

因果効果の測定 (I)

細谷 雄三

要 旨

「統計的連関は因果関係を意味しない」という命題とともに、「他の条件が同じであれば、因果関係を検出するには、対照ランダム実験が観察研究よりも優れている」ということが、経験分析を行う人々の間では共通理解をなっている。この論文は、因果性の意味とその効果測定の問題を、経験的経済分析への適用という視点から、整理展望することを目的とするものである。第I部は、実験的介入、とくに対照ランダム実験による因果効果の立証と対比しながら、計量経済学における因果性概念を論じる。因果性（原因結果の関係）の成立を判定することは、重い実際の帰結をもたらすことが多いため、経験的に観察される同時分布や条件付き分布によるデータ間の連関を調査報告すること以上の慎重さが求められる。経験分析を目的として経済現象をモデル化するとき、あらかじめ理論的に単一のモデルが与えられていて、それ以外の可能性を考える必要がないというケースは稀である。経済経験分析においては、理論モデルの多様性の上に、使用すべき統計モデルの不確定性が重畳することは避けられない。統計的推測は、モデルとデータの対の上で行われる。データは調査・実験から与えられるが、いかなるタイプのモデルにもとづいてデータを分析するかによって、結果として実行される統計的推測、データの解釈、予測に大きな乖離が生じる。経済現象の因果分析も、原因・結果変数を純粹に取り出して関係を調べることは困難であり、主題となる変数を取りまく、こうしたモデル多様性を背景として実行することが必要とされる。しかしながら、一方において、実験的介入に依拠しない観察データにもとづく統計的連関であっても、相関係数のように対称的關係ではなく、方向性が本来的に備わっている連関が存在する。因果分析において、時系列データがもつ利点は、作用とその効果の間の時間経過をとらえることができることである。因果關係の方向は時間の矢の方向と整合すると仮定すれば、時間的先行、退行を伴う変数間の統計的連関の検出は、因果關係の存在の示唆となる。クロスセクション・データはこのような情報をもたらさない。論文の第II部では、時系列モデルに焦点を合わせて、外生性、一方向効果測定、第3系列効果などのテーマを扱う。

〔キーワード〕 因果効果、対照ランダム実験、経済経験分析

1. はじめに

コウノトリが赤ん坊を運んでくると言われているが、日本周辺のコウノトリの生息数の急激な減少が、日本人の近年の出生数減少の原因となっていると考える人はいないであろう。「経験的に観察される変数間の連関は、それら変数間の因果性（原因結果の関係）を意味するとは限らない」という命題は、統計学者の意見を待つまでもなく、当然のこととして受け止められている。

一つの事象の生起がもう一つの事象の生起の原因になっていると判断することは、訴訟においての責任の追及としても、あるいは疫学的対応策の採用や経済政策の決定とそれがもたらす帰結においても、重要な影響をもたらすため、単なるデータ間の連関の観察報告を超える重さがある。このため、統計学においても、複数の変数間の同時分布や条件付き分布のあり方を調査報告することを超えた慎重な考察が求められている。

それでは、一体、事象間のいかなる連関を因果関係と見なすことができるであろうか？ この問題については、古くは、アリストテレスが、著書『自然学』において、原因には質料因、形相因、動力因、目的因の4つがあると記している。近代に入って、思想家デイヴィッド・ヒュームや経済学者ジョン・S.ミルによる因果性の規定があるが、とくにヒュームは経験的事象の間にいかに規則的な継起が観測されても、それは因果的必然を意味しないという主張をして、論理必然と経験的事実を区別した。近年の言及としてしばしば引用されるものに、疫学者ヒルによる因果性の規定がある [Hill (1965)]。ヒルは、ドールとともに、1950年代に紙巻きタバコの喫煙と肺がんの発症との連関を定量的に調査したことでよく知られている

[Doll and Hill (1952)]。彼は1965年の論文のなかで、とくに疫学研究を念頭において、統計的連関を因果性として認定できるための要件として、

- (1) 強い連関
- (2) 反復調査における一致した結果
- (3) 特定された連関
- (4) 時間関連
- (5) 服用-反応曲線の顕著な勾配
- (6) 生物学的妥当性
- (7) 諸研究で既知の事実との整合性
- (8) 実験的証拠
- (9) 類似事例

を挙げている。とくに、要件(5)に関連して、ドール-ヒル論文は、入院患者を喫煙本数で分類したとき、がん患者症例グループと非がん患者対照グループそれぞれにおける喫煙者のオッズを比較するとがん患者において高く、またこのオッズ比が喫煙本数の増加とともに単調に高くなるという傾向を発見している。

コックス-ワームスは、このヒルの要件を広い適用性をもつように次のように敷衍している：

- (1) 分析の主題となる連関について先験的な説明が存在する
- (2) 先験的説明の存在よりは劣るが、主題について納得のできる回顧的な説明が見出せる
- (3) 顕著に大きい効果が見られる場合には、非観測交絡要因による代替的説明の可能性は低い
- (4) 反応変数と説明変数の水準とが、自然な単調関係にある
- (5) 独立な諸研究において効果が、反復してとくに異なる形態で、見出される
- (6) 内因 (intrinsic) 変数との間に大きな相互作用が存在しない
- (7) 従属性がシステムへの大きな介入の結果である。

コックス等は、こうした条件が満たされる場合に、統計的従属性は因果性である可能性が高いとしている、ここで内因変数とは、属性など

調査単位に内在する特徴を表す変数である [Cox and Wermuth (2001), p. 70]。また、因果グラフ経路図 (path diagram) との関連から、コックス等は反応変数に関与する変数を、因果的従属変数、(潜在的) 原因変数、内因変数、説明変数、背景変数、中間変数、という用語で特徴づけている。

ヒルとコックス等の上述の基準のなかでは取り上げられていないが、因果性を特徴づける基準として、統計的連関において、相関係数のように対称的関係ではなく、第7、8節で論じるように、本来方向性が備わっている連関が存在する。とくに時間的前後関係は因果性を認定する場合の一つの重要な基準となる。たとえば、処理が効果に後続する場合は、その処理は原因とは考えられない。クロスセクションデータを扱う統計モデルでは、しばしば、この時間的前後関係が明示されない。

