

# 経済学とデータ分析の融合：離散選択モデル入門

2024年8月3日

@早稲田大学オープンキャンパス

講師：遠山 祐太

早稲田大学 政治経済学部

(Web配布版。一部画像を削除しています)

# 自己紹介

---

▶ 講師：遠山祐太（とおやま ゆうた）

▶ 経歴：

▶ 仙台市生まれ、仙台一高卒

▶ 京大経済→東大院修士（公共政策→経済学）→米国ノースウェスタン大学Ph.D. in Economics

▶ 帰国後、2018年9月より早稲田大学政治経済学部准教授

▶ その他のお仕事：UTEconアドバイザー

▶ 研究・教育：産業組織論、計量経済学(データ分析)、エネルギー・環境経済学

# 経済学とは？ 産業組織論とは？

- ▶ 経済学：経済主体の意思決定を分析し、社会・経済問題の解決に取り組む学問
  - ▶ ツール：数理モデル・データ分析を積極的に用いる
  
- ▶ 産業組織論：市場・産業における企業の意思決定・競争を分析する。
  - ▶ 例：どのように製品の価格を設定するか？ どんな新製品を投入するか？
  - ▶ 企業行動分析の土台としての消費者行動の分析も射程に
  
- ▶ 経営学(マーケティング・経営戦略)との違い？
  - ▶ 使うツール・分析対象はほぼ同じ。
  - ▶ 産業組織論は経済学ゆえ「社会・消費者の視点」を意識、経営学は「企業の視点」
  - ▶ 今日の講義は「企業の視点」からの話が中心です。

# 経済学の潮流：理論研究から実証研究へ

- ▶ 過去30年の進化：データを活用した実証&応用研究
- ▶ 背景1：利用可能なデータの増加（ビッグデータ）
  - ▶ 例：オンライン小売企業（楽天、Amazon）、人流データ(携帯電話)
- ▶ 背景2：技術の進化・ビジネスモデルの変化に伴う新しい社会・経済問題の発生
  - ▶ 例：巨大テック企業の独占、気候変動対策としての脱炭素
- ▶ 応用研究が広がるにつれて、社会・ビジネスへの活用も広がっている。

# テーマ：経済学とデータ分析の融合、そしてその応用

- ▶ 「**離散選択モデル**」と呼ばれる経済モデルを活用したデータ分析事例を紹介
  - ▶ 事例：ブランド価値の測定、需要予測、価格戦略、などなど
  
- ▶ 今日お伝えしたいこと：
  1. 経済学を用いることで、どのような**データ分析**ができるようになるか？
  2. 経済学の知見が**ビジネス・実務**においてどのように役立つか？
  3. 経済学は如何にして、「文系」と「理系」の知見を融合した「**総合格闘技**」であるか？

# 本日のプラン

離散選択モデル  
の導入

事例紹介

- 1 : きのこの山VSたけのこの里
- 2 : テック企業の価格戦略

まとめ：  
実務活用

# 離散選択モデルとは？

# 例：あなたはどのiPhoneを買いますか？

著作権の都合上、割愛します。ソース：<https://www.apple.com/jp/iphone/compare/>

# 離散選択モデルとは

- ▶ 離散選択モデル：「**数多くの選択肢の中からどれを選ぶか**」を表現する数理モデル
- ▶ 意思決定者は**複数の選択肢**に直面している。
  - ▶ 例：iPhone 13, 14, SE(3世代), 13 Pro, などなど
- ▶ それぞれの選択肢(例：製品)には、付随する**特徴**がある
  - ▶ 例：価格、ディスプレイサイズ、カメラの画素数、バッテリー持続時間、などなど
- ▶ 意思決定者は「どの特徴を重視するか」という**好み**をもっている。
  - ▶ 例：値段は気にしない（親が払ってくれるから）、大きい画面が良い、などなど
- ▶ 選択肢の特徴を踏まえて、**自分の満足度(効用)が最も高くなる選択**をする

# 数式を使ってフォーマルに書いてみよう

- ▶ ある携帯電話機種を購入したときに得られる効用(満足度)を以下のように書こう

$$(\text{効用}) = \alpha \times (\text{価格}) + \beta \times (\text{画面サイズ}) + \gamma \times (\text{バッテリー時間})$$

- ▶ 文字定数 $\alpha, \beta, \gamma$ は意思決定者が各特徴をどのくらい重視するかという「好み」を捉える。
  - ▶ 数理モデルにおいてはパラメータと呼ばれる
- ▶ 選択肢にあるすべての機種について効用を計算し、最も効用が高い機種を選択する。

# 簡単な例 1 : iPhoneの購入

- ▶ 効用のパラメータを以下のように考える

$$(\text{効用}) = -2 \times (\text{価格}) + 3 \times (\text{画面サイズ}) + 0.5 \times (\text{バッテリー時間})$$

