

I. 統計学とは何か？ なぜ必要なのか？

事例

イトーヨーカドー・セブンイレブンジャパンの意思決定（在庫管理・広告・価格設定・出店計画など）では統計学が利用されている。

勝見明『鈴木敏文の統計心理学—仮説と検証で顧客のこころを掴む—』プレジデント社

- ① POSシステム(Point of Sales)によるデータの収集→**標本(sample)データ**の分析
 全く情報がないと意思決定が出来ないので、データを収集して、得られた大量のデータを効率よく要約する：**統計量**（平均、分散）や**図表**の利用
- ② 標本データの背後にある**母集団(population)**を想定したうえで将来予測を行う
 データはたまたま得られたにすぎないので、不確実性を前提としたうえで意思決定（在庫予測、出店計画など）を実施する必要がある：**確率と統計的推測**（予測）

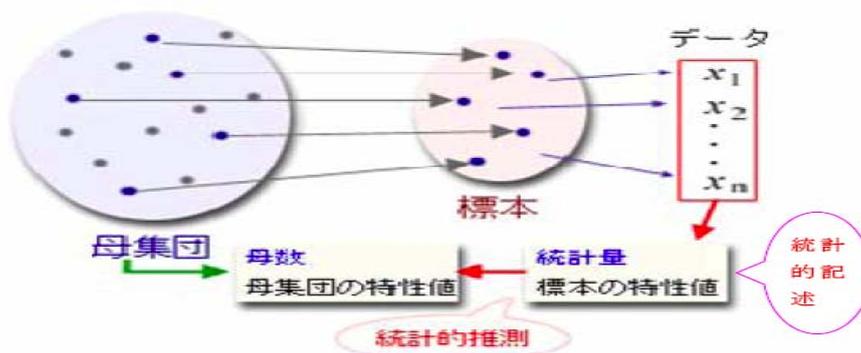
「不確実性の下にある我々の意思決定には統計的方法は不可欠な知恵である」

天気予報、株式投資、受験など

1. 統計学(statistics)の目的

- (1) **統計的記述**：標本データのもつ情報を集約する → **記述統計学**
 17～18 世紀：社会経済現象の数量的把握（ペティ、エンゲル）、国勢学←Status 国家
 「統計学とはたくさんの**標本の情報**から**統計的な規則性**を見つけ出す方法である」
- (2) **統計的推測**：不確実性を前提として標本データの背後にある母集団を推測→**推測統計学**
 不確実性(19c 貴族の賭け事)→確率論
 推測統計学の考え方は 20 世紀に Karl Pearson らにより確立された
 「統計学とは、観察者である我々が利用可能な**標本データ**の分析を行い、さらに、それを発生させた**母集団の構造(資料の発生メカニズム)**を推測する学問体系である」
- (3) **科学の文法**（The Grammar of Science by Karl Pearson）
 理論はデータを利用した統計的方法によって常に検証される必要があり、それにより**理論の一層の改善や発展**が達成される

2. 母集団(population)と標本(sample)



母集団(population) -----無作為抽出-----> **標本(sample)** $x = \{x_1, \dots, x_n\}$ 、標本の大きさ n
 有限母集団 無限母集団 統計量 statistic : $t(x_1, \dots, x_n)$
 母数(parameter): 母平均 μ , 母分散 σ^2 標本ごとに標本平均 \bar{x} , 標本分散 s^2 は異なる

2.1 母集団 (population) p-3

(定義)分析者が調べようとしている対象すべての集合→母集団は調査目的によって決まる。

- ①有限母集団：特定時点の特定選挙区に対する世論調査
 ②無限母集団(仮説的母集団)：新薬の治癒率→「日本人の 50 歳以上の男子の高血圧患者」
 (今日だけでなく将来の患者まで考えると無限に存在する)
 ※母集団は「ある母数(パラメタ)をもつ確率モデル」で表現される
 観測値 $X = \text{系統的な要因 } \mu + \text{確率的誤差 } \varepsilon$

