はじめに

データ分析の実装環境は目覚ましい速度で発展しています。近年では、オンラインで統計解析ソフトが使えるようになり、統計分析ができる環境を自前のパソコンに整えずとも、最新の分析手法を手軽に実行することができます。私が博士前期課程に在籍していた10年ほど前には、ノートパソコンのファンがうなり声を上げ、いつフリーズしてもおかしくない状況に恐れていたことを考えると、データ分析は親しみやすく身近なものとなりました。データ分析を用いて気軽に、市場調査、経営判断、仮説の検証などができるようになったのです。

これはとても喜ばしいことですが、統計学に携わる者としては、この状況を手放しに喜ぶことはできません。分析自体は簡単に実行できてしまうため、適切ではない分析手法を用いたり、分析結果を間違って解釈したりすることで、データ分析をしないよりも悪い結果が導かれてしまうことがあります。データ分析を誤用しないためには、統計理論をしっかりと理解することで、適切な分析手法の選択と正しい解釈をする能力を身につけることが求められます。

このような現状を踏まえ、本書は、統計学における基本的な考え方と、確率論に基づいた統計理論の基礎的な内容について詳述することで、統計理論に対する理解を、より多くの人に広めることを目的に書かれました。

本書のねらいと特徴

本書のねらいは、統計学の入門的な内容についての理論を丁寧に詳しく解説することです。統計学の理論は数学を用いて構成されているため、本書の序盤（2章〜7章）では、統計学の理論を理解するための数学を解説しています。中盤（8

章〜10 章）では、統計学の基本的な考え方、および、推定と仮説検定について詳述し、終盤（11 章〜13 章）において、線形回帰モデルと二項選択モデルの解釈と理論的性質を考察しています。

本書は、統計学の考え方とその理論的な正当性を詳しく学びたい人に役立つことを目指しています。応用方法を学習したことがある人はもちろん、統計学を学習したことがない人でも自力で読み進めることができるように工夫しました。読み進めるにあたって必要な数学的な知識まで取り扱っており、この意味で本書は自己完結型（self-contained）になっています。また、定義の確認や多角的な理解ができるよう多数の例題を載せていますので、手を動かして考えてみてください。すべての章末問題の解答例は本書の最後にあります。

書名にあるように、本書は経済データの分析について考える「計量経済学」の学習への橋渡しとなることを意識して書かれており、計量経済学の基礎理論が理解しやすくなるような内容を優先して採用しています。同じ日評ベーシック・シリーズの『計量経済学』とあわせることで、大学の学部レベルとして十分な内容の計量経済学の理論を学習することができます。一方で本書は、統計学の多くの教科書で取り上げられている内容を取り上げていなかったり、多変量のモデルについては本書ではなく上述の『計量経済学』で取り扱っていたりします。このため、統計学を網羅的に学習するためではなく、「計量経済学」や統計学の他の教科書を読み進めるための準備として、本書は適していると考えています。

本書の概要

1 章では統計学とはどのような学問で、どのような問題を取り扱うのかについて解説しています。2 章〜7 章は集合、写像、確率について、本書で取り扱う統計学の内容を理解するうえで必要最低限の内容を、丁寧かつ可能な限り厳密に説明しています。8 章では、母集団と標本の関係を明らかにし、推測統計学の基本的な考え方を学びます。9 章において、推定量の良し悪しの判断をするための評価方法を紹介し、10 章において仮説検定の意味と論理を紹介しています。11 章では線形単回帰モデルを定義し、モデルの解釈と推定方法（最小 2 乗法）について述べます。12 章において最小 2 乗推定量の性質を考察し、モデルを用いた仮説の検定について解説します。13 章は二項選択モデルであるプロビットモデルとロ

ジットモデルを取り扱います。モデルの定義と解釈を与え、最尤法によるモデルの推定と検定を紹介します。13 章のうち、仮定の検証と推定量の性質に関する節はいくらか発展的な内容になっており、節見出しに＊マークを付けています。これらの節は読み飛ばしても通読できるようになっています。なお、本書のより詳しい内容は 1.5 節で紹介しています。

