

目次

第 1 章	経済史	3
1.1	価格理論	3
1.1.1	一般均衡理論	3
1.1.2	現代価格理論	3
第 2 章	計量経済学	4
2.1	枠組み	4
2.1.1	過程	4
2.1.2	ミクロ消費	4
2.1.3	生産	4
第 3 章	構造方程式模型	5
3.1	構造方程式模型	5
3.1.1	歴史的経緯	5
3.1.2	枠組み	5
3.1.3	種々のモデル化	6
3.2	構造方程式模型の例	6
3.3	GLS 推定法	7
3.4	変数選択	8
第 4 章	人類と統計	9
4.1	学問と統計	9
4.1.1	National Supported Work Demonstration	9
4.1.2	Job Corps Study	10
4.1.3	PSID	10
4.2	国家と統計	10
4.2.1	米国	11
4.3	官庁統計	11
4.3.1	RadStats と蜷川統計学	11
4.3.2	英国労働党と 21 世紀の統計学	12
4.4	産業と統計	12
4.4.1	抜き取り検査	12
4.5	技術と統計	13
4.5.1	SAS	13
4.5.2	Stata	13
4.5.3	SPSS	13
4.5.4	Struts	14
第 5 章	参考文献	15

第 1 章

経済史

経済を市場の集合と定義し，市場における価格理論を考えるのがミクロ経済学の主題で，数理経済学といったときはこれを指すことが多い．

1.1 価格理論

1.1.1 一般均衡理論

歴史 1.1.1 (Léon Walras 1834-1910). エコールポリテクニックで数学に失敗して 2 浪目も入れなかったが，記述言語として数学を採用して，市場の集まりとして経済を捉える視点はその後の基礎となる．

1.1.2 現代価格理論

世界恐慌以後，シカゴ大学では Cowles 委員会設立の他に，価格理論 (ミクロの中心) の完成が Samuelson と Hicks によって見られた．

歴史 1.1.2 (Paul Anthony Samuelson 15-09 米). 『経済分析の基礎』[?] で「名声」を得て，『経済学』で「富」を得たという．シカゴの学部で価格理論を，Harvard 大学院で数学を学ぶ．

歴史 1.1.3 (John Richard Hicks 04-89 英). 『価値と資本』(1939) は『経済分析の基礎』[?] とともに，ミクロ経済学を確立した．

第 2 章

計量経済学

2.1 枠組み

2.1.1 過程

GDP, 消費, 投資などのマクロ量, 株価・為替レートなどのファイナンスの量は, トレンド成分が事前に予測困難な場合が多い. そこでこれらを計量的に扱うとき, 過程の実現として統計的モデルを指定することになる.

2.1.2 ミクロ消費

定義 2.1.1 .

- (1) \mathbb{R}^n を財空間, その実行可能な部分集合 $X \subset \mathbb{R}^n$ を**消費集合**という.
- (2) X 上の関数 $U \in \text{Map}(X, \mathbb{R})$ を**効用関数**という. 効用関数は X 上に選好関係を定める.

2.1.3 生産

定義 2.1.2 .

- (1) 財空間の部分集合 $Y \subset \mathbb{R}^n$ を**生産集合**という. 各成分は正ならば財の産出量, 負ならば投入量を表す. $0 \in Y$ な凸集合という仮定が置かれる.
- (2) 価格 $p \in \mathbb{R}^n$ に対して, $S(p) := \{y \in Y \mid \forall z \in Y \ p^\top y \geq p^\top z\}$ を**供給集合**といい, この対応 $\mathbb{R}^n \rightarrow P(\mathbb{R}^n)$ を**供給関数**という.
- (3) 一方で, \mathbb{R}^n 内にグラフを持つ関数を用いて, 生産性を捉えることもできる. $k \leq n$ に対して, 必要財と産出財との対応を表す関数 $f: \mathbb{R}_+^k \rightarrow \mathbb{R}_+$ を**生産関数**という. 生産関数は凹性を持つ.

