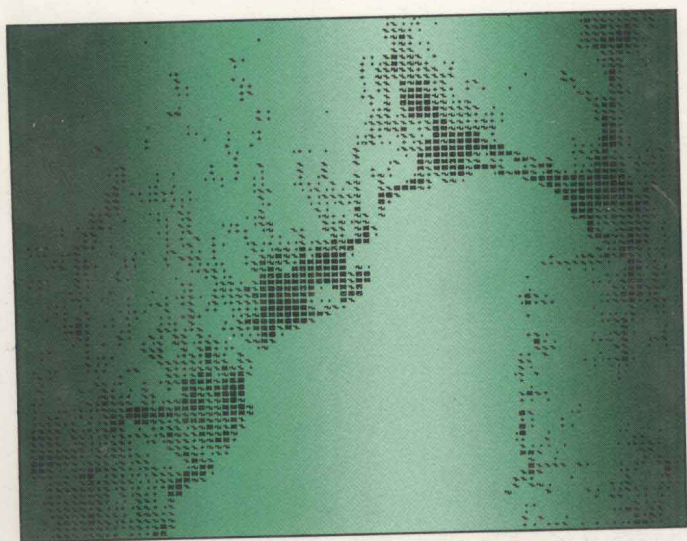


統計データの見方・使い方

上田 尚一 著



朝 倉 書 店

統計データの見方・使い方

上田 尚一 著

朝 倉 書 店

著者略歴

1927 年 広島県に生まれる
1950 年 東京大学工学部応用数学科卒業
現 在 総理府統計局統計参事官

統計データの見方・使い方

1981 年 10 月 20 日 初版第 1 刷

著者	うえだ しょういち 上田 尚一
発行者	朝倉 邦造
発行所	株式会社 朝倉書店 東京都新宿区新小川町 2-10 郵便番号 162 電話 03(260)0141 振替口座 東京 6-8673 番

<検印省略>

© 1981 <無断複写・転載を禁ず>
3041-110428-0032

中央印刷・渡辺製本

まえがき

統計的考え方

“統計的考え方は、よみかき能力と同じく、すべての人にとって、必要な常識となるであろう”

イギリスの文明批評家、SF作家である H. G. ウェールズがこういったそうです。私もそうだと思いますが、わが国の現状はどうでしょう。H. G. ウェールズのいう状態になっているのでしょうか。

まだそうっていないと思うが故に、そして、そういう状態になるのが望ましいと思うが故に、この書物をまとめてみました。

統計データの 見方・表わし方

“統計的考え方”が、すべての人の常識になっていないとすれば、数学が出てくることが一つの障害かもしれません。

数学が不必要だというのは言いすぎですが、統計的考え方を身につけるにはそれ以上に重要なこと——統計データの見方・表わし方——があり、そのことが看過されている点を問題視したいのです。統計データは数字で表わされていますから、その扱い方に数理的な手法が入ってきます。しかし、数字や数理を使って表現し、分析しようとするのは現実の社会事象や自然現象ですから、現象そのものの見方、表わし方をベースにおいた上、そのための手段として統計データの見方、表わし方を学ぶべきです。その観点で学んではじめて、統計データや統計手法が活かされるわけです。そして、統計的考え方が身につくわけです。

こういう観点で学ぶ人に役立つことを期待して、本書は、“現象を表現し、その意味をよみとる手段”として、統計データの見方・表わし方を解説します。

探索的データ解析

統計学の構成としても、データの見方・表わし方は重要な分野です。

統計学は、“記述”の方法論として生まれ、“推測”の方法論という面を包含する方向に発展してきましたが、基礎としての“記述”の重要性がかわったわ

けではありません。研究者の立場で新しい面に目を向けるのは自然の成りゆきでしょうが、統計教育あるいは実社会での適用の場面では、基礎を十分大切にしなければなりません。わが国の現状では、“記述”という基礎の部分に対する考慮が必ずしも十分でないと思うのです。

最近の傾向として“記述”の観点での分析手法(探索的手法)の重要性を再認識し、それを推測の観点での分析(検証的手法)に並ぶものとして位置づける動きが出てきました。この探索的手法では、データの見方・表わし方の論理が分析をリードすることになります。データの見方・表わし方は、こういう意味で、データ解析の基礎にある論理を学ぶ分野です。

