

因果推論の概要

2019年 ver. 1.0
倉谷 隆博

4-1. 因果推論

因果関係は原因が起きたから結果が起きたという関係、あるいは結果の原因が特定できる関係である
(注) 因子間の関係を表すものに相関関係があるが、相関関係は因果関係とは違い、原因と結果の関係ではない。

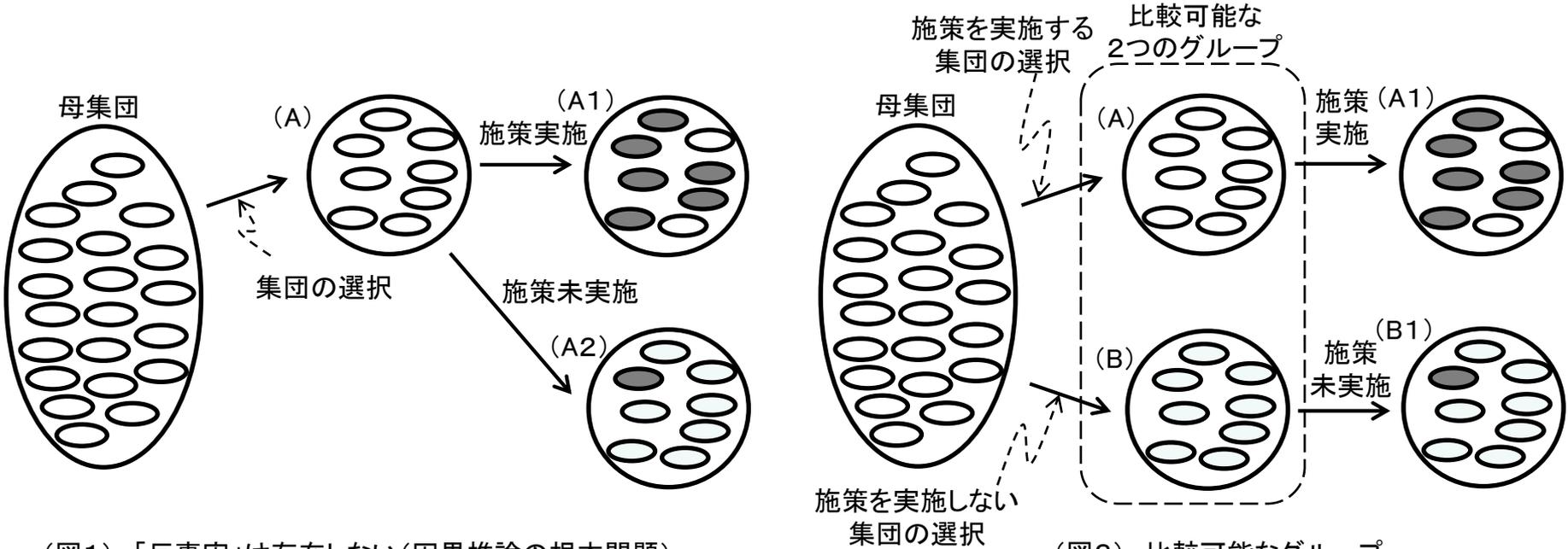
因果関係の存在を証明することを、因果推論という。因果推論はマーケティングなどの施策の効果確認に、つまり、仮説・検証に使用することができる。施策と効果との間の因果関係が明確になれば、施策の見直しなどを行うことができる。

施策の効果(因果関係の存在)を確認するには、(図1)に示すように、「事実」=原因が起こったときの結果、つまり施策を実施したときの結果(A1)と、「反事実」=原因が起こらなかったときの結果、つまり施策を実施しなかったときの結果(A2)とを比較する。差があれば施策の効果を確認できたことになる。とりもなおさず、因果関係を確認できたことになる。

しかし、同じ(A)を対象に施策の実施と未実施を同時に行うことはできない。施策未実施の結果(A2)は「たられば」の結果である。そこで、**結果に影響しそうな全ての特徴は合致していて、唯一の違いは施策を実施したか、しなかったか、だけだとする「比較可能な」2つのグループ(A)と(B)を作り、それらを使って施策実施の結果(A1)と施策未実施の結果(B1)を比較することで施策実施の効果を確認しよう、というのが因果推論の基本的なアプローチ方法である。**