第 3 章

構造方程式模型

伝統的なマクロ計量模型は「変数 Z を管理することで、 Y が確率的に管理できるとき、 Z は Y の原因である」[Wol160] として構成されている。Pearl の貢献はこれに対する計算機科学分野からの革命と認識できる。

3.1 構造方程式模型

3.1.1 歴史的経緯

歴史 3.1.1 (Jöreskog 35- スウェーデン). 潜在変数を持つ因子分析模型 (測定方程式模型という [豊田 09]) は、相関関係の有無を判別するが、因果関係は明らかにできない。パス解析はこれを可能にするが、可観測量しか扱えない (現代ではパス解析を、その双方向性を強調して**同時方程式**という)。そこで、この 2 つを兼ね備えた手法として (図示した時は端的にそう見える)、経済学と心理統計学で使われていた模型を 1 つの枠組みにまとめたというのが Karl Gustav Jöreskog であり、ETS の統計顧問かつ Princeton の客員教授であった。Sörbom と共に LISREL の著者であり、Uppsala 大学で Wold に指導を受け、そのまま教授である。

因子分析模型における最尤推定量の数値計算法を導出し、その伝道師となることを使命としていた。ETS と Princeton にいた際に、相関構造分析に用いる線型模型の枠組みを作った。一般化した理由の一つとして、LISREL なる統計プログラムの貢献は大きい [Gri16]。ただし、LISREL はその線型な特殊化であり、分かりやすさから一般化には貢献したが、構造方程式全体の広がり・裾の広さについて、誤解も生みやすい。

歴史 3.1.2 (名称の変化). Jöreskog 以前の特に重要なアイデアは [Boc66] が与えた。当時は共分散構造分析と呼ばれていたが、Sörbom (1974) が平均値も構造化して一般化したために、必ずしも良い名称とは言えなくなった。現在は SEM (Structural Equation Model with latent variables) という名称が一般的であるようである。事実、母数の推定が最終目標であり、共分散構造の分析はその手段であるという点でも、後者で呼びたいところである。

要諦 3.1.3 (抽象概念に数理的な定義を与える枠組みとしての SEM). そもそも社会科学・人文科学・行動科学は「構成概念」によって理解することを目指す。これは、複雑に込み入った社会的現象をうまく説明する虚構であるが、構造方程式における潜在変数として数学的な定義を得ることになる。すなわち、実証的な社会科学とは、SEM の構成と検証が中心になる。

3.1.2 枠組み

定義 3.1.4. 回帰関係に矢印を加える行為が、模型の因果模型化の第一段階である。

- (1) 変数間の直接的因果効果は矢印で表される。
- (2) 有向線分の終点とならない変数、すなわち、方程式の左辺に一度も現れない潜在変数を**外生変数**という。独立変数ともいう。
- (3) 変数誤差とは、各可観測量に刺さる空からの矢印で、観測誤差を含め、種々の影響を一つに代表させたものである。
- (4) 潜在変数は楕円で、その指標 (可観測量) は長方形で囲って図示される。潜在変数は因子、これに因果関係によって関連づけられる可観測量を指標ともいう。

応用例 3.1.5. SEM の射程は、従来の因子分析・テスト理論・行動遺伝学の模型・多相データ解析・実験計画・質的データ解析・多変量解析・パス解析・時系列解析を包含する。

例 3.1.6 (LISREL での変数の分類).

- (1) 潜在変数は, その内生・外生の別によって, η, ξ で表す.
- (2) 指標は, その従属性・独立性の別によって, y, x で表す.
- (3) 従属指標の誤差を ϵ , 独立指標の誤差を δ で表す.
- (4)

■

例 3.1.7. Jöreskog, Sörbom による LISREL の他に,

- (1) [McA84] による RAM (Reticular Action Model)
- (2) [Ben80] による EQS (Equations model)

などがある.

■

3.1.3 種々の模型化

模型 3.1.8 (RAM による模型化).

$$t = \alpha_0 + At + u, \quad (t = (f, x)^\top \in \mathbb{R}^{n_1+n_2}, \alpha_0 \in \mathbb{R}, u \in \mathbb{R}^n)$$

- (1) ベクトル $f \in \mathbb{R}^{n_1}$ を構成概念 (=潜在変数), $x \in \mathbb{R}^{n_2}$ が指標 (=可観測量) である. 併せた t を**構造変数**という.
- (2) u は外生変数で, 2つの残差 $u = (d^\top, e^\top)^\top$ からなる.

すなわち, 上の式は次のようにも表せる:

$$\begin{pmatrix} f \\ x \end{pmatrix} = \alpha_0 + \begin{pmatrix} A_a & A_d \\ A_b & A_c \end{pmatrix} \begin{pmatrix} f \\ x \end{pmatrix} + \begin{pmatrix} d \\ e \end{pmatrix}.$$

- (1) 観測値 t を模型の母数 A を用いて表した **plug-in** の形を通じて, $E[t], \text{Cov}[t]$ をそれぞれ A で記述したものを, 期待値構造, 共分散構造という.

