

行動経済学入門

— 人間の「不合理」を科学する —

人はなぜ、わかっているにもかかわらず損な選択をしてしまうのか。心理学と経済学が会って生まれた「行動経済学」の主要な理論を、実験の実例とともに概観する。本資料は、AI に読み込ませて「質問しながら理解を深める」練習のために、あえて少し難しめに書かれている。

この資料の使い方 HOW TO USE

この資料は、パソコン教室の「AI 活用講座」のために作られた練習用の教材です。テーマは「行動経済学」——人間の判断のクセを科学する、経済学と心理学の交差点にある学問です。

本文は、大学の教科書に近い、少しかための文章で書かれています。すらすら読めなくても、まったく問題ありません。むしろ「一度読んだだけではよくわからない」と感じるくらいが、この教材の狙いどおりです。わからない部分こそが、AI への質問の種になるからです。

STEP 1

PDF を読み込ませる

NotebookLM で新しいノートブックを作り、「ソースを追加」からこの PDF をアップロードします。

STEP 2

まず概要をつかむ

「この資料を 3 分で読める長さに要約して」「各章の要点を一覧にして」などと頼んで、全体像をつかみます。

STEP 3

質問して深掘りする

読んでいて引っかかった言葉や、腑に落ちない説明を、そのまま質問にぶつけます。各章末の「AI に質問してみよう」も使ってください。

質問にコツは要りません。「小学生にもわかるように説明して」「たとえ話で説明して」「この 2 つの違いは何?」——どんな聞き方でも、AI は資料の中身にもとづいて答えてくれます。巻末の「AI への質問アイデア集」に、質問の型を 10 種類まとめてありますので、迷ったらそこから選んでみてください。

大切なのは、正しい質問をすることではなく、たくさん質問してみることです。答えを読んでまた疑問がわいたら、続けて聞き返してみましょう。それが AI と学ばいちばんの近道です。

本資料は AI 活用学習の練習用サンプルとして作成されたものです。内容は行動経済学の標準的な教科書・文献にもとづいていますが、学術論文としての引用には適しません。出典の詳細は巻末の参考文献をご覧ください。

目次 CONTENTS

はじめに

この資料の使い方	2
----------------	---

本編

第1章 合理性の仮定とその揺らぎ	4
------------------------	---

1.1 ホモ・エコノミクスという人間像 / 1.2 期待効用理論の成立と公理化 / 1.3 アノマリー——理論と現実のずれ / 1.4 行動経済学の成立と位置づけ

第2章 限定合理性とヒューリスティック	6
---------------------------	---

2.1 限定合理性と満足化原理 / 2.2 二重過程理論——システム1とシステム2 / 2.3 代表性ヒューリスティックと利用可能性ヒューリスティック / 2.4 アンカリングと調整

第3章 プロスペクト理論	8
--------------------	---

3.1 確実性効果とアレのパラドックス / 3.2 参照点依存性と価値関数 / 3.3 確率加重関数とリスク態度の四分割 / 3.4 1979年論文の歴史的意義

第4章 保有効果と現状維持バイアス	10
-------------------------	----

4.1 保有効果とマグカップ実験 / 4.2 現状維持バイアスとデフォルトの力 / 4.3 サンクコストの誤謬 / 4.4 メンタル・アカウンティング

第5章 時間割引と自制心	12
--------------------	----

5.1 時間選好と割引効用モデル / 5.2 双曲割引と選好の逆転 / 5.3 β - δ モデルという定式化 / 5.4 コミットメント・デバイス

第6章 社会的選好と互恵性	14
---------------------	----

6.1 利己性の仮定と最後通牒ゲーム / 6.2 独裁者ゲームと不平等回避モデル / 6.3 正の互恵性と負の互恵性 / 6.4 社会規範と市場規範の衝突

第7章 ナッジと選択アーキテクチャ	16
-------------------------	----

7.1 リバタリアン・パターナリズムの思想 / 7.2 デフォルト効果——初期設定の力 / 7.3 英国の実践とEASTフレームワーク / 7.4 日本の取り組みとナッジの倫理

第8章 批判・限界・これから	18
----------------------	----

8.1 再現性の危機とプライミング研究への疑義 / 8.2 ナッジの効果量論争と外的妥当性 / 8.3 スラッジと操作の倫理 / 8.4 AI時代のナッジと学ぶことの意義

巻末資料

用語集	20
-----------	----

AI への質問アイデア集	22
--------------------	----

参考文献・もっと学びたい人へ	23
----------------------	----

合理性の仮定とその揺らぎ

経済学は長らく、人間を自己の利益を最大化する合理的な存在とみなしてきた。この仮定は理論に数学的な厳密さを与えた一方で、現実の人間の選択としばしば食い違う。本章では、合理的経済人の仮定と期待効用理論の成立過程を概観し、アレのパラドックスに代表されるアノマリーを手がかりとして、心理学と経済学の融合から生まれた行動経済学の成り立ちと位置づけを明らかにする。

1.1 ホモ・エコノミクスという人間像

伝統的経済学の中核には、ホモ・エコノミクス(合理的経済人)と呼ばれる人間像が据えられている。これは、利用可能な情報をすべて考慮し、一貫した選好に基づいて自己の効用、すなわち満足の度合いを最大化するように行動する存在である。この人間像は三つの強い仮定に支えられている。第一に完全な計算能力、第二に鉄の意志ともいえるべき自制心、第三に自己利益の追求である。感情に流されず、誘惑に負けず、あらゆる選択肢を瞬時に比較できる人間など現実には存在しないが、理論を構築するための理念型としては強力であった。

この仮定が長く支持されたのには理由がある。人間行動を効用最大化問題として定式化すれば、需要と供給、市場均衡といった現象を数学的に導出でき、経済学は自然科学に比肩する厳密さを獲得できた。ミルトン・フリードマンは1953年の論文で、仮定が非現実的であっても予測が当たればモデルは有用だと論じ、この立場を方法論的に擁護した。他方、ハーバート・サイモンは人間の認知能力には限界があるとして限定合理性の概念を提唱し、1978年にノーベル経済学賞を受賞している。合理性の仮定は、早くからその内部に批判の種を抱えていたのである。

1.2 期待効用理論の成立と公理化

不確実な状況での選択を扱う理論の出発点は、18世紀に遡る。数学者ダニエル・ベルヌーイは1738年、サンクトペテルブルクのパラドックスと呼ばれる問題を検討した。表が出るまでコインを投げ続け、表が遅く出るほど賞金が倍々が増えていく賭けは、期待値を計算すると無限大になるにもかかわらず、人々はごくわずかな参加料しか支払おうとしない。ベルヌーイは、人が評価するのは金額そのものではなく、そこから得られる効用であり、富が増えるほど追加の1単位がもたらす満足は小さくなる(限界効用の逓減)と考えることで、この謎に説明を与えた。

この着想を厳密な理論へと仕上げたのが、数学者フォン・ノイマンと経済学者モルゲンシュテルンである。両者は1944年の著書『ゲームの理論と経済行動』において、完備性、推移性、連続性、独立性という少数の公理を満たす選好をもつ人は、あたかも期待効用を最大化するかのように行動することを証明した。これが期待効用理論の公理化であり、以後、リスクを伴う意思決定を分析する標準的な枠組みとして経済学に定着した。合理性は、検証可能な公理の束として定義されるに至ったのである。

1.3 アノマリー——理論と現実のずれ

しかしその公理は、まもなく現実の選択によって揺さぶられることになる。フランスの経済学者モーリス・アレは1953年、後にアレのパラドックスと呼ばれる問題を提示した。確実に大金が得られる選択肢と、わずかな確率で何も得られない代わりにより大きな利得が見込めるくじを比べると、多くの人は確実な方を選ぶ。ところが両方の選択肢から同じ確率の当たりを取り除いて比較させると、今度は選好が逆転してしまう。これは期待効用理論の独立性公理に明確に反する行動であり、確実性が人間の判断において特別な重みをもつことを示していた。アレ自身は1988年にノーベル経済学賞を受賞している。

このように、標準的な理論の予測から系統的に外れる現象はアノマリー(変則事象)と呼ばれる。1961年にダニエル・エルトズバーグが示したパラドックスは、人々が確率のわからない曖昧な状況を嫌う傾向(曖昧性回避)を明らかにし、やはり期待効用理論では説明できなかった。重要なのは、これらの逸脱が気まぐれで無秩序な誤りではなく、一定の方向性をもって繰り返し観察される点である。誤りが系統的であるなら、それは科学的に記述し予測することができる。この認識こそが、行動経済学への扉を開いた。