- ▶ 以下の3機種を比較しよう

	価格(万円)	画面サイズ(インチ)	バッテリー(時間)	効用
iPhone 14	11.98	6.1	20	4.34
iPhone SE	6.28	4.7	15	9.04
iPhone 13	10.78	6.1	19	6.24

- ▶ iPhone SEを購入する！！

## 簡単な例 2 : もし価格をあまり気にしないと ?

- ▶ 価格の係数が小さい (= 価格をあまり気にしない)

$$(\text{効用}) = -1 \times (\text{価格}) + 3 \times (\text{画面サイズ}) + 0.5 \times (\text{バッテリー時間})$$

- ▶ 同様に 3 機種を比較しよう

	価格(万円)	画面サイズ(インチ)	バッテリー(時間)	効用
iPhone 14	11.98	6.1	20	16.32
iPhone SE	6.28	4.7	15	15.32
iPhone 13	10.78	6.1	19	17.02

- ▶ iPhone 13を購入する！！

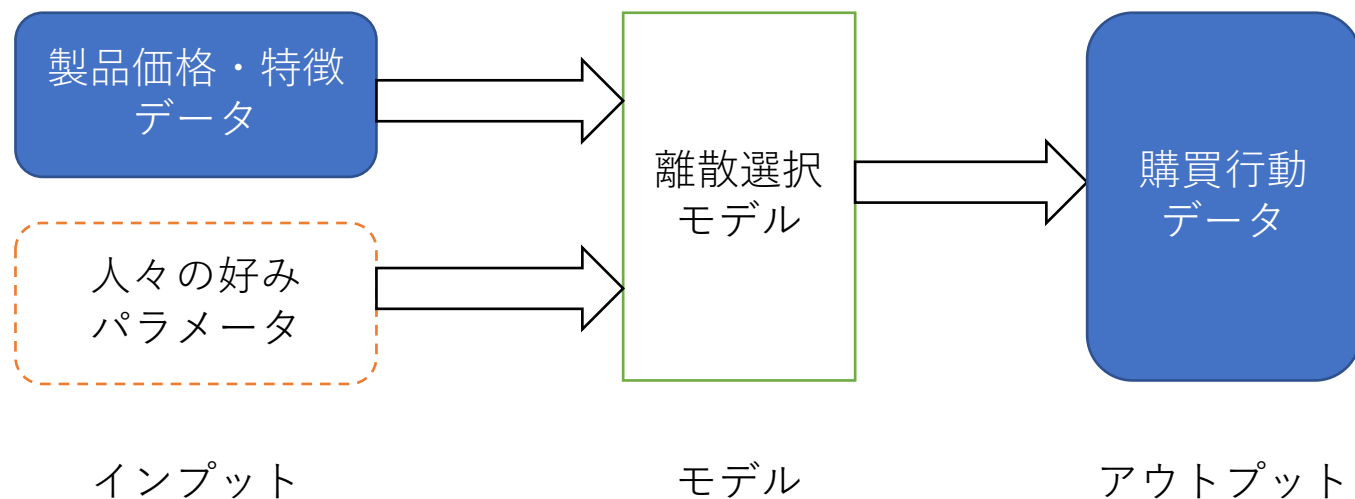
- ▶ ポイント : 人々の「好み」によって、選択は異なってくる！！

# 世の中の意思決定の多くが離散選択モデルで表現できる

- ▶ **消費者の購買行動** ← 今日の主眼
- ▶ **大学進学**：どの大学・学部に進学するか？
  - ▶ 選択肢の要素：学費、家からの距離、カリキュラム、などなど
- ▶ **交通手段**：バス、鉄道、自家用車、自転車
  - ▶ 選択肢の要素：移動費用、移動時間
- ▶ **投票行動**：どの政党に投票するか or そもそも投票に行かないか？
  - ▶ 政党の特徴（党派性）と、自分自身の政治選好を踏まえて投票先を決める。