模型 3.1.9 (EQS による模型化). RAM は構造変数に対して, 外生変数を対置した. 一方で, 内生変数と外生変数に分解するのが EQS である. その結果, 内生変数が両辺に登場し得る.

$$t_0 = \alpha_0 + A_0 t_0 + \Gamma_0 u$$

- (1) t_0 は内生変数と観測変数を並べたベクトルである.
- (2) A_0 は内生変数から内生変数への, Γ_0 は外生変数から外生変数への寄与率を表す.

模型 3.1.10 (LISREL による模型化). RAM において, 構成概念ベクトル f をさらに外生的なものと内生的なものに分ける: $f = (\xi^\top \eta^\top)$. すると, 構造方程式の特に構成概念間の部分を

$$\eta = \alpha_\eta + B_+ \eta + \Gamma \xi + d_\eta$$

と表せる.

3.2 構造方程式模型の例

例 3.2.1 (労働経済学の例: 賃金-教育の方程式). $W \in \mathbb{R}_+$ を賃金, E を教育年数とする. $LW := \log W$ に対して, 線型回帰

$$LW = \alpha + \beta E + \delta^\top \mathbf{z} + u, \quad u \sim (0, \sigma_u), \alpha, \beta \in \mathbb{R}.$$

を考えることが多い. ただし, \mathbf{z} はその他の説明変数とし, $\text{Cov}[E, u] = 0$ を満たすように取れているとする. ここからあとは統計解析の議論になるが, 果たして経済学としてこの模型で満足かといえ, 全く違う. 教育レベルがシグナルとなって, これが賃金を上げているのかもしれない (中間変数). あるいは操作変数が存在するかもしれない.

- (1) [Gri76] は全米パネル調査 (Natoinal Longitudinal Survey) のパネルデータを用いて上述研究を行った。 z から取り分け、操作変数の候補として IQ を選択した。ただし、 LW, E, IQ を同時に説明する潜在変数 A を仮定して、これが誤差とバイアス (IQ と A の変数誤差と呼ばれている) を持つて

$$IQ = \phi + A + \eta, \quad \phi \in \mathbb{R}, \eta \sim (0, \sigma_\eta).$$

とモデルした。これに対して、TSLS を適用した。

- (2) [Ang91] は census に基づく大データを用いて分析した。米国では胃一定の年齢に達するまで学校からの離脱は原則許されない州が多い。そこで、誕生日が操作変数になっており、教育年数と賃金水準に別個の効果を持っていることを発表した。上述のような模型から TSLS, Wald 推定量をはじめ、複数の推定法による推定量がほぼ等しかったことにより妥当性を主張した。
- (3) これに対して、[Bou95] は古典的な構造方程式模型は、次のいずれかを満たすと、OLS, TSLS は漸近的に強いバイアスを持つてしまうことを示した。
- (a) 操作変数・説明変数のいずれも説明力が弱い時
 - (b) 説明変数・操作変数の数が多い時
- これにより、模型の面でも、解析手法の面でも疑問が残ったわけである。

例 3.2.2 (マクロ経済学の例：マクロ消費関数は構造方程式模型である)。マクロな所得と消費の過程 $(Y_i, C_i)_{i \in [n]}$ に対して、

$$C_i = \alpha + \beta Y_i + u_i \quad (\alpha, \beta \in \mathbb{R}, u_i \sim (0, \sigma_i))$$

とモデル化する。実は、2つの過程 $(Y_i), (C_i)$ の従属性の構造 (経年の上昇トレンドと、季節・景気循環による定常的構造) により、この時点でもう線型回帰模型ではない。これが単なる大量に連立された回帰模型とみなせるには、定常性が必要である (別の時刻のデータに依らない)。定常性の検出には、「和分」の概念が用いられる。

- (1) [Eng87] は初めて本格的にこの模型を考察した。

$$\log C_i - \alpha - \beta \log Y_i = u_i$$

と変換し直した時、 u_i には定常性を仮定できることを指摘した。すなわち、この左辺が和分過程になり、組 (Y_i, C_i) が**共和分関係**にあることを見つけた。

- (2) 一般論として、次は GDP 恒等式と呼ばれる：

$$Y_i - C_i - I_i = O(\epsilon_i).$$

右辺の意味は知らない、小さいということである。 I_i は投資である。このような基本原則の集合体を**国民経済計算**という。

定義 3.2.3 (integration, co-integration).