この小論は、因果性の意味とその効果測定の問題を、経験的経済分析への適用という限定された視点から整理展望することを目的とするものであり、著者が独自に導入した概念も第7,8節で紹介している。統計的因果性については、これまでに、膨大な数の研究論文や論評が著されていて、とうてい小論の短い文章で要約できるものではない。

論文はI部、II部の2部構成となっていて、第I部は：1. はしがき、2. 介入と因果効果、3. 経済学と因果性、4. コールズ・コミッション・モデル、5. 経済時系列モデル、の5つの節から成り立っている。後続する第II部は、6. 外生性と因果性、7. 一方向効果の測定、8. 第3系列の存在、9. まとめ、の4つの節から構成される。

2. 介入と因果効果

「連関は因果関係を意味しない」という命題

とともに、「他の条件が同じであれば、因果関係を検出するには、対照ランダム実験が観察研究よりも優れている」ということが、経験分析を行う人々の間では共通理解をなっている。対照ランダム実験では、基本的には因果効果 (causal effects) の測定は、ある処理 (処置) が適用される実験グループでの効果と対照グループでの比較処理 (あるいは無処理) の効果の比較をもっておこなわれる。そして、これらのグループをランダム割り振りすることは、処理以外の要因による作用を処理の作用と確率的に独立とすることで、効果を純粋に検出する (あるいは、他の要因を中立化する) ことを可能にする。すなわち、主題となる処理の因果効果を孤立化して取り出すことが可能となり、さらに、ランダム化が測定値の標準誤差の確率分布評価を可能とするため、検定が実行可能となる。これが、通常、ランダム対照実験の正当とされる理由である。

対照的に、人の介入を伴わない観察研究においては、他の条件一定を実現することが一般に困難であり、因果効果の検出が難しい。後述するように、観察データに多く依拠せざるを得ない経験的経済分析においては、安定的な因果関係を検出する作業は、悪戦苦闘の歴史であったと言っても誇張ではない。

実験を重視する立場では、調査者の介入の結果として観察される効果のみに因果関係を見出し、介入としての処理 (処置) の効果に因果効果という用語の使用を限定する。そして、非実験・観察研究は、対照実験を範型としてそれとの類似 (模擬実験、擬似実験など) と距離から評価されるべきであると考えられる。この立場からは、因果効果が理想的に測定されるのはよく設計された実験によって観察される処理の効果である。たとえば、2種類の処理があるとして、処置と非処置とすると、同一単位への処置効果

と非処置効果の差が真の因果効果である。しかし、この実験を重視する立場では、観察単位の属性というような人的介入の余地のない変数は、原因変数には加えない。たとえば、ある個人の性別が給与に与える効果は、その人が男である場合と女である場合の給与格差が定義できないため、因果関係とはならない。すなわち、調査単位の性別というような内因変数は原因変数とはならない [たとえば、Holland (1986) を参照]。

ネイマンは、1923年の論文において、実験設計と推測を確率論モデルにもとづいて叙述し、その後、多くの研究者が因果性論議の古典的典拠としてしばしば引用している [Neyman (1923)]。論文の中でネイマンは、因果性効果の検出を、農事試験場の実験を例示として用いて、説明している。等面積の m 区画地に分割された圃場の実験を議論する。 U_1, U_2, \dots, U_m は、それぞれの区画地に作付けされた特定品種の作物の真の収穫量であるとする。このとき、平均収穫量は

$$\bar{U} = \sum_{j=1}^m U_j / m$$

として与えられる。異なる区画地からの収穫量の差異は、環境条件の差異に帰着される。異なる ν 品種が存在し、第 k 区画地からの第 i 品種の収穫量は U_{ik} として与えられているものとする。すなわち、

$U_{i1}, U_{i2}, \dots, U_{ik}, \dots, U_{im}$ ($i=1,2,\dots,\nu$) が、第 i 品種の潜在的収穫量である。ネイマンはここで、品種の作付けと収穫を壺からのランダム抽出として確率モデル化する。まず、品種に対応して ν 個の壺を用意する。 i で壺の番号を指すものとする。 i 番目の壺には、区画の数に対応する m 個の球を入れておく。球にはそれぞれの区画における i 種の潜在的収穫量のラベルと区画番号のラベルが貼られている。つまり、 i 番目の壺に入っている各ボールには

$U_{i1}, U_{i2}, \dots, U_{ik}, \dots, U_{im}$ という表現の一つが印されているとする。壺 i から球 U_{ik} が抽出されることは、第 i 品種が区画 k に作付され、収穫量 U_{ik} があったことを意味する。数量

$$\bar{U}_i = \sum_{k=1}^m U_{ik} / m$$

は、全圃場における i 番目の品種からの収穫量の真の値である。各球は壺から非復元抽出されるものと仮定する。さらに、もしある壺から一つあるいは複数のボールがランダムに抽出されると、他の壺すべてから、同じ区画ラベルをもつボールすべてが取り除かれるものと仮定する。すなわち、同一区画地で複数の品種が植えつけられることはない。最終目標は $\bar{U}_1, \bar{U}_2, \dots, \bar{U}_\nu$ の比較、あるいは標本にもとづくこれら真値の推定値の比較である。品種による因果効果は、差 $\bar{U}_i - \bar{U}_j$ ($i \neq j$) として表現されるが、異なる i と j が同一区画に適用されることがないから、この値は観察できない。

さらに、 \bar{U}_i の推定値は非復元抽出にもとづかざるを得ない。すなわち、平均値としては、 $\bar{X}_i = \sum_{k=1}^K X_{ik} / K$ を使用する。ここで X_{ik} は i 番目の壺からの非復元抽出 k 番目の試行の結果で、 K は i 番目の壺からの抽出数である。 U_{ik} は潜在量であり、観察されるとは限らないが、 X_{ik} は観察量であり、 $X_{ik} = U_{ij(k)}$ が成立する、ここで $j(k)$ は k 番目に抽出された区画の番号を表す。品種の違い (原因) の効果は、 $\bar{X}_i - \bar{X}_j$ として推定されるが、この差の分散の計算には相関係数

$$r = \frac{(1/m) \sum_{k=1}^m U_{ik} U_{jk} - \bar{U}_i \bar{U}_j}{\sigma_{U_i} \sigma_{U_j}}$$