1.4 行動経済学の成立と位置づけ

その扉を大きく開いたのは、イスラエル出身の心理学者ダニエル・カーネマンとエイモス・トヴェルスキーである。両者は1974年の論文で、人間が確率判断に用いる簡便な思考法(ヒューリスティック)と、それが生む系統的な偏り(バイアス)を整理した。さらに1979年には学術誌エコノメトリカにプロスペクト理論を発表し、人は利得と損失を参照点からの変化として評価し、同じ額であれば損失を利得よりも重く感じる(損失回避)ことを、多数の実験に基づいて定式化した。心理学者でありながらカーネマンは2002年にノーベル経済学賞を受賞している。

経済学の内部からこの流れを推し進めたのが、リチャード・セイラーである。セイラーは、心の中でお金を用途別に区分して扱うメンタル・アカウンティング(心の会計)や、自分が保有する物を高く評価する保有効果を分析し、2008年には法学者キャス・サンスティーンとの共著で、選択の自由を残したまま望ましい行動を後押しするナッジの概念を世に広めた。2017年のセイラーのノーベル経済学賞受賞は、心理学と経済学の融合が一つの分野として確立したことを象徴する出来事であった。行動経済学は伝統的経済学を否定するのではなく、その人間像を現実の姿に近づけ、理論を拡張しようとする試みなのである。

表1 伝統的経済学と行動経済学の人間観の比較

観点	伝統的経済学	行動経済学
人間像	合理的経済人(ホモ・エコノミクス)	認知能力に限界のある生身の人間
選好と判断	一貫した選好に基づき期待効用を最大化する	参照点や提示のされ方に依存し、ヒューリスティックによる系統的バイアスを伴う
自制心	無限の意志力で長期的利益を追求する	先延ばしや誘惑への屈服など意志力の限界を認める
誤りの扱い	誤りは偶然の産物であり市場で淘汰される	誤りは系統的であり科学的な記述と予測の対象となる

この章を AI に質問してみよう

- ▶ 「ホモ・エコノミクス」と実際の人間は、どこがどう違うのですか？
- ▶ アレのパラドックスを、身近な買い物の例に置き換えて説明してください。

限定合理性とヒューリスティック

第1章では、伝統的経済学が前提としてきた合理的経済人の姿を確認した。本章では、その前提を根本から問い直した2つの研究の流れ、すなわちハーバート・サイモンの限定合理性の理論と、エイモス・トヴェルスキーおよびダニエル・カーネマンによるヒューリスティック研究を取り上げる。人間の判断は無限の計算能力に支えられているのではなく、簡便な思考の近道に頼っており、そこから系統的な偏り、すなわちバイアスが生じる過程を見ていこう。

2.1 限定合理性と満足化原理

アメリカの経済学者であり認知科学者でもあったハーバート・サイモンは、1940年代から50年代にかけて、人間の認知能力の限界を経済学の理論そのものに組み込むべきだと主張した。彼が提唱した限定合理性とは、注意力、記憶力、計算能力のいずれもが有限である人間は、すべての選択肢を列挙して帰結を厳密に比較するような最適化を実行できない、という洞察である。この構想は主著『経営行動』（1947年）以来一貫して追求されたものであり、一連の業績によってサイモンは1978年にノーベル経済学賞を受賞した。

では、限定合理性の下で人間は実際にどう選んでいるのか。サイモンの答えが満足化原理である。満足化とは、あらかじめ一定の要求水準を定めておき、それを満たす選択肢が見つかった時点で探索を打ち切る意思決定方式を指す。たとえば住居を探すとき、世の中の全物件を比較し尽くす人はいない。家賃と立地の条件を満たす部屋が現れれば、そこで契約するのが普通である。サイモンは、こうした行動こそ有限の認知資源しか持たない生物にとって適応的な戦略なのだと論じ、合理性の基準そのものを問い直した。

2.2 二重過程理論——システム1とシステム2

現代の認知心理学では、人間の思考を2つの様式に分けて捉える二重過程理論が広く受け入れられている。心理学者のキース・スタノヴィッチとリチャード・ウェストが2000年の論文で導入し、ダニエル・カーネマンが著書『ファスト&スロー』で一般に広めた呼び方に従えば、システム1は速く、自動的で、努力を要しない直観的思考であり、システム2は遅く、意識的で、集中を必要とする熟慮的思考である。なお両システムは脳内の特定部位を指す実体ではなく、思考の様式を整理するための概念的な区分である。

日常の判断の大部分を担うのはシステム1である。母語の聞き取り、相手の表情の読み取り、 $2+2$ のような計算は、瞬時に、ほとんど無自覚のうちに処理される。一方、 17×24 の暗算や込み入った契約書の検討にはシステム2の関与が欠かせない。問題は、システム2が労力を節約しようとする傾向を持ち、システム1の出した答えを十分に検証しないまま採用しがちな点にある。次に見るヒューリスティックとバイアスの多くは、この構造から生じる。

2.3 代表性ヒューリスティックと利用可能性ヒューリスティック

トヴェルスキーとカーネマンは1974年に科学誌サイエンスに発表した論文で、不確実な状況での判断が少数のヒューリスティック、すなわち経験的な簡便法に支えられていることを示した。第1の代表性ヒューリスティックは、ある対象がカテゴリーの典型例にどれだけ似ているかという印象によって、そのカテゴリーに属する確率を判断する方法である。ヒューリスティックは多くの場面で素早く妥当な答えを与える適応的な仕組みだが、特定の条件下では系統的な誤り、すなわちバイアスを生むのである。

有名なリンダ問題は、その落とし穴を鮮やかに示す。31歳で独身、率直な性格で、学生時代は哲学を専攻し、差別や社会正義の問題に強い関心を持っていたという人物描写を読んだ実験参加者の多くは、「リンダは銀行員である」よりも「リンダはフェミニスト運動に熱心な銀行員である」ほうが確からしいと答えた。だが2つの条件を同時に満たす確率が単独の条件の確率を上回することは論理的にあり得ない。この誤りは連言錯誤と呼ばれる。さらに代表性への依存は、母集団に占める事前の比率である基準率を無視させ、確率判断を大きく歪める。

第2の利用可能性ヒューリスティックは、具体的な事例をどれだけ容易に思い出せるかを、頻度や確率の手がかりとして用いる方法である。想起しやすい出来事は、実際よりも起こりやすいと感ぜられる。飛行機事故や凶悪犯罪は報道で繰り返し取り上げられるため記憶に残りやすく、統計上はるかに多くの死者を出す自動車事故や生活習慣病よりもリスクが過

大に見積もられやすい。この仕組みは、メディア報道と人々のリスク認知のずれを説明する鍵として、保険や防災の分野でも重視されている。

2.4 アンカリングと調整

第3のアンカリングと調整は、数値を見積もる際に最初に与えられた値を係留点、すなわちアンカーとして出発し、そこから調整を加えて答えに至る方法である。厄介なのは、この調整がたいてい不十分に終わることだ。トヴェルスキーとカーネマンの古典的実験では、参加者の目の前で数字盤、いわばルーレットを回し、10か65のどちらかで止まるよう細工したうえで、国連加盟国に占めるアフリカ諸国の割合はその数字より大きい小さいかを尋ね、続いて割合そのものを推定させた。

結果は明瞭で、10を見た群の推定値の中央値は25%、65を見た群では45%となり、判断とは何の関係もないはずのルーレットの数字に推定が引きずられた。しかも、推定の正確さに応じて謝礼を払うと告げても、この効果は消えなかったと報告されている。アンカリングは、値引き交渉における最初の提示額や希望小売価格の表示といった形で、商業の現場でも広く作用している。直前に目にした数値が自分の判断を方向づけていないかと立ち止まって疑うことが、賢い消費者への第一歩である。

表2 代表的なヒューリスティック3種と典型的なバイアスの対応

ヒューリスティック	判断の仕組み	典型的なバイアス・現象	身近な例
代表性	典型例との類似度で確率を判断する	連言錯誤、基準率の無視	リンダ問題、血液型による性格判断
利用可能性	事例の思い出しやすさで頻度を判断する	リスクの過大・過小評価	飛行機事故の危険を過大に見積もる
アンカリングと調整	初期値を起点に不十分な調整で推定する	無関係な初期値への係留	値引き交渉における最初の提示額

