

## 文部科学省と国立大学附置研究所・センター 個別定例ランチミーティング

### 第143回 京都大学 経済研究所 (2026.2.20)

- |                    |   |                            |
|--------------------|---|----------------------------|
| 12:05 – 12:10(5分)  | : | 研究所・センターの概要                |
| 12:10 – 12:25(15分) | : | 適応的A/B テストの計量経済学的理論研究 陶 俊帆 |
| 12:25 – 12:45(20分) | : | 質疑応答                       |



# KIER

京都大学経済研究所

Kyoto Institute of Economic Research

# 京都大学経済研究所

Kyoto Institute of Economic Research, KIER

# 京都大学経済研究所



1962年4月 産業経済に関する総合研究を目的として、京都大学に附置

2005年7月 先端政策分析研究センター設置

2010年度～ 共同利用・共同研究拠点「先端経済理論の国際的共同研究拠点」認定

## 【2025年度現在】

経済情報解析研究部門

経済制度研究部門

経済戦略研究部門

ファイナンス研究部門

複雑系経済研究センター

先端政策分析研究センター（CAPS）

現代経済分析研究部門（外国人客員・公募国内客員）

# 人員構成(2025年10月現在)

- 教授 10
  - 准教授 3 (海外大学とのクロスアポイントメント含む)
  - 講師 1
  - 助教 5 (全員テニュアトラック)
- 計19名 (うち、外国人教員6名)

また、先端政策分析研究センターで、中央官庁より政策担当者5名を特定教授・特定准教授として受け入れ

# ミッション

- 国内外の経済学をリードする研究：共同利用・共同研究拠点として、内外のコミュニティに対する共同研究プラットフォーム機能も果たす
- 先端政策研究センター（CAPS）を中心に、エビデンスベースポリシー研究と研究成果の社会実装に取り組む：今後は、政策研究の国際的展開を指向
- 京都大学では、経済学研究科に協力講座を提供して、大学院教育の主要科目を担当
- CAPS主催によるシンポジウム等で、研究成果を定期的に社会へ発信

# 近年の活動状況

- 経済学分野のトップ誌に、コンスタントに論文を掲載（第4期中期目標計画期間中、Econometrica, Review of Economic Studiesなど）
- 地方自治体と連携協定を結び、地元の課題を解決する共同研究を実施（高知県、静岡県掛川市など）



# 近年の活動状況

- 直近5年間の研究費獲得状況：科研費は1年あたり2800万円、その他外部資金は1年あたり2600万円
- 共同利用・共同研究拠点における共同研究（「プロジェクト研究」）は、第4期中期目標計画期間中で約17件/年（第3期より増加、国際共同研究も増加）
- CAPSや京都大学研究連携基盤を通じた、医学研究科等との学際的共同研究



**KIER** 京都大学経済研究所  
Kyoto Institute of Economic Research

第143回

文部科学省と国立大学附置研究所・センター 個別定例ランチミーティング

適応的A/B テストの計量経済学的理論研究

---

京都大学 経済研究所

陶 俊帆 (Tao Junfan)

2026年2月20日

# 自己紹介

陶 俊帆 (とうしゅんはん)

- ・ 中国・河南省出身
- ・ 2010年9月 来日
- ・ 2018年9月 横浜国立大学 博士(経済学)
- ・ 2019年9月 日本学術振興会外国人特別研究員(京都大学) 2年間
- ・ 2022年5月 京都大学 経済研究所 講師

## 研究内容：

- ・ 経済学における 計量経済学 の理論研究
- ・ 現在のテーマ: A/B テストの理論研究
- ・ キーワード：  
A/Bテスト × 適応的意思決定 × 計量経済学



