の地名や国名を除く 98 語を選定して、外来語の認識度を調査する質問紙を作成した。中学生 500 人と大学生 500 人を被験者とし、それら外来語の意味が分かるかどうかを 2 件法 (知らない/知っている) で回答してもらった。その結果、「知らない」という回答が 20%を超えた外来語を中学生と大学生それぞれで抽出したところ、以下のような結果が得られた。

| 中学生 | 知らない | 大学生 | 知らない |
|---|---|---|---|
| 1. メセナ | 87% | 1. メセナ | 77% |
| 2. モスク | 73% | 2. インスリン | 49% |
| 3. インスリン | 68% | 3. モスク | 47% |
| 4. ブラッド | 54% | 4. キルト | 44% |

| メセナ | 中学生 | 大学生 | 知らない |
|---|---|---|---|
| 知らない | 457 | 446 | 903 |
| 知っている | 43 | 54 | 97 |
| 合計人数 | 500 | 500 | 1000 |

右上の 2×2 分割表から、中学生と大学生の間には「メセナ」という外来語に対する認知率 (知っている比率) に差があるといえるだろうか。

### 7.2 ウイルコクスンの順位和検定 (Wilcoxon rank-sum test)

K さんはストラテジー使用に関する研究をおこなった。彼は、「社会ストラテジーの使用度は、男子と女子とでは違いがある」と考えている。そこで Oxford (1990) の SILL (Strategy Inventory for Language Learning) を翻訳し、その中で述べられている 6 つの社会ストラテジーの使用度についてのアンケート調査を実施した。回答は 5 段階評定法による自己報告という形式をとり、男子学生 10 人および女子学生 10 人、計 20 人を対象とした。質問の内容は次の通りであった。

1 (全然、あるいはほとんどあてはまらない)、2 (通常、あてはまあらない)、3 (いくらかあてはまる)、4 (通常、あてはまる)、5 (常に、あるいはほとんどあてはまる)

これらの質問に対する回答結果を一覧にしたのが以下のデータ表である。はたして、男子と女子とでは「社会ストラテジー」の使用度に違いがあるといえるだろうか。

(男子)

| No.1 | 質問1 | 質問2 | 質問3 | 質問4 | 質問5 | 質問6 |
|---|---|---|---|---|---|---|
| 1 | 2 | 2 | 1 | 3 | 2 | 2 |
| 2 | 3 | 3 | 2 | 3 | 2 | 4 |
| 3 | 1 | 2 | 2 | 4 | 3 | 1 |
| 4 | 4 | 2 | 2 | 2 | 2 | 3 |
| 5 | 5 | 3 | 1 | 1 | 1 | 4 |
| 6 | 2 | 2 | 4 | 1 | 1 | 3 |
| 7 | 3 | 3 | 2 | 2 | 3 | 4 |
| 8 | 3 | 1 | 3 | 3 | 1 | 2 |
| 9 | 2 | 4 | 3 | 5 | 3 | 2 |
| 10 | 1 | 4 | 2 | 1 | 5 | 1 |

(女子)

| No.1 | 質問1 | 質問2 | 質問3 | 質問4 | 質問5 | 質問6 |
|---|---|---|---|---|---|---|
| 1 | 3 | 4 | 2 | 4 | 4 | 3 |
| 2 | 3 | 3 | 3 | 3 | 4 | 5 |
| 3 | 2 | 3 | 4 | 5 | 5 | 2 |
| 4 | 5 | 3 | 3 | 3 | 3 | 4 |
| 5 | 5 | 4 | 2 | 2 | 2 | 5 |
| 6 | 3 | 1 | 5 | 2 | 1 | 4 |
| 7 | 4 | 2 | 5 | 3 | 3 | 3 |
| 8 | 3 | 2 | 4 | 1 | 2 | 1 |
| 9 | 3 | 3 | 4 | 5 | 4 | 4 |
| 10 | 1 | 3 | 3 | 4 | 5 | 4 |

7.3 スピアマンの順位相関係数の検定（Spearman's correlation coefficient by rank test）

K さんは、英語の話し言葉のデータベースから頻度の高い単語を抽出し、それらの単語親密度を日本人高校生に調査した結果を報告した。親密度の調査に用いた英単語は COLT（The Bergen Corpus of London Teenage Language）をもとに抽出され、被験者である高校生の学習状況などを考慮し、調査する単語レベルを頻度 7 までの上位 2840 語と決定した。次に、これら 2840 語を頻度レベルによって 284 語の 10 ランクに区分し、各ランク 30 語、10 ランクで合計 300 語をランダムに並べた質問紙を作成した。被験者である高校生は、質問紙で示された英単語親密度を 1（なじみがない）から 7（なじみがある）までの 7 段階で評定するように指示された。次の表は、1 位から 50 位までの学習者の英単語親密度順位と、COLT から得られたそれらの単語の頻度順位を示している。