- (1) 過程 (X_n) が n 次**の和分過程**であるとは、 n 階の差分過程を取ると定常的になることをいう。
- (2) 過程 $(X_n), (Y_n)$ にある線型結合 $\alpha_1 X_t + \alpha_2 Y_t$ が存在して、これが和分過程となるとき、**共和分関係**があるという。

歴史 3.2.4. 1970s 年代の計量経済学の「大荒れの 10 年」を経て、共和分は時系列解析の分野で、擬似回帰を回避する目的で構想された概念で、これがマクロ計量模型にも積極的に取り入れられるようになった。こうして段々と垣根が崩れていった。

3.3 GLS 推定法

SEM においては、OLS の他に、GLS, TSLS などの母数推定法が存在する。

3.4 変数選択

歴史 3.4.1. 1970s は計量経済学者の「大荒れの 70 年代」[Hen80] であった。このとき改めて、外生・内生変数の区別 (すなわち因果連鎖に関して置かれる仮定) は多分に主観的であるという批判が高まった。Hendry は特に、計量模型が識別不能であるときに、外生変数を加えることでそれを回復しようとする試みに対して、その恣意性を批判した。そこで変数選択の枠組みで再注目されたのが Granger 因果 [Gra69] である [杉森 00]。

定義 3.4.2. 定常過程 X, Y がそれぞれ自己回帰模型に従うとする。適切な仮定を課すことで、OSL 推定量は漸近有効になる。この模型の係数の推定量を代入することで得る代入推定量 \hat{X}, \hat{Y} が、それぞれ互いのもう一方の情報を使わないとき、**Granger 因果がない**という。

第 4 章

人類と統計

4.1 学問と統計

歴史 4.1.1 (ICPSR 62-). Inter-university Consortium for Political and Social Research と呼ばれる distributor.

例 4.1.2.

MDRC(Manpower Demonstration Research Corporation) Ford 基金と 6 つの政府機関が共同で創立した (74). 2003 年からは名称を MDRC に変更して元の長い方は捨てた. 社会プログラムにおけるランダム割り当ての普及の先駆者である. NSW の政策評価が, 1990s の政策の指針になった.

司法計画室 (Office of Justice Programs) 次のように移り変わり, 84 年の Justice Assistance Act のにより計画室となる.

- (1) Office of Law Enforcement Assistance (1965 – 1968)
- (2) Law Enforcement Assistance Administration (LEAA) (1968-82) by Omnibus Crime Control and Safe Streets Act (68). Lyndon Johnson の war on crime の主な政策.
- (3) Office of Justice Assistance, Research, and Statistics (1982 – 1984)

Mathematica 政府機関を相手にデータを提供している, Princeton にあるコンサル会社. 1930s に Samuel G. Barton が創業した Industrial Surveys Company → Market Research Corporation of America からスピノフした. 初代社長は Oskar Morgenstern (von Neumann とゲーム理論の基礎を築いた人) で, 殆ど Princeton 大学の数学者・経済学者により運営されていたという. 68 年には New Jersey Income Maintenance Experiment,



4.1.1 National Supported Work Demonstration

1 万人が参加した, 1 年以上の就業体験プログラム.

表題 National Supported Work Evaluation Study, 1975-1979: Public Use Files

管轄 United States Department of Justice(アメリカ合衆国司法省). Bureau of Justice Statistics(司法統計局).

概要 前科者, 薬物中毒既往者, 被生活保護 (AFDC) 給付女性, 学校中退者からなる, 従来から雇用問題に悩んだ群への経時的プログラムである "National Supported Work Demonstration project" の政策評価.

NSWD project 政府の助成金による (subsidized) 15 ヶ所で行われたプログラムで, 1 年から 1 年半に支援を受けた状態での就業体験が, 本来の意味での就職に助けになるかの政策. 6 つの政府機関からの代表者, 実験実行者, 実行箇所の役員を含む 175 名を含む会議が 1974 に行われ, そこで計画が定まり, 6 月からの半年間の予算が 19 都市に分配され, うち 13 ヶ所が, 1975 年の 5 月から 7 月の間に実験を開始した. 1976 年に追加で 2 ヶ所実験を開始した. うち 1 ヶ所は中断されたが, 実験は合計で 4 年続き, 合計で 1 万人が参加した.

Evaluation Study MDRC(Monpower Destratation Research Corporation) と司法計画省 (Office of Justice Programs) が 1980 年に発表した研究. 結果, AFDC 受給者に顕著に, 薬物中毒既往者に大きく, 就業率の増加と所得の改善が見られたという. しかし前科者には少しの効果しか認められず, 学校中退者には長期的な効果はなかった.

8 14 が 1 7 の上位互換なのでこの後半しか使わない。そのうち DS8 の 8,11,12,14 のみ使う。

- 8 Baseline. 共変量など.
- 11 1 年経過後の結果変数.
- 12 1 年半経過後の結果変数.
- 14 3 年経過後の結果変数.

