

吉田洋一監修
新数学シリーズ

8

統計調査法

統計数理研究所員

西平重喜著

培風館

吉田洋一監修
新数学シリーズ

8

統計調査法

統計数理研究所員
西平重喜著

培風館

西 平 重 喜 略歴
にし ひら しげ き

1924 年 東京都に生まれる
1947 年 北海道大学理学部数学科卒業
現 職 統計数理研究所第2研究部第
1研究室長

主 著

世論調査 (岩波書店 吉田洋一共著)
デマの伝わり方 (東京創元社 ことばの講座 5)
日本人の国民性 (至誠堂, 共編)

© 西 平 重 喜 1957

昭和32年10月25日 初版発行
昭和39年3月30日 初版第9刷発行



新数学シリーズ 8

統計調査法

著 者 西平重喜

発行者 山本俊一

発行所 株式会社 培風館

東京都千代田区九段4丁目5・振替東京 44725

定 価 ￥ 360.

渥美活版印刷・三水舎製本

は　じ　め　に

この本の構成は、‘もくじ’からもわかるように、実際に役立つようにという立場から、まず調査の企画や方法(第Ⅰ章)から始め、サンプリングの理論は第Ⅳ章にまわして、サンプリングの実際のやり方を第Ⅱ章で述べる。つぎに第Ⅲ章では集計や計算について、これも計算を実際にする方法を中心に説明をした。第Ⅳ章のサンプリングの理論もかんたんな例をあげて説明することにつとめたが、ここでは数式を出さないわけにはいかなかった。しかし、この章をとばして読んでも、第Ⅴ章のサンプリングの結果から全体についての推定や検定をする方法についての説明を理解するのに、さまたげとならないであろう。この第Ⅴ章も実例を中心に説明を進めた。

つぎに、もうすこしこの本のねらいをはっきりさせるために、この本がどんな読者を対象として書かれたかを述べよう。第1にはある大会社の人事課のO君である。著者はかれのように統計調査の知識が実際に必要な人に役立つような本にしたいと考えた。そうして、O君に最新の技術を気にするまえに、これだけのことはどうしても知っておいてもらいたいと思うのである。第2には社会学や心理学あるいは政治、経済、法律などの文科系の学生諸君の教科書として使える本にしたいと考えた。事実、この本は著者がそのような人たちに対しておこなった講義をもとにして書かれているし、こんごはこれを教科書に使うつもりである。第3には数学科の学生諸君である。ふつう、数学科では統計が実際に役立つところまで講義がおこなわれていない。一方、世間では数学的な考え方と、実際に数学を使える人を望む声がおきているから、数学科の学生にとっては、副読本のような意味をもつ本にしたいとも考えた。

以上のような読者を対象としたので、著者自身の経験をもとにして、実例ないしは実際におこりうる例によって、書くことにつとめた。したがって、この本は統計調査法という題ではあるが、その例や解説は著者が関心をもつ社会心理学的な諸問題によっている。しかし、社会心理学的といってもべつに

むずかしい概念は持ち出してはいない。いわば新聞の社会面の記事を例にしたようなものであるから、他の領域の読者の理解も深めることになると思うのである。

さきに、実際に役立つようにと念願して、この本を書いたとはいったが、この本はけっして統計調査法の公式集でも、ハンド・ブックそのものでもない。著者が読者に知ってもらいたいと思うことは、統計調査の技術的な結果そのものよりは、その根本にある考え方である。であるから、統計調査の考え方の骨組がすいて見えるような、例をもり込むことにつとめたのである。もし、読者が統計調査の考え方に慣れることができるならば、この本では述べなかった最新の統計理論や、調査技術に出会っても、おおよその見通しがつくであろう。なお、それらの説明を与えるためには、この本でいどのことは十分にマスターしておいてもらわなければならないのである。それだけではなくて、この程度のことをマスターすれば、現実の統計調査に関する諸問題の多くを解明することができるであろう。