謝辞

本書の内容は、小樽商科大学の統計学の講義で用いた講義ノートを元にしています。講義で質問や誤植の指摘をしてくださった学生のみなさん、そして、草稿に目を通し多くのコメントと校正をしてくださった佐野学氏に感謝いたします。また、本書の大部分は大学の勤務時間外に執筆しました。家族をはじめ、私を支えてくださったすべての人に御礼申し上げます。

私が小樽商科大学に着任して間もない 2019 年 9 月、「理論をしっかりと書いた統計学と計量経済学の教科書を執筆してほしい」と、日本評論社の吉田素規氏から依頼をいただきました。本書は統計学にあたる部分であり、計量経済学については 2022 年 9 月に日評ベーシック・シリーズ『計量経済学』として刊行されています。研究者として駆け出しの私に本書を執筆する機会をくださり、内容の相談から最終稿の完成までサポートしてくださった吉田氏に御礼申し上げます。

政策やビジネスなど、社会のあらゆる分野で統計学の素養が求められる現在、本書が統計学の学習に役立ち、それを通して少しでも社会に貢献できれば、著者としてこれ以上の喜びはありません。

2023 年 3 月

岩澤政宗

目　次

第 1 章

統計学とは何か

1.1　統計学とは

「統計学」に対してどのようなイメージを持っているだろうか。近年、IT 技術の発展とともに、機械学習や AI（artificial intelligence）技術を用いた大規模データの分析手法の進展と応用が加速し、生活の様々なところでこれらの技術が応用されるようになった。機械学習や AI が多くのメディアに取り上げられたことで、「統計学 = 機械学習」や「統計学 = AI」などという漠然としたイメージを持っている人もいるかもしれない。他にも、身長や体重の平均値や大学受験の偏差値などの「指標」や、所得を表す棒グラフや死因を表す円グラフなどの「図」を思い浮かべるかもしれない。これらはまぎれもなく、統計学を構成する要素である。

これらに共通するのは、「データを取り扱っている」ことであろう。大雑把な言い方をすれば、統計学はデータの扱い方を探求する学問である。データの扱い方には、例えば、どんなデータを、どのように、どれだけ調達するのかという問題や、得られたデータはどのような特徴を持つのかという問題がある。データの収集方法や、得られたデータの質・特徴に関する問題は、統計学の中でも「標本調査法」や「記述統計学」とよばれる分野で取り扱われている。一方、得られたデータをどのように分析するのかという問題は「推測統計学」とよばれる分野で研究されている。まずは、これから学ぶ統計学の全体像を俯瞰するために、標本調査法、記述統計学、推測統計学が、統計学においてなぜ重要なテーマであるのかを紹介する。

1.2　標本調査法

標本調査法では、データをどのように何人分集めるかなど、データの収集に関することを考える。統計学では、分析対象とする集団を母集団とよび、個人、家計、企業、物、国家などの様々なものや事象が分析対象となる。例えば、消費税の増税が家計の消費行動に与えた影響を分析したいとしよう。このとき消費増税の影響を受けた全家計が母集団となる。母集団のすべてを調べる調査を全数調査という。これに対して、母集団の一部のみを調査対象とする調査を標本調査という。標本調査から得た母集団の一部を標本とよび、標本に選ばれた家計の数を標本の大きさや標本サイズという。

母集団全体のデータがあれば、母集団の特性を直接調べることができる。しかし、母集団は大きい集団である場合が多く、その場合には、費用や時間の制約から全数調査を行うことは現実的ではない。例えば、消費増税の影響を受けた家計が分析対象であるとき、増税が行われた国や地域に住む全家計が母集団であり、これら全家計をすべて調査することは容易ではないことが想像できるであろう。このため実際のデータ分析では、母集団の一部である標本を用いて、母集団の特性を調べることが多い。

