

反実仮想を用いた因果判断モデルの妥当性検証：結果の価値が判断に及ぼす影響の分析

Examining the Validity of Counterfactual-Based Causal Models : How Outcome Value Shapes Judgment

鎌田 佑[†], 中村 紘子^{†‡}, 高橋 達二[†]

Yu Kamata, Hiroko Nakamura, Tatsuji Takahashi

[†]東京電機大学, [‡]日本学術振興会

Tokyo Denki University, Japan Society for the Promotion of Science

ky84710@gmail.com

概要

本研究は、人が主要な原因を判断する過程を説明する反実仮想的因果モデルの妥当性を検討した。日本語話者を対象とした追試実験により、人は確率の極端さよりも反実仮想下での結果との相関を重視することを確認した。また、熟慮性の高い参加者ほど CESM の予測と判断が一致する傾向が見られた。さらに、結果の感情価（勝利か敗北か）によってモデルの最適パラメータが変動し、因果判断プロセスの文脈依存性が示された。

キーワード：因果推論, 反実仮想, モデリング

1. 序論

1.1 研究背景と目的

人は多数の原因候補から特定のものを重要と見なす因果選択を行う。このプロセスを説明する Counterfactual Effect Size Model (CESM) は、反実仮想世界での原因と結果の相関の強さから原因の重要性を評価する (Quillien, 2020)。Quillien&Lucas (2023) は、結果との相関が高い原因が重要だと評価されることを示し、CESM の予測を支持した。

本研究ではこの追試に加え、先行研究にはない「敗北シナリオ」を導入し、結果の感情価が因果判断やモデルのパラメータに与える影響を明らかにすることを目的とした。ネガティブな結果の後には反実仮想思考が活発になるため (Mandel&Lehman, 1996)、モデルのパラメータは感情価によって変動すると予測された。

1.2 仮説

感情価による非対称性がモデルのパラメータにどのように反映されるかについて、以下の仮説を立てた。

- ネガティブな結果では事前確率がより慎重に考慮されるため、CESM の安定性パラメータ s は先行研究の値 ($s = 0.73$) より低くなると予測した。

2. 敗因を探る文脈では「必要性」が重視されるため、NSM のサンプリング傾向 $SP(C)$ は先行研究の値 ($SP(C) = 0.15$) より低くなると予測した。

実験では、当たり確率の異なる 3 箱 (高 95%, 中 50%, 低 5%) から 1 球ずつ引き、2 つ以上当たれば勝利するくじゲームを提示した。参加者は 4 つのシナリオ ((i)3 つ全てで当たりで勝利, (ii) 中・低確率が当たりで勝利, (iii)3 つ全てはずれで敗北, (iv) 中・低確率がはずれで敗北) を観察し、各箱の試行が結果の要因としてどの程度重要かを評価した。

2. 各モデルの説明

本研究では、人間の因果判断を説明する反実仮想効果量モデル (CESM) と必要十分モデル (NSM) の妥当性を検証した。

2.1 反実仮想効果量モデル : Counterfactual Effect Size Model (CESM)

反実仮想効果量モデル (CESM) は、反実仮想シミュレーションに基づき原因の重要性を評価するモデルである (Quillien, 2020)。CESM では、より生起確率の高い反実仮想世界をシミュレートし、その世界で結果と強く相関する事象を重要な原因だと判断する。例えば、複数の当たりが必要な場合、他の当たりが期待通りだという反実仮想下で、成否を分ける一つの当たりが重要だと見なされる。

CESM は、以下の式で定義されるサンプリング確率に基づき、反実仮想シミュレーションを用いて因果効果を定量化する。

$$\text{SamplingProbability} = s\delta(x) + (1 - s)Pr(x) \quad (1)$$

- $\delta(x)$: 実際の観測結果が x である場合に 1, そうでない場合は 0 となる指示関数.
- $Pr(x)$: x の事前確率.
- s : 安定性パラメータ (観測結果をどの程度重視するかを制御) .

s=0 の場合、このモデルは Unattached CESM と呼ばれ、事前確率のみを考慮した式となる。このモデルを本研究の実験シナリオに適用すると、以下のように予測される。