💡 この章を AI に質問してみよう

- ▶ システム1とシステム2の違いを、表に整理してください。
- ▶ リンダ問題で多くの人が間違えてしまうのは、なぜですか？

プロスペクト理論

宝くじを買う人が、同時に保険にも加入するのはなぜか。期待効用理論ではうまく説明できないこの種の行動を統一的に説明する枠組みが、プロスペクト理論である。ダニエル・カーネマンとエイモス・トヴェルスキーが1979年に発表したこの理論は、人間が利得と損失をどのように評価し、確率をどのように重み付けるかを実証データに基づいて記述するものであり、行動経済学の理論的支柱となった。本章ではその中核をなす価値関数と確率加重関数を中心に解説する。

3.1 確実性効果とアレのパラドックス

伝統的経済学の意味決定理論である期待効用理論は、人が各結果の効用にその確率を掛けて合計した期待効用が最大になる選択肢を選ぶと想定する。しかし1953年、フランスの経済学者モーリス・アレは、この理論の公理に体系的に反する選択パターンが存在することを示した。アレのパラドックスと呼ばれるこの現象では、確実に得られる選択肢が絡むと人々の選好が期待効用理論の予測から一貫した方向にずれ、理論の根幹である独立性公理が破られるのである。

カーネマンとトヴェルスキーは1979年の論文で、次のような実験によりこの逸脱を再確認した。80%の確率で4000を得る賭けと確実に3000を得る選択肢を提示すると、回答者の8割が確実な3000を選んだ。ところが両方の確率を4分の1にした、20%で4000を得る賭けと25%で3000を得る賭けの比較では、多数派が前者を選んだ。期待効用理論によれば2つの問題の選好は一致するはずであり、この逆転は確率が確実から不確実へ変わる瞬間に大きな心理的断絶があることを示している。

彼らはこの現象を確実性効果と名付けた。確実な結果は確率的な結果に比べて過大に重み付けられ、当選確率が100%から95%へ下がる変化は、55%から50%へ下がる変化と確率の差としては等しいにもかかわらず、はるかに大きな心理的影響を与える。アレのパラドックスは理論の欠陥を示す孤立した反例ではなく、人間の確率認知に内在する規則的な歪みの表れとして再解釈されたのである。

3.2 参照点依存性と価値関数

プロスペクト理論の最大の革新は、人が評価するのは最終的な資産総額ではなく、参照点からの変化、すなわち利得と損失であるとした点にある。参照点とは評価の基準となる状態であり、多くの場合は現状がそれにあたる。同じ資産500万円でも、1000万円から減った人と100万円から増えた人とでは満足感がまったく異なる。これが参照点依存性であり、感覚や知覚が絶対量ではなく変化に反応するという心理物理学の知見を、意思決定の理論に持ち込んだものといえる。参照点は状況の記述の仕方によって移動しうするため、同じ問題でも表現次第で選好が変わるフレーミング効果の基盤ともなっている。

この評価の仕方を数理的に表現したものが価値関数である(図1)。価値関数には3つの特徴がある。第1に、原点である参照点を境に利得領域と損失領域が区別される参照点依存である。第2に損失回避、すなわち同じ大きさの利得よりも損失のほうが約2倍強く感じられるという非対称性であり、関数は損失側で急勾配になる。1万円を失う苦痛は1万円を得る喜びのおよそ2倍に相当するため、人は五分五分の賭けを避けようとする。第3に感応度逓減であり、参照点から離れるほど変化への感覚は鈍る。利得側では関数が凹型、損失側では凸型となり、これが特徴的なS字型の曲線を生むのである。

3.3 確率加重関数とリスク態度の四分割

プロスペクト理論のもう1つの柱が確率加重関数である。人は客観的な確率をそのまま用いるのではなく、心理的に歪んだ決定の重みに変換して意思決定に用いる。具体的には、小さい確率は過大に評価され、中程度から大きい確率は過小に評価される。飛行機事故のようなごく稀な事象が実際以上に恐れられるのは前者の表れであり、95%という高い確率が確実な100%より大きく割り引かれて感じられる確実性効果も、この関数の形状から統一的に説明できる。

価値関数と確率加重関数を組み合わせると、リスクに対する態度は確率の大小と利得・損失の別によって4つに分かれる。中程度以上の確率の利得場面では人はリスク回避的になり、確実な利得を好む。同じ確率の損失場面では逆にリスク追求的になり、確実な損失を避けて一か八かの賭けに出る。一方、小さい確率では関係が反転し、利得場面ではわずかな当選確率に賭ける宝くじ購入のようなリスク追求が、損失場面では稀な災害に備える保険加入のようなリスク回避が生じる。この四分割パターンにより、同一人物が宝くじと保険を同時に買うという一見矛盾した行動が統一的に説明される。

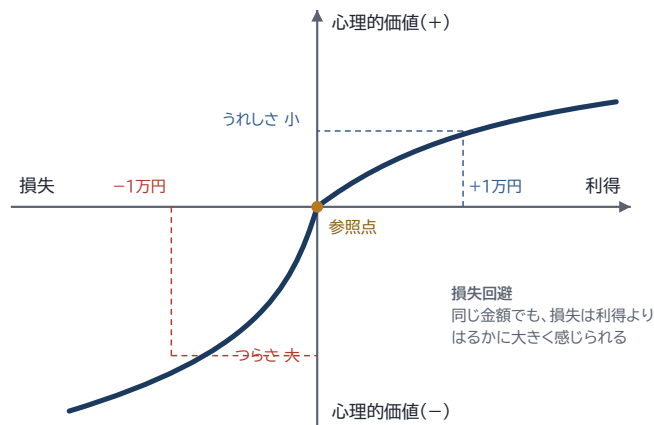


図1 プロスペクト理論の価値関数。曲線は参照点を境に形を変え、損失側の傾きは利得側より急である。

3.4 1979年論文の歴史的意義

プロスペクト理論は1979年、経済学で最も権威ある学術誌の1つとされるエコノメトリカに掲載された論文「プロスペクト理論: リスク下の意思決定の分析」によって世に出た。心理学者2人の手による論文が経済学の中核を担う専門誌に載ったこと自体が異例であり、この論文は後に同誌で最も引用される論文の1つとなった。理論の名称は、くじや賭けなどの選択対象を意味するプロスペクトに由来する。期待効用理論が人はこう選ぶべきだという規範的理論であるのに対し、プロスペクト理論は人が実際にどのように選ぶかを描く記述的理論である点も重要である。

この論文の真の意義は、期待効用理論への反例を列挙した点ではなく、アノマリー、すなわち標準理論からの系統的な逸脱を説明する代替モデルを、経済学者が扱える数理的な形で提示した点にある。心理学の知見が経済学の言語で定式化されたことにより両分野の対話が可能となり、行動経済学という研究分野の成立を決定づけた。理論は1992年に累積プロスペクト理論へと拡張され、今日では金融、保険、マーケティングなど幅広い応用領域を持つに至っている。

◆ コラム | 心理学者に贈られたノーベル経済学賞

2002年、ダニエル・カーネマンは心理学の知見を経済学に統合した功績によりノーベル経済学賞を受賞した。心理学者としての受賞は異例であり、プロスペクト理論を生んだ研究がその中心的業績とされた。しかし授賞式の壇上に、四半世紀にわたる共同研究の相棒エイモス・トヴェルスキーの姿はなかった。トヴェルスキーは1996年に59歳で死去しており、ノーベル賞は存命者にのみ贈られるという慣例のため受賞がかなわなかったのである。カーネマンは受賞に際し、この賞は2人の共同研究に対するものだとし繰り返し語った。2人の協働は、行動経済学史上最も実り多い知的パートナーシップとして記憶されている。

💡 この章を AI に質問してみよう

- ▶ 価値関数の3つの特徴を、それぞれ日常の例をつけて説明してください。
- ▶ 宝くじを買う人が保険にも入るのは、なぜ矛盾しないのですか？

保有効果と現状維持バイアス

自分が持っているものを手放すとなると、急に惜しくなる。長年使った携帯電話の料金プランを、得だと分かっているにもかかわらず変更できない。こうした日常の心理は、伝統的経済学が想定する合理的な価値評価では説明がつかない。本章では、保有効果と現状維持バイアスを軸に、サンクコストの誤謬とメンタル・アカウンティングを加えて検討し、前章で学んだ参照点依存と損失回避が、現実の意思決定をいかに深く規定しているかを明らかにする。