の知識が必要とされる。しかし、同じ区画で異なる品種が試験されることがないため、この推定値を標本から直接には得ることができない。ネイマンは、このため便法として論文においては、 $r=1$ を用いることを提案している。

以上のネイマンによる実験モデルの基本的特徴は：

(1) 第 k 区画における品種 i の収穫量は、すべての k と i について、潜在的にあらかじめ与えられている。品種の割り振りが、結果としての収穫量に影響を及ぼすことはない。すべての区画それぞれにおけるすべての品質の潜在的な収穫量は予め与えられていて、壺のボールに印されている。壺からのランダム選択は潜在収穫量を実現させるだけである。選択偏倚の不在、および、因果効果と処理割り振りの独立性がモデル化されている。

(2) ボールの抽出法から、同一の区画においては単一の品種のみが割り当てられ、複数の品種が割り当てられることはない。区画 k における品種 i と j との間の因果効果の差異を

$$U_{ik} - U_{jk}$$

で定義すると、この実験では、この効果は観察することができない。これは、標本単位において、因果効果は観察不可能であることを意味する。

(3) ボールにはあらかじめ、区画と品種から決定される収穫量が書き込まれている。すなわち、ここでは、原因が結果に時間的に先行するなどの、原因と結果の時間性はモデル化されていない。

ここで、(2)の観察不可能性を含む命題は反事実的 (counterfactual) とよばれることがある [Dawid (2000) を参照]。ネイマンの因果モデルの特徴は、反事実的命題を基本的要素として含んでいることである。

ネイマンはこの論文において、実験とその結果の産出を、非母数確率モデルを用いて表現しているが、考え方は回帰モデルを使って次のように言い換えることが可能である。区画地には、

1 から m までの通し番号が付けられているものとして、 $j(i)$ は品種 i の適用のためランダム抽出された区画の番号であるとする、すなわち、 $j(i)$ は $1 \leq j(i) \leq m$ を満たす κ 個の互いに異なる番号である。このとき

$$(2.1) \quad X_{ij(i)} = \sum_{k=1}^{\nu} \beta_k D_{ki} + \varepsilon_{ij(i)},$$

$$i=1, \dots, \nu, \quad 1 \leq j(i) \leq m$$

が成立する。ここで $X_{ij(i)}$ は第 i 品種が抽出された第 $j(i)$ 区画において実現した収穫量を表す。

また、 $\beta_i = \bar{U}_i$ であり、さらに D_{ki} はダミー変数であり、 $k=i$ ならば 1、また $k \neq i$ ならば 0 となる特性関数である。通常重回帰モデルでは、植えつけの実施された区画に通し番号 $j=1, \dots, \kappa\nu$ をつけて、 Y_j を区画 j の収穫量、 $i(j)$ によって区画 j に適用された品種番号とするとき、品種と収穫量の連関は

$$(2.2) \quad Y_j = \sum_{k=1}^{\nu} \beta_k D_{ki(j)} + \varepsilon_j,$$

$$j=1, \dots, \kappa\nu, \quad 1 \leq i(j) \leq \nu$$

と表現される。しかし、このモデルでは介入が表現できない [Rubin (1974) と Holland (1986) は、回帰モデル (2.1) と (2.2) の違いを強調した]。

コックス等は実験の有効性を認めた上で、その限界を次のように指摘している：「他の事柄が同一であるとき、ランダム実験は観察研究よりもはるかに好ましいが、それでもときには深刻な、解釈上の困難がランダム実験に残される。とりわけ重要なものは、観察されない説明変数との可能な相互作用の効果であり、また予期されない将来の介入…である」 [Cox and Wermuth (2001)]。

3. 経済分析と因果性

日常生活では、とくに、望ましくない事態の発生に際して責任のありかを問うという形で困

果関係の追及が行われる。科学研究一般においては、事態の発生や現象の継起を、よく定義された原因結果の連鎖で説明することは、科学的説明であるための、いわば、叙述の作法となっている。事態の文芸的叙述は、それがいかに現実感を引き起こしても科学とは認めてもらえないのが普通である。

マーシャルは、著書『経済学原理』[Marshall (1930), p.36]において、経済学における因果性を次のように特徴づけて、経済分析を他の科学分野と同等の地位においている。「経済学の諸法則は、しばしば、仮説的であると言われている。あらゆる他の諸科学と同様に、経済学が研究対象として引き受けているのは、一定の原因によって生み出される諸効果であるが、これは絶対的諸効果ではなく、他の事柄が等しいという条件の下においてであり、また、諸原因は、それらの効果を邪魔されずに生み出すことができるという条件のもとに置かれている。ほとんどすべての科学的原則には、注意深く、そして、とくに正式な叙述においては、他の事柄は一定であるという但し書きが含まれていることが見出されている：すなわち、問題となる原因の作用が孤立化されることが想定される。一定の結果は、それらの原因に帰着されるが、それはあくまで明確に許容される原因以外の原因が入り込むことは許されないという仮説の上に立ってである。」

当時農学研究などでは実験の設計という方法論が進行していたが、この叙述において、マーシャルが原因の孤立化を介入実験として経済諸現象に対して実現できるものと考えていたと解釈することは難しい。あくまでも理論の枠内での変数の操作を念頭においていると考えられる。他方では、引き続き文章でマーシャルは経済の因果分析が固有にもつ困難を次のように指摘する。「しかしながら、諸原因がそれらの結果を