| No | 英単語 | 親密度順位 | 頻度順位 |
|---|---|---|---|
| 1 | MOTHER | 1 | 50 |
| 2 | TV | 2 | 105 |
| 3 | GAME | 3 | 29 |
| 4 | GO | 4 | 2 |
| 5 | THIS | 5 | 7 |
| 6 | PLAY | 6 | 16 |
| 7 | EGG | 7 | 154 |
| 8 | THAT | 8 | 1 |
| 9 | FIVE | 9 | 13 |
| 10 | OR | 10 | 9 |
| 11 | LUNCH | 11 | 83 |
| 12 | ANIMAL | 12 | 162 |
| 13 | THEY | 13 | 6 |
| 14 | TAKE | 14 | 11 |
| 15 | CLASSROOM | 15 | 47 |
| 16 | MR | 16 | 32 |
| 17 | FLOWER | 17 | 232 |
| 18 | END | 18 | 35 |
| 19 | LIBRARY | 19 | 148 |
| 20 | SEA | 20 | 97 |
| 21 | DON'T | 21 | 3 |
| 22 | CLASSES | 22 | 177 |
| 23 | GAVE | 23 | 45 |
| 24 | WALK | 24 | 19 |
| 25 | BEAUTIFUL | 25 | 130 |
| 26 | BUT | 26 | 5 |
| 27 | STORY | 27 | 94 |
| 28 | JUST | 28 | 4 |
| 29 | NEXT | 29 | 22 |
| 30 | BEFORE | 30 | 24 |
| 31 | GOODBYE | 31 | 202 |
| 32 | ATE | 32 | 211 |
| 33 | WRITTEN | 33 | 85 |
| 34 | DESK | 34 | 167 |
| 35 | MAY | 35 | 51 |
| 36 | BEST | 36 | 37 |
| 37 | PAPER | 37 | 48 |
| 38 | THESE | 38 | 17 |
| 39 | HIGH | 39 | 52 |
| 40 | DIFFICULT | 40 | 86 |
| 41 | SPORT | 41 | 100 |
| 42 | HOTEL | 42 | 155 |
| 43 | STAR | 43 | 131 |
| 44 | DEEP | 44 | 203 |
| 45 | NINE | 45 | 38 |
| 46 | WORLD | 46 | 60 |
| 47 | POST | 47 | 126 |
| 48 | BECAME | 48 | 253 |
| 49 | TOTAL | 49 | 178 |
| 50 | SINCE | 50 | 76 |

英単語が使われる頻度と学習者にとってのその単語の親密度とは、いったいどのような関係があるのだろうか？

7.4 重回帰分析（multiple regression analysis）

H 氏は前期の授業終了後に、受講生 30 人を対象に、自分の授業の満足度についてのアンケート調査を実施した。質問の内容は次の通りであった。

（質問 1）授業の内容は受講目的と一致していましたか？　（７段階評定）

１　２　３　４　５　６　７

全く一致していない　完全に一致している

（質問２）授業内容はどの程度理解できましたか？　（７段階評定）

１　２　３　４　５　６　７

全く理解できない　完全に理解できた

（質問３）テキストの満足度はどの程度ですか？　（７段階評定）

| １ | ２ | ３ | ４ | ５ | ６ | ７ |
|---|---|---|---|---|---|---|
| 非常に不満 | | | | | | 非常に満足 |

（質問４）授業全体の満足度はどの程度ですか？　（７段階評定）

| １ | ２ | ３ | ４ | ５ | ６ | ７ |
|---|---|---|---|---|---|---|
| 非常に不満 | | | | | | 非常に満足 |

これらの質問に対する回答結果を一覧にしたのが以下のデータ表である。ここで、質問４の回答結果である全体の満足度を $y$、質問１の回答結果である目的の一致度を $x_1$、質問２の回答結果である内容の満足度を $x_2$、質問３の回答結果である教材の満足度を $x_3$ として、データにもとづき、次のような $y$ と $x_1$、$x_2$、$x_3$ の関係式を求めなさい。