本書はこの立場に立っており、続編として刊行する予定の『データ解析の仕方』と合わせて、探索的データ解析の入門書としても活用できると思います。

本書の構成

本書で取り上げる内容が、実世界の現象の見方・表わし方ときりはなせないものであることから、解説は、できるだけ事例、または、それに近い例を使って進めました。また、本書で取り上げる統計手法についても、問題ごとに考えるべき部分が多く、機械的に適用できる公式があるわけではありません。したがって、考え方を身につけることが必要であり、そのためにも、例題による解説を中心におく編成を採ったのです。

本書の母体は、国および地方自治体の統計職員を対象とする研修で筆者が使っていた講義メモです。研修の場合、問題(本書の例題)を研修生に提示し、種種の角度から討議した上、最後にまとめをする——こういうすすめ方を採用していました。本書は自習用ですから討議に当る部分を文章化してありますが、理解するためには、それを目でよんでいくのではなく、自問自答しつつ頭でよんでいく——そういうよみ方をしていただくことが必要です。

そうしてはじめて、例題で示したポイントを一般化して、各自が当面する問題に適用する力——それが統計的思考方——が身につくのです。

皆さんの問題を解決していくために本書がいくらかでも役立てば幸いです。

なお、目次のあとに、各章の内容要約および基本概念の関連図を示してあります。

統計学，統計教育を専門とする方へ 社会に出て実務につく人々に対する統計教育では，数理のほかに，現象を記述し，その意味をよみとる手段として，または，探索的データ解析の基本として，“統計的考え方”を取り上げることが是非必要だと思います。この面が不十分なため，統計学の活用がすすまないのではないのでしょうか。問題は取り上げ方です。本書は，特定の予備知識なしでもよめる副読本として使えるよう考えて編成しました。取り上げる内容や説明のくみ立て方など，考えればまだまだ改善を要する点があるでしょう。こういう面に関心をもつ方のコメントをいただきたいと思います。

1981年9月

上 田 尚 一

各 章 の 内 容

1. 第1章では、統計データとは何かを考えます。ただし、抽象化した一般論を展開するものではありません。“統計→数字”という短絡した理解ではなく、“表わそうとするもの”があり、それを“表現するために”数字を使うのである——こういう理解をはっきりさせておきたいのです。

2. 第2章では、“比率の使い方”，とくに，社会事象を表現するための手段としての使い方を考えていきます。分子，分母のえらび方如何が中心になりますが，事象の表現という観点でみていくと，種々，注意を要する点があります。

3. 第3章では，対比分析の手段として比率を使う場面を想定し，構成比，相対比とよばれているタイプの比率について，その形式的な定義とは別に，使い方という意味で区別されるさまざまな場合があることを示します。本来は，使い方を考えた“タイプわけ”が必要ですが，定説はないようです。

4. 第4章も第3章と同じ観点に立ちますが，とくに時系列データの場合をとり上げます。指数，変化率の見方，変化の要因をわけて示すための寄与率の求め方など，“時系列データをよむ”場面での比率の使い方を考えます。分析とよんでいい範ちゅうの使い方につながります。

5. 統計データには，二つのタイプ——ストックとフロー——とがあり，その区別は重要です。第5章では，その区別を説明した後，分子分母にこれらを組み合わせた比率の見方について，注意を要すること，および，うまく使えば有益な指標となることを述べます。たとえば，状態の継続期間を表わす指標を導出できます。

6. 第6章では、因果関係の分析に言及する場面で使う比率について、考えます。分子分母の範囲を形式的に合わせればよいというものではありません。形式的な対応でなく、因果の対応を考えに入れなければならないのです。これに関連して“発生率”を特別のタイプの比率と位置づけております。

7. 第7章では“データの求め方”と“比率の表わし方”との関連についてとくに注意を要するケースを挙げていきます。事象の見方に対応するデータがとれば問題ないのですが、データの求め方の側の制約から、必ずしも事象の見方にぴったりした数字がとれない場合があるのです。