生み出すための時間が許されるという条件が、経済学においては大きいなる困難の源泉であることは真実である。何故ならば、そうこうするうちに、それら原因が働きかける材料が、そして、原因そのものまでもが変化してしまっているかもしれないからである。経済学が叙述する諸傾向は、自己自身を十全に発揮するために十分な長期間を与えられないであろう。」ここではマーシャルは因果関係の経験的検証を念頭に置いているように見える。[イタリックは原著者による。翻訳のために著者が典拠としているのは第8版(1920年刊行)のリプリント(1930年度版)である。]前節で取り上げた実験による因果性の検証では、時間の問題には特別な考慮が示されていない。他方、グレンジャー因果性では、時間の矢印が考慮されるべき重要な要素である。

マーシャルの議論と対照するために、ウォルドの1956年の論文を取り上げる。ウォルドは、介入実験による因果分析を、段階を踏んで進まない食べ物にならないアーティチョーク(朝鮮アザミ)の料理手順にたとえて、実験において段階的手順を踏んで、因果連関複合を解きほぐしていく過程を、実験的方法のアーティチョーク原理とよんで、これを実験的方法の利点であるとしている[Wold (1956), p.37]。ウォルドは、要約的に、「最も単純化すると、因果問題を段階的に処理するための方策としては：

- (1) 一度には、他の要因は一定にして、一つあるいは少数の制御因果要因のみを変化させる
- (2) ランダム化によって非制御要因の効果を中立化する
- (3) 制御変数の効果が加法的、線形となるように準備する

を挙げることができる。」と述べている。ウォ

ルドの指摘している第1項目はマーシャルの因果効果の検出と形式的には対応しているが、実質には、大きな距離がある。ウォルドは実験の手順を述べているが、マーシャルはモデルの上での変数の操作を述べているに過ぎない。処理のランダム割り振りが実行できない観察データの場合は、(2)が実現できないため、たとえば、回帰モデルにおいて、説明変数と攪乱項との独立性を仮定することが困難となる。(3)は制御変数の変化幅を操作できる場合は、それを適度に小さく設定することで、テラー展開によって加法・線形近似が妥当となるが、そうした操作が出来ない観察研究では、(3)が保証されないため、非線形モデルの導入が必要となるかもしれない。実験の場合は、アーティチョーク原理によって、因果関連を単純な関係に分解して考察ことができるが、(経済学を含む)観察研究では、このアーティチョーク原理を用いた分解ができないため、多数の因果関係をふくむ包括的なモデルを全体として分析対象とせざるを得なくなる。

経済を対象とした初期の経験因果分析としては、ユールによる論文『イングランドにおける生活保護者数変化の原因に関する研究：主として間国勢調査期10年2期間を対象として』を挙げることができる [Yule (1899)]。論文では、1871、1881、1891年国勢調査結果をもちいてユニオンと呼ばれる600の区域について10年間の変化データをクロスセクション分析している。ユールが使用しているのは

$$(3.1) \quad \Delta \text{Paup} = \beta_0 + \beta_1 \Delta \text{Out} + \beta_2 \Delta \text{Old} + \beta_3 \Delta \text{Pop} + \text{error}$$

という重回帰モデルである。(3.1)式において、 Δ は各ユニオンにおける10年間の変化率(パーセント表示)を示し、Paupは生活保護者数、Outは救貧院外救済受給者数と救貧院収容者数の比率、Oldは65歳超の人口比率、Popは

人口である。実際の推定のためには、1871-81年と1881-91年の10年間の変化を、農村、中間部、都市、大都市の4つにグループ分けされた区域に対して最小2乗法を適用している。つまり、8本の回帰式を推定している。たとえば、1881-91年期の都市グループについて、ユールが報告しているのは

$$(3.2) \quad \Delta \text{Paup} = -16.72 + 0.344 \Delta \text{Out} + 0.767 \Delta \text{Old} - 0.369 \Delta \text{Pop}$$

という推定式とともに、

$$\begin{aligned} \text{「回帰式まわりの標準偏差} &= 16.22、 \\ \text{平均値まわりの標準偏差} &= 20.92 \end{aligned}$$

として標本変動を付記している。このときにはまだ、推定値の標準誤差の概念は存在しなかった。この回帰式が原因とその効果を識別することができるかは別にして、論文のタイトルから、ユールは貧困層人口変化の原因の究明を意図して、回帰分析を用いていることは確かである。

経験的回帰分析を経済分析に組織的に使用した研究にティンバーゲンによる景気循環分析がある。それまで生物学や農学などでもっぱら使われていた回帰分析をティンバーゲンは経済関係の分析に使用した [Tinbergen (1939)]。しかし、この回帰分析が、よく設計された対照ランダム実験によって得られたものでなく、生成過程があいまいな観察(非実験)データである経済データへの統計的手法の適用である点を、ケインズは直ちに批判した [Keynes (1939)]。これがティンバーゲンとケインズとの間の論争に発展した [Tinbergen (1940)、Keynes (1940)]。非実験観察データを用いて検出される連関から因果性を推測することにケインズは懐疑的であった。統計分析にプラグマティックな価値を見出すティンバーゲンとの対立は、その後現在にいたるまで、経済分析の相容れない対抗軸であり続けている。

経済学は、元来、経験科学のはずであるが、理論をデータで検証することはあまり熱心には行われていない。人間の主体的・合理的経済行動についての先入観によって経済現象は解明できるという考えがむしろ経済学という学問分野では根強い。データを表やグラフで示すことがあっても、それらは（無視しても構わない）記述的・付録的な役割しか与えられない。経済理論の立場からは、従来、経済モデルは理論から導出されるべきであり、計量経済学者がすべきことは、単にその経済モデルに含まれる未知母数をデータから推定することに限定されるべきと広く考えられてきている [Sutton (2001) を参照]。この理論研究の立場は、今日、実データへの適合や検定を重要視しないカリブレーション法の採用に結実している。

経験分析 (empirical analyses) を目的として現象をモデル化するとき、あらかじめ単一のモデルが与えられていて、それ以外の可能性を考える必要がないというケースは稀である。経済理論モデル自体も研究者の数だけのモデルがある。正しい理論にもとづいて経験モデルを作成することに越したことはないが、何が正しい理論なのかは、見解の分かれるところである。

