



金融マーケティングと因果推論・RCT －施策効果を正しく測れていますか？－

副島 豊（SBI 金融経済研究所 研究主幹）

効果測定できていますか？

うちの会社は、日々A/B テストや施策効果測定をやっている。スマホアプリの改修は毎週行っているし、特別金利適用キャンペーンの効果検証も高い精度でリアルタイムにできている。ポイント配布は、心が傾きかけている潜在的顧客への最後の一押しに有効に使えている一方で、確実に自社サービスを使ってくれる顧客に対しては無駄打ちをしていない。

もちろん、顧客が他社乗り換えしそうな気配を察知すれば、one to one マーケティングで臨機応変に特別オファーを打ち、離脱を抑制している。データに基づいたキャンペーンを収益最大化の観点からの確に行えているし、それを支援する分析チームが社内が存在している。マーケティングやデータ分析にかかるコストも、費用対効果のマネジメントができています。

このように言い切れる金融機関は日本にどれくらいあるだろうか。

落とし穴だらけの効果測定

効果測定は思っている以上に難しい。住宅ローンや定期預金の優遇金利キャンペーンを期間限定で行ったとしよう。前月比や前年比で伸びた新規融資分をキャンペーンの効果とみなしてよいであろうか。否である。実績値にはキャンペーン以外の様々な要因が入り込んでくる。

例えば、期間中あるいは比較対象期に住宅ローン税制ほか各種制度が変更されたり、政策金利が変更された影響が生じているかもしれない。地場産業の急回復や所得の改善など地域経済環境の変化もありうる。過去に盛り上がったローン契約の更新期が来ていたのかもしれない。競合先のキャンペーンの有無も影響する。大型の宅地開発や地域の自然災害の影響が含まれているかもしれない。

こうした様々な要素を一定の正確性をもって排除し、キャンペーンの純粋な効果を計測するのは容易ではない。

RCT という検証手法

実績値の検証だけでは正しい効果測定ができないとすると、どうすればよいのであろうか。効果測定を高い精度で比較的簡単に実行できる手法が存在している。ランダム化比較試験（RCT：Randomized Controlled Trial）である。RCT は因果推論のなかでも強力な検証手法として知られており、農業や医療分野では数十年も昔から活用されてきた。教育や社会保障、経済開発、雇用、貧困犯罪対策など多様な分野において、政策手段の選択に利用されるようになっていく。デジタルマーケティング全盛の現在では、EC や広告、SNS などでも A/B テストやレコメンデーション技術が大規模に活用されており、我々はスマホを触るたびに施策評価実験の被験者となっている。このように、RCT は日常生活の中に深く入り込んでいる。

先に述べた優遇金利キャンペーン効果の計測にも RCT が適用できる。因果推論の説明から RCT の解説に入ろう。

因果推論と RCT

因果推論とは、「あることが起きた原因は何か」「ある行動や政策をとった結果、何が起こるのか」といった原因と結果の関係を明らかにするための考え方や分析手法である。原因と結果の関係——施策効果では、施策（原因）がもたらした効果（結果）——を厳密に計測するには、「仮にその施策を行わなかったら何が実現したのか」を計測し、これと比較せねばならない。

理屈ではその通りだが、実現していない世界を計測するのは不可能である。これが「因果推論の根本問題」と言われるものである。その原因がなかったとしたら結果はどう変わっていたかという「事実に反する想定（反実仮想）」を正確に推計することは容易ではない。

こうした根本問題の最も有力な解決法が RCT である。発想は非常にシンプルである。世界を分割して結果を比較するのである。「反実仮想を実験によって作り出す」という発想に基づいている。具体的には、計測対象について、施策介入を行う集団と比較対照のために介入を実施しない集団に分割し、両者間で結果を比較する。その差が施策介入の効果とみなせる。

両集団の特徴、例えば、性別年齢や所得、居住地区特性、家族構成などが異なると、それが結果に影響を及ぼしうる。このため、集団特性に相違が生じないように実験対象をランダムに振り分けていく。ランダム化比較試験（RCT）の実験対象サンプルが少ないと、この前提を保つのが難しくなるため、集団特性の分布や実験の注目点に応じて一定のサンプル数が必要となる。