$$y=b_0+b_1x_1+b_2x_2+b_3x_3$$

| No.1 | 質問1 | 質問2 | 質問3 | 質問4 |
|---|---|---|---|---|
| 1 | 2 | 2 | 2 | 2 |
| 2 | 3 | 5 | 2 | 4 |
| 3 | 7 | 7 | 6 | 7 |
| 4 | 2 | 4 | 2 | 3 |
| 5 | 3 | 4 | 1 | 4 |
| 6 | 2 | 3 | 1 | 3 |
| 7 | 3 | 5 | 1 | 4 |
| 8 | 1 | 1 | 1 | 2 |
| 9 | 5 | 5 | 5 | 6 |
| 10 | 5 | 6 | 5 | 6 |
| 11 | 1 | 1 | 2 | 1 |
| 12 | 6 | 7 | 5 | 7 |
| 13 | 4 | 5 | 2 | 5 |
| 14 | 7 | 7 | 5 | 7 |
| 15 | 6 | 7 | 6 | 7 |
| 16 | 4 | 5 | 1 | 5 |
| 17 | 2 | 3 | 1 | 2 |
| 18 | 3 | 4 | 2 | 4 |
| 19 | 4 | 6 | 2 | 5 |
| 20 | 4 | 6 | 1 | 5 |
| 21 | 4 | 6 | 1 | 5 |
| 22 | 7 | 7 | 7 | 7 |
| 23 | 2 | 3 | 2 | 3 |
| 24 | 6 | 6 | 4 | 7 |
| 25 | 5 | 5 | 5 | 6 |
| 26 | 4 | 6 | 2 | 5 |
| 27 | 3 | 4 | 1 | 3 |
| 28 | 5 | 5 | 4 | 6 |
| 29 | 1 | 1 | 1 | 1 |
| 30 | 4 | 5 | 2 | 4 |

**他の例題の検定結果を一部公開**

例題 3. スピアマンの順位相関係数の検定結果

**相関係数**

| | | | 親密度順位 | 頻度順位 |
|---|---|---|---|---|
| Spearmanのロー | 親密度順位 | 相関係数 | 1.000 | .412** |
| | | 有意確率 (両側) | . | .003 |
| | | N | 50 | 50 |
| | 頻度順位 | 相関係数 | .412** | 1.000 |
| | | 有意確率 (両側) | .003 | . |
| | | N | 50 | 50 |

**. 相関は、1% 水準でで有意となります (両側)。

英単語が使われる頻度と学習者にとってのその単語の親密度とは、いったいどのような関係があるのだろうか？→中程度の相関が認められる（$r_s=0.412$, $p<.01$）。

例題 4. 重回帰分析結果

**ﾓﾃﾞﾙ集計**

| ﾓﾃﾞﾙ | R | R2 乗 | 調整済み R2 乗 | 推定値の標準誤差 |
|---|---|---|---|---|
| 1 | .968[a] | .938 | .936 | .47424 |
| 2 | .976[b] | .952 | .948 | .42627 |

a. 予測値:(定数)、質問1。

b. 予測値:(定数)、質問1, 質問2。

**係数[a]**

| ﾓﾃﾞﾙ | | 非標準化係数 | | 標準化係数 | t | 有意確率 |
|---|---|---|---|---|---|---|
| | | B | 標準誤差 | ﾍﾞｰﾀ | | |
| 1 | (定数) | .680 | .206 | | 3.293 | .003 |
| | 質問1 | 1.005 | .049 | .968 | 20.571 | .000 |
| 2 | (定数) | .341 | .222 | | 1.534 | .137 |
| | 質問1 | .741 | .105 | .714 | 7.057 | .000 |
| | 質問2 | .287 | .104 | .280 | 2.767 | .010 |

a. 従属変数: 質問4

**除外された変数[c]**

| ﾓﾃﾞﾙ | | 投入されたときの標準回帰係数 | t | 有意確率 | 偏相関 | 共線性の統計量 |
|---|---|---|---|---|---|---|
| | | | | | | 許容度 |
| 1 | 質問2 | .280[a] | 2.767 | .010 | .470 | .175 |
| | 質問3 | −.087[a] | −1.041 | .307 | −.196 | .318 |
| 2 | 質問3 | .046[b] | .500 | .621 | .098 | .215 |

a. ﾓﾃﾞﾙの予測値:(定数)、質問1。

b. ﾓﾃﾞﾙの予測値:(定数)、質問1, 質問2。

c. 従属変数: 質問4

質問4の回答結果である全体の満足度をy、質問1の回答結果である「目的の一致度」を$x_1$、質問2の回答結果である「内容の満足度」を$x_2$、質問3の回答結果である「教材の満足度」を$x_3$として、次のような関係式を求めなさい。

$$y=b_0+b_1x_1+b_2x_2+b_3x_3$$

↓

（モデル1） $y=0.680+1.005\ x_1$

（モデル2） $y=0.341+0.741\ x_1+0.287x_2$

（強制投入法の場合） $y=0.325+0.668x_1+0.324x_2+0.046x_3$

いかがでしたか　*Practice Makes Perfect！*