4.1.2 Job Corps Study

113265, JC (最大の RCT) の実験データ. American Economic Review の論文 Does the Job Corps Work?[Sch08] に使われた.

労働省によって 16 から 24 歳の国民に提供される職業訓練は Job Corps と呼ばれる.

表題 National Job Corps Study.

研究主体 MPR: Mathematica Policy Research, Inc. は Princeton にあるコンサル会社.

Job Corps 毎年 16 から 24 歳の 6 万人に就業支援を与える予算が労働省にある. 93 年以来, MPR がこの政策評価をしている. そしてなんと 1994 年 12 月からの約 1 年間, 9409 人の応募者を対象に CRT を行ったのである! その後 15409 人を 4 年毎に追跡調査した. すると, 「この調子で効果が持続すれば, このプログラムに掛けた予算の倍の利益が出る見通しである」と結論づけることができた. 1993 から 2003 まで MPR はこの調子で, Social Security and Unemployment Insurance のデータから追跡調査し, その結果 [Sch08] は労働省によって公開された. しかしこの結果があまり持て囃されず, たまに Web から削除されたりするのは, サーベイでのデータと tax records でのデータが一致しないことによる. MDRC も同じ問題に悩んでいる. 特に前科をもつ男の子が多いという. もし tax records を信頼するならば, 全体として Job Corps は社会に利益を与えるとは言えない. 一方で若者への Job Corps はやはり利益が大きい.

4.1.3 PSID

最長のパネルデータ.

概要 5000 家庭, 18000 人の 1968 年以来的の追跡データ. 低所得家庭に関する追加分データもある.

- (1) 子孫への追跡調査, 1997/99 調査と 2017/19 調査で移民を追加したため, 9000 家庭, 26000 人に到達.
- (2) 1968 から 97 までは毎年, 以後は 2 年毎に.

スピンオフ PSID の家庭の中から特定の小標本に注目した.

- (1) CDS (Child Development Supplement) : 97-07 に 5 年毎に子供を中心に追跡調査. 2014-からも始まっており, 全く同じ質問様式で, 今度は 17 歳以下の標本全員を対象としている. 唾液もとっている.
- (2) TAS (Transposition into Adulthood Supplement) : 05-から 2 年毎に, 18-28 の

データ構造 single-year family file 当該年度に記録された事項を, 家庭毎にまとめた. family Interview Number というのが毎年付けられる. 特に, 1968 family Interview Number は ER30001.

cross-year individual file Person Number が ER30002.

4.2 国家と統計

2005 年のデータ

4.2.1 米国

Statistical Programs of the United States Government FY 2005. Office of Management and Budget. 政府の統計に従事する人が1万人いる。日本は2000弱。

大統領府行政管理予算庁首席統計官 (Chief Statistician) が担当。

商務省 センサス局 (Bureau of the Census) にフルタイムで5000人。経済分析局 (Bureau of Economic Analysis) に500人。

労働省 労働統計局 (Bureau of Labor Statistics) に2500人。

農務省 国立農業統計サービス局 (National Agricultural Statistics Service) に1000人、経済研究サービス局 (Economic Research Service) に500人。

保健福祉省 国立保健統計センター (National Center for Health Statistics) に500人。

エネルギー省 エネルギー情報局 (Energy Information Administration) に400人。

運輸省 運輸統計局に150人。

司法省 司法統計局 (Bureau of Justice Statistics) にフルタイムで50人の職員が。

4.3 官庁統計

要諦 4.3.1. [杉森00]の第二章では Hindess から始まって、社会科学の中での統計学の認識論を展開し「社会の中での統計のメディア論的な役割の解明が必要」としている。Hindess の官庁統計論の超克として、[Des02]以来、統計とは統治のための技術的装置であると理解されていることを踏まえた上で、次のように議論する：

一般に、統計は近代主権国家の形成と国民形成のための政治的・イデオロギー的装置として機能したのである。統計の認識論的価値も、これらの機能との関連で評価されるべきものである。だとすれば、社会科学としての統計学は、《利用者》という水準にとどまることはできず、統計の生産者の水準にも関わらなければならないだろうし、統計の生産と流通と利用に効果を及ぼし、かつ統計の生産と流通と利用の諸過程によって作用を受ける社会体の諸実践に構成的に関与する社会的な場に定位し、統計的实践の社会的働きを「批評的に」評価できる統計学を我々は実践しなければならないのである。

これを「公民権のための統計学」と著者は呼ぶ。

4.3.1 RadStats と蠅川統計学

歴史 4.3.2 (Barry Hindess 39-19 英 と RadStats). 長年 Liverpool で社会学教授を務め、晩年は Australia 国立大学で。彼の booklet 『社会学における官庁統計の利用』[Hin73]はRadStatsグループを1975発足させた(70人だったという[Sha05])。その [website](#) には次の一節がある。

We believe that statistics can be used to support radical campaigns for progressive social change. Statistics should inform, not drive policies. Social problems should not be disguised by technical language.