# なぜ数理モデルが必要か？

## ▶ モデル

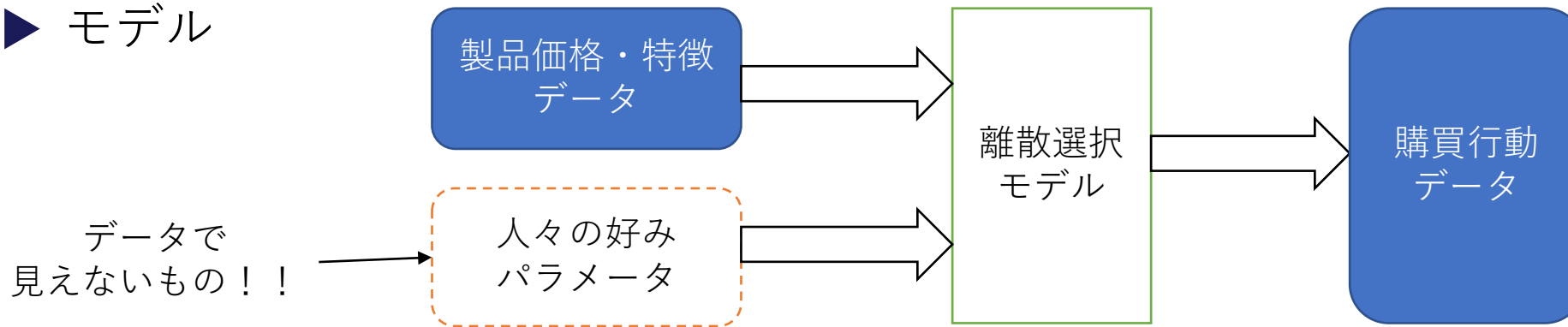


## ▶ このモデルのインプットを変えることで、様々な予測が可能となる！！

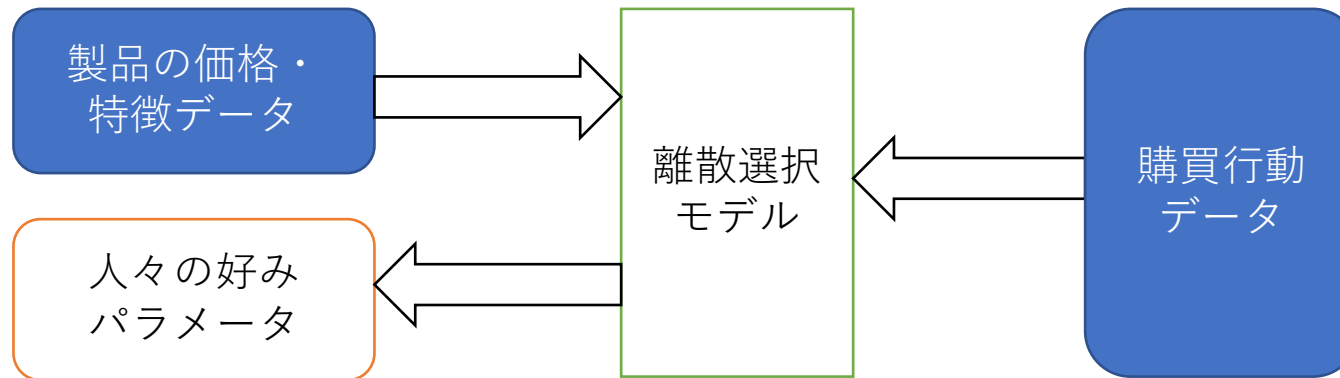
- ▶ 例1：iPhone 13を値下げしたら、売上はどう変化するか？
- ▶ 例2：新たなスペックをもつiPhone SE（第4世代）を導入したらどうなるか？

# 「モデル」から「現実(データ)」へ：データ分析の役割

## ▶ モデル



## ▶ データ分析：



- ▶ モデルが「現実」にうまくフィットするようにパラメータを決める必要→データ分析！！
- ▶ 「モデルの予測」と「実際の購買行動」が近くなるような「パラメータ」を推定する！！

# 事例 1： きのこのこの山 vs たけのこの里

# 背景：きのこのこの山 VS たけのこのこの里

割愛します。

# アンケート：きのこの山 VS たけのこの里

- ▶ 今、あなたは**300円**持ってスーパーのお菓子コーナーにいます。目の前に「きのこの山」と「たけのこの里」が並んでいるとします。以下の、**5通り**の価格に直面したとき、あなたはどの選択肢を選びますか？（なお余ったお金は手元に残ります）
  - ▶ (1)きのこの山を買う、(2)たけのこの里を買う、(3)どちらも買わない。
- ▶ Q1: (きのこ, たけのこ)=(200円, 200円)
- ▶ Q2: (きのこ, たけのこ)=(170円, 200円)
- ▶ Q3: (きのこ, たけのこ)=(240円, 200円)
- ▶ Q4: (きのこ, たけのこ)=(200円, 250円)
- ▶ Q5: (きのこ, たけのこ)=(200円, 180円)

# 皆さんのアンケート結果を見てみましょう

---

▶ アンケートアプリへGo

# 2022年春学期「産業組織論」でのアンケート結果

(きのこ, たけのこ)	きのこの山	たけのこの里	どちらも買わない
(200円, 200円)	18%	49%	33%
(170円, 200円)	47%	29%	23%
(240円, 200円)	3%	60%	37%
(200円, 250円)	36%	12%	52%
(200円, 180円)	12%	74%	15%

- ▶ ポイント 1 : 価格が上がるとシェアが下がる→**価格に反応**
- ▶ ポイント 2 : 価格差が大きくても選ぶ人もいる→**きのこ/たけのこのブランド価値**
- ▶ 離散選択モデルを使って、より詳しく分析してみよう！

