

(前編より続く)

## 1. 和泉論文

最初は、金融分野における因果推論の応用を展望した和泉論文である。計量経済学の手法を含む統計的因果推論は、金融政策の効果や金融市場の変動分析、企業行動やガバナンスの研究に早くから活用されてきた。これを第一の潮流と位置付けたうえで、金融経済研究では注目される機会がまだ少ない他の二つのトレンドを紹介している。

一つは、機械学習との融合によって登場した「因果 AI」と呼ばれる新潮流である。深層学習による AI の性能向上や機械学習の技法発展は、高次元（多変量）で複雑な挙動を示すデータに潜む因果構造を発見し、因果効果の推定や、反実仮定の生成を可能にした。市場変動の予兆や与信判断、政治経済情勢の先行き予測などにおいては、ベテランの経験則や直感が有効であり、そこには実務を通じて蓄積されてきたパターン認識や因果性認知が隠されている。こうした職人芸を AI や機械学習に取り込もうとする動きは過去からあった。森羅万象のデジタルデータ化が進展しつつあり、AI や機械学習の手法を適用することで、膨大なデータを網羅的に探索し、強力な因果発見力を活用することが可能になった。これらは既に、人間には到達不可能な領域に到達している。分野は異なるが、将棋や囲碁において AI が人間を超越したことがその一例であり、レントゲン画像の読み取り、創薬、物流最適化、異常検知などでも人間以上の高い精度やパフォーマンスを示している。

もう一つの潮流は、自然言語処理に基づく因果推論である。金融レポート、決算短信、ニュース、SNS 投稿といったテキストを対象に、因果関係の抽出と分析を試みる領域であり、言葉として人間に直接理解可能な因果構造を扱う分野である。本アプローチは統計的因果推論や因果 AI にはない特徴と重要性を有している。この点を本論文では、「経済現象のように人間の行動が深く関わる領域では、統計的因果だけでは事象を説明することが困難である。その主な理由は、人間が出来事をどう認知し、それにどう反応するかという人間の認知と行動ルール自体が因果構造の中核をなすためである」と端的に指摘している。また、具体例を示してテキストから因果を抽出する手法を紹介している。上場企業の決算短信は、企業を巡る外生環境や企業の戦略選択が決算という結果にどのような因果性をもって影響したかを示している。この因果連鎖の探索を行うシステムが実装・公開されており、その利用事例が示

されている。

本稿は単に手法の紹介にとどまらず、因果推論を「統計的因果」と「経験的因果」という二つの概念枠に整理し、それらの補完性にも焦点を当てている。統計的因果は反実仮想モデルや構造方程式に基づく形式的手法であり、再現性と数量性を強みとする。一方、経験的因果は人間の認知や文脈的理解を含む因果の捉え方であり、稀少事象や未知の状況において直感的判断を支える。この両者の統合的視座が、現代の複雑化・非定常化する金融環境においてますます重要になることを指摘している。

取り上げられている技術分野の幅広さも本稿の特徴である。AI、機械学習、自然言語処理、統計学、計量経済学、マルチエージェント・シミュレーション、オルタナティブデータ、金融のドメイン知識といった異分野の知見を繋ぐ知的格闘技が因果推論のフロンティアを拡げていくことを期待させる。本特集号の編集企画にも、そうした視点が反映されている。

## 2. 市瀬他論文

展望論文で紹介された第三の潮流に属する研究であり、筆者らが開発した最新技法を紹介する論文である。取り上げられている具体例から入ろう。有価証券報告書には MD & A (Management Discussion and Analysis、経営者による財政状態および経営成績の説明) が含まれている。経営者の視点からの業績や財務状況の説明である。企業経営者を、株主重視、取引先や従業員などステークホルダー重視に二分したとすると、二つのグループで因果性の捉え方に相違はあるのだろうか。例えば、コーポレートガバナンスの改善が収益力の改善をもたらしたのか、収益面での余力が生じたのでコーポレートガバナンスの改善に取り組む機会をもてたのか、逆方向の因果性がそれぞれ存在しうる。株主重視の企業経営は前者であり、ステークホルダーを優先する企業経営は後者であるかもしれない。はたして MD & A のテキスト情報から、経営者の頭の中にあるこうした因果性を抽出することはできるのだろうか。「できる」というのが本論文のエッセンスであり、筆者らが開発したツールが紹介されている。