これは、1968年に、生物兵器の開発研究への懸念から科学の責任を問うという理念で Bertrand Russell と Click, Bragg と Julian Huxley ら科学者が BSSRS (British Society for Social Responsibility in Science) を創設したことに刺激を受けてのことである(が、現在も続くのはRadStatsのみ)。実際、多くのメンバーは被っていたという[Sha05]。メーリングリスト、メディア、出版物、年次会議などを通じて、官庁統計と政策決定について考え議論する場を持つ[Eva16]。そもそも1970sは、第二次世界大戦後から準備しつつあった統計による政策決定の試みなどが花を咲かせた時期で、学術界では統計的なポストが増え、また官庁での統計部門が大きく増え、さらに産業界では製薬会社を中心にして統計学者は引っ張りだこであった時期であった。英国らしいというべきか、同時に統計の使われ方、政治家による引用のされ方に強く疑問を持つ流れも拡大した。Radical Philosophy, Radical Midwivesなども同時期にできた。ただし、Radical Statisticsとは何か、あるいはその歴史の振り返りなどは、今の所ないという

[Eva16]. ただし, RadStats Issue 90 [Sha05] では

歴史 4.3.3 (蜷川虎三 97-81). 日本での官庁統計とその問題点の認識は, 蜷川統計学が大きい. [蜷川 35] で統計利用の基本的問題を議論し, 京都大学教授を務めるも, 45 年に戦争責任を自覚し辞職. その後政治家に転身して京都府知事も務める. ここで蜷川が経済学者という事実は肝要である. Keynes 経済学以後は, 国家の経済・社会的施策の総過程に対するマクロ的視点からの管理という文脈が前面に出たことにより, 官庁統計に関する関心が高く, 経済・社会の全体論的な把握の最も肝心な手段との位置付けから不備や誤りは厳しく批判された. 一方で, より普遍的な立場から分析・批判する見解が出てこず, Hindess 官庁統計論との大きな違いとなっている [杉森 00]. もう一つの大きな違いはドイツ社会統計学 (社会学の中の統計学) の批判的読解から形成された点であるという. ドイツでは経済は社会の中に埋め込まれた形で理解されることが多く, 資本主義後進国としてのドイツが国家主導の上からの資本主義化のための思想的準備・国家に奉仕する官学としての側面が強かった. 蜷川は, 官学としての側面を批判し, いわば統計利用を民主化する形で社会統計学を再構築しようとした.

4.3.2 英国労働党と 21 世紀の統計学

歴史 4.3.4. (1) 英国の社会統計学者 Claus Moser により, 1960s 年代に政府と一体になって統計制度の機能の再点検がなされた. その結果, 中央統計局 (CSO) は統計管理に関してより強い権限を持ち, 政府統計サービスなる教育の機能も持つ機関が設立された.

(2) 一方で, 1980s にサッチャー政権下で統計制度が見直されたとき, 結果的には「統計抑圧政策」となった. Derek Rayner なる Marks & Spencer CEO に見直しの責任権が渡され, 「情報は公表を第一義的な使命として作成されるべきではない. それは, 何よりも政府が自らの業務に必要なために集められるべきである」ということを旨とする Rayner Doctorine が主軸となり, CSO では予算も人員も 3 割以上削減された [森 99].

(3) すると統計の質が落ちた. 86 年から 87 年にかけて特に問題とされたのが, 国民経済計算の各勘定の計数の間の不突合であった. 88 年度上半期には, 生産統計が年率 6% の成長を計上したのに対し, GDP 統計によればわずか 2.5% しか成長していないという信じがたい乖離が発生した.

(4) これを反発力として 1990s に復活する. 今度は Stephen Pickford らが review をまとめ, 「データを提供する各省庁やイングランド銀行との連携に欠ける」「そもそも内部組織にて国民経済計算の作成に対応できていない」「利用者が許容できる統計の誤差の範囲についての認識が CSO に欠けている」として, 対応として「統計業務の CSO の移管とそれに伴った人員移動」を提示し, CSO は 170 人から一気に 1000 人を超える. 続いて大蔵大臣も Chancellor's initiative として, 統計の質の改善に大きく予算をつけた. また CSO は英国で 57 番目の執行庁 (Executive Agency) となった.