2.1.1 勝利シナリオの予測

シナリオ (i) (3つの当たりで勝利)：参加者は、高確率箱は当たり、低確率箱はずれるという、より起こりやすい反実仮想をシミュレートする。この世界では中確率箱の成否が勝敗を分けるため、これが最重要原因だと予測される。

シナリオ (ii) (中確率と低確率が当たりで勝利)：高確率箱が当たると期待される中、勝利には中・低確率箱両方の当たりが必要となる。より達成が困難な低確率箱の当たりが決定的な要因と見なされ、最重要原因だと予測される。

2.1.2 敗北シナリオの予測

敗北シナリオでは、「はずれを引くこと」が原因、「2つ以上のはずれで敗北」が結果となる。CESM の予測は「期待からの逸脱」を反映する。

シナリオ (iii) (3つ全てがはずれで敗北)：「成功するはずだった」箱からの失敗が、重要な敗因として認識される。そのため、はずれ確率が最も低い（当たり確率が高い）箱からの失敗が最重要で、次いで中確率箱、はずれが当然と見なされる低確率箱の順で評価されると予測される。

シナリオ (iv) (中確率と低確率の箱がはずれで敗北)：敗因候補である中・低確率箱を比較すると、「もし中確率箱が当たってれば勝利できた」という反実仮想の方が、「もし低確率箱が当たっていたら」という反実仮想より生じやすい。そのため、中確率箱の失敗がより重要な敗因だと判断される。

2.2 必要十分モデル：Necessity-Sufficiency Model (NSM)

必要十分モデル (NSM) は、原因の「必要性」と「十分性」に基づき因果強度を評価するモデルである (Icard et al., 2017)。因果強度は以下の式で表される。

$$\text{因果強度} = (1 - SP(C)) \cdot N(C \rightarrow E) + SP(C) \cdot S(C \rightarrow E) \quad (2)$$

- $N(C \rightarrow E)$ ：必要性（もし C がなかったら、 E は生じなかったか）。
- $S(C \rightarrow E)$ ：十分性（ C があれば、 E を引き起こすのに十分か）。
- $SP(C)$ ：要因 C のサンプリング傾向（0~1）。0 に近いほど必要性を、1 に近いほど十分性を重視する重み。

2.2.1 勝利シナリオの予測

このモデルでは、当たりを引く事象が単体で勝利を最も引き起こしやすい（＝十分性が高い）高確率の箱が、シナリオを問わず最重要だと予測される。

2.2.2 敗北シナリオの予測

敗北シナリオでは、「必要性」は「もしその箱が当たれば敗北を避けられたか」、「十分性」は「その箱がはずれるだけで敗北するか」と解釈される。

シナリオ (iii) (3つ全てがはずれで敗北)：どの箱か1つが当たっても敗北は避けられないが、「もし高確率の箱が当たっていたら」という反実仮想は最も勝利に近づく。そのため、この箱の失敗が必要性が高いと判断され、重要な敗因と予測される。

シナリオ (iv) (中確率と低確率の箱がはずれで敗北)：このモデルは極端な確率の事象を重視するため、中・低確率箱の比較では、より極端な低確率箱の失敗が必要性が高いと見なされ、重要な敗因だと予測される。

CESM が統計的相関を重視し中間確率要因も評価するのに対し、NSM は必要・十分という概念的関係を重視し極端な確率要因を評価する点が、両モデルの主な違いである。

3. 実験

3.1 実験1

本実験では、Quillien&Lucas(2023)で示された因果判断パターンが日本語話者でも再現されるかを検証し、因果判断における文化差や熟慮的思考傾向が与える影響を検証した。

3.1.1 参加者

クラウドソーシングサイトで募集した602名（男性：281人、女性：320人、その他：1人、平均年齢41.92歳）のうち、回答に不備のない547名のデータを分析に用いた。

3.1.2 実験材料と手続き

実験はオンラインアンケート調査ツール (Qualtrics) で実施した。デザインは2 (シナリオ：全箱当たり vs 部分当たり) × 3 (評価対象箱：高・中・低確率) の混合計画で、シナリオを被験者間、評価対象箱を被験者内要因とした。

参加者には、当たり確率の異なる3つの箱（高:95%、中:50%、低:5%）から1つずつボールを引き、2つ以上当たれば勝利というゲームルールが説明された。ルール学習と練習の後、参加者は因果判断課題を行った。

