

經濟演習双書 9

統計學

全訂版

寺尾琢磨 監修

學文社刊



統 計 学

寺尾 琢磨 監修

浜田 文雅 佐野 陽子
佐藤 保 鳥居 泰彦 著

東京 学 文 社 発 行

著 者 紹 介

寺尾 琢磨 慶応義塾大学名誉教授
浜田 文雅 慶 応 義 塾 大 学
佐野 陽子 慶 応 義 塾 大 学
佐藤 保 慶 応 義 塾 大 学
鳥居 泰彦 慶 応 義 塾 大 学

経 済 演 習 9 統 計 学

昭和43年 1 月 30 日 改訂第一版発行

昭和55年 6 月 20 日 改訂第六版発行

著 者 寺 尾 琢 磨 監 修

印刷者 亨有堂印刷株式会社

発行所 株式 学 文 社
会 社

代表者 北 野 登

東 京 都 目 黒 区 中 目 黒 1-2-6

電話(715)1501(代) 振替東京 3-98842 番

(落丁・乱丁の場合は本社でお取替します)

3333—102004—1023

は し が き

大学ともなれば、どんな科目をとってみても多かれ少かれ難解で、やさしいといえるようなものはあるはずもないが、経済学部で特に学生の頭痛の種となっているのは、統計学という科目であるらしい。慶応の経済学部では統計学は第二学年の必修科目だが、その学年度の不合格者の可成りの部分は、統計学でつまづいた連中で占められるのが毎年の慣わしである。これは統計学なるものが実際にむずかしい学問だからだが、わたしはもう一つの理由があると思う。それは講義の形式がこの学問の真の理解に必ずしも適当でないということである。現在、慶応に限らず、恐らくすべての大学の経済学部で、統計学は必修科目となっている。これは今日の経済学の性格から当然のことで、統計を離れて経済現象を理解することは到底できない相談だからである。ところが必修科目ともなれば、学生数は当然ふえるはずだが、統計学のような特殊の学問に関しては、講義を担当できるような専門家はざらにあるものではない。勢い大きな教室でマスプロ的講義を行なわざるを得ないのが実情である。しかしこれは邪道も甚だしい。統計学の学修に最も必要なことは、実際の問題を実際に計算してみることである。空粗な理論や抽象的な公式を積み重ねただけでは統計学は無意味であろう。統計学は複雑な現実を正確に理解するための手段でなければならないからである。

そこで学生は自ら進んで実習してみる努力が肝要である。小人数の教室なら実習を交えた講義も可能だが、上にのべたように、今日の大学ではまず以って望めないからで、講義担当者からは甚だ申しわけない次第である。そこで慶応の当事者達は幾分でも学生の負担を軽減しようという殊勝な考えから、今回こ

の小著を編集したわけである。

本書は最も多く現われるであろうような問題を選択し、その一つ一つに最も新しい理論と計算を示したもので、形はコンパクトだが、ほとんど統計学の全域を網羅している。もちろんこれで統計学そのものを卒業したことにはならないが、実用には充分で、少なくとも諸君は、統計学を学んだと広言して差支えないと信ずる。

今から数十年前には、慶応では統計学の担当者はわたし一人であった。いま優秀な後継者が続々と輩出し、高度の研究に、また学生の指導に精進している姿を見うることは、わたしにとってこれ以上のよろこびはない。本書を執筆したのはこれら新進学徒の一部でしかないが、しかし他の人達も多かれ少かれ本書の完成に関与している。わたしは、本書がわが慶応義塾ばかりでなく、広く他の大学においても活用され、統計学の普及と進歩に大きな役割を果たすであろうことを希望し、また確信するものである。

昭和 43 年 1 月

寺 尾 琢 磨

目 次

第 1 章 統 計 学 史

1. 1	統 計 学 史	1
1. 2	計量経済学史	4

第 2 章 母 集 団 と 標 本

2. 1	母集団の考え方	7
2. 2	グ ラ フ	10
2. 3	標本の選び方	12

第 3 章 度 数 分 布

3. 1	度数分布の描き方	13
3. 2	累積度数分布	19

第 4 章 標 本 平 均 と 標 本 標 準 偏 差

4. 1	平均, 分散, 標準偏差の計算	25
4. 2	組分けした資料の計算	29
4. 3	記号の説明	33
4. 4	算術平均と幾何平均	35
4. 5	標準偏差の性質	36

第 5 章 確 率 論 の 初 歩

5. 1	順列, 組合せ	37
5. 2	確率の計算	40
5. 3	数学的期待値	44
5. 4	独立の概念 (1)	45

5. 5	独立の概念 (2).....	48
------	----------------	----

第 6 章 確 率 分 布

6. 1	離散型の分布.....	51
6. 2	連続型の分布.....	55
6. 3	分布の平均・分散の計算及びグラフ(1).....	58
6. 4	分布の平均・分散の計算及びグラフ(2).....	61
6. 5	標 準 形.....	63