(5) 2000s は英国労働党 "Tony" Blair の政権である. 議会に提出された [Tre98] では, 全国民的な (national), そして気高さ (integrity) の 2 つを原則として, 公開性と信頼性を重視する政策方針を統計についても発表した.

4.4 産業と統計

4.4.1 抜き取り検査

歴史 4.4.1 (抜き取り検査統計化の歴史). 大量生産が加速するにつれて, 大量検査は納期・原価の面で不可能になった. そこで抜き取り検査 (Sampling inspection) が開始された.

(1) OC (Operating Characteristic) 曲線を描いて, 消費者危険・生産者危険を指定する方式であった. Bell 電話研究所では 25 年以来研究されており, 平均出検品質 AOQ (Average Outgoing Quality) が主に関心のある統計量であった.

(2) しかしこの古典的な取り組みでは, 経済計算の観点がなく, 継続的取引や顧客重視という経営実践を反映していなかった. そこでまず, [Dod98] の「数値表」によって統計化されたが, 不合格ロットの全数検査では同様のコスト面で難点を持ち, 工程平均不良率の推定には検査に先立つ実験計画が必要であった [杉森 00]. しかし, 統計的な品質保証の面では大事であった.

(3) ここから前進するには, 検査方式・手法の改良よりも, 真の問題は生産者の工程管理であることに注目する必要がある. 管理図は統計的管理状態を達成するために統計的管理を用いる, という二重の図式が特徴的で, 対象をロットではなく工程そのものに抽象化されたのが特徴である. 現在も, さらに経営管理そのものに管理図の考え方を応用する流れがある

歴史 4.4.2 (Hawthorne 実験を通じた経営学との交差). Western Electric は 1905 年から Hawthorne 工場を持っており, 電気

業界の当時の懸念「タングステン製のフィラメントの発明による効果的な照明が電力需要の低下を招くかもしれない」から、同工場で作業中の照明と効率の関係について MIT に 1923 年から委託研究を行っていた。なんとこの研究の結果得た仮説は「労働者の作業能率は、客観的な職場環境よりも職場における個人の人間関係や目標意識に左右されるのではないか」ということであった！この実験に途中から参加していた Harvard 大学経営大学院の George Mayo は「組織における人間関係論」を展開した。

歴史 4.4.3 (Walter Shewhart 91-67 米). 1926 年に、Bell 研究所での研究結果を現場に適用するために、Dodge と Shewhart は Hawthorne 工場に視察に行く。Shewhart にとってはこれが初めてであった。このときに 2 人を案内したのが検査部にいた Juran である。Deming は 1925 年から Yale の博士過程生で、Hawthorne 工場での短期間研究活動で Shewhard の知遇を得る。

4.5 技術と統計

4.5.1 SAS

拡張子 .sas7bdat, .xpt のバイナリデータ。仕様が公開されているのは後者。

Python との連携 (1) [sas7bdat パッケージ](#)で行える。

```
1         from sas7bdat import SAS7BDAT as SAS
2         with SAS('test.sas7bdat') as file:
3             b = file.to_data_frame()
4             print(b)
5             print変数型 ("-----")
6             print(b.dtypes)
```

(2) pandas で読み込める。

```
1         import pandas as pd
2
3         a=pd.read_sas("class.xpt",encoding='UTF-8')
4
5         print(a)
6         print変数型 ("-----")
7         print(a.dtypes)
```

開発者 SAS Institute が販売。

4.5.2 Stata

拡張子 .dta.

Python との連携 (1) pandas に pd.read_stata() がある。

(2) [StatsModels パッケージ](#)は安定している。

```
1         import statsmodels.iolib.foreign as smio
2         from pandas import DataFrame
3         arr = smio.genfromdta('input.dta')
4         frame = DataFrame.from_records(arr)
```

R との連携 gretl は Stata, SPSS のデータを、R, LaTeX に出力することが出来る。

開発者 StataCorp が開発 (85)。比較的安価な製品でフル GUI で使える。日本の代理店は LightStone 社から購入可能であるが、直接 Stata 本社のホームページから購入も可能である。

4.5.3 SPSS

開発者 IBM の統計ソフトウェア。コミュニティ版が PSPP。

4.5.4 Struts

拡張子 .do. ソフトウェア Struts で作成されたファイル.

開発者 Java ベースのソフトウェアを開発している「Jakarta プロジェクト」が開発したアプリケーションフレームワーク.