# 離散選択モデルをデータに当てはめる

- ▶ 各消費者は(1)きのこ (2)たけのこ (3) その他 から、最も効用が高い選択肢を選ぶ。
- ▶ それぞれの選択からの効用 $U$  :

$$\begin{aligned}U(\text{きのこ}) &= \alpha(\text{きのこ}) - \beta \times (\text{きのこの価格}) \\U(\text{たけのこ}) &= \alpha(\text{たけのこ}) - \beta \times (\text{たけのこの価格}) \\U(\text{買わない}) &= 0\end{aligned}$$

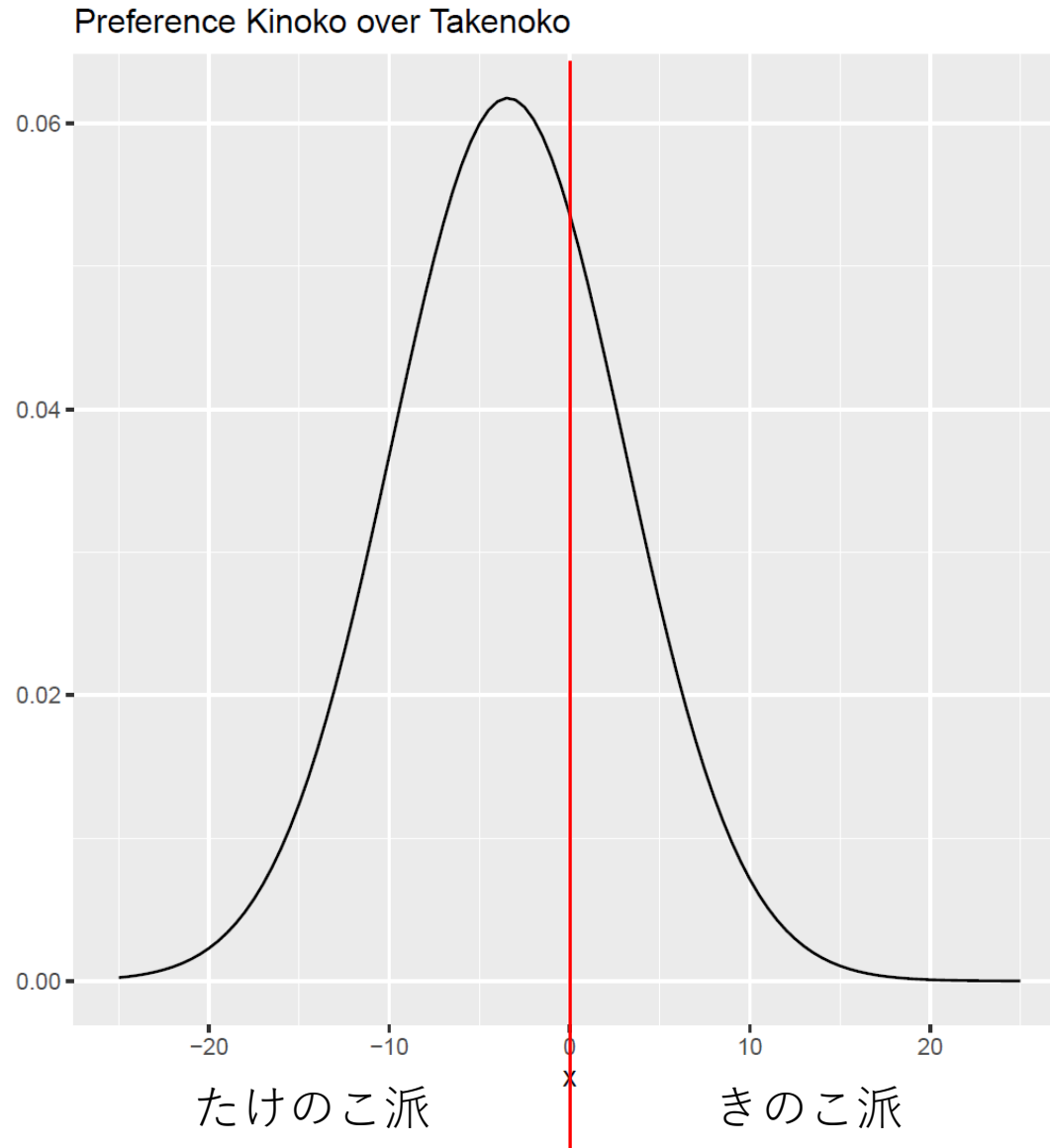
製品の価格を  
どれだけ気に  
するか？

きのこ／たけのこ自体への好み

- ▶ このモデルのパラメタは $\alpha(\text{きのこ})$ 、 $\alpha(\text{たけのこ})$ 、 $\beta$ の3つ
- ▶ アンケートで得られたデータを用いて、パラメタを推定する！！
  - ▶ ここでは人によってパラメタの値が異なると考え、その分布（バラツキ）を推定する。