参加者は以下の2つの勝利シナリオのいずれかにランダムに割り当てられた。

- シナリオ (i) 全箱当たりで勝利
- シナリオ (ii) 部分当たり (中・低確率箱のみ) で勝利

各シナリオの結果を観察後、「(特定の) 箱から当たりを引いたので勝利した」という文章への同意度を9段階で評価した。

課題後、参加者は各箱の確率理解度と、認知的熟慮性を測る Cognitive Reflection Test (CRT-7) に回答した。

3.1.3 シミュレーション結果

CESM と NSM, さらに事前確率のみに基づく Unattached CESM ($s=0$ の条件) について, 先行研究で報告されたパラメータ (CESM: $s=0.73$, NSM: $SP(C)=0.15$) を用いて実験条件ごとのシミュレーション結果を算出した (図1)。

3.1.4 因果判断の行動データ

実験参加者の因果判断は, CESM の予測と一致するパターンを示した。

分散分析の結果, 参加者の因果判断は, 3点条件では3つの確率条件 (0.05, 0.5, 0.95) 間で, 2点条件では2つの確率条件 (0.05, 0.5) 間で有意な差が見られた (3点条件: $F(2, 554) = 6.659, p < .001$, 2点条件: $F(1, 145) = 36.94, p < .001$). 3点条件では中確率が低・高確率より因果判断の評定が高く, 2点条件では低確率が中確率より評定が高かった。

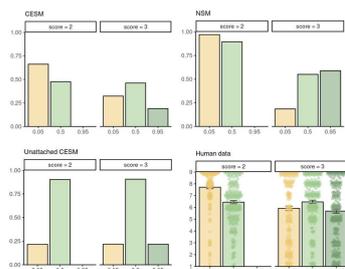


図1 各モデルのシミュレーション結果と人間の因果判断

また, CRT (認知的熟慮の検査) の成績の高群と低群で因果判断が異なるかを検討した (図2). 熟慮的傾向が高いグループ (High) は, CESM の予測と一致する判断パターンを示した. 一方, 低群 (Low) の参加者は, 特に3点条件において, 全ての箱からの当たりを同程度に重要と評価する傾向が見られた. 分散分析の結果, 低群では3点条件において, 各箱の重要度判断に有意な差が見られなかった ($F(2, 429) = 1.606, p > 0.1$). 一方, 2点条件については有意差が見られた ($F(1, 296) = 16.46, p < .001$).

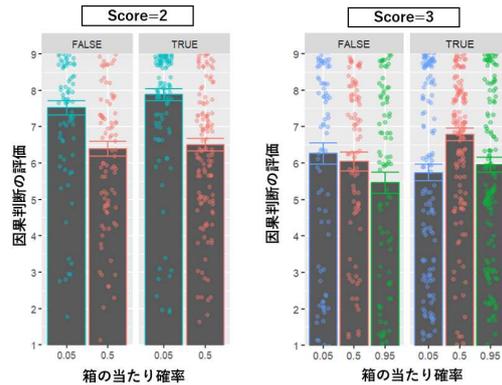


図2 CRTの各成績グループごとの因果判断

3.2 実験2

実験2では, 結果の感情価 (勝利 vs 敗北) が因果判断に与える影響を検証し, 各因果モデルのパラメータが変動するかを検討した. 実験1と同様の課題設定のもとで, 勝利シナリオに加えて敗北シナリオ (iii)3つ全てがはずれて敗北, (iv) 中確率と低確率の箱がはずれて敗北) を提示した。

3.2.1 参加者

参加者は, クラウドソーシングサイト (Crowd-Works) を用いて募集した141名であった (男性: 83人, 女性: 51人, 平均年齢40.9歳). このうち, 回答に不備のない129名のデータ分析を用いた. 実験は, オンラインアンケート調査ツール (Qualtrics) を用いて行い, 参加者はブラウザ上に提示されるシナリオを読み回答した。

3.2.2 実験材料と手続き

実験1と同様のくじ引きゲームを用いて, 勝利と敗北の両シナリオを検証した. 参加者は以下の手順で課題を行った:

1. 実験1と同様のルール学習, 理解度確認, 練習試行を実施
2. 参加者をランダムに以下の2群に分割した:
 - G1群: 全箱当たりで勝利 (シナリオ i) + 全箱はずれて敗北 (0点条件: シナリオ iii)
 - G2群: 部分当たりで勝利 (シナリオ ii) + 部分はずれて敗北 (1点条件: シナリオ iv)