経済経験分析においては、理論モデルの多様性の上に、さらに、用いられる統計モデルの多様性が重畳する。統計モデルは、実際に観測されるデータ系列そのものが生成される過程をモデル化したものであり、現実についてわれわれがあらかじめ十全たる理解をもつことがない以上、分析に際して、使用すべき統計モデルの不確かさがつきまとうことは避けられない。統計的推測は、モデルとデータの対の上で行われる。データは調査・実験から与えられるが、モデルは通常モデル作成者が選ぶ。いかなるタイプのモデルにもとづいてデータを分析するかによって、結果として実行される統計的推測、

データの解釈、予測に大きな乖離が生じ得る。

自然科学の諸分野では、因果性に関して、「経験データの深層構造として、原因から結果へいたる（中間変数を介した上で）諸作用の系列を理論的に想定することができて、統計分析の役割は、こうした理論的想定を立証することである」 [Cox and Wermuth (2001), p.69] という考え方が支配的であるとも言えよう。こうした、データにもとづく因果性の立証 = 検定という統計的推測本来の役割も、経済学では（ケインズに典型的に見られるように）認知されないことがしばしばあるように思われる。

4. コールズ・コミッション・モデル

経済諸現象を観察データにもとづいて分析する方法を体系化し、計量経済学の理論的基礎を築いたのは、1935年に発足してその後シカゴ大学を主たる拠点としたコールズ・コミッションに集まった研究者達である。メンバーの1人であるハーベルモーは、経済データに観察される変動は確率モデルで表現することができ、また、経済諸関係を連立（多変量）回帰式によって有効に表現することができるという基本的な考え方にもとづいて、経済経験モデル作成とその統計的推測を多変量解析の手法で実行する骨組みを体系化した [Haavelmo (1944)]。またメンバーの1人であるマーシャックは、政策効果の評価に構造方程式モデルを用いることを示唆した [Marschak (1953)]。コールズ・コミッションの研究プロジェクトは、独自の貢献として、経済諸関係が構造方程式の中でいかにして識別されるか、識別可能であるための条件は何かを定式化し、とくに、過剰識別された構造方程式の統計的推定法を開発した。こうした一連の方法論は、その後の経験分析にとっての研究範型（パラダイム）を与えることになっ

た。

マクロ経済の経験分析においては、当初、コールズ・コミッションの範型に従って、同時(連立)方程式法にもとづく統計的推測法の適用の有効性が信じられた。1950、60年代において広く世界的に、ケインズのヒックス解釈である IS-LM モデルを具体化するという方向でマクロ経済の経験モデルの特定化の作業が活発に行われ、コールズ・コミッションの多変量解析手法による経済データ解析が実行された。ケインズとヒックスは、ともに経済分析に統計学的手法を用いることに対して、つまり計量経済学に対して、批判的であった。また IS-LM モデルは必ずしもケインズの考え方を正確に理論化したものではなかった。したがって、IS-LM モデルにもとづく計量マクロ経済分析に対する批判が、その後、ケインズ批判と同等な意味に取られたことは、研究史における 2 重のねじれである。

1960年代に主流であった同時方程式法による統計的推測というコールズ・コミッションの研究プログラムにもとづく経済分析からは独立して(あるいは、孤立して)、経済時系列データに対して、もともとはコルモゴロフやウィナーを中心とした研究者たちが開発した定常過程確率論を用いたスペクトル解析などの時系列統計解析を適用するという研究プログラムが進行していた [Granger-Hatanaka (1964) を参照]。1970年中頃になると、コールズ・コミッション範型に代わる新しい方法の模索の時期に入った。IS-LM マクロモデルと同時方程式法による経験分析の実績(とくに予測)が悪化する。また同時方程式モデルの実際の作成において必要な識別条件を満たすために、外生変数が各方程式にいかにも含められるかが便宜的、恣意的に決定される慣行に対しての批判が起こった [たとえば、Sims (1980) を参照]。

外的諸条件の変化にもかかわらず安定して普遍的に成立する経済変数間の関係を検出し利用することが、経済学の主要な課題である。分析主題となる経済変数は内生変数 (endogenous variables)、観測可能な外的条件は外生変数 (exogenous variables) とよばれる。 p 個の内生変数に関する因果的仮定と確率的仮定を線形モデル化しようとする場合には、内生変数を p -ベクトル $y(t)$ 、外生変数を q -ベクトル $z(t)$ 、 A 、 B を $p \times p$ 、 $p \times q$ の行列であるとして

$$(4.1) \quad Ay(t) + Bz(t) = \varepsilon(t), \quad t=1, \dots, T$$

という線形同時(連立)方程式体系が考えられる。外生変数 $z(t)$ は (4.1) の方程式の外部において決定される変数であり、内生変数 $y(t)$ は (4.1) の連立方程式からその値が定まる。

(4.1) のモデルは構造方程式または構造形 (structural form) とよばれる。行列 A が可逆であるとき (4.1) の構造方程式は完備であるという。 $\varepsilon(t)$ は平均 0、共分散行列 Ω の確率的 p -ベクトル攪乱項 (誤差項) とする。(4.1) を内生変数 $y(t)$ について解いた形

$$(4.2) \quad y(t) = \Phi z(t) + \eta(t)$$

を誘導形 (reduced form) とよぶ(ここで $\Phi = -A^{-1}B$ であり、また誘導形攪乱項 $\eta(t)$ の平均は 0、共分散行列は $\Sigma = A^{-1}\Omega(A^{-1})'$ である)。

同時方程式モデルの推定及び検定は、構造形に対する先験的制約が誘導形母数 Π 、 Σ にもたらず制約と誘導形から導出される尤度関数(またはその代理関数)を基礎として行われる。同時方程式にもとづく内生変数間の関係の統計的推測のためには、まず、関係そのものが識別されなければならない、構造形モデル化における先験的制約の妥当性を前提としている。この先験的制約の妥当性と $\{\varepsilon(t)\}$ のホワイトノイズ過程を仮定すると、統計的推測の観点からは、コールズ・コミッションの同時方程式モデルに