第 7 章 母集団の分布と標本の分布

7. 1	二 項 分 布.....	65
7. 2	二項分布の計算.....	68
7. 3	ポアソン分布.....	71
7. 4	正 規 分 布.....	78
7. 5	確 率 紙.....	81
7. 6	正規分布のあてはめ.....	84
7. 7	χ^2 分布, t 分布, F 分布.....	100
7. 8	正規分布を用いる計算.....	108

第 8 章 標本抽出の初歩

8. 1	標本平均の平均と分散.....	109
8. 2	標本抽出の練習.....	112
8. 3	標本抽出の計算(1).....	115
8. 4	〃 (2).....	118
8. 5	〃 (3).....	119
8. 6	〃 (4).....	121

第 9 章 母数の信頼限界

9. 1	二項母集団の p の信頼限界.....	123
9. 2	標本数の問題.....	125

9. 3	母平均の信頼限界(1).....	127
9. 4	" (2).....	130
9. 5	母分散の信頼限界.....	132
9. 6	有限母集団の場合(1).....	133
9. 7	" (2).....	135

第 10 章 統計的有意性の検定

10. 1	母平均の検定.....	139
10. 2	母分散の検定.....	139
10. 3	平均の差の検定 個体の比較と群の比較(等分散の検定を含む).....	140
10. 4	百分率の検定(1).....	145
10. 5	" (2).....	149
10. 6	大標本の場合と小標本の場合.....	150
10. 7	計算の注意.....	153

第 11 章 回帰分析と分散分析

11. 1	最小自乗法.....	155
11. 2	回帰係数の検定.....	161
11. 3	二変数のグラフ.....	164
11. 4	二変数のパラメーターの計算.....	165
11. 5	独立変数と誤差の相関.....	168
11. 6	従属変数と誤差の相関.....	169
11. 7	回帰係数の信頼限界.....	170
11. 8	相関係数の検定.....	171
11. 9	回帰係数の性質.....	172
11. 10	マルチコリニアリティー.....	176
11. 11	自己相関.....	180
11. 12	最尤法.....	184
11. 13	常数のない場合の回帰.....	193
11. 14	二種類の過誤.....	195

11.15	順位相関と一致係数	202
11.16	季節変動	206
11.17	分散分析の手法（二つ以上の平均の一様性の検定）	208
11.18	回帰係数の差の検定	214

第 12 章 経済分析への応用例

12.1	消費函数の計測	218
12.2	二段階最小自乗法	227

第 13 章 Bayes（ベイズ）推定

13.1	ベイズの定理	234
13.2	ベイズの定理による計算	236
13.3	不良率の計算	237
13.4	主観的確率	238
13.5	連続分布の場合	242

〔付録〕	行列の説明	245
	付 表	250

第1章 統計学史

1.1 統計学の歴史

〔解答〕 統計的ないろいろな調査は国家の成立と共に国家運営の必要上行なわれた。そこで統計学も最初は国状記述の方法として生まれたのである。その意味ではエジプト時代から統計調査があったわけであるが、国勢記述の学として確立したのはずっとあとでドイツであった。ヘルマン・コンリングによって国家の記述がまとめられたが、その門弟アッヘンワルははじめて Statistik という文字を使った。したがって普通統計学はアッヘンワルをもってそのはじめとするのである。そしてこの学派は国勢学派と呼ばれている。国勢学派は国家の記述を四原則にしたがって記述するというような方法をつくりだしたが、その記述は数量的ではなかった。一方英国では政治算術という別の一派が生まれていた。そしてこの表や数を取りあつかう人達とそれをこぼむ国勢学派の間にはじめは論争が行なわれたが、次第に政治算術が行なう方法の優位が認められるようになり 国勢学派もそれを取り入れるようになった。1850 年 クニースが「独立の学問としての統計学」をあらわし、政治算術こそ統計学の名に値するものと説くに及んで、国勢学は統計学として終りをとげた。そして賭の理論から端を発した確率論の発展と結びついて、政治算術は統計学の主流となっていたのである。政治算術の中ではジョン・グラントとウィリアム・ペティがもっとも有名である。ジョン・グラントは 1662 年「死亡表に関する自然的政治的観察」を著した。彼の研究をペティが受けつぎ政治算術という名称はペティの著書に基づくのである。ついでハレーがさらに科学的に人口統計を取扱い生