A/B テストは RCT をビジネスに応用したものである。アプリのボタンの配置は A 案と B 案のどちらかよいかを検証するために、期間を決めて一部のユーザに A 案、他のユーザには B 案を実施し、結果を比較して反応やパフォーマンスがよかったものを採用する。キャラクタ

ーA と B、価格割引 A と B、キャンペーン A と B の比較など応用範囲は広い。子供の教育と大人になってからの所得の関係を調査する RCT では、結果が分かるのに数十年を要するが、ビジネス利用が活発な A/B テストは、実施と計測、改善のサイクルが早いという特徴がある。A 案が介入案で B 案が現行通りとする A/B テストもありうる。

応用事例

先述の金利優遇キャンペーンをダイレクトメールで行う場合への応用を考えてみよう。潜在的な対象者のデータベースがなんらかのかたちで存在しており、そこには種々の属性情報が含まれているであろう。そのデータベースからダイレクトメールを送る（あるいはアプリでお知らせを告知する）施策介入群をランダムに選択し、残りが比較対照群となる。属性で層化した上で、層内で無作為に介入群と対照群を割り当てると、サンプルの均質性をより確実に担保できる。両群の均質性は重要であり、ランダム化試験後に両群の属性分布を比較し、大きな偏りが生じていた場合には、分析時に属性を共変量としてコントロールするといった補正が必要となる。

介入群は半分である必要はなく、両群の同質性が保たれていれば全体の一部であっても構わない。コストを要する実験では、小さい介入群から実験を始め、有効性の確認を段階的に行っていくことも効率的なアプローチである。また、潜在的な対象者のすべてを対照群にする必要もない。施策を打ちたい対象に特性がある（一定の所得以上など）場合、むしろ特性に合せて対照群を選定する必要がある。そうでないと、所得の違いが結果の違いをもたらした可能性が否定できなくなる。

介入群と対照群が経験する外部要因（税制や経済環境の変化、競合先キャンペーン、大型宅地開発や自然災害など）は平均的には同一になると期待されるため、実験期間の結果を比較すればキャンペーンの効果を計測することができる。

他の金融サービスとのクロスセル効果を狙ったキャンペーンであった場合、RCT の設計においてクロスセルへの誘導を含む介入・対照群の設定も同時に取り込むことができる。例えば、A 群を通常金利、B 群を優遇金利、C 群を優遇金利＋クロスセル導線とすると、B 群と C 群の比較によって導線の追加効果を優遇金利効果と分離して計測することが可能となる。その際、新規口座開設からのクロスセルであったのか、既存口座からのものであったのかを分別計測すると経路別のキャンペーン経路が分析可能となる（詳細には触れないが Intent-to-Treat と Treatment-on-the-Treated の区別が必要となる）。

なお、RCT の実験内容によっては倫理的問題が伴う場合もある。介入が有益もしくは害悪をもたらすことが事前に一定の確度で判明している場合などは、なおさらである。ただし、施策の有効性を検証せず現状維持を続けるという判断もまた、指摘・批判こそ受けにくい、改善の可能性を放置したという点で倫理的問題になりうる。ある施策が有益もしくは害悪で

あることが実験で判明すれば、広く全体に適用することが可能となり、社会全体の便益が向上するのである。ワクチンの臨床試験がその典型例であろう。

見落としやすい内部要因にも有効

施策評価を単純に実績値の前年比などで行う場合であっても、外部要因の影響は気が付きやすい。優遇金利キャンペーンで実績が伸びているように見えるが、実は大型宅地開発のせいであったというのは融資の現場での気づきとしてありえよう。あるいは、自然体で契約が伸びる環境にあるので、むしろキャンペーンを打たないほうがよいという判断が事前になされるかもしれない。

見落としやすい情報として、自社の過去の施策の影響があげられる。以前に同様なキャンペーンを打った顧客とそうでない顧客を考えてみよう。ある集団に住宅ローンのマーケティングを行うと、その一部が反応し成約に至る。結果として、その集団の潜在的顧客は少なくなる（次のキャンペーンでは効果が薄れたように見えてしまう）。もっと単純な事象として、一度目に反応しなかった顧客は二度目にも反応しにくいことも考えられる。