もとづく分析は多変量回帰分析の特殊な事例に帰着する。

マーシャルにおける因果性の観点は、経済モデルが (4.1) 式の形で表現されるとき、一定の内生変数の値が他の内生、あるいは外生変数の変動によって、それ以外の変数の値を一定としたとき、いかに影響を受けるかを問題とする。この効果の大きさは (4.1) では係数 A または B によって決定される。(4.2) のモデルでは、外生変数の変化が内生変数に与える効果は誘導形係数 Φ によって決まる。具体的には、係数行列 Φ の (i, j) 要素を Φ_{ij} 、ベクトル $y(t)$ 、 $z(t)$ の第 i 要素、第 j 要素をそれぞれ $y_i(t)$ 、 $z_j(t)$ と表わすとき、外生変数 $z_j(t)$ の変化 $\Delta z_j(t)$ に対して、(攪乱項を含めて) 他の条件を一定とすると

$$(4.3) \quad \Delta y_i(t) = \Phi_{ij} \Delta z_j(t)$$

として、内生変数 $y_i(t)$ への因果効果を測ることができる。この因果効果は乗数効果とよばれる。しかしながら、内生変数間の因果関係をこのように測ることはできない。構造係数 A の第 i, j 列ベクトルをそれぞれ $A_{(i)}$ 、 $A_{(j)}$ と表わすとき、他の条件を一定とすると、内生変数 $y_k(t)$ の $y_i(t)$ に対する因果効果は

$$(4.4) \quad A_{(i)} \Delta y_i(t) + A_{(k)} \Delta y_k(t) = 0$$

という関係から導かれるが、構造形が完備であるとするとき、(4.4) が成立するのは $\Delta y_i(t) = \Delta y_k(t) = 0$ の場合しかない。つまり、構造方程式 (4.1) では、内生変数相互間のマーシャルの因果性を測定することができない。

いずれにしろ、因果効果は係数が識別されて始めて測定可能となる。一般に誘導形 (4.2) では係数は識別可能であるが、構造形 (4.1) の各方程式の係数が識別可能であるためには、係数行列 A または B に十分な数の制約 (たとえば十分な数の要素が 0) が先験的に課されている必要がある。一般に、誘導形から構造形は

一意的に導出されない。言い換えると、構造方程式が識別されるためには、構造方程式を構成する各一本一本の方程式に、異なる内生変数と外生変数が都合よくちりばめられていることが必要である。しかし、これは経験的な状況で保証されるわけではない。たとえば需要供給の価格決定モデルにおいて、需要側、供給側の経済主体がそれぞれの意思決定において異なる外生的要因しか考慮しないと仮定することは識別のためには好都合であるが不自然である。たとえばブローカーは売り手にもなり得るし、買い手にもなる。いかなる場合に識別 (あるいは過剰識別) が可能となるかを解明したことはコルズ・コミッション研究プロジェクトの貢献であるが、実経済データの分析で識別条件を課することが妥当であるかは疑問のあるところである。この点、とくにリユーは、構造方程式は基本的に過少識別であると主張している [Lieu (1960)]。

これに対して、識別について妥当な市場モデルとしては、バーテン-ベッテンドルフによる魚の価格決定の経験分析を挙げることができる [Barten and Bettendorf (1989)]。

5. 経済時系列モデル

1970年代に入ると、それまでのマクロ計量経済分析に対してさまざまな批判がわきあがった。各国のマクロ経済予測において、同時方程式法がよい結果を生まなくなったこと以外にも、IS-LM モデルにマクロ価格水準の決定を説明するフィリップス曲線を追加したマクロ経済モデルがそれまでよく用いられたが、スタグフレーションという現象がフィリップス曲線と矛盾することからこのアプローチが批判された (フィリップス仮説はそもそもケインズの理論に含まれていなかった)。また、計量経済分析という点からは、系列相関が支配的なマクロ

時系列データの解析に（比較）静学的均衡を表現するのに適した同時方程式法を用いることには無理があった。この頃から、時系列解析として独立に発展してきた定常過程の理論と統計的推測法が経済時系列モデル分析に積極的に取り入れられるようになった。こうして、比較静学のマクロ計量経済分析＝多変量解析から、経験的動学分析＝時系列解析への移行が始まった。

コールズ・コミッションが打ち出した研究範型は、経済経験分析の理論的な明確化へ大いに貢献したにもかかわらず、実際の経済の動きをとらえるという点では、成功したとは言えない。経済経験分析について、クラインは次のように回顧している：「1945年シカゴ大学コールズ・コミッションに属していたわたしたちは、改良された統計的方法、とくに同時方程式推定法の適用が、わたしたちのモデルを質的に向上するために重要であろうと考えていた。この同時方程式法の経験を通してわたしたちは研究テーマについて多くを学んだ。また、いかにモデルを解釈すべきかを学んだ。しかしながら、当時発達していた高級な推測方法を適用した結果として、予測精度が大きく向上することはなかった。後年、モデルの推定に対して、大量のコンピュータ入力を伴う分布ラグを用いるという新しいアプローチの導入が、モデルダイナミックスを把握することに大いに寄与し、予測の改善に貢献した」[Klein (1981)]。

誘導形 (4.2) 式の外生変数として先決内生変数（時間の遅れをもつ内生変数）も含めると、誘導形は外生変数を含む多変量 AR（自己回帰）モデルとなる。それは、 $\{\varepsilon(t)\}$ をホワイトノイズ過程であるとするとき、

$$(5.1) \quad y(t) = \sum_{j=1}^a \Gamma(j)y(t-j) + \Phi z(t) + \varepsilon(t)$$

という形式の多変量 AR (VAR) モデルである。さらに攪乱項部分に弱い従属性 (depend-

ency) を許容するため、MA 表現を取り入れたものが

$$(5.2) \quad y(t) = \sum_{j=1}^a \Gamma(j)y(t-j) + \Phi z(t) + \varepsilon(t) + \sum_{j=1}^b \Psi(j)\varepsilon(t-j)$$

という形の多変量 ARMA (VARMA) (自己回帰移動平均) モデルである。最近の研究では、マクロ経済予測にとって ARMA モデルが AR モデルよりも精度が高いという指摘がされている [たとえば、Athanasopoulos and Vahid (2008) を参照]。モデル (5.1) または (5.2) によって生成される内生変数 $y(t)$ について平均値調整した値 $\{y(t) - E y(t)\}$ が定常である場合は、有限標本にもとづく推定・検定に関して定常時系列解析の統計的推測理論はほぼそのまま適用できる。

