

統計学

山根太郎 著

東洋経済新報社



統計学

山根太郎 著

東洋経済新報社

著者紹介

1923年 アメリカに生れる。

1949年 一橋大学卒業。

1955年 Wisconsin 大学大学院卒業, M. A., Ph. D.

1955~57年 Columbia 大学研究員。

1957~61年 New York 大学助教授。

現在 青山学院大学経済学部教授。この間, San Diego State College 客員教授 (1963年), New York 大学客員教授 (1965~67年) にむかえられる。

専攻 計量経済学, 数理経済学, 統計学。

著書 *Mathematics for Economists*, Prentice-Hall, 1962.

Statistics, An Introductory Analysis, Harper & Row, 1964.

Elementary Sampling Theory, Prentice-Hall, 1967.

訳書 D・A・バワーズ, R・N・ベアード『初級数理マクロ経済学』(監訳), 第三出版, 1967.

現住所 東京都渋谷区代々木 4-62-3.

統計学

昭和53年12月8日発行

著者 ^{やまね たろう} 山根太郎

発行者 宇梶洋司

発行所 東京都中央区日本橋本石町1の4 東洋経済新報社

郵便番号 103 電話 03(270)4111(大代表) 振替口座 東京3-6518

©1978 (検印省略) 落丁・乱丁本はお取り替えいたします。 3033-3955-5214
Printed in Japan

は じ め に

昭和 21 年に大学に入ったときには筆者の関心はもっぱら理論経済学にあり、有名な先生が統計学の講座を担当しておられたにもかかわらず、たいして重要な科目ではないだろうと一人合点して受講しなかった。ところが、昭和 25 年にアメリカの大学の大学院に留学して、理論経済学を専攻したところ、なんとこの統計学が理論経済学とともに必修科目であった。仕方なしに統計学を受講することになったのであるが、数週間受講しているうちに、なるほどこれは重要であり、かつまた面白いとわかってきた。そして講義が進むにつれてこの認識がますます深くなっていったのである。

それからは、毎学期いろいろな統計学の講座を取るようになり、修士課程を終えて博士課程に入ってから、統計学の助手をするようになってしまった。ちなみに、たいていのアメリカの大学の修士課程と博士課程では、統計学は理論経済学とともに中心的な必修科目になっており、修士の資格試験および博士の資格試験の必須科目の一つになっている。

今日では、この統計学は経済学のみならず、経営学、社会学、心理学、政治学などきわめていろいろな分野において重要な分析道具であり、これなしには専門的な研究はできないといってよいであろう。

ことに経済学と経営学においては、貨幣という価値単位があるので、いろいろな結果を数量的に表現することができ、また電子計算機の発達によって経済学や経営学のいろいろな問題を数量的に処理することが著しく発達した。そしてこの傾向は将来ますます発展し、統計がますます重要になることは明らかである。

はじめて統計学を学ぶ多くの文科系の学生は、統計学は数学をふんだんに使う数理的な科目であると思っているようではあるが、統計学の中の数理統計学

を勉強するときは確かに数学をふんだんに使っている。しかし、統計学と数学とは異なるのであり、統計学には独自のものの見方、考え方があり、これを会得することが統計学を学ぶときに最も大切なのである。そして、この統計的なものの見方、考え方を会得することは、あまり数学に頼らずにできる。このテキストではむずかしい数学を使わないで、統計学における“いかにして” how と“なぜ” why のうちの why に重点をおいて、はじめて統計学を学ぶ人がわかるように解説することに心がけた。

統計学を勉強するについてのもう一つの特徴は、理科系の科目と同じように、いきなり、たとえば、第 15 章から読みはじめることはできない。第 1 章から第 14 章までのことを理解していないと、第 15 章はわからない。すなわち、コツコツと初めから勉強していかなければならず、途中で勉強しないとその後のことがわからなくなってしまう。

また、それぞれのトピックをよく理解するためには、練習問題を解くことが必要である。筆者の考え方は、このような初歩の統計学を勉強するときは、30 分もかかるむずかしい問題を一題解くよりも、5 分くらいで解ける簡単な問題を数題解いて、統計的な考え方を反復練習する方がよいと思っている。このテキストの章末練習問題はこのような主旨に沿って作られている。

学生諸君が、このテキストを読むことによって統計学に親しみを感じ、興味を覚えることをねがうものである。

最後に、この本の出版にあたって足利学園理事長の上岡一嘉氏、東洋経済新報社の佐野佳雄氏、山口正氏、須永政男氏に一方ならぬお世話になったことに深い謝意を表したいと思います。