4.1 保有効果とマグカップ実験

保有効果(授かり効果)とは、自分が保有している財を、同じ財をまだ保有していない場合よりも高く評価する傾向を指す。リチャード・セイラーが1980年の論文で命名した概念であり、経済学的には、財を手に入れるために支払ってもよい最大額(WTP、支払意思額)と、保有する財を手放す代償として要求する最小額(WTA、受取意思額)との間に生じる系統的な乖離として定式化される。標準的な経済理論では、取引費用や所得効果が小さい限り、両者はほぼ一致するはずである。

この乖離を鮮やかに示したのが、カーネマン、クネッチ、セイラーが1990年に発表したマグカップ実験である。コーネル大学の学生を売り手と買い手に無作為に分け、売り手にだけ大学のロゴ入りマグカップを配って取引させたところ、売り手が要求する価格は買い手が支払ってもよいとする価格のおよそ2倍に達し、理論が予測する水準よりはるかに少ない取引しか成立しなかった。カップは無作為に配られたのだから、好みの差では説明できない。保有という事実そのものが評価を引き上げたのである。

保有効果の背後にあるのは、前章で扱ったプロスペクト理論の損失回避である。いったん保有した財は参照点に組み込まれ、それを手放すことは損失として、同じ財を新たに得る利得よりも心理的に大きく感じられる。中古品を売る際に自分の品物へ市場価格より高い値を付けてしまう心理や、不動産の売り手が購入時の価格に固執して値下げをためらう現象は、この参照点の移動によって統一的に理解できる。

4.2 現状維持バイアスとデフォルトの力

現状維持バイアスとは、現在の状態を変更したほうが有利な場合でも、現状にとどまることを不釣り合いに好む傾向をいう。サミュエルソンとゼックハウザーは1988年の研究でこの現象を体系的に示した。遺産を相続したという仮想状況で投資先を選ばせる実験では、同じ選択肢でも、すでに保有している投資先として提示されたものが選ばれる確率が明確に高まり、選択肢の数が増えるほどこの傾向は強まった。

彼らはまた、医療保険プランの選択に関する実際のデータでも、既存の加入者が新規の加入者に比べて古いプランにとどまり続けることを確認している。現状維持バイアスの背後には、変更が失敗に終わった場合の後悔を避けたいという心理、多数の選択肢を比較検討する認知的負担、そして現状を参照点とする損失回避が複合的に働いていると考えられる。

この性質は、あらかじめ設定された選択肢、すなわちデフォルトが人々の行動を強力に方向づけることを意味する。ジョンソンとゴールドスタインが2003年に示したように、臓器提供の意思表示を提供するを初期設定とする国と、提供しないを初期設定とする国とは、同意率が劇的に異なる。デフォルト設計の詳細はナッジを扱う後の章で改めて論じるが、その理論的基盤が現状維持バイアスにある点を、ここで押さえておきたい。

4.3 サンクコストの誤謬

サンクコスト(埋没費用)とは、すでに支出され、いかなる決定によっても回収できない費用を指す。標準的経済学の規範によれば、意思決定は将来の費用と便益のみに基づくべきであり、埋没費用は無視されなければならない。しかし現実には、人はここまで注ぎ込んだのだからという理由で、不採算な事業や見込みのない行動を継続してしまう。これがサンクコストの誤謬である。

その代表例が、英仏両政府が共同開発した超音速旅客機コンコルドである。開発の途中で商業的採算が見込めないことが明らかになった後も、それまでの巨額の投資を理由に計画は続行された。この故事から、サンクコストの誤謬はコンコルド効果とも呼ばれる。アークスとブルーマーが1985年に行った実験でも、劇場のシーズンチケットを正規価格で購入した観客は、割引価格で購入した観客よりも足しげく劇場に通うことが示された。支払額の大きさが、本来は無関係なはずの将来の行動を左右したのである。

4.4 メンタル・アカウンティング

メンタル・アカウンティング(心の会計)とは、セイラーが提唱した概念で、人が企業の会計係のように、お金を出所や用途ごとに心の中の別々の勘定に振り分けて管理する傾向をいう。経済学の想定では貨幣は完全に代替可能であり、1万円はどこから来ようと同じ1万円のはずである。しかし実際には、給料で得た1万円と臨時収入の1万円は異なる扱いを受ける。人はお金に色をつけるのである。

カーネマンとトヴェルスキーが1984年に示した例では、映画館で10ドルのチケットを紛失した人はチケットを買い直すことをためらうのに対し、現金10ドルを紛失した人の多くは予定どおりチケットを購入すると答えた。失った金額はどちらも同じだが、前者では観劇という勘定に合計20ドルが計上されるため、割高に感じられるのである。ギャンブルで得た金銭を浪費しやすいハウスマネー効果も、同じ枠組みで説明できる。

もっともメンタル・アカウンティングは、非合理の源泉であると同時に、給与を生活費と貯蓄に袋分けするような自己コントロールの装置としても機能する。本章で見た4つの現象は、いずれも参照点と損失回避という共通の土台の上に立っており、セイラーが2017年にノーベル経済学賞を受賞した業績の中核でもある。次章では、これらの知見を社会に応用するナッジと選択設計を扱う。

💡 この章を AI に質問してみよう

- ▶ マグカップ実験は、何を明らかにした実験ですか？
- ▶ サンクコストにとらわれてしまう行動の例を、5つ挙げてください。

時間割引と自制心

締切前夜の一夜漬け、明日から始まるはずのダイエット、増えない貯蓄——これらに共通するのは、将来の利益より目先の満足を優先してしまう人間の性向である。本章では、この性向を「時間割引」という概念で科学的に分析する。まず伝統的経済学が仮定してきた指数割引を確認したうえで、実際の人間に観察される双曲割引との違いを検討し、現在バイアスが引き起こす選好の逆転を具体例で示す。さらに、その対策として人々が自ら編み出してきたコミットメント・デバイスの有効性を、実証研究に基づいて考察する。

5.1 時間選好と割引効用モデル

人は同じ大きさの利益であっても、遠い将来に受け取るものより今すぐ受け取るものを高く評価する。この傾向を時間選好と呼び、将来の価値が現在から見て目減りして評価される現象を時間割引という。たとえば1年後の1万円を現在の9000円と同等とみなす人は、およそ年10%の率で将来を割り引いている計算になる。この割引率には大きな個人差があり、割引率の高い人ほどせっかちで、低い人ほど辛抱強く将来の利益を待てる。割引率の高さが貯蓄額の少なさや喫煙・肥満などの健康行動と相関すること、数多くの調査で報告されている。

伝統的経済学は、この割引をポール・サミュエルソンが1937年に定式化した割引効用モデルによって表現してきた。このモデルでは、将来の効用は每期一定の割合で、すなわち指数関数的に割り引かれると仮定される。指数割引の最大の特徴は時間整合性である。時間整合性とは、どの時点から評価しても選択肢の優劣が入れ替わらない性質を指し、今日立てた計画を明日になって覆す理由が原理的に存在しない。合理的で一貫した経済人を描くうえで都合のよい仮定であり、長らく経済分析の標準とされてきた。

5.2 双曲割引と選好の逆転

ところが現実の人間の選択は、この仮定をしばしば裏切る。「今日1万円をもらう」か「明日1万100円をもらう」と問われれば、多くの人は待たずに済む今日の1万円を選ぶ。ところが「1年後に1万円」と「1年と1日後に1万100円」の比較になると、同じく1日待つだけで100円増える構造は変わらないにもかかわらず、今度は多くの人が後者を選ぶのである。2つの選択肢の時間差はどちらも1日で同一なのに、選好の順序が入れ替わってしまう。この選好の逆転は、指数割引のもとでは決して起こりえない現象である。

この現象が示すのは、割引率が一定ではなく、現在に近い時点ほど急激に高くなるという事実である。心理学者ジョージ・エインズリーはハトを用いた実験で、割引関数が指数関数よりも双曲線によく当てはまることを見だし、リチャード・セイラーらは人間の選択でも同様の傾向を確認した。これが双曲割引である。図2に示すように、双曲割引の曲線は直近の期間で価値が急落し、遠い将来ではなだらかに減少する。目先の報酬を過大評価するこの性質は現在バイアスと呼ばれ、先延ばしや衝動買い、依存症といった自制の失敗の根源にあると考えられている。