分析から特徴を明らかにしたいのは母集団であることに注意してほしい。分析に使えるのは、母集団の一部である標本だけであるから、標本が母集団を代表するようなものになっており母集団の特徴を引き継いでいなければ、標本から母集団の特徴を調べることはできない。そこで、母集団から標本をどのように抽出すればよいのかという問題が生じる。

標本は母集団の一部であるため、その取り方により「ばらつき」が生じる。ばらつきは、分析の精度に影響する。標本サイズを大きくして、標本を母集団に近づければ、分析の精度を高めることができるが、標本サイズを大きくするにはコストがかかる。それでは、納得できる分析精度を保つことができる標本サイズはどのように決めることができるだろうか。

このように、標本を用いて母集団の特性を分析するためには、標本の抽出方法や標本サイズの決め方が重要な問題であり、標本調査法では、この問いに対して学術的にアプローチする。標本調査法は重要なテーマであるが、本書ではこれ以上取り扱わない。標本調査法について詳しく知りたい場合には、参考文献リスト

の土屋（2009）やその他の関連図書を参照することをお勧めする。

1.3 記述統計学

記述統計学では、標本の特徴を明らかにする手法を取り扱う。統計学を本格的に学習したことがなくても、平均値や偏差値という言葉は耳にしたことがあるだろう。例えば、$1, 2, \cdots, 10$ の平均値は、

$$\frac{1+2+\cdots+10}{10} = 5.5$$

である。それでは、平均値をより一般的に表してみよう。まず、n 個の数値を記号 $x_1, x_2, \cdots, x_n$ で表す。先ほどの例に当てはめれば、$n = 10$ で、$x_1 = 1$, $x_2 = 2, \cdots$, $x_{10} = 10$ となる。このとき、$x_1, x_2, \cdots, x_n$ の平均値とは、これら n 個の記号をすべて足し合わせ、n で割ることで得られる値に他ならない。これを式で表せば、

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) \tag{1.1}$$

である。ただし、$\sum_{i=1}^{n} x_i$ は、和記号であり、x_i を $i = 1$ から $i = n$ まで足し合わせることを意味する。記号 Σ はギリシャ文字で、「シグマ」と読む。

平均値の一般公式を表す (1.1) 式において、$x_1, x_2, \cdots, x_n$ は標本サイズ n のデータ１つ１つの点を表す記号である。このように考えるとき、(1.1) 式はデータから平均値を求める公式を示していることが分かるだろう。データの関数として表された公式のことを統計量という。平均値は、標本の重心を表す統計量である。また、標本のばらつきを表す統計量には、分散や標準偏差がある。平均値と分散から、その標本がどのように分布しているのかをおおよそ知ることができる。

データ分析をする前には、標本がどのような特徴を持ち、どのように分布しているのかを必ず調べる。繰り返しになるが、分析対象としているのは母集団であり、分析に使う標本は母集団の一部である。このため、標本が適切に分布しているのかを確かめることは非常に重要なタスクとなる。標本の分布を表す平均値や分散を表や図にまとめたものを、記述統計という。記述統計学は、データの特徴を表す様々な統計量の開発や、それらの意味と性質を明らかにする学問である。

1.4　推測統計学

統計学の歴史は長く、すでに 17 世紀にはデータを用いて物事の規則性を知ろうとする試みが行われていた。初期の統計学では、手元にあるデータの特徴を明らかにすることを主な目的としていた。一方、近代的な統計学とよばれているものは、Karl Pearson（1857〜1936）や、Ronald Fisher（1890〜1962）らの功績により 19 世紀後半から急速に発展したもので、標本から母集団特性を推測することを目的としている。初期と近代の統計学では、手元にあるデータを分析するという点では共通しているが、何を知りたいのかに関して両者には大きな違いがある。近代統計学では、興味の対象は母集団であり、標本から母集団特性を明らかにすることを考えるのである。

