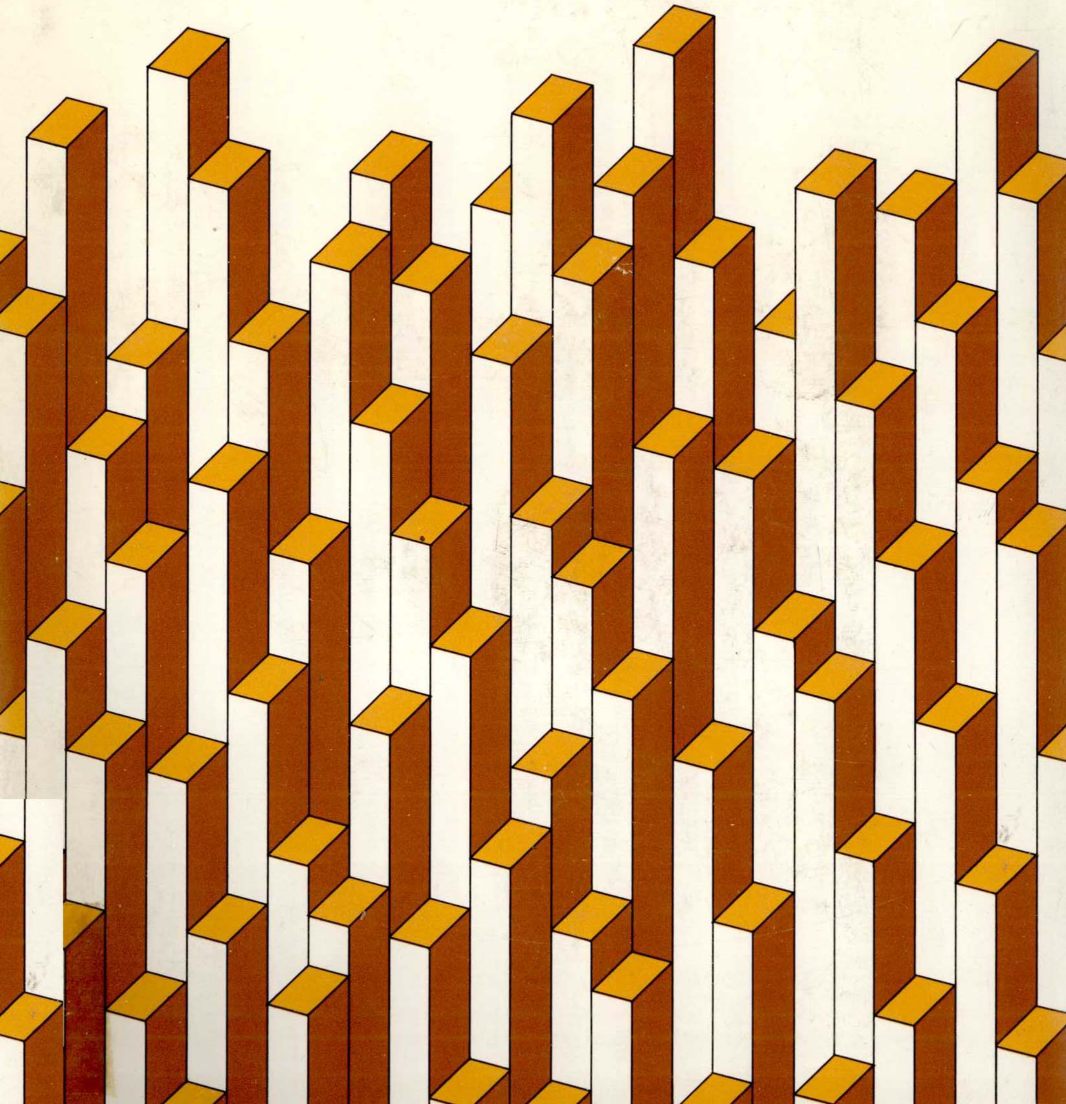
数理統計学 2

Introductory Mathematical Statistics

E. クライツィグ＝著

近藤次郎・出居 茂＝共訳

培風館



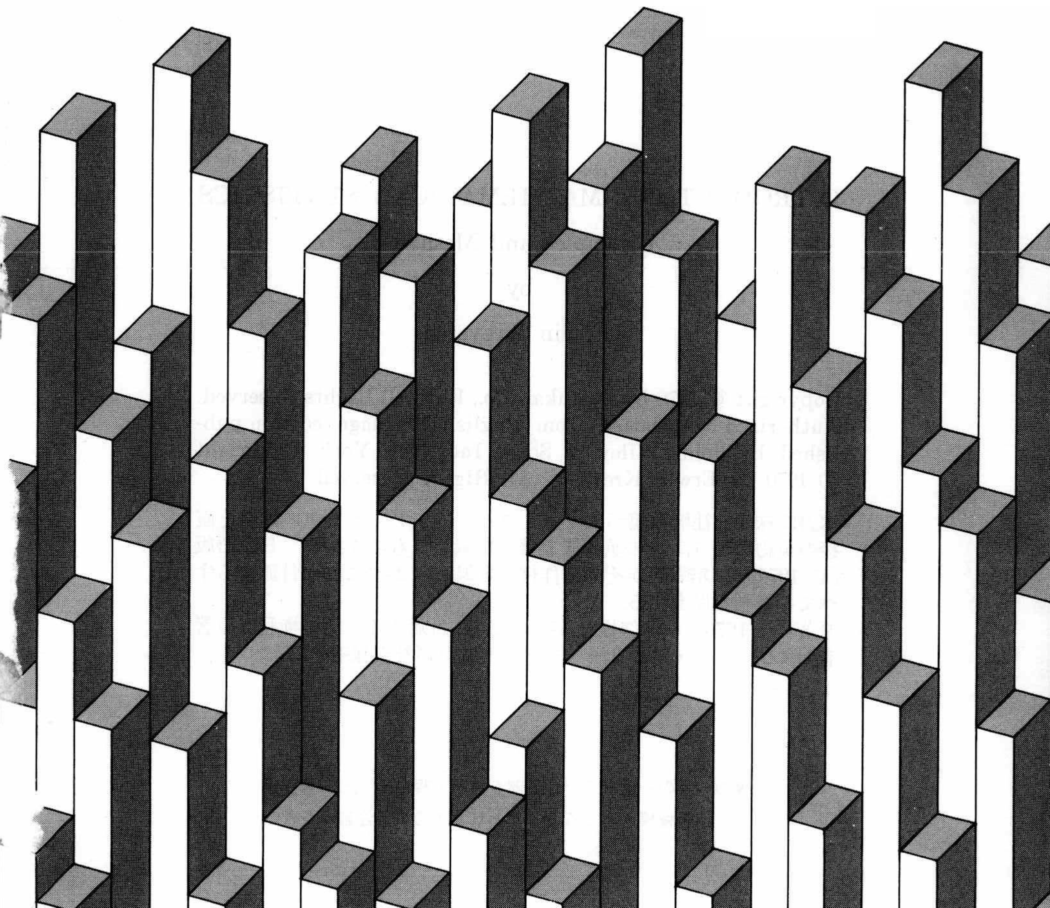
数理統計学 2

Introductory Mathematical Statistics

E. クライツィグ＝著

近藤次郎・出居 茂＝共訳

培風館



訳者略歴

近藤次郎（こんどう じろう）

1940年　京都大学理学部数学科卒
1945年　東京大学工学部航空学科卒
1958年　工学博士
1977年　東京大学名誉教授
現　職　国立公害研究所副所長

出居　茂（いでい しげる）

1954年　東京大学工学部応用物理学科卒
現　職　早稲田大学システム科学研究所
　　　　特別研究員・講師

昭和52年9月30日　初版発行
昭和54年11月15日　初版第2刷発行

数理統計学2

著　者　E.クライツィグ
訳　者　近藤次郎
　　　　出居　茂
発行者　山本健二
発行所　株式会社 培風館
東京都千代田区九段南 4-3-12・郵便番号 102
電話 (03) 262-5256 (代表)・振替東京 4-44725

定価 ¥ 2200.