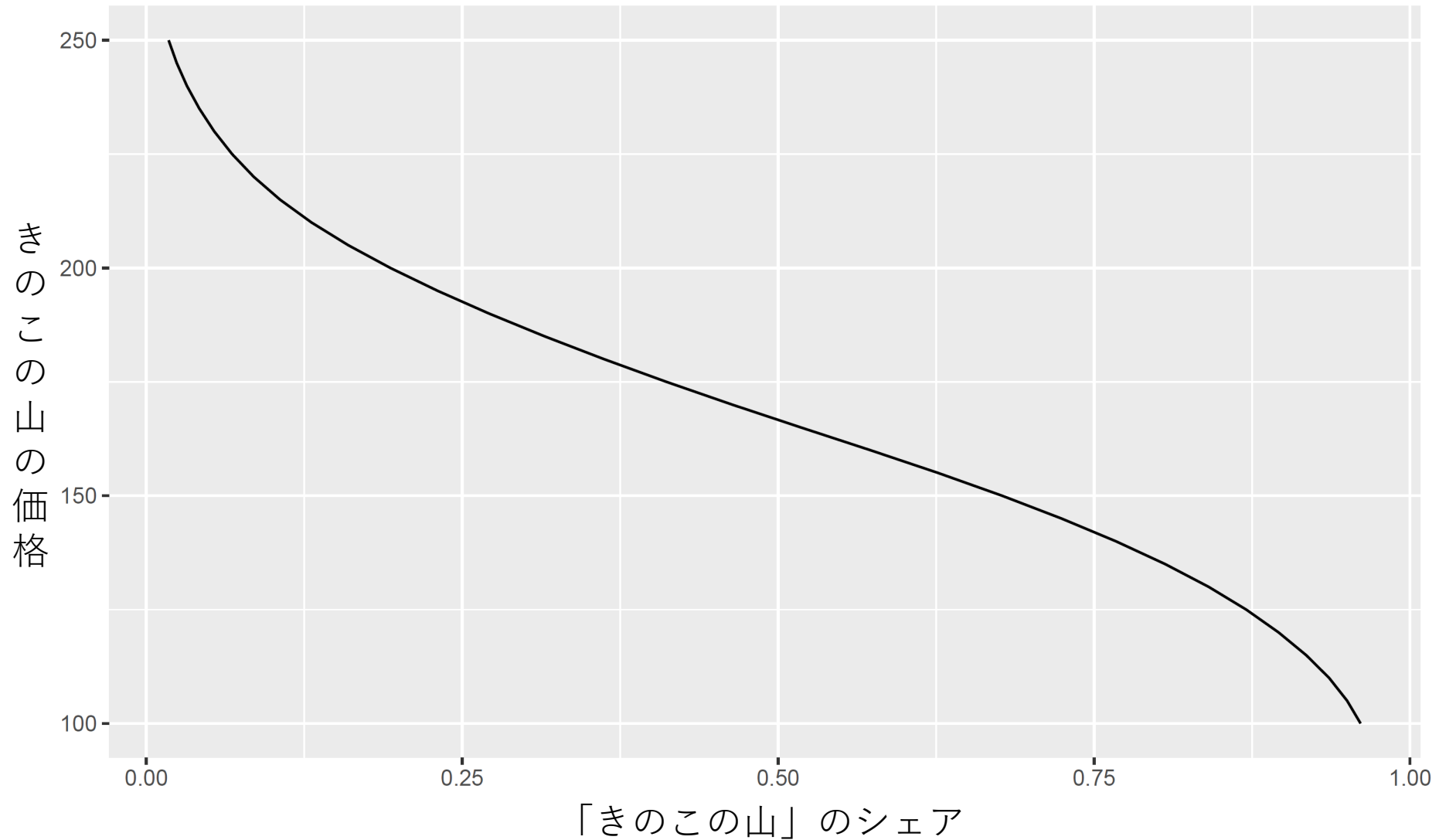
# $\alpha(\text{きのこ}) - \alpha(\text{たけのこ})$ の分布(バラツキ)

約7割が  
たけのこ派！！



# きのこのこの山のシェアは価格にどう反応するか？(需要曲線)

「たけのこの里」の価格が200円の際の「きのこのこの山」の需要曲線



# 事例 2： テック企業の価格実験

# 背景：テックビジネスの流行

## ▶ 例：GAFA

割愛します。（GAFAのロゴが入ってます）

## ▶ 日本でも数多くのテックビジネス

▶ 楽天、PayPay、じゃらん、などなど

# テックビジネスの肝：大量のデータを経営に活用

- ▶ テック企業は個人に関する**大量のデータ**を保有している
  - ▶ Google: 検索履歴、メールのログ
  - ▶ Amazon: 購買履歴
- ▶ **データに基づく意思決定**を、オンラインプラットフォームで即座に実行可能
  - ▶ 例1：個人の嗜好を反映したターゲティング広告
  - ▶ 例2：需給に応じて、価格を臨機応変に変更する（ダイナミックプライシング）
- ▶ データ利活用は**ビジネス・政策両方の観点から重要**なトピック
  - ▶ ビジネス：利益を上げるためにどう活用するか？
  - ▶ 政策：データの活用が消費者の不利益になるのでは？ プライバシーの問題も？