各シナリオ観察後, 各箱の重要度を9段階で評価し, 勝利シナリオでは「勝因度合い」, 敗北シナリオでは「敗因度合い」の回答を求めた。

3.2.3 シミュレーション結果

CESM と NSM, さらに事前確率のみに基づく Unattached CESM ($s=0$ の条件) について, 先行研究で報告されたパラメータ (CESM: $s = 0.73$, NSM: $SP(C) = 0.15$) を用いて実験条件 (敗北シナリオ) ごとに

とのシミュレーション結果を算出した (図 3)。

3.2.4 因果判断の行動データ

分散分析の結果, 参加者の因果判断はモデルの予測とは異なる傾向が示された。

- 0点条件(確率: 0.05, 0.5, 0.95): 3つの確率条件間で有意な差が認められた ($F(2, 554) = 6.659, p < .001$). 特に, 中確率の箱が, 低確率および高確率の箱よりも因果判断の評定が高い傾向にあった。
- 1点条件(確率: 0.05, 0.5): 2つの確率条件間で有意な差が認められた ($F(1, 145) = 36.94, p < .001$). 低確率の箱が, 中確率の箱よりも因果判断の評定が高い傾向にあった。



図 3 敗北シナリオにおける各モデルのシミュレーション結果と人間の因果判断

4. 考察

本研究は, 反実仮想思考モデルの妥当性と, 結果の感情価や認知的熟慮性が因果判断に与える影響を検証した。

実験 1 では Quillien&Lucas (2023) を追試し, 日本語話者でも CESM の予測が支持されることを確認した。3点勝利条件では中確率箱が, 2点勝利条件では低確率箱が最重要と判断され, 人が確率の極端さより反実仮想下での結果との相関を重視することを示唆した。また, 認知的熟慮性 (CRT) が高い参加者ほど CESM の予測と一致する傾向も見られた。

実験 2 の敗北シナリオでは, 人間の判断は「期待からの逸脱」を重視する傾向を示した。例えば 3つのはずれで敗北した際, 参加者は「失敗するはずではなかった」高・中確率箱を重要な敗因と見なした。この論理は勝利時と対称的だが, 判断の「強度」は敗因評価の方が勝因より極端であり, 非対称なネガティブ・バイアスの影響が示唆された。

モデルと人間の判断との適合度を定量的に評価するため, 各モデルの予測値と参加者の平均判断値との間の最小二乗和 (SSE) を算出し, それを最小化する最適パラメータを探索した (表 1)。

表 1 各シナリオにおけるモデルの最適パラメータと最小二乗和 (SSE)

モデル	パラメータ	勝利時	敗北時
CESM	最適 s 値	0.56	0.58
	SSE	0.37	2.15
NSM	最適 $SP(C)$ 値	0.40	0.09
	SSE	0.02	6.21

表 1 の通り, 勝利シナリオでは NSM (SSE=0.02) が CESM より良好に判断を説明した。一方, 敗北シナリオでは NSM の最適 $SP(C)$ 値が 0.09 と, 勝利時の 0.40 から大きく低下し, 敗因の探索では「必要性」が強く重視されるという仮説を支持した。こうした文脈に応じたパラメータの変動は, 人間の因果判断が固定的でなく, 柔軟なプロセスであることを示している。

結論として, 人間の因果判断は「期待からの逸脱」という対称的論理と, 感情価で強度が変わる非対称な心理的側面を併せ持つ。既存モデルが人間の判断を捉えるには, 結果の感情価を考慮したパラメータ調整が今後の重要な課題となる。

文献

- Quillien, T. (2020). When do we think that X caused Y?. *Cognition*, 205, 104410.
- Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength. *Cognition*, 161, 80-93.
- Quillien, T., & Lucas, C. G. (2023). Counterfactuals and the logic of causal selection. *Psychological Review*.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4), 25-42.
- Quillien, T., & Barlev, M. (2022). Causal judgment in the wild: Evidence from the 2020 us presidential election. *Cognitive Science*, 46(2), e13101.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4), 323-370.
- Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of personality and social psychology*, 71(3), 450.