ここで誘導形 (5.2) に再度構造を導入したモデルとしては、構造方程式が以下で与えられている場合が考えられる：

$$(5.3) \quad \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \begin{bmatrix} y(t) \\ z(t) \end{bmatrix} = \sum_{j=1}^a \begin{bmatrix} \Gamma_{11}(j) & \Gamma_{12}(j) \\ \Gamma_{21}(j) & \Gamma_{22}(j) \end{bmatrix} \begin{bmatrix} y(t-j) \\ z(t-j) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{bmatrix} + \sum_{j=1}^b \begin{bmatrix} \Psi_{11}(j) & \Psi_{12}(j) \\ 0 & \Psi_{22}(j) \end{bmatrix} \begin{bmatrix} \varepsilon_1(t-j) \\ \varepsilon_2(t-j) \end{bmatrix}$$

ここで $y(t-j)$, $j=0, 1, \dots$, a は p -ベクトル、 $z(t-j)$, $j=0, 1, \dots$, b は q -ベクトル; $\Gamma_{mn}(j)$, $\Psi_{mn}(j)$, $m, n=1, 2$, はすべて $(p+q) \times (p+q)$ 行列の p 対 q , p 対 q 分割行列である。また $\varepsilon_1(t-j)$, $\varepsilon_2(t-j)$, $j=0, \dots, b$ はそれぞれ p -ベクトル、 q -ベクトルのホワイトノイズであり、 $\varepsilon_1(t)$ と $\varepsilon_2(t)$ は直交すると仮定する。このモデルでは、先決変数の生成構造がモデル化されていることが特徴である。このモデルは、コールズ・コミッション・モデルと時系列モデルとの特徴を合成している。ただし、構造係数行列 A の識別問題が再び発生する。この

モデル (5.3) は、ヘックマンの示唆したモデルを若干修正したものである、ヘックマンは、MA 部分を含めておらず、また $\varepsilon_1(t)$ と $\varepsilon_2(t)$ との直交性を仮定していない [Heckman (2000)]。

経済時系列データに共通して見られる特徴は、異時点間観測値の間に正の相関があり、長期的トレンド、中期的循環、季節変動、非系統的変動といった周波数 (周期) を異にする変動から構成されていることである。ただし循環成分は正確な確定的周期性 (線スペクトル) をもつわけではなく、また長期的トレンドは差分 $\Delta y(t) = y(t) - y(t-1)$ [$y(t)$ としては観測値そのものであるよりは、その対数変換したものを使用することが多い] によって除去され、差分過程 $\{\Delta y(t)\}$ が定常な系列となる。 $\{\Delta y(t)\}$ が定常過程であることをモデル化の出発点とすると、その和過程である $\{y(t)\}$ は確定的トレンド (deterministic trend) に定常過程が加わった過程 (トレンド定常過程) になるとは限らず、定常過程の和を含む可能性も出てくる。これは誘導形 (5.1) において、固有方程式

$$\det(I_p - \sum_{j=1}^a \Gamma(j)z^j) = 0$$

の根として単位根 $z=1$ が含まれていることを意味する。このように単位根が含まれているとき (5.1) または (5.2) によって生成される過程を単位根過程 (unit-root process) とよばれる。単位根過程におけるトレンド部分は確定的過程ばかりではなく確率トレンド (stochastic trend) をも含む。ランダムウォークは持続的な上昇あるいは下降過程を生み出すため、トレンドが含まれているように見える。与えられた経済時系列データにおけるトレンドが確率的か確定的かの判別は経済データ解析の基礎的な部分をなしている。

確率トレンドが生成される典型的な経済メカニズムがある。経済時系列 $\{y(t)\}$ が VARMA モデル (5.2) によって生成されていると仮定

する。経済量 $y(t)$ 間に一定の長期的均衡関係 $\beta'y(t) \equiv \mu$ が成立している (β は $p \times r$ 行列でランク r 、 μ は p -ベクトル)、経済メカニズムが働いて $y(t)$ がこの均衡から乖離するとき、均衡へ引き戻す力が働き $\beta'y(t)$ は定常過程 ($\beta'y(t)$ という関係が動学的に安定している) となると仮定する。均衡の安定性が $\beta'y(t)$ の定常性として表現されるとする (このとき行列 β の各縦ベクトルを共和分ベクトルとよぶ)。このとき $\beta'y(t)$ が定常であっても、行列 β が $p \times p$ の非退化行列、すなわち $\text{rank}(\beta) = p$ 、でない限り、 $y(t)$ は単位根過程となる。このとき、 $\{y(t)\}$ は共和分過程 (cointegrated process) とよばれる。この共和分関係を均衡からの乖離に対する変数の調整メカニズムとして明示したものが誤差修正 (error correction) モデル

$$(5.4) \quad \Delta y(t) = \alpha \beta' y(t-1) + \sum_{j=1}^a \Gamma(j) \Delta y(t-j) + \Phi z(t) + \sum_{k=0}^b \Psi(k) \varepsilon(t-k)$$

である。ここで α は $p \times r$ のランク r の行列である。モデル (5.3) は構造形と (本来誘導形である) ARMA モデルを合成したものであるが、モデル (5.4) も ARMA モデルに $\beta'y(t) \equiv \mu$ という構造形を含んだモデルとなっていることが特徴である。

共和分 VAR モデルの統計的推測理論については、ヨハンセンが尤度関数にもとづく推定・検定法を与えている。定常時系列の標準推測理論では尤度比検定統計量は、漸近的に、 χ^2 分布に従うが、共和分モデルではこれが成立しない。(5.4) において、 $z(t)=1$ であり、また $b=0$ である場合、標本サイズ T の共和分ランク r の最大対数尤度を $L_T(r)$ と表すとき、対数尤度比 $2(L_T(p) - L_T(r))$ は漸近的に確率変数

$$\text{tr} \left\{ \int_0^1 dB \cdot F' \left[\int_0^1 FF' du \right] \int_0^1 F(dB) \right\}$$