# 事例：Ziprecruiter.comの価格実験

- ▶ アメリカのオンライン求人サイト Ziprecruiter.com とシカゴ大学の共同研究
- ▶ Ziprecruiter.com とは？
  - ▶ ホームページ：<https://www.ziprecruiter.com/>
  - ▶ 求職者は無料で使える。
  - ▶ 求人企業は月額99ドルのサブスクリプション制。解約は自由

割愛します。(Ziprecruiterのトップページ)

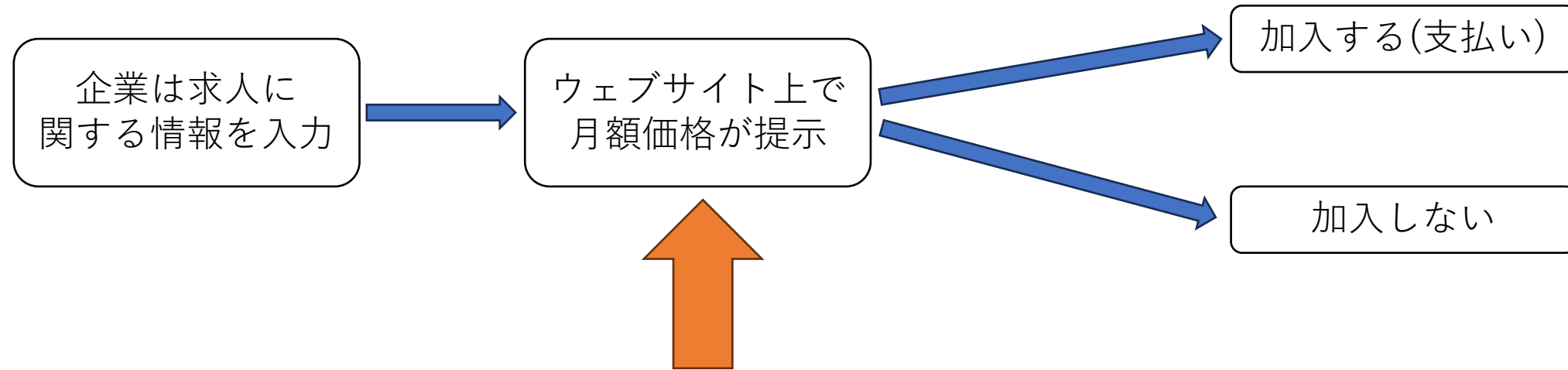
# 検索結果の例：Economist job in Chicago, IL

割愛します。