興味対象を母集団に移したことで発展したのが統計的推測である。標本は母集団の一部であるため、その取り方によりばらつきが生じる。このばらつきは、標本から推測した母集団特性にも反映される。例えば、母集団における所得の平均を知りたいとしよう。母集団から無作為に選んだ 100 人の所得データの平均値を計算する。次に、別の 100 人を無作為に選び、所得の平均値を計算する。この 2 つの平均値は、ともに母集団の平均を推測するために計算されたものであるが、両者は一致しないだろう。したがって、1 つの平均値を得たときに、それが母集団の平均の推測として、どの程度妥当であるのかを考えることが必要になる。標本を用いた推測から得られる母集団特性の妥当性を、標本のばらつきを考慮して評価するのである。推測統計学では、標本のばらつきを考慮した手法により、母集団の特徴を明らかにすることを目的とする。標本のばらつきを表すツールとして使われるのが確率である。したがって、近代的な統計学を学習するためには、確率論の習得が必要不可欠となる。

1.5　本書の内容

統計学は、応用される分野によって異なる名前でよばれている。経済データを扱うための統計学は計量経済学とよばれているし、医学データを取り扱うための統計学は医療統計学や生物統計学などとよばれている。経済学、医学、疫学、ビジネスなど、分野によってデータの特徴やデータを分析する目的が異なる。この

ため、統計学は、応用される分野の特徴にあわせて様々な形に発展してきたのである。本書では、応用分野を限定しているわけではないが、経済データの統計分析である計量経済学を学習するための橋渡しを意識した内容になっている。

本書は、近代的な統計学の理論を理解するための基礎知識を学ぶことを目的としている。前述のとおり、近代的な統計学である推測統計学の土台は確率論であるため、2 章から 7 章では確率論を学習する。確率論の内容は、本書で取り扱う統計学の内容を理解するうえで必要最低限のものにとどめており、丁寧かつ可能な限り厳密に紹介している。さらなる学習のためには、確率論の専門書を読むことをお勧めするが、本書の内容を理解しておくことで、専門書を理解しやすくなることが期待できる。

8 章では、推測統計学の基本的な考え方を学ぶ。基本的な考え方を知らずに推定や統計的仮説検定に進んでしまうと、推定や検定を考える理由が分からなくなってしまうため、8 章はしっかりと理解してほしい。推定と仮説検定の方法は、それぞれ、9 章と 10 章で紹介する。9 章では、主に、推定量の良し悪しを評価するための性質や、仮説検定の根拠となる性質について考える。また、10 章において、仮説検定のロジックを丁寧に解説している。

11 章以降は、線形単回帰モデルと二項選択モデルのパラメータの統計的推測（推定と検定）を取り扱う。本書では、確率論を土台とした推測統計学の基礎理論の理解に焦点を当てているため、線形回帰モデルと二項選択モデルは説明変数が 1 つだけの場合に限定している。説明変数が複数あるモデルを考える場合には、線形代数（ベクトル・行列）を用いると、モデルの表記が簡単になり、複数のパラメータの漸近的な性質を一度に示すことができる。このため、大学院レベルの教科書では、線形代数を用いることが一般的である。一方で、ベクトル・行列の演算や性質などの線形代数の知識が十分でない場合には、ベクトル・行列による表記は、推定量の性質の直感的な理解の妨げになってしまうことがある。本書では、説明変数が 1 つだけという限定的なケースのみを取り扱うことで、線形代数の知識なしで読み進めることができるようになっている。この意味で、本書では狭く深い内容を取り扱うことになるが、基礎的なモデルにおいて十分な理解ができていれば、本書以外で発展的な内容を学ぶ際にも有用であると考える。統計分析の手法を広く知ることを目的にしている場合には、本書は適さない。

1.6　記号

1.6.1　ギリシャ文字

統計学ではギリシャ文字が多く使われる。表 1.1 にギリシャ文字とその読み方をまとめた。本文にギリシャ文字が現れたら、その都度この表を確認して、ギリシャ文字に慣れるとよい。