命保険の基礎となる生命表をつくった。ジュース・ミルヒは人口統計の規則性を追求し、そこに神の秩序を見出そうとした。政治算術は数量を取扱って規則性を見出そうという点で進んだ考え方であったが、主として人口統計にとどまっていた。一方賭博の理論に興味をもった数学者、パスカル、フェルマ、ベルヌーイ等が確率論を発展させたが、ベルヌーイは「大数の法則」の原理を確立した。しかし統計数値を扱う統計学者の方が数学的知識がなかったため、すぐには応用されなかった。ラプラス、ラグランジ、ガウス等の数学者によって数学理論は益々進歩したが、この応用を開拓したのがケトレーであった。ケトレーは統計学を社会現象に応用した点で大きな功績があったが、あまりに統計的法則を強調した結果、社会現象も物理的法則の下にあり、すべての攪乱の要素をとり除いた平均人を考え、この平均人の性質を求めることに重点がおかれたので社会物理学ともいわれている。そして統計万能時代が出現したのである。これに対する反省は一つは社会統計学、他は方法論派によって行なわれた。ドイツの社会統計学派は国勢学派の流れをくみ、大量観察によって社会現象の規則性を見出そうとし、これが統計学であると考えた。エンゲル、ワグナー、マイル等の学者がある。方法論派は統計的方法すなわち資料の蒐集・整理および解析の方法を研究することを目的とする。したがってこの方法はあらゆる集団現象に適用されなければならない。この派の基礎をきずいたのはフランシス・ゴルトンでありその門弟のカール・ピアソンであった。RA・フィシャー、ネイマン、ワルト等によって、さらに厳密に集団の性質を標本から推定する理論が展開された。もう一度簡単にくり返すと、統計学は17世紀の半ばより実質的に登場し、18世紀になってその名称を獲得したが、単なる国勢の記述と人口統計であった。18世紀の後半になって確率論と結びついた社会物理学が盛んとなり統計万能時代となったが、19世紀はその反動と反省が起り数学的理論の

一層の進歩は20世紀にいたって近代統計学といわれる推測の理論をつくりあげた。しかしそれは主として、生物学に応用されたので経済現象に適用されるようになったのは、この30年間ぐらいである。

1.2 計量経済学の歴史

〔解答〕 景気変動の周期性についての注目とその予測をしようとする試みは19世紀の半ばよりあったが、20世紀に入るとその組織的資料が豊富になったので、ハーバード大学の景気予測が行なわれた。ジュボンズからジュグラー、コンドラチュフをへてミッチェルにいたってこの方法が大成した。世にハーバード法といわれるもので、各種の経済指標の動きをとらえてそれらの関連から景気を予測しようとするものであった。理論的な分析は行なわれなかったもので、のちに理論なき測定といわれるようになった。なぜそのような変動が起るかという経済機構についての認識がなかったのである。しかし単に予測という点からだけいえば使うのに価いするので、現在でもデフュージョニンデックスという形で用いられている。その後理論的分析にもとづく予測も行なわれたが、必ずしも満足すべき結果が得られなかったためでもある。理論の不備にも原因を求められるが、景気の予測ということ自体そう簡単なものでないであろう。ムーアは「総合経済学」という名称を用いて理論的分析と統計的分析を結合した。特に農産物の需要曲線の測定が有名である。価格がどのくらい上ると需要量はどのくらい下がるか、所得が増減した場合はどうなるか、所得と価格では需要量に対してどちらの影響が大きいかなど、現在でももっぱら行なわれる分析方法が使われた。ダグラスは今日、ダグラス函数と呼ばれる函数形を発見した。それは $P = bL^k C^{1-k}$ の形をとる。 P = 生産量、 L = 労働量、 C = 資本量である。今日まで非常に多く用いられている形である。

1930年計量経済学会が設立された。フリッシュ・ルース、I・フィッシャーが首唱者であった。1930年代はようやく統計的方法を用いた経済分析が盛んになってきた。シュルツが需要分析を数多く行ない、ある時点をきめて所得水

準と家計支出の関係を求める家計の支出線の研究はアレンとポーレーによって行なわれた。そして支出線が近似的に直線になるということを見出した。これがアレン・ポーレーの法則と呼ばれるもので、所得中、食料費のしめる割合をしめすエンゲル係数と共に有名である。フリッシュは貨幣の限界効用の測定や家計の労働供給にすぐれた分析を行なったが、ダークラスの生産函数にもいろいろな批判が起った。その一つは1本の方程式で経済すべての構造を反映できるかどうかということであった。そこで以後何本も方程式を置いて、それから全体を推定しようという方法が盛んになった。ティンバーゲンは多くの方程式を連立させて景気の予測をしようとした。この一つの方程式でよいか連立方程式が良いかについては、その後も論争が重ねられているが、いまだに結論はでていない。