# 価格実験及びデータ分析の概観

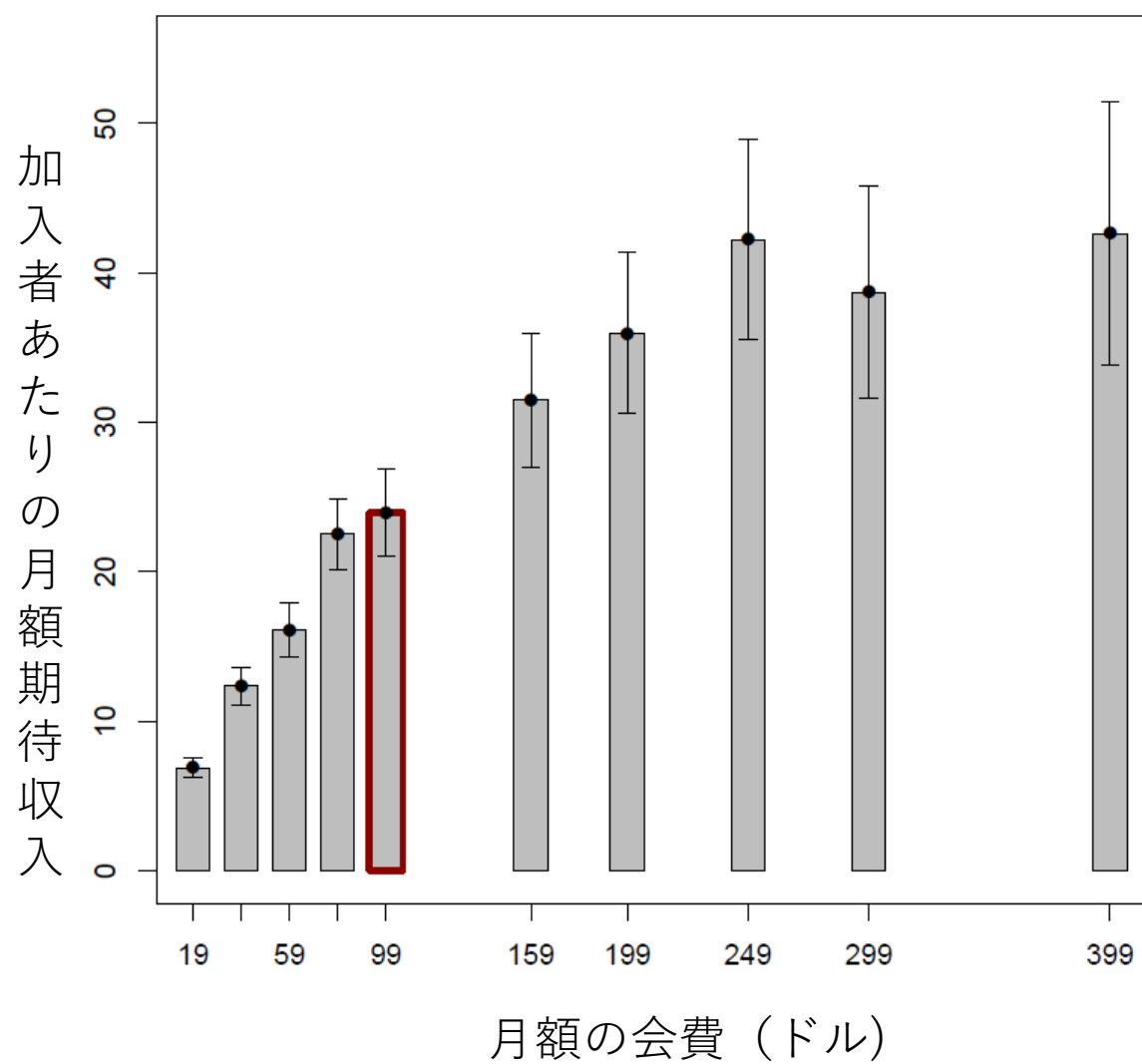
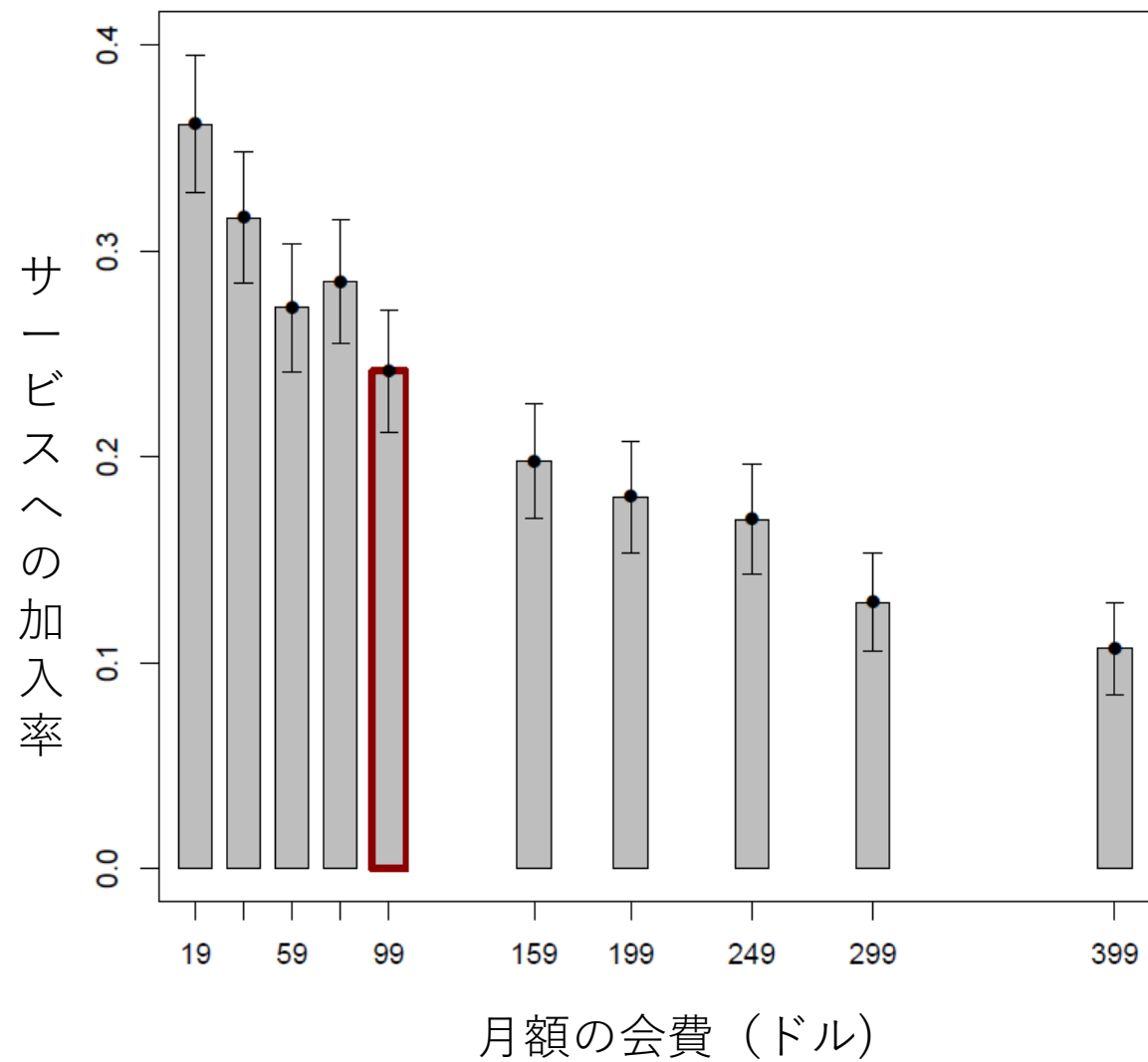
- ▶ 2015年春：経営陣がサービス価格の再検討を始める。
- ▶ 目標：自社の利潤を上げるような価格をつけたい
  - ▶ でも、価格を上げたらお客さんが逃げるかもしれない。。。
  - ▶ 自社のサービスに対する需要曲線を知る必要がある！！
- ▶ 分析ステップ：
  1. ウェブサイト上でデータを収集する。
  2. データから消費者の需要曲線を推定
  3. シミュレーション分析によって、「利益を最大にするような価格」を求める。