表 1.1　よく使われるギリシャ文字

小文字	大文字	読み方
α		アルファ
β		ベータ
γ	Γ	ガンマ
δ	Δ	デルタ
ϵ		イプシロン
ζ		ゼータ
η		イータ
θ	Θ	シータ
ι		イオタ
κ		カッパ
λ	Λ	ラムダ
μ		ミュー
ν		ニュー
ξ	Ξ	クシー
o		オミクロン
π	Π	パイ
ρ		ロー
σ	Σ	シグマ
τ		タウ
υ	Υ	ユプシロン
ϕ	Φ	ファイ
χ		カイ
ψ	Ψ	プサイ
ω	Ω	オメガ

1.6.2 数の集合を表す記号

自然数全体の集合を $\mathbb{N}$、整数全体の集合を $\mathbb{Z}$、有理数全体の集合を $\mathbb{Q}$、実数全体の集合を $\mathbb{R}$、複素数全体の集合を $\mathbb{C}$、偶数全体の集合を $\mathbb{E}$、奇数数全体の集合を $\mathbb{O}$ で表す。

第 8 章
統計学の考え方

2 章から 7 章にわたり、統計学を学ぶための準備として確率論の入門的な内容を学習した。本章より先は、統計学の基礎となる推定と仮説検定について学習する。手元にデータがあるときに、データをどのように分析するのか、分析結果をどのように解釈するのかなど、統計学の核心的な問いに触れていく。推定と仮説検定を学習した後は、それらの応用として線形回帰モデルの推定と検定を取り扱う。

本章では、推定・仮説検定で何をしているのかを理解するために、背景となる問題意識を紹介する。次章以降で推定と仮説検定の具体的な内容に入るための準備として、初めに、「なぜ推定や仮説検定をするのか」について考えるのである。この問いは、統計理論の根拠の理解につながる重要な問いであるため、本章以降で推定と仮説検定のアイデアを理解するための大きな手助けとなるだろう。

8.1 母集団と標本

統計学は、「多くの場合に正しい結論を導く、意思決定のための判断基準」を提供することを目的の 1 つとしている [*1]。「多くの」場合に正しい結論を導くことを考えているため、統計学は必ず正しい結論を導くことができるわけではない

---

*1 実は統計学にはいくつかの主義があり、それぞれの主義において、統計学のあり方について異なる哲学を採用している。本書では、一般的な統計学の入門書と同じく、頻度主義とよばれる立場から統計学を紹介する。頻度主義の他にも、ベイズ主義、尤度主義などがある。統計学の発展については竹内（2018）が詳しい。また、ベイズ主義の立場からこれらの主義の違いを解説しているものとして、ソーバー（2012）がある。

ことに注意しよう。統計学が提供する判断基準は、確率を土台として構成されているため、高い確率で正しい結論を導く手法を考えることはできるが、それが必ず正しい結論を導くとは限らないのである。それでは、その確率はどこから生じているものなのであろうか。

実例で考えてみよう。大学生が希望する職業を分析するために、大学生 100 人に対してアンケート調査を行ったとする。分析対象の全体の集合を**母集団**（population）といい、分析に使うために母集団から取り出した一部（部分集合）を**標本**（sample）という。この例では、一般的な大学生の希望職種を興味対象としているため、全国の大学生が母集団であり、アンケート調査の対象となった 100 人の大学生が標本である。標本に含まれる大学生の数は、標本の大きさ（サイズ）を表すため**標本サイズ**（sample size）とよばれる。多くの場合、標本サイズは n で表される。大学生の希望職種の例では、$n = 100$ である。標本は集合であるため、標本サイズのことを「標本の数」や「標本数」とはいわない。標本の数や標本数という言葉は、標本という集合自体の個数を意味するため、標本の大きさを表す言葉として適していないのである。

大学生 100 人に対してアンケート調査を行ったのは、全国の大学生が希望する職業について知るためであるから、分析したい対象は標本ではなく母集団である。調査に参加した特定の 100 人の大学生の希望職種について知りたいのではなく、100 人分の情報を使って、どうにか母集団の特性を知ろうとしていることに注意しよう。