# A/Bテスト = RCT のシンプルな実装形態

- 計量経済学の中の一つ重要な課題は、経済現象の「因果関係」をデータから明らかにする。

「本当に効果があるのか？」をデータで検証する

- 研究の多くは、さまざまな政策や経済現象の「因果関係」に関心を持っている。例えば、

- 「小規模なクラスサイズは、生徒のテストのスコアを向上させるでしょうか？」

教育分野

- 「最低賃金の引き上げは、特定の産業 (例:飲食業) の雇用にどのような影響を与えるか？」

労働経済分野

- 「政策金利を0.25%引き上げた場合、物価上昇率にどれくらい影響を与えるか？」

金融政策分野

- データの種類：

- 観測データ (Observational Data) → 因果推論は一般に難しい。

- 実験データ (Experimental Data) → 因果効果をより信頼性高く識別できる。

因果推論の本質的困難は、反事実 (counterfactual) が観測できないことである。

- では、どうやって因果を確かめるのか？ **ランダム化比較実験 (Randomized Controlled Trial: RCT)**

**A/Bテスト = RCT のシンプルな実装形態**

→ インターネットやプロダクト改善で広く利用



# A/Bテスト

- ・ A/Bテストは、2つの異なる選択肢をランダムに割り当てて比較し、効果の差をデータに基づいて測定する方法。
- ・ なぜ有効なのか？ 「もし変えなかったら？」という状況を人為的に作り出すことができるから。

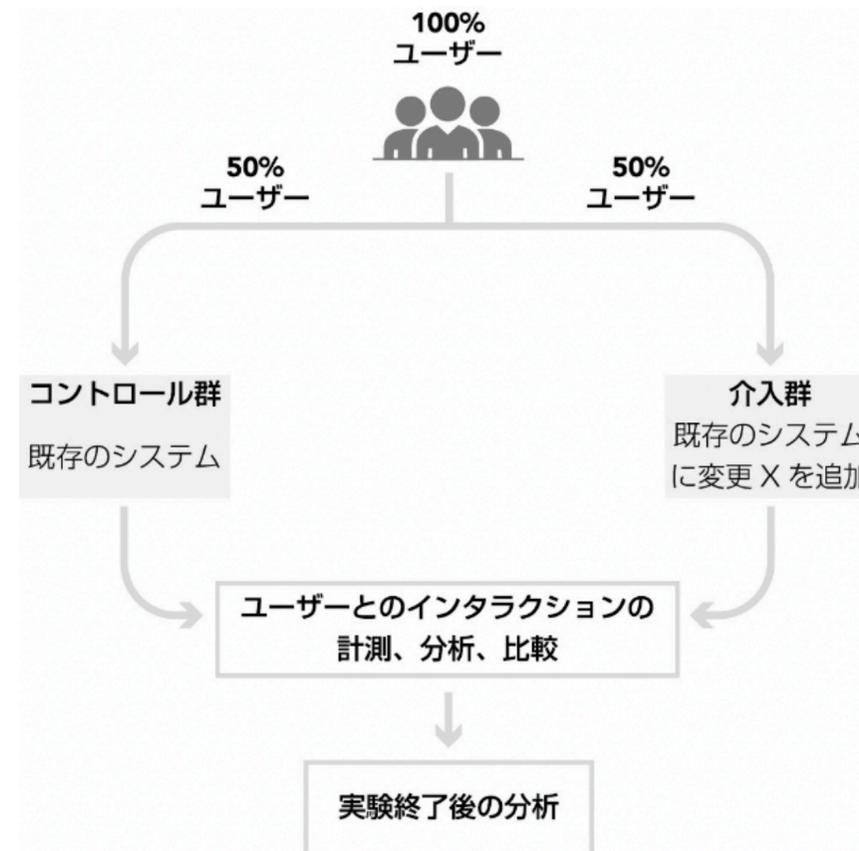


図 1.2 ●最も単純なコントロール実験 (A/B テスト)

# 日本の政府の政策評価

## 電気料金の変更と省エネ要請の因果効果を、地域ランダム化実験で識別した政策評価

- 背景:** 日本では、原発事項以降、電力不足は問題となっている。  
経済学的な観点から、節電を促す政策として、以下の2つがあり得る。
- ① 価格変動による節電誘導 (電気料金の変更)
  - ② 自発的な省エネを呼びかけ (政府や電力会社の節電要請)