1978 年 8 月

山根 太郎

目 次

はじめに

第 1 章	序 論	3
--------------	-----	---

第 2 章	度数分布	5
--------------	------	---

2・1 変 数 5

2・2 度数分布 7

2・3 度数分布のグラフ 10

2・4 累積度数分布 13

2・5 相対度数 15

問 題 17

第 3 章	代 表 値	19
--------------	-------	----

3・1 算術平均 20

3・2 メディアン 23

3・3 モード 27

3・4 幾何平均 28

問 題 29

第 4 章	散らばりの代表値	33
--------------	----------	----

4・1 範 囲 33

4・2	平均偏差	34
4・3	標準偏差	34
4・4	母集団, 標本と標本分散	39
4・5	相対偏差(変動係数)	40
	問 題	41

第5章 確 率 論43

5・1	集合と標本空間	44
5・2	客観的確率	57
5・3	主観的確率	66
5・4	確率論の公理	67
5・5	条件付き確率	69
5・6	確率の乗法定理	73
5・7	統計的独立性	74
5・8	確率変数と数学的期待値	76
5・9	要 約	82
	問 題	82

第6章 正規分布曲線と正規分布表87

6・1	序 論	87
6・2	正規曲線と正規分布表	88
6・3	正規分布の特性	93
6・4	連続修正	95
	問 題	98

第7章 標 本 分 布 101

7・1	可能な全標本数	102
7・2	単純無作為標本抽出	105
7・3	標本分布	108
7・4	所与の母集団から，標本平均の標本分布 を使って標本を選ぶ確率の求め方	116
7・5	標本比率の標本分布	121
	問 題	126

第8章 推 定 135

8・1	推定と推定量	135
8・2	不偏性	138
8・3	一致推定量	139
8・4	有効性	141
8・5	十分性	143
8・6	最尤推定量	143
8・7	誤差，誤差の起こる確率と標本の大きさ の関係	149
8・8	区間推定	156
8・9	数学補注	162
	問 題	163

第9章 仮説検定 167

9・1	はじめに	167
9・2	第1種と第2種の過誤	168
9・3	仮説検定	171
	問 題	183

第10章 決定理論..... 187

10・1 単純な決定問題 187

10・2 ベイズの手法 191

問 題 197

第11章 指 数..... 199

11・1 序 論 199

11・2 加重価格比物価指数 201

11・3 加重総和物価指数 205

11・4 連鎖指数 206

11・5 指数の接続と基準の移動 208

11・6 デフレーター 210

11・7 物価指数に関する諸問題 212

11・8 消費者物価指数と卸売物価指数 214

11・9 数量指数 219

問 題 224

第12章 時系列——傾向線..... 231

12・1 傾向線を求める簡単な方法 235

12・2 最小二乗法 236

12・3 単位と原点の変化 240

12・4 移動平均法 244

問 題 247

第13章 季節変動と循環変動..... 251

13・1	総 説	251
13・2	用語の定義	252
13・3	季節変動	254
13・4	単純平均の方法	256
13・5	移動平均比法	258
13・6	センサス局法	261
13・7	季節調整済みデータ	262
13・8	循環変動	264
13・9	経済時系列の分析の例	264
13・10	時系列についてのコメント	269
	問 題	274

第14章 線型回帰分析 279

14・1	序 論	279
14・2	回帰の問題	280
14・3	A と B の推定——類型1の母集団	292
14・4	回帰関数の求め方	296
14・5	回帰の標準偏差	298
14・6	$\hat{\sigma}_{YX}^2$ と決定係数 r^2 の関係	301
14・7	σ_a^2 と σ_b^2 の推定	313
14・8	要 約	318
14・9	A と B の推定——類型2の母集団	319
14・10	a と b に関する検定	321
14・11	μ_{YX} の信頼区間	324
14・12	Y の信頼区間	329
14・13	回帰分析に関して留意すべき点	332

14・14	独立変数が二つ以上の場合の回帰分析	333
-------	-------------------	-----

問 題	339
-----	-----

第15章 相 関 分 析 345

15・1	2変量分布	346
15・2	相関係数：第1の場合	350
15・3	2変量分布と回帰分析	369
15・4	相関係数：第2の場合	371
15・5	要 約	378
15・6	相関係数の仮説検定	382
15・7	順位相関	387
	問 題	393

付 表	399
-----	-----

参 考 文 献	409
---------	-----

索 引	411
-----	-----

統計学

第 1 章 序 論

過去数十年の間に統計学が目覚ましい発展をとげて非常に重要になってきたのは、いわゆる情報化時代に移行しつつあるからであろう。長い文明社会の発展を見るとき、いわゆる“力”の所在は、土地を所有するものから物(貨幣)を所有するものへと過去数百年の間に移行した。そして、今後は物(貨幣)中心から情報中心の時代へと次第に移行しようとしていることに気づくであろう。

この情報化時代への発展の一つの大きな要因となったのは、いうまでもなく膨大な情報量を瞬時にかつ正確に処理できる電子計算機の出現である。

しかし、単に多量の情報を処理するだけでは、あまり意味がない。この多量の情報の中に、社会現象や自然現象を説明するいわゆる“法則性”あるいは“規則性”があれば、これらを知ることによって社会現象に対しては政策を立案する手掛りをえ、また自然現象に対してはそれらを制御する手段を発見することができる。これによってはじめて多量の情報を処理する有用性が生まれる。

では、社会現象や自然現象によって醸成される多量の情報の中に法則性や規則性があるのであろうか。本書の第5章“確率論”および第6章“正規分布”で説明するが、多量のデータの中に、このような法則性や規則性を確率現象としてとらえることができる。