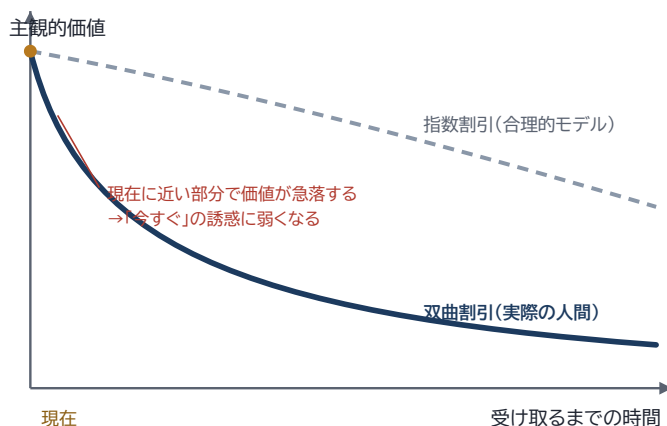


図2 指数割引と双曲割引。人間の価値の目減り方(実線)は、遠い将来ではゆるやかだが、現在の直前で急激になる。

5.3 β - δ モデルという定式化

双曲割引を扱いやすい形で経済モデルに組み込んだのが、ハーバード大学のデイヴィッド・レイブソンが1997年の論文で普及させた β - δ モデル、いわゆる準双曲割引である。このモデルでは、将来の効用は通常の割引因子 δ で每期割り引かれることに加え、現在以外のすべての期に一律の係数 β ($0 < \beta < 1$) が掛けられる。 β の存在によって現在とそれ以外の期のあいだに特別な断絶が生じ、現在バイアスがわずかに2つのパラメータで簡潔に表現される。 β が1に等しければ、モデルはサミュエルソンの指数割引にそのまま一致する。

β - δ モデルが描く人間は時間非整合的である。今日の自分は「来月から貯蓄を増やす」と計画するが、来月になればそのときの自分にとっての「現在」が再び特別扱いされ、計画は先送りされる。テッド・オドノヒューとマシュー・ラビンは、自分の現在バイアスを自覚する「賢明な人」と自覚しない「ナイーブな人」を区別し、後者ほど先延ばしの罠に深くはまることを理論的に示した。自制心の問題とは、要するに現在の自分と将来の自分とのあいだの利害対立であり、意志の強さだけでなく自己認識の正確さが結果を左右するのである。

5.4 コミットメント・デバイス

現在バイアスへの有力な処方箋が、将来の自分の選択肢をあらかじめ縛っておく仕掛け、すなわちコミットメント・デバイスである。セイラーとシュロモ・ペナルツィが設計した貯蓄プログラム「Save More Tomorrow」はその代表例で、加入者は昇給のたびに貯蓄率が自動的に引き上げられる契約に事前に同意する。手取りの減少という痛みを将来に配置し、現在の負担を感じさせないこの設計により、導入企業では従業員の貯蓄率が数年で大きく上昇したことが報告されている。

発展途上国でも効果は確認されている。ナヴァ・アシュラフらが2006年に発表したフィリピンの銀行での実地実験では、目標額が目標期日に達するまで引き出せない貯蓄口座を提供したところ、通常口座のみの場合と比べて利用者の貯蓄残高が有意に増加した。またダン・アリエリーとクラウス・ヴェルテンブロッホによる2002年の研究では、レポート課題の締切を学生に自分で設定させると、多くの学生があえて拘束力のある早めの締切を自らに課した。ただし成績が最も良かったのは教員が等間隔の締切を課したクラスであり、自分で設定した締切の効果はそれには及ばなかった。

人は自分の弱さを見越して、あえて自分を縛る手段に価値を見いだす。この洞察は、セイレーンの歌声に備えて自らを船のマストに縛らせたオデュッセウスの故事になぞらえて語られることも多い。ただしコミットメントには、事情が変わっても方針を変更できないという柔軟性喪失の代償が伴う点にも注意が必要である。将来の自分を信用しすぎず、かといって縛りすぎない仕掛けをどう設計するか。これは行動経済学が貯蓄政策や商品設計に応用される際の中心的な課題となっている。

◆ コラム | マシュマロ実験と自制心の神話

心理学者ウォルター・ミシェルが1960年代末からスタンフォード大学で行ったマシュマロ実験は、自制心研究の古典である。4歳前後の子どもの前にマシュマロを1つ置き、15分ほど待てば2つももらえると告げて実験者は退室する。追跡調査では、待てた時間が長い子どもほど後年の学業成績や社会的適応が良好だと報告され、幼児期の自制心が人生を左右するという通説が広まった。しかし2018年にタイラー・ワッツらが約900人のより大規模で多様な標本を用いて行った追試では、家庭環境などの要因を統計的に調整すると相関は大幅に縮小し、当初の主張ほど強い効果は確認されなかった。自制心は重要だが、それだけで人生が決まるわけではないのである。

💡 この章を AI に質問してみよう

- ▶ 「選好の逆転」はなぜ起こるのですか？図2と関係づけて説明してください。
- ▶ 三日坊主を防ぐためのコミットメント・デバイスを、3つ提案してください。

社会的選好と互惠性

伝統的な経済学は、人間が自己の利益のみを追求する存在、すなわちホモ・エコノミクスであると仮定してきた。しかし過去数十年の実験研究は、人々が他者の利益や公平性に配慮し、時に自らの利益を犠牲にしてまで不公正を罰することを繰り返し示している。本章では、最後通牒ゲームと独裁者ゲームの頑健な実験結果を出発点として、不平等回避モデルと互惠性の概念を解説し、社会規範と市場規範の衝突がもたらす逆説的な現象までを検討する。

6.1 利己性の仮定と最後通牒ゲーム

最後通牒ゲームは、1982年にドイツの経済学者ヴェルナー・ギュートらが考案した実験である。提案者と応答者の2人に一定額、たとえば1000円が与えられ、提案者がその分配案を提示する。応答者が受諾すれば提案どおりに分配され、拒否すれば両者とも1円も受け取れない。利己的な合理性を仮定する標準理論の予測では、応答者はわずかな額でも拒否するより受け入れる方が得であり、それを見越した提案者はごく少額のみを提示するはずであった。

ところが実験結果はこの予測を明確に裏切った。世界各地で繰り返された追試において、提案者の平均的な提示額は総額の4〜5割に達し、総額の2割を下回るような低い提案は応答者によってしばしば拒否された。応答者は自らの取り分を放棄してでも、不公正な提案者に損失を与える道を選ぶのである。この傾向は報酬額を数か月分の収入に相当する水準に引き上げた実験でも大きくは変わらず、経済実験のなかでも特に頑健な知見として知られている。

こうした行動は、自己の物質的利得だけでなく、他者の利得や分配の公平性そのものが人の効用に影響することを示唆している。他者への配慮を組み込んだこのような選好は社会的選好と呼ばれ、利他性、公平性への関心、互惠性などを包括する概念として行動経済学の中心的な研究対象となった。利己性の仮定は理論を単純化する有用な道具であったが、それだけでは説明できない体系的な行動が確かに存在するのである。

6.2 独裁者ゲームと不平等回避モデル

最後通牒ゲームの寛大な提案は、純粋な公平感によるものか、それとも拒否を恐れた戦略的計算によるものか。この問いを切り分けるために用いられるのが独裁者ゲームである。ここでは応答者の拒否権が取り除かれ、提案者は与えられた金額を一方的に分配できる。1994年に発表されたフォーサイスらの実験によれば、分配額は最後通牒ゲームより明確に低下するものの、ゼロにはならず、多くの提案者が総額の2割前後を相手に渡した。提案行動には戦略的動機と公平性への配慮の双方が働いているのである。

これらの結果を統一的に説明する理論として、1999年にエルンスト・フェールとクラウス・シュミットが提唱した不平等回避モデルがある。このモデルでは、人は自分の利得だけでなく、自分と他者の利得の差からも不効用を感じると仮定される。自分が相手より少ない不利な不平等は強い苦痛をもたらし、自分が相手より多い有利な不平等もまた、程度は弱いものの効用を下げるのである。

この単純な仮定から、最後通牒ゲームにおける低額提案の拒否は、不利な不平等を解消するための合理的な行動として導かれる。また有利な不平等を嫌う心理は、独裁者ゲームでの自発的な分配を説明する。このモデルは利己的な個人と公平性を重視する個人が混在する状況の帰結を予測でき、競争的な市場では利己的な結果が、二者間の交渉では公平な結果が生じやすいという実験知見とも整合的である。

6.3 正の互惠性と負の互惠性

社会的選好のなかでも行動の動学を理解する鍵となるのが互惠性である。互惠性とは、相手の好意的な行動には好意で報い、敵対的な行動には報復で応じる傾向を指し、前者を正の互惠性、後者を負の互惠性と呼ぶ。フェールらが1993年に行った贈与交換ゲームの実験では、雇用主役が市場の相場を上回る賃金を提示すると、労働者役は強制されていないにもかかわらず高い努力水準に応じた。好意が好意を呼ぶこの関係は、契約で縛りきれない職場の信頼関係の基盤とみなすことができる。