### 実験の目的

- ・ 価格変動 (電気料金の変更) と省エネ要請 (行動変容の呼びかけ) のどちらがより効果的を検証
- ・ 消費者がどの条件で節電行動を取るのかを分析を明らかにすることを目的とした。

### 実験の実施

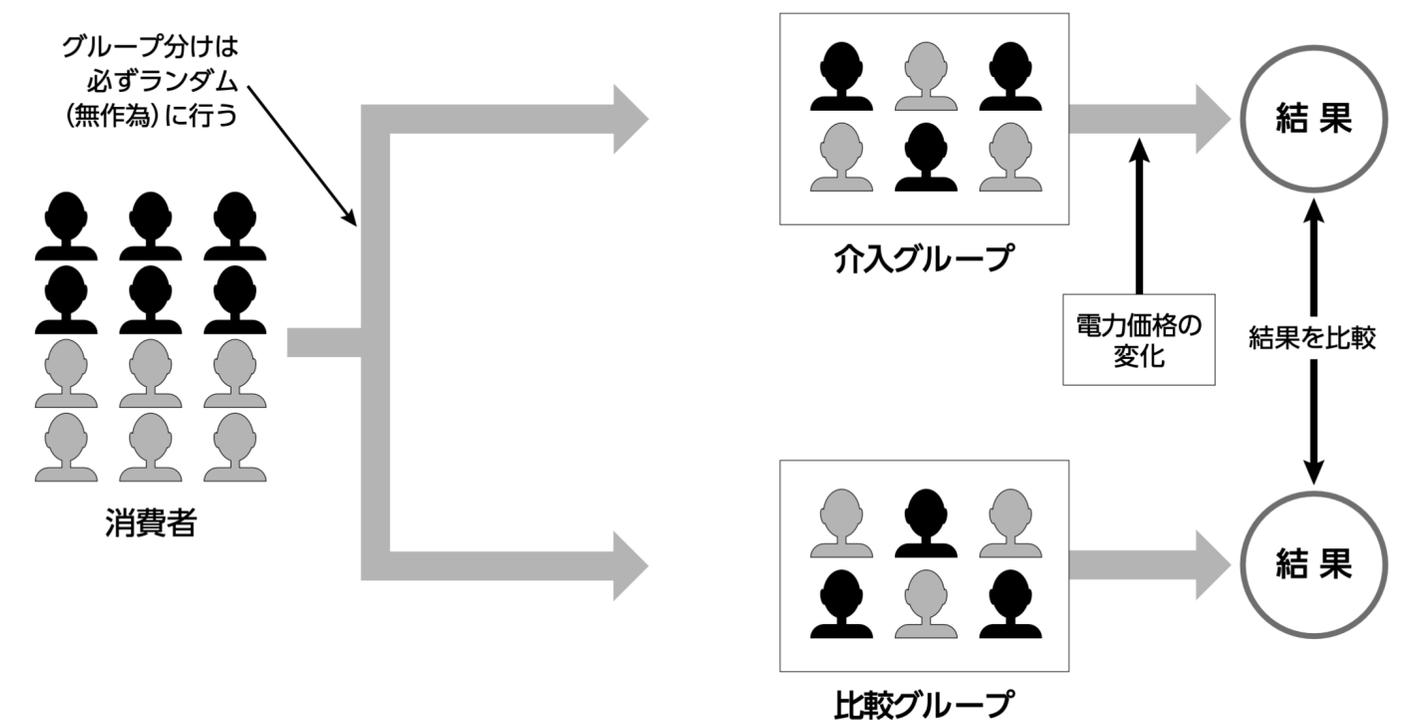
大学の研究者が**実験設計**を行い、以下の4つの地域で**ランダム化比較試験を実施**。

- 1) 横浜 2) 京都 (けいはんな) 3) 豊田 4) 北九州

### 実験の結果

この実験により、どの政策がより効果的に節電を促すかが分析された。

図表2-3] ランダム化比較試験 (RCT) (A/Bテストの一般化の形)



ランダム化比較試験(RCT)では、介入グループと比較グループのグループ分けをランダム(無作為)に行います。

出典: 『データ分析の力 因果関係に迫る思考法』伊藤公一朗 (光文社新書、2017年)

論文: Ito, Koichiro, Takanori Ida, and Makoto Tanaka. "Moral suasion and economic incentives: Field experimental evidence from energy demand." American Economic Journal: Economic Policy 10.1 (2018): 240-267.