このように特性要因の一つに、過去の施策の有無という内部要因が含まれることは施策評価の際に見落とされやすい。一度すでに薬を飲んだ人と、全く飲んだことがない人を区別せずに、今回薬を飲んだ人の反応をみても本来の効果は見えてこない。また、薬が効いてよくなった人が母集団から抜けていくことの影響を考慮せずに薬の効果を検証すると、効き目が薄いように見えてしまう。

RCT は、こうした内部要因の調整の難しさに対しても効果的である。過去の施策実施が特性要因になっていることに気が付きさえすれば、これをコントロールして実験を行うことができる。具体的には、介入群と対照群の選定において、過去にキャンペーンを受けていない、あるいはX回受けたという前提条件を加えれば、その影響を排除できる。

金融業界での活用の遅れ

広告や小売などで常識になっている効果検証や RCT（A/B テスト）は、金融業界では必ずしも広く普及しているわけではない。データ利活用の掛け声は 10 年前からあっても、社内でデータサイエンティストを組織的に育成していく体制や、ビジネス戦略の定量評価に基づく立案や実施、連続的改善を行う体制の整備、またそれを支える情報インフラへの投資は遅れている。

外注するにしても、ビジネスの現場の問題意識や経験知と擦り合わせて効果的に推進するには、現場スタッフにも因果推論や効果検証の知識がある程度必要となる。金融は情報産業・情報装置産業であるにもかかわらず、そこに経営リソースを大規模投入する意思決定がなかなか進まず、IT やデータ検証が経営戦略部門の中枢に置かれていない。一方で、リスク計量

や経営への活用は 1990 年代より進展しているため、素地がないわけではなく、単に気づきや普及時間の問題かもしれない。実際、近年では DX 推進のためのデータ活用やそのインフラ整備が加速している。

所報特集号の狙い

上述の金融業界事情に加え、因果推論の技法も進化しつつある。自然言語処理や生成 AI 技術と融合することで、数値に基づく因果推論や効果検証だけでなく、人間が因果性をどう認知しているかをテキストというオルタナティブデータから分析しようという試みが登場している。因果性の認知に癖やバイアスがあるとしたら、それが経営や家計行動、金融投資、世論形成、選挙結果などにどう影響を及ぼしているかは重要な検証課題となる。

そこで、当研究所の機関誌「[所報](#)」の最新号（8号）では、「因果推論の最前線」を特集テーマとして、先端的な研究事例や展望論文、レコメンデーション技術の基礎にある数理モデルなどを取り上げた。研究所が行っている[暗号資産等のアンケート調査](#)では 2024 年度調査より RCT を一部に含めており、[集計レポート](#)や[分析論文](#)も公表している。

所報収録の個々の論文は技術的な詳細に踏み込んだものであるため、冒頭に[解題](#)としてエグゼクティブサマリーを掲載している。因果推論や効果検証、RCT の基本には触れていないため、近年出版が相次いでいる教科書や実践書、啓蒙書のリストを解題中の Box 記事に掲載した。

最初の一冊

最初の一冊をお勧めするなら、「RCT 大全：ランダム化比較試験は世界をどう変えたのか」（アンドリュー・リー）である。RCT が如何に多様な分野で活用され、常識や先入観とは反する意外な実験結果を含めて様々な政策効果検証が行われ、社会を善くすることに貢献してきたのが多数の事例をもって示されている。インドのグラミン銀行に代表されるようなマイクロクレジットや少人数学級の効果、職業訓練プログラム、途上国での教科書配布支援、薬物依存対策、薬害サリドマイド禍、選挙活動戦術、寄付促進方策、固定給と歩合制比較、選択のパラドックス（消費財の種類が多すぎると購買促進にむしろマイナス）の真偽などである。また、実験結果を一般化する際の留意点や、倫理的観点からの反対運動に向き合ってきた歴史なども紹介されている。

RCT に限らず効果検証の実施には落とし穴が多数ある。実践に際しては Box 記事にあげたような教科書や実践書に触れることをお勧めする。我流で行うと誤った検証結果になりやすい。解題の執筆後にも良書の出版が続いている。大規模書店で手に取り、あるいは EC 書店のプレビューを眺めて、金融ビジネスへの応用可能性を感じてほしい。