一方、負の互惠性は協力の維持に決定的な役割を果たす。フェールとジモン・ゲヒターが行った公共財ゲームの実験では、罰の仕組みがない場合、協力率は回を重ねるごとに低下していった。しかし自らコストを払って非協力者を罰する機会を導入すると、多くの参加者が実際に罰を行使し、協力水準は高く維持された。自分の利益にならなくても規範の違反

者を罰するこの行動は利他的処罰と呼ばれ、2002年にネイチャー誌で報告されて以来、人間社会における協力の進化を説明する鍵と考えられている。

6.4 社会規範と市場規範の衝突

社会的選好の働き方は、その場がどのような規範で律されているかに強く依存する。友人の引っ越しを手伝う、隣人に贈り物をするといった行為は、金銭を介さない社会規範の世界に属する。他方、労働と賃金の交換は市場規範の世界に属し、対価の計算が支配する。行動経済学者ダン・アリエーらは、この2つの規範が衝突すると、金銭的な報酬や罰則がかえって望ましい行動を減らすことがあると指摘した。

その代表例が、Uri GneezyとAldo Rustichiniがイスラエルの保育園で行った実験である。彼らはハイファ市内の保育園を対象に、子どもの迎えに遅刻した保護者へ少額の罰金を科す制度を一部の園に導入した。標準理論の予測に反し、罰金の導入後に遅刻はむしろ大幅に増加した。罰金は遅刻を、保育士に迷惑をかけるという規範的な問題から、料金を払えば買えるサービスへと変質させ、罪悪感という心理的な抑止を消し去ってしまったのである。

さらに注目すべきは、実験の後半で罰金を撤廃しても、遅刻の水準が元に戻らなかったことである。一度市場規範に置き換えられた関係は、容易には社会規範へ復元されない。2000年に発表されたこの研究の題名が罰金は価格であると述べるとおり、金銭的インセンティブの設計は既存の規範を破壊しうる。報酬や罰則で人を動かそうとする制度設計は、内発的動機や規範意識を締め出す危険、いわゆるクラウディングアウトを常に考慮しなければならないのである。

この章を AI に質問してみよう

- ▶ 最後通牒ゲームで低い提案を断ることは、なぜ「不合理」とされるのですか？
- ▶ イスラエルの保育園の罰金実験から、何がわかりますか？

ナッジと選択アーキテクチャ

政府や企業が人々の選択の自由を奪わずに、より望ましい行動へとそっと後押しする——これがナッジの発想である。行動経済学が明らかにしてきた人間の系統的なバイアスは、裏を返せば、選択肢の提示方法を工夫することで行動を予測可能な形で変えられることを意味する。本章では、リバタリアン・パターナリズムの思想と選択アーキテクチャの概念を学び、デフォルト効果の代表的な実例を確認したうえで、英国と日本における政策応用の展開を概観する。

7.1 リバタリアン・パターナリズムの思想

2008年、シカゴ大学のリチャード・セイラーと法学者キャス・サンスティーンは著書『Nudge』（邦訳『実践 行動経済学』）を刊行し、ナッジという概念を世に広めた。ナッジとは、選択を禁止することも経済的インセンティブを大きく変えることもなく、人々の行動を予測可能な形で変える働きかけの総称である。彼らは自らの立場をリバタリアン・パターナリズムと呼んだ。選択の自由を完全に保障する点でリバタリアン（自由尊重主義）であり、人々を本人にとって望ましい方向へ導こうとする点でパターナリズム（温情主義）だからである。セイラーはこうした一連の業績により2017年にノーベル経済学賞を受賞した。

鍵となるのが選択アーキテクチャという概念である。これは、選択肢がどのような順序・形式・初期設定で提示されるかという選択環境の設計を意味する。セイラーらが挙げる有名な例がカフェテリアである。料理をどの順番で並べても生徒の選択に影響が及ぶ以上、中立的な提示方法は原理的に存在しない。だとすれば、選択環境を設計する立場にある者、すなわち選択アーキテクトは、どうせ影響を与えるのであれば人々の厚生を高める方向に設計すべきだ、というのが彼らの主張である。伝統的経済学が想定する合理的経済人であれば提示方法に左右されないはずだが、現実の人間はそうではないのである。

7.2 デフォルト効果——初期設定の力

ナッジの中で最も強力とされるのがデフォルト（初期設定）の設計である。エリック・ジョンソンとダニエル・ゴールドスタインが2003年に発表した研究は、臓器提供の意思表示方式の国際比較によってこれを鮮やかに示した。提供したい人が自ら登録するオプトイン方式のドイツでは同意率が約12%にとどまるのに対し、拒否したい人だけが登録するオプトアウト方式の隣国オーストリアでは99%を超える。文化的にも経済的にも近い両国でこれほどの差が生じるのは、多くの人が初期設定をそのまま受け入れるからにほかならない。

企業年金でも同様の効果が確認されている。ブリジット・マドリアンとデニス・シーが2001年に発表した米国企業の401(k)プランの研究では、加入手続きを自動加入方式（加入がデフォルトで、外れたい人だけが手続きする）に変更したところ、新入社員の加入率が約37%から86%へと大きく上昇した。加入した方が有利であることは制度変更の前後で何も変わらないのに、初期設定の違いだけで行動が劇的に変わったのである。

さらにセイラーはシュロモ・ベナルチとともに、Save More Tomorrow（明日はもっと貯めよう）というプログラムを考案した。これは、いま貯蓄率を上げるのではなく、将来の昇給のたびに掛金率を自動的に引き上げることを事前に約束させる仕組みである。現在の手取りが減らないため損失回避に触れず、面倒を先送りしたがる現在バイアスをむしろ利用している点が巧妙である。2004年に報告された導入企業での実践では、参加者の平均貯蓄率が約40か月の間に3.5%から13.6%まで上昇した。

7.3 英国の実践とEASTフレームワーク

ナッジはやがて各国の政策に取り入れられていった。先駆けとなったのが、2010年に英国キャメロン政権が内閣府に設立したBehavioural Insights Team（行動洞察チーム、通称ナッジユニット）である。心理学者デイヴィッド・ハルバーンが率いた同チームは、ランダム化比較試験(RCT)によって施策の効果を厳密に検証する姿勢を徹底した。有名な成果が納税催促状の実験であり、「10人中9人が期限内に納税しています」といった社会規範メッセージを文面に加えるだけで納付率が向上することを示した。

同チームが2014年に公表したのがEASTフレームワークである。効果的なナッジを設計するには、行動をEasy（簡単に）、Attractive（魅力的に）、Social（社会的に）、Timely（タイムリーに）すればよいという実務的な指針であり、行動経済学

の膨大な知見を現場で使える形に凝縮したものといえる(表参照)。シンプルであるがゆえに世界中の政府機関や自治体に普及し、政策立案の共通言語となっている。

7.4 日本の取り組みとナッジの倫理

日本でも2017年、環境省が中心となって日本版ナッジ・ユニット(BEST)が発足し、産学官連携でナッジの社会実装が進められている。代表例が家庭向けの省エネレポートであり、近隣の類似世帯と自宅の電力使用量を比較して示す社会規範フィードバックによって省エネ行動を促す。この手法は米国オパワース社の実践に由来し、大規模な検証で効果が確認されてきた。自治体でも、大腸がん検診の受診勧奨はがきに損失フレームを用いて受診率を高めた東京都八王子市の事例など、応用が着実に広がっている。

一方でナッジには、人々を無自覚のうちに操作しているのではないかという倫理的批判も存在する。サンスティーンは、ナッジは透明で公開に耐えるべきであり、仕組みを説明されてもなお機能するものが望ましいと応じている。また近年では、煩雑な手続きなどによって望ましい行動をかえって妨げる摩擦をスラッジと呼び、その削減自体が政策課題とされるようになった。人間が選択アーキテクチャの影響から逃れられない以上、問われるべきは、それを誰が、何のために、どのように設計するかなのである。

表3 EASTフレームワークの4要素

要素	意味	例
Easy(簡単に)	手間や手続きの摩擦を減らし、望ましい行動を最も簡単な選択肢にする	企業年金の自動加入方式、申請書類の簡素化
Attractive(魅力的に)	注意を引く提示や損得の枠組みの工夫で行動の魅力を高める	個人名入りの通知、損失フレームを用いた検診案内
Social(社会的に)	多数派の行動や社会規範を示し、人が同調する力を利用する	「10人中9人が期限内に納税」という催促状の文面
Timely(タイムリーに)	人が最も反応しやすい節目のタイミングで働きかける	昇給時に貯蓄率を引き上げるSave More Tomorrow