なお、この本を書くにあたっては、多くのかたがたのおかげをうけている。ここで、そのお名前をあげつくせないことを残念に思う。

一般的、科学的あるいは数学的な考え方については、吉田洋一先生のお教をうけている。それだけではなく、先生は古屋茂氏、赤撰也氏とともに原稿を何度も読んで、かずかずの注意を与えてくださった。さらに統計学、技術的諸問題やその処理方法については林知己夫氏をはじめ統計数理研究所の諸氏に負っている。釣谷純子さん、石川温子さんはこの本で扱った実例のほとんどすべての集計をしてくださった。また美恵子は相談相手となり、計算をしてくれた。みなさんに深く感謝の気持を表わすしだいである。

1957年春

著 者

も く じ

は じ め に

I. 調査の企画と方法

§ 1. 実験・観測・調査	1
§ 2. 調査の構成	3
§ 2.1. 調査の進め方	3
§ 2.2. パネル調査	7
§ 3. 調査実施の方法	10
§ 3.1. 面接調査法	11
§ 3.2. 留め置き調査法	13
§ 3.3. 郵便調査法	15
§ 3.4. 集合調査法	16
§ 4. 質問のつくり方	17
§ 4.1. 質問および調査票をつくるとき的一般原則	18
§ 4.2. 質問の形式	22
§ 4.21. 自由回答法	22
§ 4.22. 二項選択法	23
§ 4.23. 多項選択法	24
§ 4.24. 順位を決定するための方法	26
§ 5. 調査票のていさいと調査に必要な書類	28

II. サンプルニングのやり方

§ 1. サンプルニング	30
§ 2. 乱数表	32

§ 3. 乱数表の使い方, 単純ランダム・サンプリング	34
§ 4. 等間隔サンプリング	36
§ 5. 2 段サンプリング (サブ・サンプリング)	39
§ 6. 層別サンプリングのかんたんな場合	42
§ 7. 多段サンプリングの場合の層別	48
§ 8. 2 相サンプリング (ダブル・サンプリング)	53

III. 集 計, 計 算

§ 1. 度 数 分 布 表	55
§ 2. グ ラ フ	58
§ 3. 代 表 値	60
§ 4. ちらばりの度合	63
§ 5. モーメント, 歪度, 尖度	67
§ 6. 平均値, 分散, 標準偏差の計算方法	69
§ 7. 相 関 係 数	80
§ 7.1. 問 題 の お こ り	81
§ 7.2. 回 帰 直 線	82
§ 7.3. 相関係数の意義	86
§ 7.4. 相関係数の計算方法	88
§ 8. 相 関 比	96

IV. サンプルング理論

§ 1. 例によるサンプリング理論のあらまし	100
§ 2. サンプルの平均値	104
§ 2.1. サンプルの平均値の理論分布	104
§ 2.2. 分 布	107
§ 2.3. ガウス曲線の性質	109

§ 2.4. 平均値のサンプリング誤差	111
§ 3. 比率(パーセント)のサンプリング誤差	113
§ 4. サンプルの分散などについて	121
§ 5. 層 別 に つ い て	123
§ 5.1. 層別の場合のサンプリング誤差	123
§ 5.2. 各層へのサンプルの割当法	128
§ 5.3. 層 別 の 効 果	132

V. 推 定 と 検 定

§ 1. サンプルでの平均値から, 母集団での平均値を推定すること	138
§ 2. サンプルでの比率(パーセント)から 母集団での比率を推定すること	140
§ 3. サンプル数の決定方法	143
§ 4. 平均値の差の検定	144
§ 4.1. 平均値の差の分布	145
§ 4.2. 有 意 差	149
§ 4.3. 平均値の差の検定の手順	151
§ 5. 2組のサンプルの間の比率(パーセント) の差の検定	152
§ 5.1. その 考 え 方	152
§ 5.2. そ の 手 順	153
§ 6. 1組のサンプルにおける比率(パーセント) の差の検定	155
§ 6.1. その 考 え 方	155
§ 6.2. そ の 手 順	157
§ 7. 分布の間の差の検定— χ^2 検定	158
§ 7.1. どんな場合に必要か	158