# 身近のA/B テストの例

---

実は、インターネットの広告は人によって広告が違う。

「YouTube や Instagram、Yahoo ニュース」

- ▶ 同じサイトでも、表示される広告は人によって違う!!

その理由は：企業の戦略で、「どの広告が効果的か」をA/Bテストで実験している。

- ▶ どの画像がクリックされるか？
- ▶ どの言葉が購入につながるか？

そして、テストの結果が反映される → より効果的な広告が表示される。

- ▶ 以前の行動や好みに基づいてより効果的な広告が表示される。

つまり、私たちは日常的にA/Bテストの世界にいる。

# 私の研究

- ・ 今まで説明したA/B テストは実験開始したら、実験が終わるまでデータを収集し、分析する方法。



- ・ **良いところ:**

1. **客観的な評価:** データに基づいて、**客観的な評価**ができる。
2. **高いエビデンスの質:** 医学や政策決定など、重要な意思決定の根拠として広く用いられている。

- ・ **問題点:**

1. **時間とコスト:** 大規模なサンプルサイズが必要となるため、時間やコストがかかる。
2. **倫理的な問題:** 実験実施中に、**経済的損失** 或いは **薬の有害作用** など、参加者に不利益を生じる可能性がある。

👉 この問題点を克服するため、近年では、**適応的サンプリング**が注目されている。

私は研究: **適応的サンプリング + A/B テスト = 適応的A/B テスト** に対する計量経済学の理論研究。

# 適応的サンプリングとは？

---

## サンプリングとは

- ・ サンプリングとは、全体（母集団）の特徴を、一部の標本を調べることで推測する手法。
- ・ 多様な手法があり、目的に応じて、適切な手法を選択することが重要。

## 適応的サンプリングとは

- ・ 実験の進行状況に応じて、実験の途中でサンプルサイズや実験のやり方を変える方法。
- ・ 簡単に言うと「様子を見ながら、臨機応変にやり方を変える」(柔軟性あり)。
- ・ 「良いものは伸ばし、悪いものは早く止める」という考え方。
- ・ 元々は医療の薬開発、特に臨床試験の分野で発展してきた手法。
- ・ IT技術の発展に伴い、経済学やマーケティングなど、幅広い分野で活用されるようになってきている。

# サンプリングデザイン

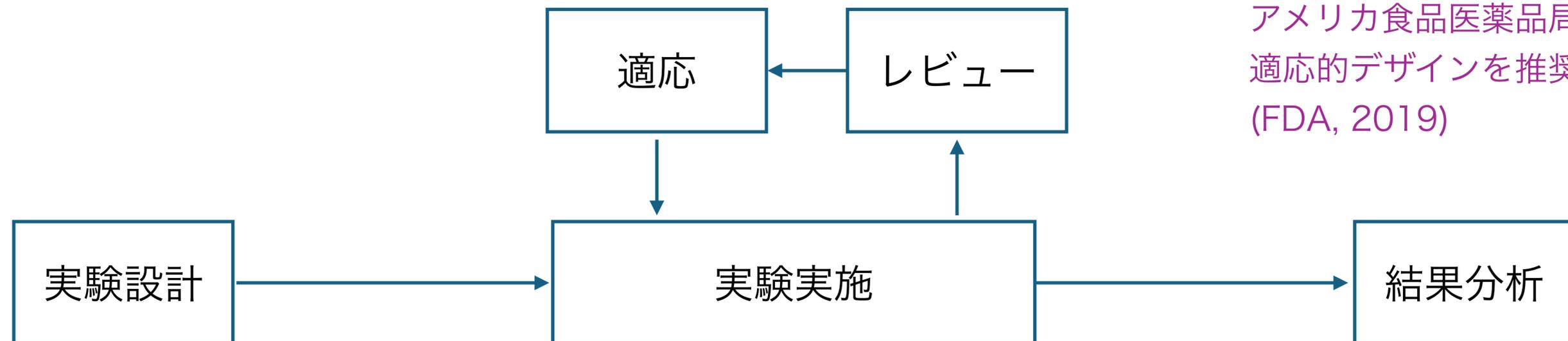
- ・ 通常のサンプリングデザイン (固定標本数 (Fixed Sample Size))