中央印刷・坂本製本

訳者の承認をえて検印を省略しました

3033-0831-6955

序

本書の目的　この本は学生に数理統計学の基礎概念，原理および方法とその応用を紹介する意図をもって書かれたものである。

この本は，週3～4時間の1学期の講義または週2時間の通年講義の教科書として適している。もっと短い講義のために内容を取捨選択してもよいが，それはこの序の終わりに示しておく。

前提科目としては基礎的な解析学(微積分)で十分であって，確率論や統計の知識は前提にしていない。

書物の内容は，確率論，統計的方法，一般の数理統計学とその応用などの講義と，統計学の実際の仕事などの影響を受けたものである。これらの講義はいろいろな数学の素養の差や応用上の興味の違いをもった欧米の受講者に対してなされたものである。

内容および配列　この本はⅢ部からなる。

Ⅰ. 記述統計

Ⅱ. 確率論

Ⅲ. 統計的推測

第Ⅰ部は短く，数学的には非常に単純である。

第Ⅱ部は，分布の数学模型および第Ⅲ部にある統計的方法のための基礎を与える。また，第Ⅱ部は，統計的推論の考え方が理解される範囲の概念構成を与えるとともに，数理統計に不可欠な概念と方法を強調する。書き方はかなり詳しい。それは確率についての十分な背景が，統計的方法の深い理解を得る手助けになるからである。

第Ⅲ部は，点推定と区間推定，仮説検定，品質管理，抜き取り検査，逐次解

析の初歩，適合の良さ，分散分析，回帰分析と相関分析，物理計測における誤差理論の初歩，いくつかの母数によらない検定，決定理論の諸要素など，一般に興味がもたれる統計的方法を含んでいる。

できるだけ簡単に書き表わすことと，記号を単純にすることにはとくに注意をはらった。混乱を避けるために，確率変数を大文字で記し，確率変数がとる値を小文字で記した。各章はできるだけ独立に読めるようにした。また，この本の程度をこえた証明のいくつかは，章から離して付録1に入れた。

理論と応用 ここ数十年の間に，数理統計学は目をみはらさせる速さで発展してきているので，一冊の本の中に入れるべき話題の選択はおそらく他の分野におけるよりいっそうむずかしいであろう。この本では，一般に応用上大切な方法にまず選択の目を向け，理論と応用との調和を適切に保つよう心がけた。この選択方法を用いたことが，読者に早くから応用上の有用性に気づかせる助けになることを望んでいる。実際，統計理論の多くは応用から離れては完全な評価を受けられないものであるし，それはこの理論が実生活の中での問題を解決しようという必要から発展してきたことからも明らかである。

ごく少数の例外を除いては，この本の中に出てくるデータは，実際の観察や実験にもとづくものである。いくつかの例や問題はきわめて初等的であるが，それは読者自身が実験や観測を始めようと思ってもらうためにほかならない。

電子計算機と電卓 統計上の仕事における数値処理に電子計算機はますます重要になってきた。そして多くの大学で統計学の入門の講義においても計算機を使うことが試みられている。このやり方はいっそう広まっていくであろう。統計学者が電子計算機を使うか電卓を使うかということにかかわりなしに，数値解析の標準的手法を知り，理解していなくてはならないことはもちろんである。この本で数値計算に十分な注意をはらったのは，この理由からである。

問題と答 とくに統計学においては，実際問題の訓練なしに道に長じることはできない。この本には読者を鍛えるために，また数理統計に関するカンを養い，その本質の理解と把握を助けるために，1000題ほどの問題を入れてある。問題はいろいろな領域から注意深く選んだ。奇数番号の問題の答が付録2に与えられている。