もう一つ紹介されている事例をあげよう。自信過剰な経営者は、自身の意思決定の正確さを過大評価し、確率的事象における不確実性を過小評価する傾向があるという指摘が先行研究によって行われている。こうした経営者は、良好な業績を自身の能力に起因させ、不振の結果を外部要因に帰する傾向もあると推測され、自己奉仕的帰属バイアスという名前が付けられている。上述のような経営者の認知バイアスや過剰な自信は、企業の戦略決定や資本構成、投資行動、業績や株価の予測などに強い影響を及ぼしうる。MD & A のテキスト情報から、どの企業の経営者にそうした傾向が強いかを炙り出すことができる。筆者らの研究によれば、自己奉仕的帰属バイアスと自信過剰・楽観予測は正相関している。また、同バイアスは、CEO の年齢が高いほど、在任期間が長いほど、理工系教育を受けているほど低下することを発見している。同バイアスが高い企業は、配当減少、自社株買い増加、高財務レバレッジ、企業価

値棄損に繋がる M&A の実施という特徴を持つことも検証されている。

金融業で企業評価にかかわる者であれば、筆者らが開発した手法に強い関心を持つであろう。人間にしかできなかった、かつ労働集約的であり、サイエンスとして実行し難かった経営者評価を、定型化し自動処理化することが可能となり、人間による情報処理との比較検証も可能となる。論文では、テキストから因果関係を抽出する方法、これを整理して因果知識グラフにする方法、因果推論における視点の相違を抽出する方法が解説されている。また、機械学習を活用し、経済専門家の視点から因果知識グラフを自動生成するフレームワーク FinCaKG（筆者らによる開発）も紹介されている。計量経済学の視点からは、因果性検証に用いられる操作変数をテキスト解析によって膨大な候補の中から自動探索する技法（ETE-FinCa）も興味深い。

本手法のより一般的な貢献は、因果推論は用いるデータやモデル以上に「観測者の視点」に依存するという指摘である。「金融商品を推奨された/金融教育を受けた個人は、その推奨/教育の情報に基づいてどう意思決定を変えたのか」といった問題設定では、個人の認知や理解プロセスそのものが因果推論の対象になる。そこでは、人間がどのように世界を理解しているか（因果認知モデル）を踏まえたモデル構築が必要となる。このようなアプローチは、行動経済学や認知科学の成果とも接続し、因果推論をより人間中心的な視点から再設計することを可能にする。

### 3. 三内他論文

LLM と統計的因果探索の統合が本論文のテーマである。統計データのみに依存する因果推論では、データを機械的に処理すると直感的にありえないような因果性を提示することがある。アイスクリームの売上が増えると溺死事故が増加するという因果は、気温と水辺のレジャー機会というチャンネルを無視したことによって導き出されうる。分析対象となる統計データにこれら4者が適切に含まれている保証はない。

実際の分析でこうした事態が生じにくいのは、分析者が自分の事前知識に基づき、注目すべき因果性の仮説設定を意識的・無意識的に行っているからである。このため、分析者の持つ情報次第ではバイアスが発生する可能性がある。因果のストーリーの誕生起点がどこにあるのかという問題である。アイスクリームの事例は誰もが持つ常識で解決できる（仮説として取り上げられることすらない）。では、財政支出による経済成長刺激が不足していたから長期にわたる低成長がもたらされたという因果仮説はどうであろうか。考察に入れるべき事象が非常に多く、1枚のグラフで検証できるような仮説ではない。しかし、そこに因果性を導き出したいという強いバイアスを持つ者が少なからず存在する。なぜそうしたバイアスが人の心の中に生じるのかも、社会学に関する広義の因果推論の研究対象として重要であろう。