§7.2. 適合度の検定	159
§7.3. 分布の一樣性の検定	164
§7.4. 項目の独立性の検定	168
§8. χ^2 分布, t 分布	173
§9. ノン・パラメトリックな検定	174
§10. 順位相関係数, 一致係数	176
あ と が き	181~182
参 考 書	183~184
問 題 の 答	185~186
付 表	
付録第1表 乱数表の一部	187
付録第2表 パーセントのサンプリング誤差の早見表	188
付録第3表 $1/n = 1/n_1 + 1/n_2$ の n の表	189
付録第4表 χ^2 の 表	190
索 引	191~195
略字, ギリシア文字	196

1. 調査の企画と方法

この本では、調査とは統計調査、しかも社会現象の諸問題についての調査をさすことにする。統計調査というからには、多くのものを適当に、まとめあげることになるから、例外的なことを知ろうとするためには、能率のよい方法ではない。また、統計調査ではあまりにもこまかいことは問題にすべきではない。それには特別な技術が必要であるが、それはこの本では扱わない。

まず第1章では調査全体の計画や運営方法について述べる。ただ、被調査者のきめ方については、すこしくわしく話をするために、第11章にまわすことにするが、第1章でも出てくることばの意味だけはここで述べておこう。結果を出したい対象者全体を実際に調査できないとき、一部分をえらび出して被調査者をきめることがしばしばある。そのえらび出すことを**サンプリング**といい、えらび出された被調査者を**サンプル**という。とくにサンプリングのうち、最もすぐれている方法である**ランダム・サンプリング**というのは、クジ引きと同じ原理で、被調査者をきめることである。この本でとりあつかうのは主としてランダム・サンプリングによる調査法なのである。

§1. 実験・観測・調査 調査の意義とか価値ということを考えだすと、大きな問題になりそうである。ここでは実験について、ポアンカレ (H. Poincaré, 1854-1912) がいっていることを引用してみよう*。

“実験は真理の唯一の根源である。実験のみが我々に何か新しいことを教える、実験のみが我々に確実性を与える。

“観測するだけでは充分ではない。これらの観測を利用しなければならないし、それには一般化を行わなければならない。

“よい実験とは何か。それは一つの孤立した事実とは別のことを我々に知らしめるものであり、我々に予見することを得しめる、換言すれば我々に一般化することを得しめ

* 科学と仮説 (河野伊三郎訳) 第9章物理学に於ける仮説、実験及び一般化の役割。

るものである。なるほど、一般化がなければ予見は不可能である。人が仕事をした環境はいつになっても、一度に全部がそのまま再現することはない。だから一度観測した事実はいつになっても再び起ることはあるまい。人の断定しうることは、類似の環境においては類似した事実が生ずるだろう、ということだけである。だから予見するためにすくなくとも類似に助けを藉りなければならぬ。換言すればこれはもう一般化することである。”

これらの引用はどこまでつづけてもきりがない。実験とか観測ということばを調査ということばに読みかえれば、調査の意義や役割がはっきりとすることであろう。ただ実験と調査とは多少ちがう点がある。実験とは十分に——完全に——条件を整えておいてこれに人為的な操作をくわえ、その前後の変化と操作との関係をみようとするものである。自然科学は実験をくりかえすことによって進歩してきたといえるであろう。しかし自然科学のなかでも、とくに天然現象ととりくんでいる天文学、気象学、地震学あるいは海洋学という部門では、実験はごく限られた場合にしかおこなえない。これらの科学は観測によって進歩してきたといつてよいであろう。この観測というのは各種の現象を計量したり、現象があらかじめ考えた理論のとおりに起るか否かを、実地にしらべることである。たとえば、南極における観測という場合、南極で漠然と自然現象を眺めていて、珍しいことが起ったら、それをしらべようというのではない。あらかじめ起るべきことを予想しておいて、それに関する観測をつづけているわけである。

こう考えてみると、実験も観測もあらかじめ仮説をたてて、その仮説が成り立つか、成り立たないかを検証する点では、同じである。ただその検証を人為的な変化によってためすか、自然におこる変化を待ちうけて、ためすかのちがいにあるようである。