参考書 より進んだ，参考にするための本や論文のリストが付録3にある。引用は本文のいろいろな所にあげた。

表 付録4は，第II部と第III部で必要な分布の表のほかに，乱数表および(指数関数，2項係数，ガンマ関数などの)補助的な関数の表を含んでいる。

短い講義のための材料の選び方　章は(とくに第Ⅲ部において)比較的独立に書かれているので，短い講義のための材料の選び方は困難ではない。1つの示唆を与えておく。

標本の度数分布(2章)
標本平均と標本分散(3.1-3.3節)
確率論の基礎概念(4.1-4.9節)
確率分布(5.1-5.6節)
分布の平均と分散(6.1-6.3節)
2項，ポアソン，超幾何分布(7.1, 7.2, 7.4, 7.7節)
正現分布(8.1-8.4節)
いくつかの確率変数の確率分布(9.1-9.3, 9.5節)[以上訳書，第1巻]
母数の推定(13.1-13.4節)
信頼区間(12.1, 12.3, 12.5節)
仮説検定(13.1-13.4節)
カイ2乗検定(15.1, 15.2節)
回帰分析(17.1-17.5節)
相関分析(18章)
母数によらない方法(20章)

謝辞　原稿のいろいろな部分が謄写印刷の形で学生に配布され，より良くするための意見がついて戻されてきた。私の同僚や以前の学生たちで，とくにこの本を準備する上で助言と手伝いとをしてくれたのはR. Bradley, H. O. Hartley, J. S. Hunter, H. Klinger, R. Roeloff, J. P. Spencerの諸教授，および，W. F. Taylor博士である。

最後に，John Wiley and Sons社に対して緊密な協力と出版の注意深い諸準備に負うところが多大であった。

この本をより良くするためのご意見やご教示は，ありがたくお受けするつもりである。

Erwin Kreyszig

INTRODUCTORY MATHEMATICAL STATISTICS

Principles and Methods

by

Erwin Kreyszig

目　次

1 巻 目 次

第Ⅲ部　統計的推測

第Ⅰ部でわれわれは，統計的なデータを適当な表にまとめたり，グラフの形に表わしたりするやり方を論じた。第Ⅱ部では，母集団の重要な数学模型を考えた。

この第Ⅲ部では，理論と観測可能な現実の間の関係を求めることにする。われわれは与えられた標本から，母集団に関してどのような結論を引き出すことができるであろうか，ということを考え，そしてそのような結論が，どれほど信頼しうるものであるかを考える。これらの結論は量的なものである。現実的な興味のあるいくつかの典型的な状況や問題が，p. 3～4 の表に示されている。

数学模型が重要なことは明らかであり，そして そのことは前に強調されていたことを思いだしてほしい。現在の方法は，模型を選んだり，つくり上げたり修正したりして，それらが現実の世界に合致するかどうかを検定をすることである。

これらの方法において，われわれは標本を使わなければならない。いままで**母集団からの標本**は，母集団から取り出された1つの選択であることを知っていれば十分であったが，この概念を精密な形式に述べ直すことが必要になってくる。定義によって標本は，次のようないくつかの性質をもたなくてはならない。

標本は，それに対応する母集団に関する情報を提供するものと考えられているが，その母集団に含まれるすべてのものを観測したり，測定したりすることは，費用がかかりすぎたり，時間がかかりすぎたり，あるいは現実に不可能であるような場合もある。典型的な例は，1.2 節で述べた。そのような情報が得られるためには，標本は**無作為選択**によらなければならない。このことは，母

集団の1つ1つの要素が抜き取られる確率が既知でなければならないこと，すなわち，個々の要素が標本の中に取り入れられる確率を知っていなければならないことを意味する。最も簡単で，かつ最もありふれた場合は，この確率が，母集団のすべての要素に対して同一である場合である。この要請が満たされたとき(少なくとも近似的に)，かつそのときにかぎって，考えられている統計的方法は，筋の通った役に立つ結果を生み出すものと考えられる。

さらに，われわれが n 個の標本値を得るために行なう n 回の無作為実験は**独立**であることが必要である。すなわち，実験のどの結果も，他の結果がどのようになるかについて，何の影響も与えてはならない。別の言い方をすると，母集団のどの要素についても，それが標本の中に取り入れられる確率は，その母集団のそれ以外のどの要素が標本の中に取り入れられている(あるいは取り入れられていない)かということには，まったく依存しないのである。