💡 この章を AI に質問してみよう

- ▶ ナッジと「強制」や「禁止」は、どう違うのですか？
- ▶ EASTの4要素を使って、健康診断の受診率を上げるアイデアを考えてください。

批判・限界・これから

本書はこれまで、人間の判断が体系的なバイアスに満ちていることを示す数々の研究を紹介してきた。しかし科学は自己批判によって前進する。最終章では行動経済学そのものに向けられた批判と限界を検討する。心理学の再現性の危機、ナッジの効果量をめぐる論争、実験室と現実の乖離、スラッジと呼ばれる悪用、そして操作をめぐる倫理である。これらを直視することは行動経済学を捨てることではなく、その知見をより賢く、より誠実に使うための条件にほかならない。

8.1 再現性の危機とプライミング研究への疑義

2015年、オープン・サイエンス・コラボレーションという国際共同プロジェクトは、心理学の主要誌に掲載された100の研究を追試した結果をサイエンス誌に発表し、統計的に有意な結果が再現されたのは4割に満たないことを明らかにした。いわゆる再現性の危機である。少数の被験者しか集めない検定力の低い実験、都合のよい結果だけが公表される出版バイアス、有意な結果が出るまで分析をやり直すpハッキングといった研究慣行が、その背景として指摘されている。行動経済学は社会心理学の実験に多くを負ってきただけに、この危機と無縁ではいられなかった。

とりわけ疑義が集中したのがプライミング(先行する刺激が無意識のうちに後の行動を変える現象)研究である。バージャーが1996年に報告した、高齢者を連想させる単語を見た学生の歩行速度が遅くなるという有名な実験は、2012年のドワイヤンらによる追試で再現に失敗した。カーネマン自身も2012年にプライミング研究者への公開書簡でこの分野の信頼の危機を警告し、後に自著『ファスト&スロー』の該当章が検定力の低い研究に依拠していたことを認めた。意志力は使うと減るという自我消耗の研究も、大規模な事前登録追試で効果が確認できなかった。

8.2 ナッジの効果量論争と外的妥当性

ナッジの効果はどれほどのものか。2022年、メルテンスらは200件を超えるナッジ研究を統合したメタ分析(複数の研究結果を統計的に統合する手法)を米国科学アカデミー紀要に発表し、平均効果量をコーエンのdで0.43という中程度の値と報告した。しかし直後にマイヤーらが出版バイアスを補正して同じデータを再分析すると、平均効果はゼロと統計的に区別がつかないと指摘された。効果の出なかった研究は論文になりにくいいため、公表済みの研究だけを集めれば効果は過大に見積もられるのである。

同じ2022年、デラヴィーニャとリノスは米国のナッジ・ユニット(行動科学を政策に応用する政府組織)が実施した大規模実験群を分析し、学術誌に載ったナッジの平均効果が約8.7%ポイントであるのに対し、行政実務での平均効果は約1.4%ポイントにとどまると示した。さらに実験参加者が西洋の高学歴で豊かな民主主義社会の人々、いわゆるWEIRDに偏っているとヘンリックらが2010年に指摘したように、実験室の学生で観察されたバイアスが異なる文化や現実の重い意思決定でも同様に現れる保証はない。実験結果を現実へ一般化できるかという外的妥当性の問題である。

8.3 スラッジと操作の倫理

セイラーは2018年、望ましい行動を妨げるために摩擦を増やす設計を、ナッジの対概念としてスラッジ(ヘドロ)と名づけた。ワンクリックで契約できるのに解約には電話が必要なサブスクリプション、気づきにくい自動更新、煩雑な還付申請の書類などが典型である。利用者を欺くウェブ画面の設計はダークパターンと呼ばれ、欧米や日本で消費者保護規制の対象になりつつある。選択アーキテクチャーの技術はそれ自体では中立であり、使い手の目的次第で人を助けもすれば搾取もするのである。

倫理面では、ナッジは人の自律性を損なう操作ではないかという批判が続いてきた。哲学者のハウスマンらは、熟慮を経ない心理的傾向を利用する介入は、たとえ本人のためであっても理性的な説得とは質的に異なると論じた。ギーゲレンツァーはバイアスを迂回するのではなく人々の判断力そのものを鍛えるべきだと主張し、ヘルトウィヒらは統計リテラシー教育などによって意思決定の能力自体を高めるブーストを対案として提示した。これに対しセイラーとサンステイーンは、ナッジは透明に公開され、いつでも選び直せるものであるべきだと応答している。

8.4 AI時代のナッジと学ぶことの意義

ビッグデータとAIの発達は、ナッジを個人ごとに最適化する段階へ進めつつある。法学者のカレン・ヤンは、アルゴリズムが個人データをもとに選択環境をリアルタイムで動的に設計し直す手法をハイパーナッジと呼び、その不透明さに警鐘を鳴らした。2018年に発覚したケンブリッジ・アナリティカ事件は、性格プロフィールに基づく政治広告の配信が民主主義を揺るがしうることを世に示した。動画サイトの自動再生や通知の設計も、行動科学の知見を使って人々の注意を奪い合っている。個人の弱点を精密に突く介入は、公衆衛生や貯蓄支援に役立つ可能性と、監視と操作の道具になる危険とを同時に抱えている。

では行動経済学を学ぶ意義はどこにあるのか。第1に、効果量をめぐる論争が示すのはナッジが無効だということではなく、効果が文脈に強く依存するため事前登録や追試による厳密な検証が不可欠だという教訓である。第2に、スラッジやダークパターンの仕組みを知ることは、消費者として自分を守る武器になる。第3に、自らのバイアスを知ることは、完全な合理性という到達不可能な理想ではなく、誤りやすい自分と上手につきあう技術につながる。批判とともに学び続けるこの姿勢こそ、本書が読者に手渡したい最良の道具である。

この章を AI に質問してみよう

- ▶ ナッジの効果をめぐる論争を、賛成・反対の両方の立場から整理してください。
- ▶ 「スラッジ」の例を、身の回りのサービスから探してください。

本文に登場した主要な専門用語を章ごとにまとめた。意味があいまいなまま読み進めてしまった用語があれば、この一覧からAIに「〇〇についてもっと詳しく」と質問してみしてほしい。