社会現象の研究や、社会の実態を知るためにも、実験や観測の態度が必要である。すなわち、漠然と事件の起ることを待っていて、新聞記事のようなものしか書けないし、急にある場所に出かけて行っても、ルポルタージュ文学にしかならないであろう。はっきりとした結論を得るためには、あらかじめ起りうべきこと、ありうべきことを予想し、それを調査するのでなければ、効果的でない。いいかえると、やはり仮説をたてて、その真偽をたしかめるという態度

でのぞむべきである*。

§ 2. 調査の構成 まず、一般的な社会調査の場合の調査の進め方を、順を追って説明し、それから、時間的変化を追う調査の例を述べよう。しかし、被調査者のきめ方（サンプリング）や調査の方法については、あとからくわしく述べるから、それらについては順序としてあけておくだけで、説明ははぶいてある。

§ 2・1. 調査の進め方

1° 調査のねらいをはっきりさせる。いままでの文献や資料、または知識を活用して、問題点をうきばりにしてゆき、質問の形に持ち込むのが第 1 段階の仕事である。調査の範囲や対象、サンプリングの台帳（被調査者をきめるための名簿）、時期、その他の準備をととのえることも、この段階できめてしまうことになる。

2° サンプリング（被調査者の選定）。広い地域——たとえば、東京都 23 区とか日本全国——のような場合には、サンプル（被調査者）個人を中央でえらぶことはできない。したがってサンプリングは 2 段階（以上）に分かれ、中央では第 1 段のサンプリング——たとえば、投票区とか市区町村のサンプリング——をおこない、調査地点を決定する。このとき、必要があれば、サンプリングの精度をあげるために、等質なものからなるグループをつくること——これを層別するという——もある（層別については第 II 章をみよ）。そうして調査地点で第 2 段のサンプリング——個人のサンプリング——をおこなう。第 2 段のサンプリングは、東京での調査のような場合は、企画者があらかじめ区役所などをまわってやってしまうが、全国調査のような場合は、4° の現地作業にふくまれる。

3° 調査員への指示(instruction)。調査票 その他必要書類を調査員に渡し、また、調査のしかたや、ときにはサンプリング（第 2 段）などについて、調査員に指示を与える。このとき、口頭で説明するほかに、簡単明瞭な「調査員のてびき」(interviewer guide) を渡したほうがよい。指示事項は

* 仮説については、ポアンカレの前出書、第 9 章のうち、仮説の役割、をみよ。

できるだけ少なくすむように計画し、スムーズに説明できるようにしておかなければいけない。この指示は、調査の企画者が直接調査員に接しておこなうのがよいわけである。ある全国調査では各地（大学）に調査員を集めてもらって、中央から数名の指示者（instructor）を派遣し、指示者が調査員に直接に指示を与えた。しかし別の全国調査では、各地の責任者（調査員を何人が監督する人、大学の先生）を東京と大阪に集め、調査員への指示は責任者を通じて与えた。方法論的にいっても前者のほうがよいし、費用もほとんどかわらない。

4° 現場での作業（field work）. 指示を受けた調査員は、必要があればまずサンプリングをおこない、つぎにそれらのサンプルを調査する。

5° 回収. 調査がすんだら、指示に従って調査員に調査票や必要書類を返させる。指示者が調査員といっしょに泊り込んでいるような場合には、毎日、調査がすんだ分を回収し、不明をただしたり、必要があれば再調査を命じることもある。しかし、そうでないときは、たんに誤字や読めない字をなおすことにとどめるべきで、調査票を回収する人の主観や調査員の記憶にたよって、調査結果をなおすのはよくない。たとえ答の記録にむじゅんしていることがあるとしても、中央で統一ある処理をするようにしたほうがよい。