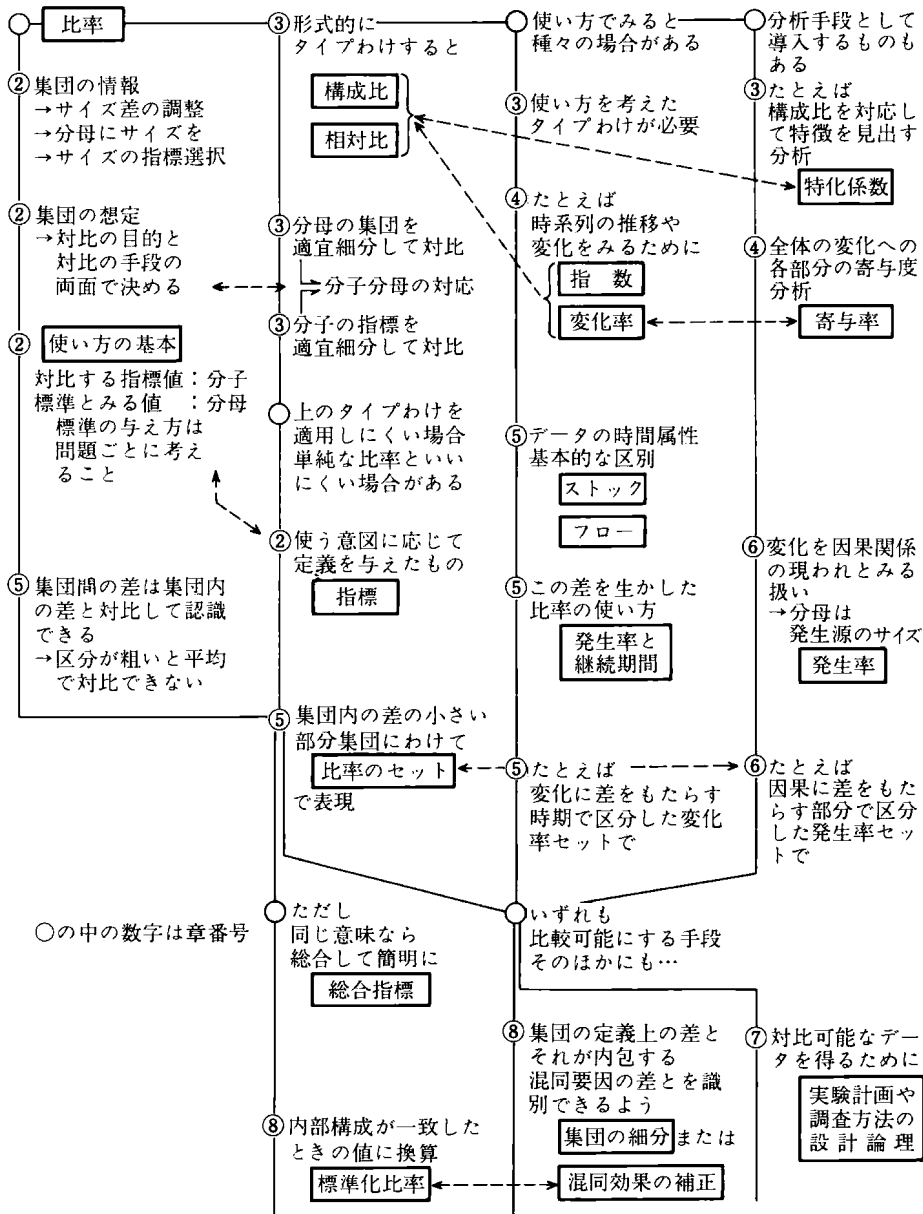
8. 第8章は、第6章のつづきです。因果関係に言及しようとするとき、“説明したいこと”が“説明されている”とは限りません。たとえばある集団区分で対比したつもりであっても、第3の要因の差がその区分に重なっており、結果がどちらの効果か判別できない場合があるのです。そういう場合に対処するためのデータの表わし方・求め方を述べます。

9. 第9章は、本書でとり上げた“データの見方・表わし方”の意義について述べます。統計は“統計数学”ではなく、社会事象を表現し、分析するための手法ですから、数理的な側面だけでなく、その基礎の考え方として身につけておくべきこと——それが、統計データの見方・表わし方です。