施策効果の検証手順(図2)

- 1) 比較可能な実在する、施策を実施するグループと実施しないグループを作る。
- 2) 2つのグループの結果の差で、施策効果を検証する。



(図1) 「反事実」は存在しない(因果推論の根本問題)

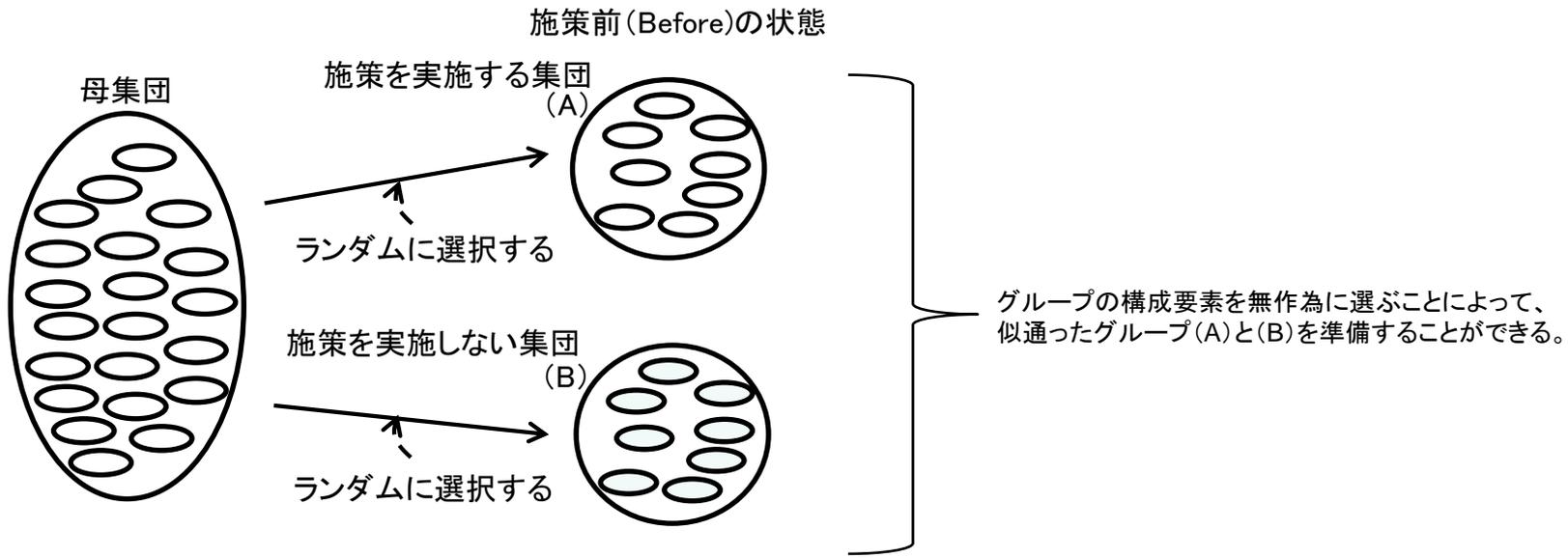
(図2) 比較可能なグループ

4-2. 比較可能なグループ(ランダム化比較試験)

因果推論を行うにあたっては、結果に影響しそうな全ての特徴は合致していて、唯一の違いは原因が起こったか、起こらなかったか、つまり施策を実施したか、しなかったか、だけだとする「比較可能な」2つのグループを準備することが肝になる。

比較可能なグループ(A)と(B)を作る方法の一つとして、ランダム化比較試験(RCT: Randomized Controlled Trial)がある。この方法は(図3)に示すように母集団から無作為(ランダム)に要素を選択して、比較可能なグループ(A)と(B)を構成する方法である。**要素を無作為に選ぶことによって、似通った、比較可能なグループを作り出すことができる。**なお、施策の実施にあたっては、準備した2つのグループの平均値、分散値などの基本統計量が近い値であることを事前に確認しておく。つまり似たグループを準備できたことを確認しておく。