第1章 合理性の仮定とその揺らぎ

■ ホモ・エコノミクス

完全な情報処理能力と一貫した選好をもち、自己の効用を最大化するように行動すると仮定された理論上の人間像である。

■ 期待効用理論

不確実な状況での選択を、各結果の効用に確率を掛け合わせた期待効用の最大化として説明する意思決定理論である。

■ アレのパラドックス

確実な利得が特別に重視される結果、期待効用理論の独立性公理に反する選好が観察される現象である。

■ アノマリー

標準的な理論の予測から系統的に外れて繰り返し観察される現象を指し、理論見直しの重要な手がかりとなる。

■ プロスペクト理論

カーネマンとトヴェルスキーが1979年に提唱した、参照点からの損益評価と損失回避によって選択を記述する理論である。

■ 限定合理性

サイモンが提唱した概念で、認知能力の限界ゆえに人間は最適化ではなく一定水準で満足する選択にとどまるとする。

第2章 限定合理性とヒューリスティック

■ 限定合理性

人間の認知能力には限界があり、完全な最適化ではなく限られた情報処理の下で意思決定が行われるとするサイモンの概念である。

■ 満足化

あらかじめ定めた要求水準を満たす選択肢が見つかった時点で探索を打ち切る意思決定方式である。最適化と対比される。

■ ヒューリスティック

厳密な計算の代わりに用いられる経験的な簡便法である。判断を速める一方で、系統的なバイアスの原因ともなる。

■ システム1

速く自動的で努力を要しない直観的な思考様式である。日常の判断の大部分を担うが、錯誤も生みやすい。

■ 基準率の無視

母集団に占める事前の比率を軽視し、個別の記述の典型性だけから確率を判断してしまうバイアスである。

■ アンカリング

最初に提示された数値が係留点となり、その後の推定がその値の方向へ引きずられて偏る現象である。

第3章 プロスペクト理論

■ プロスペクト理論

カーネマンとトヴェルスキーが1979年に提唱した、リスク下の意思決定を記述する理論。価値関数と確率加重関数を柱とする。

■ 参照点依存性

人は資産の絶対水準ではなく、現状などの基準となる参照点からの変化を利得や損失として評価するという性質である。

■ 損失回避

同じ大きさの利得と損失を比べると、損失のほうが約2倍強く感じられるという利得と損失の非対称性をいう。

■ 感応度逓減

参照点から離れるほど、利得や損失の変化に対する感覚が鈍っていく性質。価値関数のS字型の曲線を生む要因である。

■ 確実性効果

確実な結果が確率的な結果に比べて過大に重み付けられる傾向。アレのパラドックスを説明する鍵となる概念である。

■ 確率加重関数

客観的な確率を心理的な決定の重みに変換する関数。小さい確率は過大に、大きい確率は過小に評価される。

第4章 保有効果と現状維持バイアス

■ 保有効果

自分が保有する財を、保有していない場合よりも高く評価する傾向。損失回避に由来し、WTAとWTPの乖離として観察される。

■ WTPとWTA

財を得るために支払ってもよい最大額が支払意思額(WTP)、財を手放す代償として要求する最小額が受取意思額(WTA)である。

■ 現状維持バイアス

変更したほうが有利な場合でも、現在の状態にとどまることを不釣り合いに好む傾向。サミュエルソンらが体系的に示した。

■ サンクコスト

すでに支出され、いかなる決定によっても回収できない費用。合理的な意思決定では無視すべきものとされるが、人は囚われやすい。

■ メンタル・アカウンティング

お金を出所や用途ごとに心の中の別々の勘定に振り分けて管理する傾向。セイラーが提唱し、心の会計とも訳される。

■ デフォルト

利用者が何も選ばなかった場合に自動的に適用される初期設定。現状維持バイアスにより人々の行動を強く方向づける。

第 5 章 時間割引と自制心

■ 時間割引

将来に得られる利益や損失の価値を、現在から見て低く評価すること。その目減りの程度は割引率で表される。

■ 指数割引

每期一定の割引率で将来の効用を割り引く仮定。どの時点から見ても選好が変わらない時間整合性をもつ。

■ 双曲割引

割引率が現在に近いほど高く、遠い将来ほど低くなる割引の形。時間の経過による選好の逆転を生み出す。

■ 現在バイアス

遠い将来の選択では辛抱強いのに、目先の報酬を前にすると過大に評価してしまう傾向。先延ばしの根源とされる。

■ β - δ モデル

レイブソンが普及させた準双曲割引の定式化。係数 β が現在と将来の断絶を表し、現在バイアスを簡潔に記述する。

■ コミットメント・デバイス

将来の自分の選択肢をあらかじめ制限し、誘惑に負けるのを防ぐ仕掛け。引き出せない貯蓄口座などが典型例である。

第 6 章 社会的選好と互恵性

■ 社会的選好

自己の物質的利得だけでなく、他者の利得や分配の公平性をも考慮に入れる選好。利他性や互恵性を含む包括的な概念である。

■ 最後通牒ゲーム

提案者が金額の分配案を提示し、応答者が拒否すれば両者の取り分がゼロになる実験。公平性への関心を測る代表的な手法である。

■ 独裁者ゲーム

応答者の拒否権を取り除き、提案者が一方的に分配を決める実験。戦略的動機と純粋な公平性への配慮を切り分けるために用いられる。

■ 不平等回避

自分と他者の利得の差そのものから不効用を感じる傾向。1999年にフェールとシュミットが理論モデルとして定式化した。

■ 互恵性

相手の好意には好意で、敵意には報復で応じる傾向。正の互恵性と負の互恵性に区別され、協力の維持を支える基盤となる。

■ クラウディングアウト

金銭的インセンティブの導入が内発的動機や規範意識を締め出し、かえって望ましい行動を減らしてしまう現象である。

第 7 章 ナッジと選択アーキテクチャ

■ ナッジ

強制や大きな経済的インセンティブによらず、選択の自由を残したまま人々の行動を予測可能な形で変える働きかけである。

■ リバタリアン・パターンリズム

選択の自由を保障しつつ、人々を本人の厚生が高まる方向へ緩やかに誘導すべきだとするセイラーらの政策思想である。

■ 選択アーキテクチャ

選択肢の並べ方や初期設定など、人々の意思決定に影響を与える選択環境の設計を指す概念である。

■ デフォルト効果

あらかじめ設定された初期選択肢が変更されずにそのまま受け入れられやすく、結果を大きく左右する現象である。

■ EASTフレームワーク

英国の行動洞察チームが2014年に提唱した、簡単・魅力的・社会的・タイムリーの4原則によるナッジ設計の指針である。

■ スラッジ

煩雑な手続きや分かりにくい情報提示によって、人々の望ましい行動をかえって妨げる摩擦のことである。

第 8 章 批判・限界・これから

■ 再現性の危機

公表された実験結果の多くが追試で再現されないことが2010年代に判明し、心理学を中心に研究の信頼性が問われた問題である。

■ 出版バイアス

統計的に有意な結果ほど論文として公表されやすく、公表研究のみを集計すると効果の平均が実際より大きく見える偏りである。

■ 外的妥当性

実験で得られた結果が、実験室の外の異なる集団・文化・現実の状況にもあてはまる度合いを指す概念である。

■ スラッジ

解約手続きの煩雑さなど、摩擦を増やして人の望ましい行動を妨げる悪しき選択設計を指すセイラーの造語である。

■ ハイパーナッジ

アルゴリズムが個人のビッグデータを用いて選択環境を動的かつ個別的に設計し直す手法を指すヤンの用語である。

■ ブースト

バイアスを迂回して行動を誘導するのではなく、統計リテラシーなど人々の判断能力自体を高めようとする介入である。

AI への質問アイデア集 QUESTION PATTERNS

NotebookLM に読み込ませたこの資料に対して使える、質問の「型」を 10 種類紹介します。例文はそのままコピーして使えます。〇〇の部分、自分が気になった言葉に入れ替えてみましょう。

型 01 やさしく言い換えてもらう

例 ▶ 「プロスペクト理論」を、中学生にもわかる言葉で説明してください。

型 02 たとえ話にしてもらう

例 ▶ 「双曲割引」を、日常生活のたとえ話で説明してください。

型 03 違いを聞く

例 ▶ 「ヒューリスティック」と「バイアス」はどう違うのですか？ 資料の説明にもとづいて教えてください。

型 04 具体例を追加してもらう

例 ▶ 「アンカリング」の例を、買い物の場面で 3 つ挙げてください。

型 05 要約してもらう

例 ▶ 第 3 章の内容を、300 字以内で要約してください。

型 06 自分の生活に引きつける

例 ▶ この資料の内容を、スーパーでの買い物に活かすとしたら、どんな点に注意すればよいですか？

型 07 理由を聞く

例 ▶ なぜ人は損失を利得より大きく感じるのですか？ 資料ではどう説明されていますか？

型 08 クイズを作ってもらう

例 ▶ この資料の内容から、3 択クイズを 5 問作ってください。答えは最後にまとめて教えてください。

型 09 反対意見を聞く

例 ▶ ナッジには批判もあると書かれていますが、どんな批判ですか？ 賛成派と反対派の意見を整理してください。

型 10 資料にない話を確かめる

例 ▶ 「サンクコスト」について、この資料に書かれていないことで、知っておくとよい知識はありますか？ 資料の内容と区別して教えてください。

本資料の内容は、以下の書籍・文献における標準的な解説にもとづいている。いずれも一般の読者に開かれた良書であり、本資料で興味を持った読者の次の一冊として推薦できる。

- ダニエル・カーネマン(村井章子訳)『ファスト&スローー あなたの意思はどのように決まるか?』(上・下)早川書房、2012年
- リチャード・セイラー、キャス・サンスティーン(遠藤真美訳)『実践 行動経済学 — 健康、富、幸福への聡明な選択』日経BP、2009年
- リチャード・セイラー(遠藤真美訳)『行動経済学の逆襲』早川書房、2016年
- 大竹文雄『行動経済学の使い方』岩波新書、2019年
- 友野典男『行動経済学 — 経済は「感情」で動いている』光文社新書、2006年
- 依田高典『行動経済学 — 感情に揺れる経済心理』中公新書、2010年
- Kahneman, D. & Tversky, A. (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2).
- Thaler, R. H. & Sunstein, C. R. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press.

※ 本資料は、パソコン教室の AI 活用講座で NotebookLM などの AI ツールに読み込ませて使用することを目的に作成されたサンプル教材である。専門的な学習・研究には上記の原典を参照されたい。