100 人の大学生を適当に選び直して再調査をした場合、元の 100 人の大学生の調査結果と同じ結果が得られるであろうか。恣意的に同じ結果が得られるように選ぶということを考えなければ、全く同じ結果を得ることは考えにくい。つまり、標本は、どの部分集合が母集団から選ばれるかということから生じる「ばらつき」を持つのである。この標本のばらつきを考慮するために確率を導入する。

例えば、表 8.1 に示されているデータを得たとする。公務員を希望する学生は 32 人で最も多く、次いで、銀行を希望する学生が 30 人いる。ここで、「学生に最も人気がある職業は銀行員である」と主張する人がいるとしよう。手元にあるデータは、この人の主張と一致していない。しかしデータは、この主張に反するものであるといえるだろうか。あるいは、実際には反しておらず、銀行を希望す

る学生がたまたま少なく抽出され、かつ公務員を希望する学生がたまたま多く抽出されたために、偶然に、主張に反するような結果を得ただけなのであろうか。

表 8.1 大学生の希望職種

職種	公務員	銀行	商社	$\cdots$
希望人数	32	30	23	$\cdots$

このような疑問に答えるために、統計学では、標本を確率変数として捉えることでデータのばらつきを考慮した分析手法を考える。ばらつきがあるため、「学生に最も人気がある職業は銀行員である」という主張が正しいか、あるいは間違っているのかを確実に判断することはできない。その代わりに、「学生に最も人気がある職業は銀行員である」という主張が間違っている場合には、多くの場合にその主張が間違いであるという結論を導くような統計的手法を提案するのである。

標本の 1 つ 1 つの要素を、確定した値ではなく確率変数として取り扱うことを考えてみよう。100 人の学生の希望職種を 1 番目から 100 番目まで適当に並べる。このとき、$i = 1, \cdots, 100$ として、i 番目の学生の希望職種を X_i で表すことにする。つまり、X_1 は 1 番目の学生の希望職種、X_2 は 2 番目の学生の希望職種である。この表記方法を用いると、標本を、

$$\{X_1, X_2, \cdots, X_{100}\} \tag{8.1}$$

で表すことができる。$X_1, X_2, \cdots, X_{100}$ は決まった職種を指しているのではなく、学生 100 人分のデータがあることを想定したときに、i 番目の学生が希望するであろう職種を X_i と表している。つまり X_i は、実際に標本を採取した後にその実現値が決まる変数であり、どのような標本が得られたかによってその実現値が変わる確率変数であると考えるのである [*2]。

確率変数 $X_1, X_2, \cdots, X_{100}$ はどんな分布に従うだろうか。このことを考えるために、標本が母集団から無作為に選ばれているとしよう。この「選ばれる」という言葉は、**抽出**されるともいう。標本を無作為に抽出することを**無作為抽出**（random sampling）といい、無作為抽出により得られた標本を、**無作為標本**

*2 多くの場合、統計学において確率変数は、X, Y, Z, W などの大文字のアルファベットを用いて表される。標本を大文字の X を用いて $\{X_1, X_2, \cdots, X_{100}\}$ と表しているのは、X_i が確定した値ではなく、確率変数であることを表すためである。

（random sample）という。確率変数 $X_1, X_2, \cdots, X_{100}$ が、それぞれ、母集団から無作為に抽出されている場合には、これらは互いに独立（定義 6.9）となる。つまり、任意の i と j について、$i \neq j$ であれば、

$$X_i \text{ と } X_j \text{ は独立}$$

である。また、$X_1, X_2, \cdots, X_{100}$ は、全国の大学生の希望職種という共通の母集団から選ばれたものである。母集団である全国の大学生の希望職種は、何 % が公務員、何 % が銀行という具合で、その分布が定まっているとしよう。このとき、その母集団から無作為に 100 人を選ぶことを考えているため、$X_1, X_2, \cdots, X_{100}$ は、母集団における希望職種の分布に従ってその実現値が決まる[*3]。つまり、確率変数 $X_1, X_2, \cdots, X_{100}$ はすべて同一の分布に従っているのである。一般に、確率変数 $X_1, X_2, \cdots, X_n$ が独立かつ同一の分布に従うことを、**独立同一に分布する**（independent and identically distributed）という。独立同一に分布することは i.i.d や iid などと略して表記されることが多い。