この仮定に基づく標本抜き取りの理論は，次の理由で，**無限母集団**からの標本抜き取りとよばれる。すなわち，標本値 $x_1, \cdots, x_n$ は，その独立性がいつまでも保証されているような1つの母集団から抜き取った結果であるとみなすのである。

もし，もとに戻しながら抜き取るならば，毎回抜き取る前に初めの状態に戻しているわけだから，このことは，もちろん成り立っている。たとえば，1つのサイコロを n 回投げるならば，その標本は n 個の目の数からなるが，この母集団は，そのサイコロを無限に多くの回数投げて得られる無限に多くの1, 2, …, 6からなる無限母集団である。もちろんこの母集団は仮想のものである。

もとへ戻さずに抜き取る場合には，母集団が無限である場合，かつその場合にかぎって独立性が保証される。実際には，われわれは母集団が標本の大きさに比べて“非常に大きい”ならば，母集団は無限であるとみなす(例：ある大きな都市の母集団から選ばれた1000人の人たちの身長)。そのときわれわれは，もとに戻さずに抜き取っているにしても，標本値は独立であると仮定してよい。

ある標本が無作為選択である，という要請を満たすことは容易なことではない。なぜならば，標本抜き取りの結果にかたよりをもたらすような，多くの微妙な要因が存在するからである。たとえば，1人の興味をもった購買者が，どれを買うかを決める前に，800個の品物の中から10個の品物を標本として抜き取り，検査をしたいという場合に，すべての可能な $\binom{800}{10}$ 個の大きさ10の標本が同様に起こりやすいことを確実にするために，その10個を実際にどのよう

に選択すればよいであろうか。

無作為に選択するという問題を解決する方法が発展してきたが，ここで，しばしば用いられる1つの手続きを述べることにしよう。

800個の品物が与えられたとして，無作為にそれから10個を選択したい。この目的のために，われわれは品物に1から800までの番号をつける。選択は付録4の表5，**無作為選択の表**を使って行なう。われわれは，まず2つの整数，たとえば1728と219を選ぶ。最初の数を100でわり，次の数を10でわる。剰余は28と9で，それぞれ行と列の番号になる。表5の第28行，第9列に，数83814がある。われわれは下に読んでいって

83814　87802　67056　61340　など

を見いだす。それぞれの数字の下2けたを取り除く。そうすると

838　878　670　613　など

を得る。800を超える数と，2度目に出てきた数を取り除く。そうして10個の数

670　613　783　528　90　658　765　242　324　600

を得る。これらの番号のついた10個の品物が，要求される無作為選択を表わしている。

基礎的な統計の問題

状　況，問　題	出てくる章，節
分布の未知な母数の推定	
母数の近似値を得るには	11章
その近似の精度の判定には	12章，とくに 正規分布は12.1，12.3，12.5節 2項分布は12.6節 任意の分布は12.7節
母数がある特定な値をもつかどうかの検定には	13章
品質管理，受け入れ検査	
生産工程が適正に動いているかどうかの検定には	14.1節
ロットが仕様通りの品質要請を満たしているかどうかの検定には	14.2-14.5節
平均や分散が異なるいくつかの正規母集団を含む問題	
平均が等しいかどうかの検定には	13.4節，16章
分散が等しいかどうかの検定には	13.6節

状　況，問　題	出てくる章，節
中央値についての検定	
中央値がある特定の値をもっているかどうかの検定には	20.1 節
分布関数についての検定	
母集団が特定の型の分布をもっているかどうかの検定には	15 章
2つの母集団が同じ分布をもっているかどうかの検定には	20.4 節
無 作 為 性	
標本値の無作為性を検定するには	20.3 節
対象を無作為に選ぶには	p.3
物 理 計 測	
物理計測の精度の判定には	19 章
普通の変数 x と x に依存しているであろう変数 Y を含む問題	
Y が x に線形依存しているという仮説の検定には	17.10 節
その回帰直線を求めるには	17.1 節
その直線の傾きの精度の判定には	17.5 節
傾きがある特定の値をもつかどうかの検定には	17.7 節，17.9 節
回帰放物線を決めるには	17.11 節
回帰曲線がある型であるかどうかの検定には	17.12 節
Y の値がトレンドを表わしているかどうかの検定には	17.7，17.9，20.2 節
互いに関連しあっている2つの確率変数を含む問題	
相関係数を求めるには	18.1 節
相関係数の値の精度の判定には	18.4 節
母集団の相関係数の値が0かどうかの検定には	18.4 節