と同じ分布に分布収束する。ここで B は平均

0 の $(p-r)$ 次元標準ブラウン運動であり、 F は $(p-r)$ -ベクトル過程であり、

$$F_i(u) = B_i(u) - \int_0^1 B_i(u) du,$$

$$i=1, \dots, p+r-1; F_{p-r}(u) = u - \frac{1}{2}$$

として定義される確率過程である [Johansen (1991) を参照]。

因果分析において、時系列データがもつ利点は、作用とその効果の間の時間経過をとらえることができることである。因果関係の方向は時間の矢の方向と整合すると仮定すれば、時間的先行、退行を伴う変数間の統計的連関の検出は、因果関係の存在の示唆となる。クロスセクション・データはこのような情報をもたらさない。また時系列データからは、長期・短期など周波数（周期の長さ）の異なる構成要素間の連関を分析することが可能であるため、因果性についても、たとえば、経済成長率とイールド・スプレッドの因果関係を長期と短期に分解して検出することが可能である。また、非定常時系列である共和分モデルについて第7節で叙述する一方向効果を測定することには、基本的な困難はなく、一方向効果測度を統計的に推定・検定することは可能である [Granger and Lin (1995)、Hosoya (1997) および Yao and Hosoya (2000) を参照]。

経済の因果分析は基本的にモデルにもとづいて行われるが、経済時系列解析において、われわれが直面する困難は、異なる代替的モデルの多様性に足をとられ易いことである。線形モデルにおけるモデルの多様性は説明変数として何を使うかの問題に限定されているが、非線形モデルではこれにいかなる非線形関数を用いるかという底なしの多様性が重層化される。あまりにも多くのモデルを前にして、われわれはいかなるモデル群を選択して、用いたらよいかについてその方法論を必要としている。利用可能

なデータが増え、計算技術（とくに計算速度と記憶容量）が向上することは、それとともに利用可能なモデルも増加するので、問題の解消には繋がらない。必要とされるのは、モデルを比較する方法論である。開けてしまったパンドラの箱は閉じることができない。

参考文献

- Athanasopoulos, G. and Vahid, F. (2008). VARMA versus VAR for macroeconomic forecasting, *Journal of Business and Economic Statistics*, **26**, 237-252.
- Barten, A.P. and Bettendorf, L.J. (1989). Price formation of fish: An application of an inverse demand system. *European Economic Review*, **33**, 1509-1525.
- Cox, D.R. and Wermuth, N. (2001). Some statistical aspects of causality, *European Sociological Review*, **17**, 65-74.
- Doll, R. and Hill, A.B. (1952). A study of aetiology of carcinoma of the lung, *British Medical Journal*, **2** (4797), 1271-1286.
- Dawid, A. P. (2000). Causal inference without counterfactuals, *Journal of the American Statistical Association*, **95**, 407-448.
- Granger, C. W. J. and Hatanaka, M. (1964). *Spectral Analysis of Economic Time Series*, Princeton University Press, Princeton.
- Granger, C.W.J. and Lin, J.L. (1995). Causality in the long run, *Econometric Theory*, **11**, 530-536.
- Haavelmo, T. (1944). The probability approach in econometrics, *Econometrica*, **12**, supplement.
- Heckman, J. J. (2000). Causal parameters and policy analysis in economics: A twentieth century retrospective, *The Quarterly Journal of Economics*, **115**, 45-97.
- Hill, A. B. (1965). The environment and disease: association or causation. *Proceedings of the Royal Society of Medicine*, **58**, 295-300.
- Holland, P. (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, **81**, 945-970.
- Hosoya, Y. (1997). Causal analysis and statistical inference on possibly non-stationary time series, in *Advances in Economics and Econometrics: Theory and Application*, Seventh World Congress Vol.III, Chapter 1.1-33, eds D.M. Kreps and K.F. Wallis,

- Cambridge University Press, Cambridge.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, **59**, 1551-1580.
- Keynes, J. M. (1939). Professor Tinbergen's method, *The Economic Journal*, **49**, 558-570.
- Keynes, J. M. (1940). Comment on Tinbergen's response, *The Economic Journal*, **50**, 154-156.
- Klein, L.R. (1981). Econometric Models as Guides for Decision-Making. The Free Press, London.
- Liu, T.-C. (1960). Underidentification, structural estimation and forecasting, *Econometrica*, **28**, 855-865.
- Marschak, J. (1953). Economic measurements for policy and prediction, in William Hood and Tjalling Koopmans, eds., *Studies in Econometric Method*, 1-26, John Wiley, New York.
- Marshall, Alfred (1920). *Principles of Economics*, Eighth Edition, Macmillan and Co., London (Reprinted 1930).
- Neyman, J. (1923). Sur les applications de la theorie des probabilites aux experiences agricoles: Essai des principes, *Roczniki Nauk Rolniczki*, **10**, 1-51, in Polish. English translation by D. Dabrowska and T. Speed (1990), *Statistical Science*, **5**, 463-80 (with discussion).
- Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, **66**, 688-701.
- Sims, C. A. (1980). Macroeconomics and Reality, *Econometrica*, **48**, 1-47.
- Sutton, J. (2000). *Marshall's Tendencies*, The MIT Press, Cambridge.
- Tinbergen, J. (1939). *Statistical Testing of Business Cycle Theories*, vols. **1** and **2**, League of Nations, Geneva.
- Tinbergen, J. (1940). Reply to Keynes, *The Economic Journal*, **50**, 141-154.
- Wold, H. (1956). Causal inference from observation data, *Journal of the Royal Statistical Society*, **119**, 28-60.
- Yao, F. and Hosoya, Y. (2000). Inference on one-way effect and evidence in Japanese macroeconomic data, *Journal of Econometrics*, **98**, 225-255.
- Yule, G.U. (1899). An Investigation into the causes of changes in pauperism in England, chiefly during the last two intercensal decades, *J. Roy. Statist. Soc.*, **62**, 29-295.