2.2 標本 (sample) の集め方：標本調査 (⇔全数調査) の方法 p-60

- ①偏りのないランダム標本(Random Sampling)
 母集団の代表性を維持(母集団の姿を反映)する偏りのない無作為標本
 効率の良い標本抽出法として層化 2 段階抽出がよく行われる
 無作為抽出 random sampling ために乱数(p-60)が利用される
 ②十分な標本の大きさ(sample size)
 多ければ多いほど、母集団の特徴を正確に反映するようになるが、調査費用が高くなる

2.3 ランダム性を利用した調査方法

- Q_1 「あなたは期末試験でカンニングをしましたか？」 → YES の割合 P_1
 Q_2 「あなたの携帯番号の末尾は偶数ですか？」 → YES の割合 $P_2 = \frac{1}{2}$
 コイン投げをして表なら Q_1 に裏なら Q_2 に正確に答える → YES の割合 P
 $P = \frac{1}{2}P_1 + \frac{1}{2}P_2$ だから $P_1 = 2P - P_2$ より P_1 を求める！

3. 春学期の予定

1. 統計学とは何か
統計的記述
2. 標本データの要約と記述
確率論
3. 母集団の確率モデル(確率と確率分布)
4. 標本の確率分布
統計的推測
5. 統計的推測(母平均の推定)
6. 統計的推測(検定)

(練習問題)

- 1 今年の 2 年生の数学力を調査しようとして、今日、教室にいる学生にテストを行った。
 - (1) 母集団は何か。またそれは有限か無限か
 - (2) この調査方法では偏った結果が得られる恐れがあるのはなぜか
 - (3) 標本の大きさを 100 人(2 年生総数が 1200 人とする)程度とすると、どのような標本調査を行うのがよいと思うか
- 2 以下の調査の母集団は有限か無限か？ 標本調査と全数調査どちらを行うべきか
 - (1) アサヒビールの品質調査 = 無限母集団、標本調査
 - (2) 今年の 2 年生の喫煙率 = 有限母集団、全数調査/標本調査
 - (3) あるテレビ番組の視聴率調査 = 無限母集団、標本調査
- 3 クラスの学生に万引きとカンニングの経験割合を調べてみよう

A. 講義にあたっての注意

(1) class page

<http://www.econ.keio.ac.jp/staff/hk/stat/>

①レジュメ+練習問題のダウンロード ②質問掲示板 ③リンク

(2) テキスト

岩田暁一『経済分析のための統計的方法』東洋経済新報社

レジュメは予めホームページから各自ダウンロードしておくこと

(3) 電卓

平方根が計算できる電卓（1000 円程度の関数電卓で十分）を手に入れてください。

(4) コンピュータ

授業では MS Excel(できれば VBA も)について紹介します。レポート作成では必須です。

B. 成績評価について

絶対評価：[80~100]=A、[60~79]=B、[50~59]=C、[0~49]=D

(1)試験(90%)+レポート(10%)+即レポ(不定期)

(2)レポート（春学期2回、秋学期2回を予定）：Excelの実習。提出しないと単位なし。

(3)試験：春学期末、秋学期末それぞれ実施。電卓のみ持ち込み可。50分。

C 参考文献

(1) 講義のペースについていけないという人

鳥居泰彦『はじめての統計学』日本経済新聞社

(2) 練習問題をこなしたい人

ホーエル『入門数理統計学』培風館

マンズフィールド『統計学入門-ビジネスマンと経済学のために-』多賀出版

(3) 上級

永田靖『統計的方法のしくみ』日科技連 Tips 集

稲垣宣生『数理統計学』裳華房 数学好きに

(4) 統計学の歴史などに興味がある人

蓑谷千風彦『統計学のはなし』、『推定と検定のはなし』、『回帰分析のはなし』東京図書

サルツブルグ『統計学を拓いた異才たち』日本経済新聞社

安藤洋美『確率論の生い立ち』、『多変量解析の歴史』現代数学社

天野徹『統計学の想像力』ハーベスト社

(5) 表計算ソフトを使っての統計学

内田治『すぐわかる EXCEL による統計解析』東京図書

長谷川勝也『EXCEL で学ぶ統計学入門』技術評論社

縄田和満『Excel による統計入門』朝倉書店

縄田和満『Excel VBA による統計データ解析入門』朝倉書店