なお、この方法を適用するにあたっては、ランダム選択ができるようにすること、母集団を十分大きくすることなどがポイントになる。

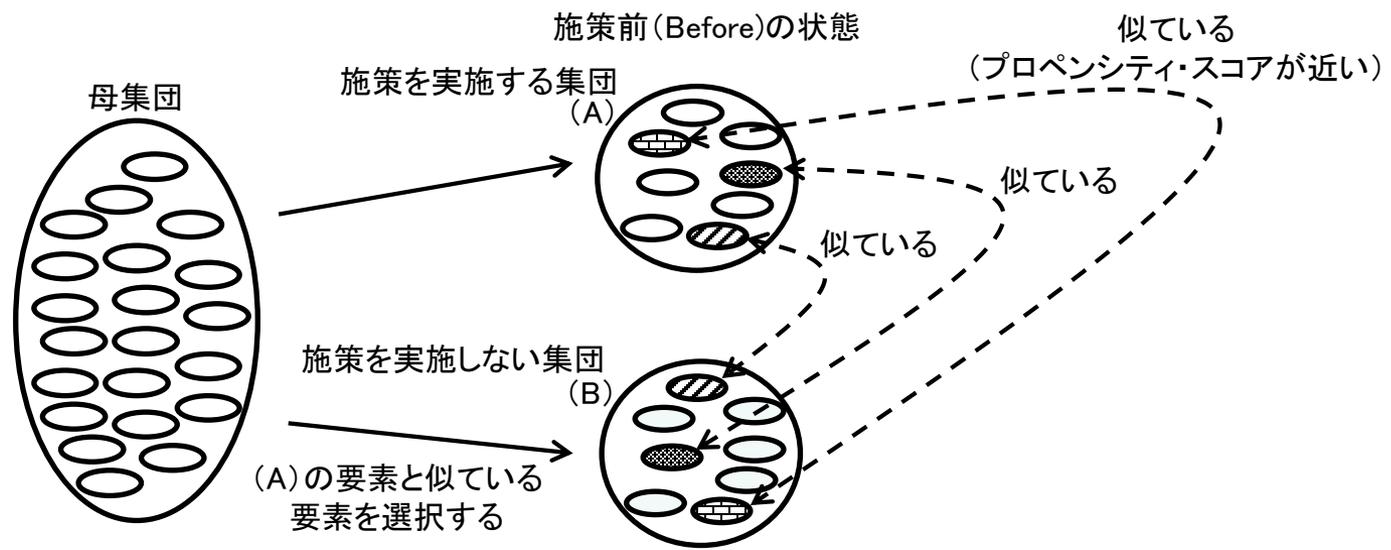


(図3) ランダム化比較試験

4-3. 比較可能なグループ(マッチング法)

比較可能なグループ(A)と(B)を作る別の方法として、マッチング法がある。
施策を実施するグループの個々の要素と良く似た要素を母集団から選び出し(マッチングし)、施策を実施しないグループを構成する。
これによって、比較可能な2つのグループを作る。

似た要素を選び出す(マッチングする)ときに、プロペンシティ・スコア(傾向スコア)を使用する。プロペンシティ・スコアとは結果に影響しそうな全ての特徴をまとめて1つの得点にしたものであり、プロペンシティ・スコアが近い値の要素同士を似た要素とみなす。

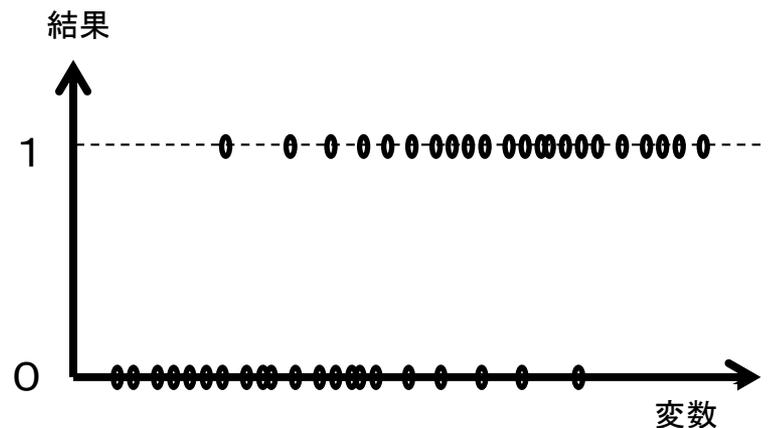


(図4) マッチング法

4-4. プロペンシティ・スコア(傾向得点)