本論に戻ると、本論文の元となっている研究論文の筆者らは、LLM が膨大な分野の情報を

大量に学習し、多様な専門家の知識やビジネスドメインの常識を備えている点に注目し、因果推論において統計的手法と LLM を融合させる方法を開発した。その着目点を順に整理してみる。

1) 最初に統計データに基づく因果探索を行い、因果知識グラフを作成しておく。各種の統計的分析手法も実施しておく。

2) 次に、検出された因果性、例えば A と B の事象について LLM が有している知識を引き出す。具体的には、「A が B に影響を与えるメカニズムについて、専門知識に基づいて説明してください」と問いかける。その際、各種の統計的手法の分析結果も与え、LLM の推論能力を引き出す CoT (Chain of Thought) の技法を用いる。すなわち、段階的な推論過程を明示させるようなプロンプト (LLM への指示文) を与える。

3) ここまでの分析で得られた因果推論を LLM に与え、これが正しいかどうかを Yes/No で LLM に回答させる。いわば、LLM の知識や推論能力を多重に活用する。確率的言語生成モデルである LLM の出力には揺らぎがあるため、同じ質問を繰り返して Yes/No の平均確率を採用する。

4) LLM から得た情報を統計的手法に活用し、全事象 (変数) 間の因果性を再検証する。以上の分析過程により、統計データに基づく因果探索や統計的手法の情報処理と、LLM が有する情報や推論能力を相互参照的に統合させる。

論文では具体例として、因果探索の研究で広く利用されているデータセット (自動車の燃費に関連した変数セット、気候地形関連の変数セット) や、LLM がデータセット情報を有していない (LLM の学習に利用されていない) 非公開の健康診断データセットを用いた検証を行っている。上記の手法を適用することで因果探索の精度がどれだけ改善するかを計測し、LLM の持つ情報と推論能力が改善に寄与することを確認している。

ちなみに、最新の LLM は専門家が有する医療情報も学習済みであり、健康診断データセットそのものに関する情報はなくとも、医療専門知識や推論能力が貢献していると考えられる。解題筆者は、自からの実証分析研究で大学院クラスの計量経済学の知識やその数式展開、プログラミング実装 (コードの書き方やライブラリの存在・用法を含む) を LLM から引き出しているが、専門教科書と突き合わせてもハルシネーションが生じることが殆どなくなっている。LLM の劇的な進化を因果推論に統合した本手法は、LLM のさらなる進化によってパフォーマンス改善が期待できる。

## 4. 齋藤論文

統計的因果推論では、RCT のような実験介入や、他の環境が同一という仮定のもとで介入前後の変化を比較する手法 (差の差分法)、介入処理群と非処理群で介入以外の要素が類似し

たペアを見つけて比較する手法（マッチング法）、偶然の環境差に依存する方法（回帰不連続デザイン）、操作変数法などがよく利用される。しかし、検証したい内容によってはこうした手法が利用できないケースも少なくない。例えば、EC や音楽映像配信の推薦システムなどパーソナライズドサービスでは提示候補が膨大にあり、単一の施策介入を前提とした分析手法では全く対応できない。ほかにも、「他の環境や要素が類似している」という仮定が成立していなかったり、介入が生じる前後の状況にしか適用できない、適切な操作変数が存在しないといった限界を各手法が抱えている。RCT や A/B テストがユーザー体験を損ねる場合もあるし、新薬テストや治療法選択のように倫理的に実行が難しいケースもある。

こうしたケースに対しては、過去のある環境・ある介入状況のもとで採集されたデータのみに基づいて、複数の介入方策のどれが適切なかの選択を迫られることがある。これを実現する技術はオフ方策評価（Off-Policy Evaluation）と呼ばれており、推薦システムを用いている企業などで実践されている。