標本は確率変数の集合であり、調査の結果として得られた実現値とは異なる。この実現値のことを**観測値**（observed value）や**データ**（data）ともいう。分析対象は母集団であり、標本から母集団の特性を知ることが分析の目的である。この目的は、観測された確率変数の実現値を用いて、確率変数が従う分布を明らかにすることであると言い換えることができよう。

本書において取り扱う標本は無作為標本であると仮定する。つまり、標本 $\{X_1, X_2, \cdots, X_n\}$ について考えるとき、特に断りがなくても X_i と X_j は $i \neq j$ で独立であり、X_i はすべての $i = 1, \cdots, n$ で同一の分布に従うと考えることになる。

8.2　推定とは

標本で表された関数のことを**統計量**（statistic）という。定義から明らかであるが、統計量は、その標本に観測値を当てはめれば、実際に値を計算することがで

*3　例えば、母集団において、30% が公務員、25% が銀行員、45% はその他を希望しているとき、X_i は確率 0.3 で「公務員」、確率 0.25 で「銀行員」、確率 0.45 で「その他」となる確率変数であると考える。

きる。このように考えると、統計量は計算方法（アルゴリズム）、あるいは、公式を与えている関数であると考えてよい。例えば、**標本平均**（sample mean）は統計量の 1 つである。$\{X_1, X_2, \cdots, X_n\}$ を無作為抽出により得た標本とする。このとき、標本平均は、

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{8.2}$$

である。標本平均は確率変数 $X_1, X_2, \cdots, X_n$ の関数であり、観測値が定まれば実際に値を計算することができる[*4]。

母集団の分布の性質や特徴を表す未知の値のことを**パラメータ**（parameter）という。例えば、母集団の分布の期待値や分散は、母集団の分布の特徴を表すパラメータである。母集団分布に関する未知のパラメータを、標本を用いて言い当てることを**推定**（estimation）という。統計量のうち、推定に用いるものを**推定量**（estimator）という。また、推定量に観測値を当てはめて実際に計算された値を**推定値**という。

推定量は、標本を用いて未知のパラメータを言い当てるための計算公式であるから、計算して得られる推定値は、その未知のパラメータの良い近似になっていることが望ましいだろう。そのためには、良い近似とは何かを具体的に決め、推定量が持つべき性質を考える必要がある。ところで、推定量は標本の関数であるから、推定量自体も確率変数である。統計学では、標本を確率変数として捉えることで、標本から母集団分布を推定する際の誤差（ばらつき）を考慮することを説明したが、その確率的な振る舞いは、推定量にも受け継がれているのである。そのため、推定量が持つべき性質として、推定量の確率的な振る舞いに対する性質を考えることになる。

例えば、(8.2) 式で与えられた標本平均は、母集団分布の期待値の推定量として、ある種の良い性質を持つことが知られている。その性質について少し詳しくみてみよう。標本平均は、確率変数 $X_1, X_2, \cdots, X_n$ で構成された関数であるから、標本平均自体も確率変数である。このため、標本平均の期待値や分散を考えることができる。期待値の線形性（定理 5.1, p.65）を用いると、標本平均の期待

---

*4 平均という言葉を聞くと、テストの平均点や平均時給額などの値を想像するかもしれない。観測値を当てはめて実際に計算された値は標本平均値という。統計量としての標本平均と、実際に計算されて得られる標本平均値は異なるものを指していることに注意されたい。

値は、

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

と表せる。1 つの母集団から無作為に抽出された標本であれば、すべての $i = 1, 2, \cdots, n$ について、X_i は母集団分布と同じ分布に従う確率変数であり、$E(X_i)$ は母集団分布の期待値を表す [*5]。これを $\mu_X = E(X_i)$ で表そう。以上の議論から、