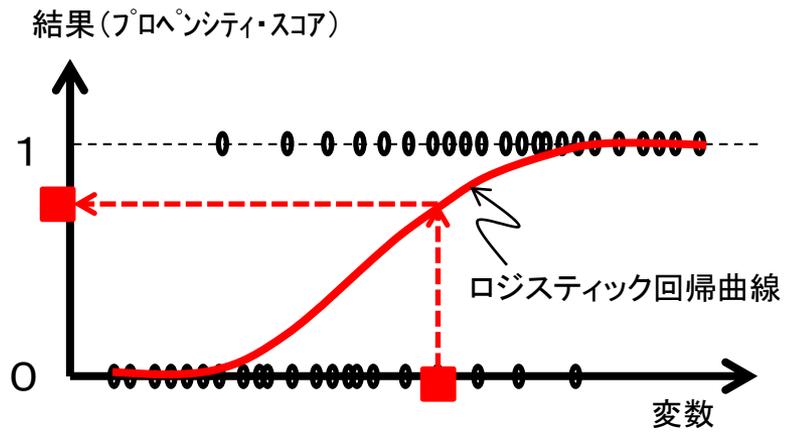
プロペンシティ・スコアは、似ている程度を表す点数(0~1)である(0=「全く違う」、1=「その通りである」)。
 「全く違う」と「その通りである」をロジスティック回帰することで、その中間の状態、つまりどの程度似ているかを点数化することができる。

例	変数	結果
契約	訪問回数+面談時間+...	成約できた(1) 成約できていない(0)
投薬	病歴+生活習慣+...	投薬した(1) 投薬していない(0)
広告	店主年齢+通行人+...	広告出した(1) 広告を出していない(0)
改装	建築年齢+客数+...	改装した(1) 改装していない(0)



ロジスティック回帰

例	変数	結果(プロペンシティ・スコア)
契約	訪問回数+面談時間+...	契約できる可能性
投薬	病歴+生活習慣+...	投薬が必要な可能性 (病気の重さの程度)
広告	店主年齢+地域人口+...	広告を出す可能性 (効果がありそうな程度)
改装	建築年齢+客数+...	改装する可能性 (老朽化の程度)



(図5) プロペンシティ・スコア

参考にした書籍

- 河本薫：「会社を変える分析の力」、講談社現代新書
小山昇：「数字は人格」、ダイヤモンド社
永野裕之：「ビジネス×数学＝最強」、すばる舎
竹内薫：「数学×思考＝ざっくりといかにして問題をとくか」、丸善出版
中西達夫：「統計データをすぐに分析できる本」、アニモ出版
中西達夫：「すぐれた判断が統計データ分析から生まれる」、実務教育出版
豊田裕貴：「マンガでわかる ビジネスを成功に導くデータ分析」、ナツメ社
向後千春、富永敦子：「統計学がわかる」、技術評論社
石井俊全：「意味がわかる統計学」、ベレ出版
涌井良幸、涌井貞美：「中学数学でわかる統計の授業」、日本実業出版
涌井良幸、涌井貞美：「統計学の図鑑」、技術評論社
西内啓：「統計学が最強の学問である」、ダイヤモンド社
西内啓：「統計学が最強の学問である(実践編)」、ダイヤモンド社
西内啓：「統計学が最強の学問である(ビジネス編)」、ダイヤモンド社
森岡毅、今西聖貴：「確率思考の戦略論」、角川書店
デビッド・マクアダムス：「世界一流企業はゲーム理論で決めている」、ダイヤモンド社
河村真一ほか：「本物のデータ分析力が身に付く本」、日経BPムック
末吉正成、末吉美貴：「Excel ビジネス統計分析 この分析できますか?」、翔泳社
谷岡一郎：「社会調査のウソ リサーチ・リテラシーのすすめ」、文藝春秋
林知己夫：「調査の科学」、ちくま学芸文庫
伊藤公一朗：「データ分析の力 因果関係に迫る思考法」、光文社新書
中室牧子、津川友介：「原因と結果の経済学」
八谷大岳：「データ解析」シリーズ 全15回、(例)[データ解析 第1回 ベクトルの復習 – YouTube](#)