挙動 Struts で作成されたこの拡張子「.do」のファイルは、URL を指定してアクションサーブレットを呼び出し、アクションサーブレットが URL からアクション名を取得することで関連したアクションフォーム・アクションクラスを呼び出し（コール）ます。ブラウザから直接 Struts のアプリケーションを開きたい場合は、「.do」が明示的に指定される必要があります。

開発環境 Apache Tomcat は Jakarta specification に沿ったオープンソースの実装. なお, Jakarta EE は Java EE の進化版.

第 5 章

参考文献

参考文献

- [Ang91] Angrist, J. D. and A. B. Krueger: Does Compulsory School Attendance Affect Schooling and Earnings?, The Quarterly Journal of Economics, Vol. 106, No. 4, pp. 979–1014, 1991.
- [Ben80] Bentler, P. M. and D. G. Weeks: Linear structural equations with latent variables, Psychometrika, Vol. 45, No. 3, pp. 289–308, 1980.
- [Boc66] Bock, R. D. and R. E. Bargmann: Analysis of covariance structures, Psychometrika, Vol. 31, No. 4, pp. 507–534, 1966.
- [Bou95] Bound, J. D. A., John and R. M. Baker: Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak, Journal of the American Statistical Association, Vol. 90, No. 430, pp. 443–450, 1995.
- [Des02] Desrosières, A. T. b. C. N.: The Politics of Large Numbers: A History of Statistical Reasoning, Harvard University Press, 2002, フランスの社会学者・科学史家.
- [Dod98] Dodge, H. F. and H. G. Romig: Sampling Inspection Tables: Single and Double Sampling, Wiley Classics Library, Wiley, 2 edition, 1998.
- [Eng87] Engle, R. F. and G. C. W. J.: Co-Integration and Error Correction: Representation, Estimation, and Testing, Econometrica, Vol. 55, No. 2, pp. 251–276, 1987.
- [Eva16] Evans, J. and L. Simpson: Beyond Economic Interests, in The Radical Statistics Group, pp. 205–211, SensePublishers Rotterdam, 2016.
- [Gra69] Granger, C. W. K.: Investigating Causal Relations by Econometric Models and Cross-spectral Methods, Econometrica, Vol. 37, No. 3, pp. 424–438, 1969.
- [Gri76] Griliches, Z.: Wages of Very Young Men, Journal of Political Economy, Vol. 48, No. 4, pp. S69–S86, 1976.
- [Gri16] Grimm, L. G. and P. R. Yarnold: 研究論文を読み解くための多変量解析入門 応用篇, 北大路書房, 2016.
- [Hen80] Hendry, D. F.: Econometrics-Alchemy or Science?, Econometrica, Vol. 47, No. 188, pp. 387–406, 1980.
- [Hin73] Hindess, B.: The Use of Official Statistics in Sociology: A Critique of Positivism and Ethnomethodology, Studies in Sociology, Palgrave Macmillan London, 1 edition, 1973.
- [McA84] McArdle, J. J. and R. P. McDonald: Some algebraic properties of the Reticular Action Model for moment structures, British Journal of Mathematical and Statistical Psychology, Vol. 37, No. 2, pp. 234–251, 1984.
- [Sch08] Schochet, P. Z., Burghardt, J., and McConnell, S., : Does Job Corps Work? Impact Findings from the National Job Corps Study, American Economic Review, Vol. 98, No. 5, pp. 1864–86, 2008.
- [Sha05] Shapiro, J.: Our First Thirty Years, Radical Statistics, Vol. 90, , 2005.
- [Tre98] Treasury, H.: Statistics: A Matter of Trust, published under the 1997 to 2001 Labour government, 1998, 英国の policy paper である.
- [Wol60] Wold, H. O. A. and R. H. Strotz: Recursive vs. Nonrecursive Systems: An Attempt at Synthesis (Part I of a Triptych on Causal Chain Systems), Econometrica, Vol. 28, No. 2, pp. 417–427, 1960.
- [森 99] 森：戦後イギリス統計機構の展開, 法政大学日本統計研究所『研究所報』, No. 25, pp. 265–276, 1999.
- [杉森 00] 杉森滉一, 木村和範：統計学の思想と方法, 統計と社会経済分析, 第 2 巻, 北海道大学図書刊行会, 2000.
- [豊田 09] 豊田秀樹：共分散構造分析 [実践編], 統計ライブラリー, 朝倉書店, 2009.
- [蜷川 35] 蜷川虎三：統計利用に於ける基本問題, PhD thesis, 京都帝国大学, 1935.