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} \mu_X = \mu_X \tag{8.3}$$

を得る。この式をじっくり眺めてみよう。標本平均の期待値は、母集団分布の期待値に一致しているのである。

次に、標本平均の分散を考えよう。分散の性質（定理 5.2, p.74）から、

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\,\mathrm{Var}(X_1 + X_2 + \cdots + X_n)$$

である。標本は無作為に抽出されているから、$X_1, X_2, \cdots, X_n$ は互いに独立である。したがって、(7.3) 式と定理 7.3 [3]（p.105）から、

$$\frac{1}{n^2}\,\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \frac{1}{n^2}[\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)]$$

が成り立つ。$X_1, X_2, \cdots, X_n$ は同一分布に従うため、これらの分散は一致する。分散を $\sigma_X^2 = \mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \cdots = \mathrm{Var}(X_n)$ で表せば、

$$\mathrm{Var}(\bar{X}) = \frac{1}{n}\sigma_X^2 \tag{8.4}$$

である。分散 σ_X^2 は定数であるから、標本サイズ n が十分に大きければ $\mathrm{Var}\,(\bar{X})$ は小さい値になることが分かるであろう。

以上の議論から、無作為抽出により得た標本 $\{X_1, X_2, \cdots, X_n\}$ の標本平均について 2 つのことが明らかになった。1 つ目は、標本平均の期待値が母集団分布の期待値に一致していることで、2 つ目は、n が十分に大きいときに、標本平均の分散は小さい値をとることである。分散が小さいということは、ばらつきが小さいということを意味するため、標本平均は、母集団分布の期待値に近い値をとる確率が高いことを意味する。このことを、チェビシェフの不等式（定理 5.3,

---

*5 母集団分布の期待値は母平均ともよばれる。

p.76）を用いて確認してみよう。確率変数 Z の期待値を μ_Z、標準偏差を σ_Z で表す。このとき、任意の定数 $k>0$ について、チェビシェフの不等式は、

$$P(|Z-\mu_Z|\geq k)\leq\frac{\sigma_Z^2}{k^2}$$

と表すことができる[*6]。Z を標本平均として、上式に当てはめてみよう。つまり、$Z=\bar{X},\ \mu_Z=\mu_X,\ \sigma_Z^2=\frac{1}{n}\sigma_X^2$ とみなせば、任意の定数 $k>0$ について、

$$P(|\bar{X}-\mu_X|\geq k)\leq\frac{\sigma_X^2}{nk^2} \tag{8.5}$$

が成り立つ。$|\bar{X}-\mu_X|$ は $\bar{X}$ が μ_X からどれくらい離れているのかを表す。定数 $k>0$ は任意の値でよいから、ここでは小さな値としておこう。n が十分に大きければ、右辺は 0 に近い値をとるため、この不等式は、$\bar{X}$ と μ_X が離れている確率が低いことを表している。これは、$\bar{X}$ が μ_X に近い値をとる確率が高いと言い換えることができる。この意味において、確率変数 X_i の標本平均は、X_i の母集団分布の期待値に対する良い近似になっており、良い推定量であるといえる。

標本サイズが大きいときに、推定量が推定したい量に近い値をとる確率が高いという性質は「一致性」とよばれる。標本平均は、一致性の意味で、母集団分布の期待値の良い推定量になっているのである。標本サイズが大きいときに、推定量がどのような確率的振る舞いをするのかを調べることで、その推定量の性質を評価し、統計分析の根拠とするアプローチを**漸近理論**（asymptotic theory）、あるいは、**大標本理論**（large sample theory）という。一致性は、漸近理論に基づいた推定量の性質の 1 つである。一方で、標本サイズが有限である場合にも成り立つ性質を用いて、推定量を評価する理論は、**小標本理論**（small sample theory）という。9 章では、一致性を含む、推定量の評価方法について述べる。

---

*6 確率変数 Z を、

$$Z=\begin{cases}1 & (|X-\mu_X|\geq k)\\ 0 & (|X-\mu_X|<k)\end{cases}$$

と定義する。定理 5.3 の証明で用いた Z の代わりに、上記の Z を用いて証明をすることで、不等式がこの形でも成立することを確かめることができる。