設計 → 実施 → 分析という直線的な流れ



- ・ 適応的サンプリングデザイン (Adaptive Sampling)

「実施」の途中で「適応」と「レビュー」を繰り返しながら、試験や研究の進行に応じて調整できる仕組み。



アメリカ食品医薬品局 (FDA) も  
適応的デザインを推奨する。  
(FDA, 2019)

# A/Bテストの2つのサンプリング手法

**目的:** 2種類の蜜柑（品種Aと品種B）の甘さを比較し、甘い方だけを来年生産する。

- **固定標本数 (fixed):** あらかじめ決めた数のデータを取る。

- AとBを平均20個ずつ食べて、最後に比較して決定（途中では変更しない）
- 公平で統計的に安定、しかし時間とコストがかかる。

- **適応的 (adaptive):** 食べたデータを使って、次にどの品種を取るか決める方法。

- A、Bを少しずつ食べて様子を見る、甘そうな方に多く配分する、良くなさそうなら早めにやめる。
- 有望な選択肢に資源を集中できる、より迅速な意思決定が可能。

▶ 経済学の視点では、どちらの方法も目的は同じ: 「**最小のコストで、最適な選択肢を見つける**」

▶ しかし現実では、長期間データを集める余裕がない、**意思決定のスピード自体に価値がある**場面が多い。

▶ 適応的A/Bテストは、限られた資源（時間・予算）を、有望な施策に動的に配分する仕組み。これは  
**経済学の「利益最大化」と「損失最小化」**

**「探索と活用 (exploration vs exploitation) のトレードオフの問題」**

と深く関係している。

# 適応的A/Bテストの例

マクドナルドのアプリでのコーヒー割引キャンペーン

## 通常のA/Bテスト(固定標本数)

- ・ 1か月間データを収集し、期間終了後に効果を評価.
- ・ その後に割引の有効性を判断する.
- ・ 判断は一度だけ.

## 適応的A/Bテスト

- ・ データを収集しながら分析、効果が見えれば即座に拡大、多くの顧客に適用する.
- ・ 逆に、効果が弱ければ早期に中止、無駄なコストを抑える.

## 経済的メリット

有望な施策を早く拡大できる、無駄なコストを抑えられる。 経済合理性を追求できる。

意思決定のスピードが向上する直感的には、適応的な方法の方が経済的に合理的に見える。

# 適応的A/Bテストの課題

適応的A/Bテストは強力だが、統計的なリスクを伴う。適応的なデータ収集では、将来の観測が過去に依存するため、

$$X_t \sim P(\cdot | \text{past data})$$

i.i.d. 仮定(「データ収集がデータに依存しない」, 古典的推論の基礎)が崩れる。

→ 固定標本サイズに基づく古典的推論は適切でない

何が起きるのか？

1. 早期データへの依存 初期の偶然の変動を過大評価する可能性

例：初週がたまたま好調だった → 偶然を「本当の効果」と誤認するリスク。

2. 推定の偏り 配分を動的に変えることで、標本が偏る可能性がある。

・例：特定の顧客層（リピーター）に効果が集中 → 全体効果を正しく測れない可能性

3. 長期効果の見落とし 短期的な指標だけで判断すると、長期的な影響を過小評価する可能性

**Adaptive な意思決定下では、推論もまた逐次的でなければならない。**

→ **Sequential Inference (任意の時点で停止しても有効な推論枠組み)**

本質的なトレードオフ：経済的効率性 vs 因果推定の正確性

# 私の研究

適応的サンプリング + A/B テスト = 適応的A/B テストに対する計量経済学の理論研究

適応的サンプリング × A/Bテスト = 動学的資源配分のもとでの因果推定

核心の問い

経済的効率と統計的正確性をどのように両立させるか？

具体的研究課題

1. 因果効果の推定： 動的配分のもとで、因果効果をどのように識別・推定するか？
2. 推論の妥当性： 通常の推論はそのまま使えない、誤判定確率をどのように制御するか？
3. 理論的保証： 上記の推定・推論に対して、どのような理論的保証を与えるか？