11

母数の推定

2 項分布における p や，正規分布における μ や σ のように，分布関数の中に出てくる量を母数という。与えられた標本から，母数の推定値をどのようにして得るかを説明しよう。これは重要な実際問題である。

そのような推定として，2 つの種類のものがよく使われる。1 つは点推定とよばれ，1 つの数であって，与えられた標本から計算され，母数の未知の正しい値の近似値として使われる。この種の推定をこの章で取り扱う。もう 1 つの種類は区間推定とよばれ，1 つの区間である。それは 12 章で学ぶ。

11.1 平均値，分散，モーメントの方法

標本の平均 $\bar{x}$ を，対応する母集団の平均 μ の近似値とみなすのは自然である。このことから，われわれは，μ に対する推定値 $\hat{\mu}=\bar{x}$ を得る。すなわち

$$\hat{\mu} = \bar{x} = \frac{1}{n}(x_1 + \cdots + x_n) \tag{1}$$

ここに n は標本の大きさである。

同様に，標本の分散 s^2 は，母集団の分散 σ^2 の近似値とみなすことができる。それゆえ σ^2 の推定値 $\hat{\sigma}^2=s^2$ を得る。すなわち

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2 \tag{2}$$

いくつかの分布においては，μ は母数としてはっきりしており，そのとき(1)は母数の推定値である。たとえば，正規分布やポアソン分布がそうである。

2 項分布の場合には，$p=\mu/n$ [7.2 節(3)参照] である。この場合，(1)の x_j

は，それが起こる確率が p である事象が，j 番目に実施された実験で起これば値 1 をとり，起こらなければ値 0 をとる。(1)から p に対する次のような推定値 $\hat{p}$ を得る。

$$\hat{p}=\frac{\bar{x}}{n}$$

(1)は**モーメントの方法**とよばれる方法の，非常に特別な場合であるということに触れておこう。この方法では，推定されるべき母数を，分布のモーメントを使って表現する。導かれた公式では，これらのモーメントは，標本の対応するモーメントでおきかえられる。このようにして望ましい推定値が得られる。ここに**標本** $x_1, \cdots, x_n$ **の** k **番目のモーメント**は，公式

$$m_k=\frac{1}{n}\sum_{j=1}^{n}x_j{}^k \qquad (k=1, 2, \cdots) \tag{3}$$

で定義される。

グループ分けは，標本のモーメントの値に影響を及ぼすであろう。この効果は Sheppard (1898) (付録 3 参照) によって導入された**シェパードの補正**を使うことによって，ある程度直すことができる。m_1 に対してはこの補正はゼロである。m_2 は

$$m_2-\frac{l^2}{12} \qquad (l=\text{クラスの幅})$$

によっておきかえられなくてはならない。詳しくは，Cramér (1961) と Wold (1934) を見よ。

ポアソン分布の場合には，$\mu=\sigma^2$ である。ゆえに(1)も(2)もともに，その分布の母数の推定値である。これは数学のいろいろな分野で，ある特定の量に対していくつかの公式があるという事実に照らせば，めずらしい状況とはいえない。しかし，そのような近似公式の“良さ”を特徴づける上で本質的な性質は何であるかを見いだすようにしなければならない。このことを次の節で述べる。

11.2 推定量，不偏推定量，有効推定量

われわれは 1 つの未知のパラメータ θ を含む分布を考え，与えられた標本 $x_1, \cdots, x_n$ から，θ の近似値 $\hat{\theta}$ を計算する公式を知っているとしよう。明らかに $\hat{\theta}$ はその標本値に依存する。すなわち，標本値の関数である。この関数を g で表わすと，公式は次の形をしている。