なお、このとき調査不能となったサンプルに対しては、その理由を明瞭にしておかなければならない。

6° 集計への準備. 調査が完了すれば、いよいよ調査票に記入された結果をまとめることになる。これを集計するというが、調査票が1枚ですまないようなときには、集計の作業が便利のように、集計用のカードに調査票から必要事項を転記する。このカードには調査票に記載されていることを一言一句正確に書き写すようなことはしないで、内容によって分類し、各分類ごとに記号（符牒，コード，code）を定めておき、その記号だけを転載する。この記号化することをコーディング（coding）という。（たとえば、ある意見に‘賛成’と答えても、‘それがいいですね’といっても、同じくその意見に肯定的であるから、同じ記号をつけるわけであるが、この肯定的、否定的あるいはその中間

的などに分けて、それぞれに与える記号がコードである。)調査のときあらかじめ予想される答の主なものを印刷しておき、そのそばにコードもそえておくと、コーディングの手数ははぶける。これをプリ・コーディング (pre-coding) といい、これに対して調査がすんでからコードを定めることをアフター・コーディング (after-coding) という。

集計には集計用機械を使うもの (機械集計) と、手でカードを扱うもの (手集計) とがある。機械集計では 0 から 9 まで、またはそれに x , y という 10 あるいは 12 のコードが、80 ないし 90 ならんでいるカードに、パンチをして集計をする。このためコードは数字 (x , y をふくむ) にしておかなければならない。これに対して手集計の場合は、コードをアルファベットやかんたんな漢字など、てきぎに使うことができる。(手集計のカードはハガキの 2 倍前後のものが扱いやすい。そして大切な項目はなるべくカードの縁に近いほうにしておく、見やすい。さらに男は青、女は赤などの色をぬると便利である。)

また、機械集計では 80~90 の列 (カラム) があるが、すこし複雑な調査になると、これだけでは足りなくなる。そうして 1 度パンチするとあとから別のコードに打ちなおすことはできない。それにカードが痛みやすい。さらに、まちがいが案外多い。これらの点を考えると、機械集計が威力をあらわすのは、よほど枚数の多いときである。手集計は集計の途中で現状を考慮しながら、適切な分類をしてゆけることなど、研究調査では捨てることができない。

7° 集計. まず各項目ごとに、どういう答 (もの) がいくつずつあったかをしらべることを、単純集計という。そうして 2 つ (以上) の項目の関係をみるために、こちらでこう答えたものが、あちらでどうなっているか、という関係がわかるような集計をすることを、クロス集計 (相関集計, cross-tabulation) という。クロス集計はどの項目間でおこなうかをあらかじめ計画しておいて、順序よく進めてゆかなければならない。

以上が調査実施の各段階での、一般的な仕事および注意であるが、さらにつぎのような吟味調査をおこなう必要もあろう。

8° 調査不能者の分析. ふつうの集計は調査できたものについてだけおこなうが、調査不能者にかたよりがあると、全体に対してゆがんでいる結果になってしまう。そのため、調査不能者がどんな性質のものか、どんなゆがみを持っているかを、なんらかの方法でしらべておかなければならない。

ただし、一般的には調査不能者の意見と、調査できるものの意見に大きく違いがないことが、検証されている。しかし、問題によってはそうはいえないであろう。

また、調査員が不正行為——調査票にかつてな書きこみをしたり、別の人を調査してくること——をしてしまうこともあるが、その影響は案外小さい。しかし、この点にも注意を払う必要がある。

9° 妥当性 (validity) の調査. 調査しようとしたことを、たしかに調査しているかどうかを反省してみる必要がある。たとえば、知能検査は知能をはかるものであるが、検査の結果、高く出たもののほうが、ほんとに本来の知能も高いかどうか心配である。この心配に答えるための調査は当然必要なことであろう。また、‘こんなときは、ふだんどうしていますか?’ というような質問の答は、そのままのみににはできない——たとえば、気取って答えたり、‘どうすべきだ’ という答をするものもあるだろう。収入の調査など、税金をおそれて、安くいうことはしばしばおこることである。そこで、適当な方法で吟味調査をおこない、その妥当性あるいは修正率をしらべることが必要である。