本論文は、オフ方策評価の基本概念、主要な推定量、近年の研究動向に焦点を当てて解説したものである。概念解説だけでなく、数式を用いて具体的にどのような計算がなされているのかが平易に説明されている。前述の「過去のある環境」とは、例えばユーザーの視聴・購買履歴、過去の株価推移などであり、機械学習分野の用語法で特徴量と呼ばれる。それが観察されたもとで、ある介入（方策の選択）が行われている。例えば特定の推薦アルゴリズムの実行、株式ポートフォリオの選択などである。これを行動と呼ぶ。所与の特徴量と行動のもとで、報酬（購買、視聴時間、投資損益など）が決定され、これは機械学習などの関数で推計される。期待報酬の最大化を目指すのが、意思決定方策の役割である。

オフ方策評価の目標は、ある特徴量に対して、期待報酬を最大化するような方策が採られ、その結果、報酬が観察されたという「過去の状況に依存した」ログデータだけにに基づき、いまだ試したことがない方策の性能（期待報酬）をより正確に推定する方法（推定量）を構築することである。これは難しい作業であるが、過去データのみから算出可能で、試験コストを要しない、迅速な実行が可能というメリットを有している。それゆえ、実務で利用されており、推定量改善に向けた努力もなされている。

論文では、三つの代表的な推定量が紹介され、その長所短所を実際の分析事例を用いて定量的に解説している。各推定量の精度は、期待報酬を推計する関数の精度や、観測値に含まれるノイズ、利用可能な観測値の数（多いほど精度が安定する）、検証すべき方策の数（多いほど精度にマイナスに効く）などに左右される。実務ではこうした要因を考慮しながら、各問題に適切な推定量が選択されている。

この推定量には、期待値としての正確さと、分散（誤差の振れ幅）の小ささという二つの評価軸がある。資産ポートフォリオ選択問題における平均分散アプローチと同様な性質である。様々な推定量が開発されるにつれ、対処する問題の性質にあわせて、どの推定量を選択すればよいかという実務的な課題が生じている。一つの評価法は、平均分散のトレードオフ

に関する自らの選好を明らかにしたうえで、選択を行うものである。例えば、医療手法の選択のように大きなリスクをとれない場合、分散の最小化に重きが置かれる。こうした発想は、資産ポートフォリオの運用方針の選択と類似している。本稿のもう一つの貢献は、この方針選択に関する新しい指標の提示である。シャープレシオを参考にリスク調整後の性能評価指標を明示的に算出し、その利用方法を解説している。

## 5. 増島・難波論文

最後の論文は、当研究所が実施しているアンケート調査を用いた RCT と操作変数による因果推論の事例である。これらは統計的因果推論の典型的な活用法であり、当研究所でも因果推論の王道的な手法の実践が行われていることを紹介したものである。本論文の詳細は、研究所のワーキングペーパーシリーズで公表されている。

アンケート調査は、株式などの金融資産に加え、暗号資産やステーブルコインなど新しい金融資産を対象としており、米国や中国、ドイツでも同じ調査を実施しているため国際比較が可能となっている。暗号資産などについては日本の企業や組織が実施しているものの中では最大規模のものとなっている。過去 3 回のアンケート調査の集計結果も公表されている。

本論文では、暗号資産を保有している人の特徴分析に焦点が当てられている。一般的な傾向として、男性、若年層、高所得層、低学歴層、リスク回避度が低い層ほど暗号資産に投資する傾向が強いことを明らかにしている。また、インフレ期待が高いほど、インフレ期待と成長期待の不確実性が高いほど、投資傾向が強いことも判明している。