10. 付録として、統計データの定義をよむ上の注意点、統計データを求めるための手引きなどを説明してあります。

11. このほか、各章末にいくつかの補注を示してあります。これは、必ずしも各章に対する補足ではなく、補足全体で“統計データの表現および対比の論理”を説明したものになっています。本書で取り上げたのは、そのうち、主として分類区分として表現されたデータの扱いです。

基本概念関連図



- 比率の見方・使い方は、本書で取り上げる内容の主要部分です。したがって、各章で、順次、視点をひろげつつ話を発展させております。その体系を関連づけて示すのが左ページの表です。
- 一つの章の説明を十分よんだ上この体系図上の位置づけを確認して次の章へすすむとよいと思います。図で○でかこんだ数字が、その事項を説明してある章の番号です。
- なお、比率のタイプわけについては、通常採用される形式上の区分に、使い方の観点での区分をおり込む必要があります。そこを含めた体系づけは必ずしも定説になっていないようですが、図示した体系づけが考え方をフォローするために最も効果的だと思います。

現代人の統計

本にはいろいろの型がある。理論体系をしっかりと明確に書いた理論的な本、実用書的な本といった分類もその一つである。理論的に書いた本を読んでも、直ちに右から左へと役立てることは難しい。これは当然のことであると思う。こうした本は、基礎をきっちり身に付けるのに役立ち、ポテンシャルを蓄えるという機能を持つものであるが、問題意識の不明確な書き方のものでは、冷い空々しさのみが残る。実用書はその書く通りやって行けば何かが出来るようになってはいるが、これは読むだけではやはり解るものではなく、その通りやって行くうちに実用のコツを自ら覚えるという努力をしなくてはならない。統計学のようなもので、実用書的な書き方を希望する人も多いが、やはり本質を見失って、知らず知らずのうちに誤りを犯すことになる。しかし折衷というのは、二兎を追うことになってうまくない。理論的にしっかりと書くが、その視点を「役に立つ統計」というところに置くという行き方が望ましい。つまり役に立てるという考え方を軸にして、統計理論をしっかりと書き分けるという立場である。

これはなかなか難しいことであるが、いままでの統計学の書にあきたらなく思い、何とかしようという意欲のある新進気鋭の統計研究者が集ってこの叢書をまとめることになった。出来上ったものは、これまでにない新鮮な内容で随分ユニークなものとなったと思う。

しかし、本はあくまでも本であって、これをポテンシャルにして冷暖自知、工夫を凝らしてデータ解析に生かすのは読者諸氏にあると思う。その種子はこの叢書の中にあるので育てあげて欲しいものである。 [林 知己夫]

- | | | |
|----------------|-----------------|-----------|
| 1. 統計解析法の原理 | 鈴木義一郎著 | 定価 2300 円 |
| 2. 多変量解析法 | 柳井 晴夫著
高根 芳雄 | 定価 2000 円 |
| 3. 社会調査の基礎 | 杉山 明子著 | |
| 4. 意志決定の基礎 | 松原 望著 | 定価 2000 円 |
| 5. 実験計画法の基礎 | 早川 毅著 | 定価 2000 円 |
| 6. データ解析と管理技法 | 大隅 昇著 | 定価 2500 円 |
| 7. 多元的データ分析の基礎 | 駒沢 勉著 | 定価 2300 円 |

上記定価は 1981 年 10 月現在のものです。

人間科学の統計学

本シリーズは、米国において非常な好評を博し、多くの版を重ねている
“Series: Quantitative Applications in the Social Sciences” (SAGE
Publications) の日本版です。各種の数理的手法を、統計学や数学の専門的な知識を必要としないで理解できるよう、簡潔にわかりやすく、またきわめて実的にまとめられております。多忙な現代人にとって待望のシリーズの誕生です。