目指しているもの

リアルタイム意思決定に応じた推論理論の構築

# まとめ

---

- ▶ A/Bテストは、マーケティング・政策・医療など、社会の様々な場面で広く使われている。
- ▶ しかし、適応的な動的な配分のもとで、正しい推論を行うことは容易ではない。
- ▶ リアルタイム意思決定の時代において、それに対応した推論理論が必要である。

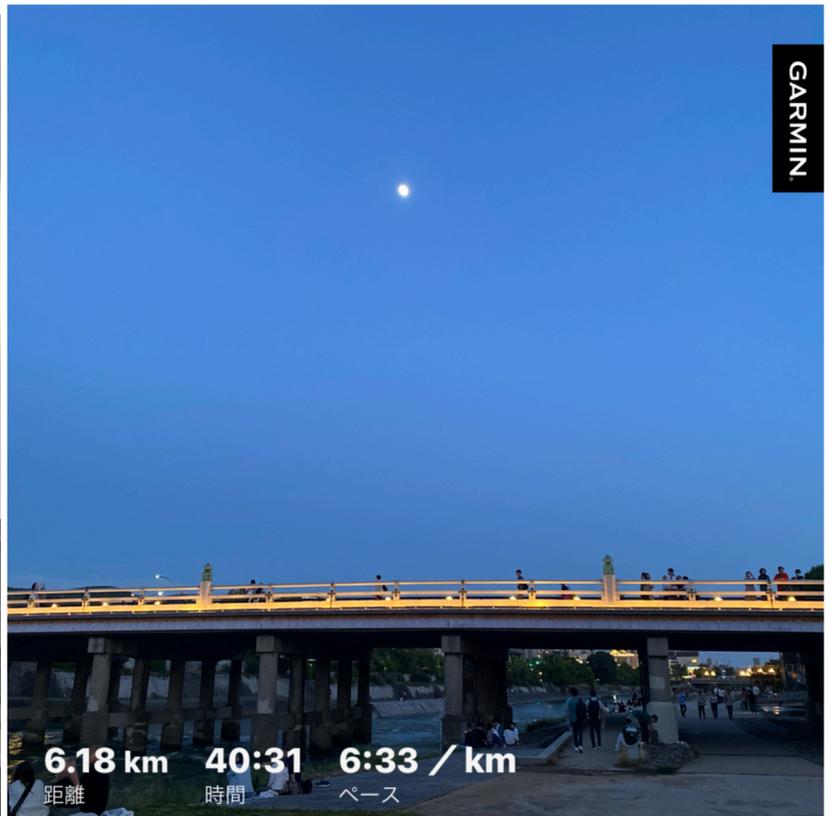
そして、私たち個人にとっても、私たちは日々、適応的なアルゴリズムの中で意思決定をしている。

- ・自分に合う情報ばかり見ていないか？
- ・それは本当に自分の選択か？
- ・それともアルゴリズムによる配分の結果か？

- ▶ 適応的な仕組みは便利である一方、情報の偏りや公平性の問題とも無関係ではない。
- ▶ だからこそ、その背後にある推論理論が重要である。
- ▶ 本日の報告は、Karun Adusumilli (The University of Pennsylvania, USA) 及び Jiaying Gu (University of Toronto, Canada) との共同研究の一部に関するものである。
- ▶ この共同研究は 科研費および 2025年度京都大学経済研究所 プロジェクト研究の助成を受けている。



どうもありがとうございます。



GARMIN

6.18 km 40:31 6:33 / km  
距離 時間 ペース



GARMIN

6.60 km 44:19 6:43 / km  
距離 時間 ペース