本論文の白眉は、金融リテラシーが暗号資産投資に与える影響について、逆方向の因果性や欠落変数バイアスから生じる内生性を考慮した分析がなされている点である。本研究に先立つ研究（他の研究所スタッフによるもので、研究所ワーキングペーパーとして公表済み）では、金融リテラシーと暗号資産投資には非線形関係があることが指摘されていた。具体的には、金融リテラシースコアが中程度のグループにおいて暗号資産の投資経験割合が高く、高・低グループでは同割合が低いという関係である。本論文の分析は操作変数法を用いることにより、こうした関係が見せかけのものであり、金融リテラシーが高いほど暗号資産の投資経験割合が高まり、金融資産ポートフォリオに占める暗号資産のシェアが高まる可能性を示唆している。

これは、表面的な関係性や通常の回帰分析からは特定できない因果関係を識別したものであり、操作変数法の有効性を示した実証分析となっている。なお、計量分析に必要な条件を満たした適切な操作変数が利用できたのは、アンケートの設計段階からリサーチプランが検討されていたためであり、適切な操作変数が見つからないという同手法の弱点をリサーチデザインで克服している点も本研究の特徴である。

もう一つの特徴は RCT の活用であり、こちらもアンケート設計段階からリサーチデザイン

が実施されている。具体的には、情報提供が暗号資産の購入行動に与える影響を検証している。調査対象者を無作為に二つのグループに分け、何の情報も提供しない対照群とビットコインの収益率に関する情報、「ビットコインの価格は過去 5 年で 6 倍以上、過去 10 年で 100 倍以上になった」を提供した処置群について、1 年後の望ましいポートフォリオを尋ねている。その結果、情報提供が暗号資産の購入意欲や保有割合を高めることを示唆する結果を得ている。この分析は、金融資産に関する情報提供が投資行動を促進する可能性を示唆するものである。

## 6. 研究会報告書

本号の最後には、当研究所が主催している「次世代金融インフラの構築を考える研究会」の第二次提言「例示」を掲載した。この研究会の問題意識は、本号のテーマと関連している。提言書の冒頭にある問題意識の提示では、「金融 API やブロックチェーン技術、ビッグデータの活用によって代表される情報技術の革新によるデジタル化社会の進展に伴い、新しい決済・送金手段、暗号資産などのデジタル金融資産、分散型金融サービス(DeFi)の登場など、金融サービスの提供主体・手段等に変化が生じており、今やデータが収益の源泉となり、情報生産機能の高度化が金融機関の経営課題となっている」と指摘している。

提言書の中でも、以下のような言及がなされている。

- ・「金融・非金融領域を跨ぐデータの活用による利用者ニーズの可視化」
- ・そのための「データ基盤等の整備や IT システムへの実装」
- ・「非金融分野における利用者動向のデータ（会計情報・購買履歴・生産活動情報等）の利活用を通じた金融サービスへの展開」
- ・「業務サービスの一環として派生した顧客データを他の用途に活用することで、需要予測やダイナミックプライシング、推薦システムなどのパーソナライゼーションサービスを進化させて高収益化に繋げるデータ・マネタイゼーション」
- ・「データサイエンス（AI/機械学習/因果推論等）と新しい IT システムの構築手法が活用され、ビジネス実装と運用結果のフィードバックに基づく連続的改善という高速サイクル」
- ・これにより「短期間で効率的にサービスの高度化や新サービスを発見」
- ・「利用者データの権利（CDR = Consumer Data Rights）の法制化」

これらを一言で要約すると、金融サービスは非金融サービスとのデータ連携によって統合され、「情報の循環回転ビジネスモデルやデータ駆動型システムに発展していく」というメッセージとなる。

ここまで解題を読まれた読者は、提言書の内容と因果推論の関連性に気付かれたであろう。因果推論は次世代金融インフラの重要なピースである。また、因果推論のポテンシャルを引き出すためには、1) データの観測・蓄積・活用に必要な IT インフラやこれを支える人材、2) データや IT インフラの標準化・相互互換性の推進と情報のセキュアな流通、3) 分析推進のためのデータサイエンティストとビジネスドメイン知識を持った人材の連携、4) 全体推進役としての経営者や政策当局者の理解が重要となる。

今号の特集、「因果推論の最前線」がその一助となれば幸いである。

(終わり)