1. **多次元尺度法** クラスカル・ウィッシュ著 高根芳雄 訳 価1400 円
2. **因果分析** アッシャー著 広瀬弘忠 訳 価1400 円
3. **生態学的推論** ラングバイン・リヒトマン著 長谷川政美 訳 価1300 円
4. **探索的データ解析の方法** ハートウィグ他著 柳井晴夫・高木広文 訳
価1800 円
5. **多重指標分析** サリバン・フェルドマン著 広瀬弘忠・石塚智一 訳
価1900 円
6. **統計的検定** ヘンケル著 松原 望・野上佳子 訳 価2300 円

[以下続刊]

テストの信頼性と妥当性 カーマイン他著 水野欽司・野嶋栄一郎 訳

パネルデータの分析 マルクス著 水野欽司 訳

コーホート分析法 グレン著 藤田英典 訳

時系列解析—ダイナミックプロセスの分析 オストローム著 松原 望 訳

多変量相関分析の方法 レビン著 柳井晴夫 訳

質的データの解析法Ⅰ—名義尺度データ レイノルズ著 中村 隆 訳

質的データの解析法Ⅱ—順序尺度データ ヒルデブランド他著 大隅昇 訳

定価は 1981 年 10 月現在のものです。

東工大教授
工学博士 日野 幹雄著

統計
ライブラリー **スペクトル解析**

A 5 判 312頁 価3800円

〔内容〕ランダム変動の表現・自己相関関数とスペクトル、相互相関とクロス・スペクトル、定常性、情報エントロピーとスペクトル、線型システムの理論、スペクトル計算の誤差理論、データ処理の手法、さらにすすんだスペクトルの概念。

京大助教授
経済学博士 佐和 隆光著

統計
ライブラリー **回 帰 分 析**

A 5 判 196頁 価2300円

統計手法のうち最も実用に供されることの多い回帰分析の方法を、豊富な実例と厳密な数学的証明をおりまぜて解説。〔内容〕回帰分析への誘い、ベクトルと行列、多変量正規分布、線形回帰モデル、仮説検定・区間推定・予測、説明変数の問題

名大教授
医学博士 柏木 力著

統計
ライブラリー **医学統計解析**

A 5 判 248頁 価2800円

医学・生物学の領域においてその統計的方法はここ数年来の電子計算機の著しい普及によってますますその応用範囲を広げている。本書は臨床医学出身の著者が具体的な例題を中心に、きわめて明快平易に实际的にまとめあげた医学統計の入門書

後藤昌司・畠中駿逸・田崎武信訳
統計ライブラリー

コックス **二 値 データ の 解 析**

—医学・生物学への応用—

A 5 判 236頁 価3000円

医学・薬学・生物学・農学系の研究者、技術者のための適用例が数多く収められ、さらに日本版にはプログラムも加えられ、実際に役立つようまとめられている。〔内容〕線形ロジスティックモデル、単一パラメータの解析、複雑な解析、他。

北大助教授 齋藤 堯幸著

統計
ライブラリー **多次元尺度構成法**

A 5 判 256頁 価2700円

心理学、社会学、政治学をはじめとした社会科学から、医学、生物学を主とした自然科学に至るまで広い範囲にわたってその応用範囲を拡げている多次元尺度構成法について応用例を中心にしてみわめて实际的にかつ懇切丁寧にまとめられている

東工大教授
Ph.D 池田 央 立教大助教授
工学博士 岡太彬訓訳
統計ライブラリー

アブトン **調査分類データの解析法**

A 5 判 180頁 価2500円

統計調査に必然的に生ずる2元分類および多元分類表の統一的分析法を初学者にもわかるように平易に解説。クロス分類表の古典的連関分析から始まって近年急速に発展した対数線形モデルを詳解調査の実験家や統計調査に係る研究者に好適の書

脇本和昌・後藤昌司・松原義弘著

多 変 量 グ ラ フ 解 析 法

A 5 判 208頁 価2500円

多次元データのもつ意味を総合的に伝達また解析する手法として開発された顔形グラフ、体形グラフ、星座グラフ、木形グラフ、パイプロット法、非線形変換グラフ、樹形図等の多変量グラフ解析法を平易に解説。巻末にプログラムを付した。

大山 正・武藤真介・柳井晴夫著

現代の心理学

行動科学のための統計学

A 5 判 228頁 価2500円

心理学を中心に行動科学の領域において、その統計的方法はますます応用範囲を広げている。本書は具体的な例題を中心に、直観的・感覚的の把握ができ、現実の問題に的確な適用ができるよう細かい配慮がなされ極めて实际的にまとめられている