10° 信頼性 (reliability) の調査. 調査の結果というものは日によってちがったり、不安定なものであってはいけな。たとえば、同じことを 2 度調査してみたとき、同じ結果が出なければ、その結果は信頼するにたりない。一般に世論調査や態度調査では、個人個人の意見や態度はあまり安定していないが、全体の構成 (パーセント) はあまり変化しないものである。たとえば数週間をへだてておこなった世論調査で、2 度とも同じ政党を支持しているものは約 70 % であったが、全体としてみた支持政党の % はほとんど変化がなかった。他の項目では同じ答をした人の率はこれより低くなるが、やはり全体の % には変化がない。すなわち、全体的にみた世論の構成 (%) は一般

に信頼性をもっているが、個人個人の意見はそうとうゆれるものである。

また、2, 3 の中学校である知能検査（よく使われているもの）を約 1 年後にくりかえしてみると、相関係数* は 0.8 前後を示しており、そうとうに信頼できるものと考えられる。しかし、だからといってただちにその知能検査は妥当なものである——ほんとに知能をはかっている——ということにはならない。このように、信頼性と妥当性とは別のものではあるが、操作的にはどちらとも区別できない場合もおこってくる。ただ、いつでもこのふたつの立場から、調査の吟味をするように心がけるべきである。

すこし複雑な調査になると、本調査のほかの吟味調査や付帯調査に、費用の半分以上を費やすことさえある。

§2・2. パネル調査 調査の目的によっては、時間的な変化をみるために、くりかえして調査をしなければならないことがある。ただ時間的な変化を全体として——世論としてみるだけならば、各回、別別にそのときの全体をあらわすようなグループをとり、比較をすればよいわけである。第 1 図は日本とソ連の交渉の進行中に、世論がどう変化してゆくかを A, B, C, D の 4 回にわたって調査をしたものである。すなわち：

(1) ‘最近とくに気を配って読む記事はなにか?’ という質問に対して、日ソ交渉の記事をあげたものは、A 調査のときは 35 % あったが、しだいにへって、共同宣言の批准案が衆議院を通過して、参議院にまわった D 調査の時期には、わずか 4 % にすぎない。

(2) ‘アメリカが沖縄を返せば、ソ連も南千島ぐらい返すだろうか?’ という質問を A 調査のときにしたが、その後、逆に‘南千島をゆずると、沖縄は返さない’というようなことが、アメリカ側から伝えられた。その結果、沖縄と南千島とが関係あるというものは、C 調査では 42 % に達したが、さらに政府やアメリカ側の弁明などがあつたけれど、D 調査では 38 % であつた。

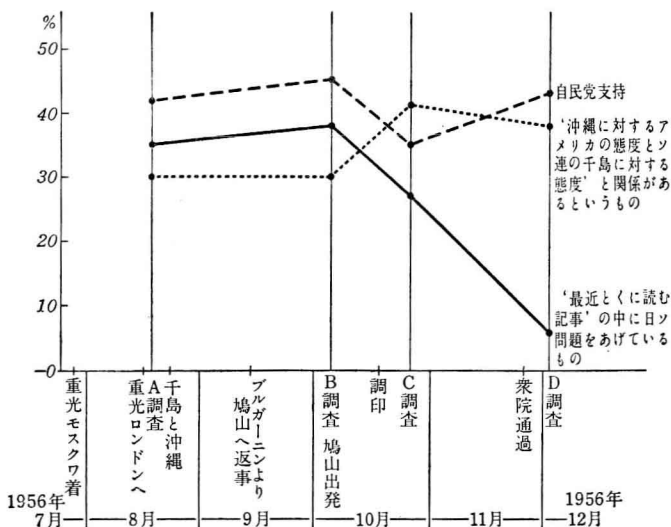
(3) ‘支持政党はなにか?’ という質問に対して、自由民主党を支持

* 第 III 章をみよ。ふたつの量の関係の深さをしらべるときにつかう。絶対値が 1 にちかいほど関係が深い。

第1図 日ソ交渉の世論調査

【統計数理研究所】

台東区、目黒区 各回約300サンプル



第2図 参戦前のアメリカの世論