一方、確率理論によって種々の統計手法が導出され、これらの統計手法によって上記の法則性や規則性を統計的な現象として整理把握できる。

かくして、社会現象や自然現象の分析に統計学がきわめて重要な分析道具となった。経済学および経営学においては貨幣という計算単位があるので、統計学は分析道具としてはことに重要であり、また、これらの分野への応用においてその力をいかに発揮している。

したがって、経済学や経営学を研究しようとする者にとっては統計学の修得

は不可欠であるといっても過言ではない。

統計学は通常、記述統計と推測統計とに大きく分けられる。記述統計においてはデータの収集、整理、および分析を行なう。推測統計においては、たとえば、全国の有権者から 3000 人を抽出して、この 3000 人という標本から全国の有権者という母集団について、何パーセントがどの政党を支持しているかを推測するというように、部分のデータから全体に関して推測する統計操作を行なう。近年、この推測統計の発展が特に目覚ましい。

本書では、第 2 章～第 4 章において記述統計を説明し、第 5 章～第 10 章で推測統計の基本的な概念や手法を解説する。第 11 章～第 15 章では、経済や経営の学生にとって特に重要な指数、時系列、回帰分析、および相関分析について研究する。

統計学を勉強するにあたって、いかにして(how)と、なぜ、なにゆえ(why)の二つの面があるが、how は統計手法に関することであり、why は、その how の理論的根拠である。この why の研究と理解が重要であることはいうまでもなく、本書では特にこの点に留意して説明している。

第 2 章 度 数 分 布

最近の経済や経営に関する諸問題の分析や処理には多量のデータを取り扱う場合が多い。デパートやスーパーマーケットなどが、顧客についてマーケット・リサーチを行なう場合は、万単位の客について調査する必要がある。自動車メーカー、レコード会社、食料品メーカー等はいずれも多くのデータを分析して、それから有用な情報を抽出して、経済や経営についてのいろいろな決定を行なっている。特に最近、コンピュータの利用が普及するにつれ、また、統計学や計量経済学の発達によって、多量のデータを調査や分析に利用する傾向が一層強くなった。

これらの一見すれば雑然としている多量のデータを分析してみると、多くの場合そこに経済や経営に関する諸行動や現象の安定した、あるいは規則的なパターンを見いだすことができる。たとえば、所得の分布状況、毎年のビールの売上げのパターン、受験生の入学試験の点数の分布など、多くの例をあげることができよう。

ここで当然、これら多量のデータを使いやすい形に整理して、安定したあるいは規則的なパターンを見いだしたり、その他有用な情報が抽出しやすいようにすることが必要になってくる。

これからの第 2 章～第 4 章では、調査や測定を行なって得たデータを有用な情報が抽出しやすいように組織的に整理するいくつかの方法について述べる。このためにまず変数について説明し、ついで度数分布について述べることにする。

2・1 変数(Variable)

いろいろな値を取りうる量を変数という。この変数を二つの観点から考えて

みよう。一つは変数の値(あたひ)についてであり、他は変数の性質による分類である。

変数の値については、変数の取りうる値と測定値とに区別する。たとえば、経済学の試験があり、その得点 X の範囲は0点から100点であり、点数は整数で表わされるとする。得点 X が変数で、0から100までの101個の値が変数 X の取りうる値である。いま5人の学生がいて、それらの得点が、 $x_1=55$, $x_2=89$, $x_3=45$, $x_4=60$, $x_5=80$ とすると、それらが変数 X の測定値である。

表 2-1

X	度 数
1	3
2	5
3	4
4	4
5	3
6	5
	24

例 1 さいころを投げたときに出る目の数を X とし、24回投げたとする。そのとき変数 X は、1, 2, 3, 4, 5, 6の取りうる値を持つことになる。その分布は表 2-1 のようになった。この表は、1 という目が3回、2の目が5回、3の目が4回、…起こったことを表わしている。この24個の値は変数 X の測定値である。

表 2-2

X	出た回数
1	3
0	2
	5

例 2 硬貨を投げたときに出る表と裏を X とし、表は1裏は0とする。変数 X の取りうる値は1か0である。5回投げたとき、その結果は表 2-2 のようになった。このときの変数 X の測定値は五つの値、1, 1, 1, 0, 0である。

次に、変数はその性質によって連続変数と離散変数に分類できる。たとえば、あるクラスの生徒の体重を X とすると、いちばん軽い体重が40kgでいちばん重い体重が72kgのとき、変数 X の取りうる値は40kgから72kgの間で連続しており、その間にとぎれはない。すなわち、二つの値(たとえば、60.19と60.20)が接近していても、その間には X の取りうる値がある。このような変数は連続変数といわれる。

また、ある地方の農場の牛の数を X とすると、このとき取りうる値は、0, 1, 2, 3, …である。このため、二つの取りうる値、たとえば15頭と16頭の間には、15.6というような値はなく、続いていない。このように取りうる値の間にとぎれがある変数を離散変数という。