定価は 1981 年 10 月現在のものです。

目 次

0. はじめに	1
1. 統計データ	5
問 題	10
2. 比率, 指標	11
2.1 比率の機能——サイズをそろえる	11
2.2 比率のつくり方(1)——集団区分のとり方	12
2.3 比率のつくり方(2)——分母データのえらび方	18
2.4 指標——表現したいことを表わす尺度	20
2.5 比率——傾向値に対する相対比	21
問 題	25
3. 構成比と相対比	26
3.1 構成比と相対比	26
3.2 サイズ効果を除去するための相対比	27
3.3 指標を細分してみる	30
3.4 内訳区分どうしを対比するための相対比	31
3.5 内訳区分の構成比の比較	31
3.6 標準構成比に対する相対比——特化係数	33
問 題	35
4. 指数と変化率	37
4.1 指数, 変化率	38

4.2	変化率のよみ方	41
4.3	寄与率	44
4.4	地域変化を表わす指数, 変化率	48
	問 題	50
5.	ストックとフロー	52
5.1	ストックとフロー	52
5.2	フロー/ストックの形の比率	56
5.3	モデル	59
5.4	滞留期間の情報	61
5.5	遷移確率	65
	問 題	70
6.	因果関係の表現	72
6.1	発生率と変化率	72
6.2	タイムラグに対する考慮	80
	問 題	83
7.	データの求め方との関係	85
7.1	実験群と対照群	85
7.2	追跡調査と回顧調査	90
7.3	クロスセクション分析とコーホート分析	94
7.4	OD 調査	96
	問 題	101
8.	比率の解釈	103
8.1	混同要因	103
8.2	集団の細分とクロス集計	107
8.3	クロス表の効果	109

8.4 比率の標準化	112
8.5 DK の扱い	115
8.6 データの解釈と分析	119
問 題	120
9. なぜ統計データの見方、表わし方を学ぶか	123
9.1 統計データの性格	123
9.2 統計情報の体系的整備のために	123
9.3 問題解決の手段として	126
9.4 むすび	129
付 録	
A. 統計データ利用上の注意	130
B. 統計データ利用の手引	142
C. 補足資料	152
参 考 資 料	169
索 引	171

①。 は じ め に

① 統計データの見方は、よみかき能力と同じく、誰にとってもひとしく必要な知識です。決して、“特定の専門家だけが知っていればよい”というものではありません。誰にとっても身近なものであり、世の中の現象を理解するために必要な情報、あるいは方法です。

② 図0.1をみて下さい。女性の一生を戦前と比べたものです。現在の女性のライフサイクルでは、平均でいって、35歳に末子が就学し育児から解放されます。その後、それまで以上に長い40年もの期間が残っております。現在の女性が、このめぐまれた期間をどうすごすかは、当然、重要な関心事です。

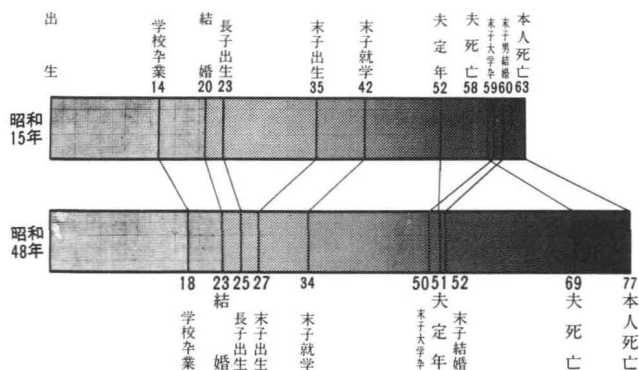


図 0.1 女性の一生いま・むかし¹⁾

- 1) 死亡時における年齢は、男 25 歳，女 20 歳時における平均年令を使って計算

これだけ大きい変化があるのですから、個人レベルの問題としてすましておけません。たとえば、かなりの女性は再就職を望むことでしょうが、これを受入れる社会の側としては、高齢化社会への対応としての定年延長の問題もから