レオンチェフは産業連関表というものをつくって、各産業間の関係を求めようとした。1940年代になるとハーベルモが近代確率論を導入した連立方程式形を發展させ、シカゴ大学を舞台とするカウルス・コミッションがこの方法を展開して、この方法の全盛時代となるのである。クラインは先のティンバーゲンの方法をうけついで、連立方程式形をケインズ理論から導出して具体的にアメリカ経済の模型をつくった。1950年代になると、消費支出を説明するのに単に所得ばかりでなく、その他の要因によって変動されることからいろいろな仮説がたてられた。これらの仮説をめぐって、デューセンベリー、トービン、クライン、ブラウン等の人々が消費函数について論争した。投資決定の要因を探る投資函数は、チェナリーによって進められた。クラインとゴールドバーガーは前の予測に若干の改訂を行ない、レオンチェフの産業連関表は経済計画分析に用いられることから世界各国で作られるようになった。

線型計画法と呼ばれる方法は、一次の不等式を条件として目的函数を極大

(小)にしようとするものであるが、企業の経営に応用される面が多い。フリッシュは産業連関表に線型計画法による最適配分の技術を応用し、経済計画を国民経済の最適計画をもってゆこうとしている。計量経済学もそのときどきの流行といったものはあっても、次第により厳密に国民経済、或いは企業や家計の行動をとらえようとしてゆくのである。

以下の章で、主として経済資料や経済に関係ある問題で統計的方法をのべてゆくことにする。

第2章 母集団と標本

2. 1 わが国の鉄鋼業における法人企業 1,000 の社長のなかから 100 名を選んで、来年度の景気動向に関する見通しをたずねようとする。この標本を選ぶにはどうしたらよいか。また、この場合の母集団はなにに当るか。また、100 名の社長に質問表を郵送するとして、その中の 60 名が回答に応じたとすれば、その際の標本の母集団はなにを指すことになるか。

〔解答〕 調査しようとする特定の標識が定められると、その標識をもつすべての単位の集合全体が、その調査のための母集団となる。つまり、具体的な調査方針が決定された後に、はじめて、その調査対象としての母集団が決定されることになる。この問題では、標識は質問表の内容である「景気動向の見通し」であるが、この標識をもつ調査単位（これを抽出単位という）は、わが国の鉄鋼業に属する法人企業 1,000 社の社長 1,000 名ということになる。したがって、母集団は 1,000 名の社長全員といってもよい。厳密には、これらの 1,000 名の社長たちがもっている来年度の景気動向に関する見通しの 1,000 個の回答である。この場合、回答の仕方として、「よくなる」、「今年度と変らない」、「悪くなる」等々に、それぞれ a 、 b 、 c ……と記号をつけるとすれば、 a の数、 b の数、 c の数……の全体の集合が母集団と考えられるであろう。

この母集団から標本を選ぶには、いろいろな方法が考えられる。標本として抽出される抽出単位は、わが国の鉄鋼業における法人企業 1,000 社の各社長であるから、その標本抽出法としてはつぎのような方法が考えられよう。

1) 単純任意抽出法

1,000名の社長に0番から999番までの番号を付け、3桁の乱数を矩形乱数表から100組だけ抽出する。たとえば、050, 828, 139, 900, 108等々の000から999までの数字が100個だけ抽出されることになる。そして、これらの番号に相当する社長に質問表を送付すればよい。通し番号が特定の順位にしたがって付けられたものではなく、全く無差別に付けられている場合には、標本抽出率が1/10であることを利用して、000から999までの数の一組だけを矩形乱数表から抽出し、その数字が、たとえば109であれば、抽出される標本を、109, 119, 129, …999, 010, 020, …090というように、10番目毎の社長を抽出してもよい（系統的抽出法）であろう。1,000社の企業の間にも明瞭な企業規模の格差がみられる場合には、全企業を大企業から順に番号を付け、上記の系統的抽出法を適用するとよい（系統的確率比例抽出法）。

2) 層化抽出法

1,000社の企業を、規模、地域等でいくつかのグループに分類し、そのグループ内の企業が大体において等質なものになるようにし（層化）て、それらの各層内では単純任意抽出することができる。各層に属する企業数がいちじるしく異なる場合には、各層の抽出率が、層の大きさに比例するように抽出率を決めることもきでる（ウェイトづき抽出法）。また、層内での標識（前記のa, b, c, ……等）についての分散が予め推定できる場合には、各層の抽出率が、その分散に比例するようにする（ネイマン法）こともできる。

以上、ここでは二つの方法についてのべたが、どのような方法によって標本を抽出するかは、上に規定した母集団の性質にも依存しているわけである。したがって、標本抽出を行なう前に、母集団の性質を十分に検討しておくことが必要であろう。

100名の社長が選出され、それらの人々に質問表を送付した結果、60名が