# ステップ1：Ziprecruiter.com加入の仕組みとデータ収集



- ▶ 月額価格について、ランダムな提示を行った。(ランダム化比較試験)
  - ▶ 対照群：99ドル (実験前の価格)
  - ▶ 処置群：19ドルから399ドルまでの価格をランダムに提示した。
- ▶ 提示された価格によって、企業の加入パターンがどのように変わるかを分析した。

# 実験結果：加入率(左)と収入曲線(右)



# ステップ 2 & 3 : 価格をいくらにするか？ → 経済学の出番！

- ▶ 企業は以下の利潤を最大にするように価格 $p$ を決めたい

$$(\text{利潤}) = (p - c)D(p)$$

需要関数：価格 $p$ の時の予測加入人数

↑  
サービスのコスト。  
Webサービスなのでゼロと考える。

- ▶ 需要関数 $D(p)$ がわかれば、利潤最大化する価格がシミュレーションでわかる！！
- ▶ 価格実験から得られたデータを用いて、離散選択モデルの推定を行った。
  - ▶ 「加入する」か「加入しない」という2択
  - ▶ 意思決定において「価格」が重要な要因として入ってくる。


# モデルに基づく最適価格

▶ 元の価格：99ドル

▶ 平均加入率は**25%**、潜在的な消費者一人あたりの期待収入は**25ドル**

▶ 一時点における利潤最大化の価格：**327ドル**

▶ 平均加入率は**12%**、潜在的な消費者一人あたりの期待収入は**39ドル**



収益が40%増加！！

▶ ただし、価格が高いと継続率は下がる。

▶ 中長期的な継続率を加味した収入を計算すると、**261ドルが最適**という結果が得られた。

▶ 実験後、実際に経営陣は249ドルへと値上げを行った。

# まとめ：経済学の実務活用

# ケース：バッファローのnasne（ネットワークレコーダー）

割愛します。

# ケース：バッファローのnasneの価格付け

- ▶ 東京大学エコノミックコンサルティング(UTEcon)によるnasneの価格付けプロジェクト

割愛します。(記事参照)

参考記事：<https://www.buffalo.jp/contents/topics/special/nasne-pricing-int/>

## ——それまで、販売価格はどのように決まっていたのですか？

---

**松崎：** 新商品の価格には、ユーザーニーズやブランド価値、発売タイミングなどさまざまな要素を反映する必要がありました。しかし、従来のバッファローの商品価格決定は、どちらかというところ「KKD（経験、勘、度胸）」によるもの。

「nasne」は多くのユーザーが愛着を持つブランドで、実際にいくらが適切なのかという点で、明確な理論や根拠があるわけではなく、判断に困る状態でした。そこで白羽の矢が立ったのがデータサイエンスでした。過去のデータや数値から、新商品の価格を科学的に算出できないか、UTEconに相談しました。

## ——発売後の反響について教えてください

---

**松崎：** バッファロー版「nasne」は2021年3月25日からamazon内バッファロー公式ストアで予約開始となりました。初回分は4時間以内で完売し、2次・3次入荷分も24時間以内に完売と想像以上の反響が得られました。これほどの反響は過去にほとんど経験したことがなく、販売が好調な反面、お客様をお待たせする結果となり、ご迷惑をおかけしています。

反響の大きさは、29,800円（税込）という販売価格がお客様のスイートスポットに合致した結果と考えます。そうした価格を科学的な分析の裏付けによって数値化してもらったことに大きな意味があると思います。

# おわりに

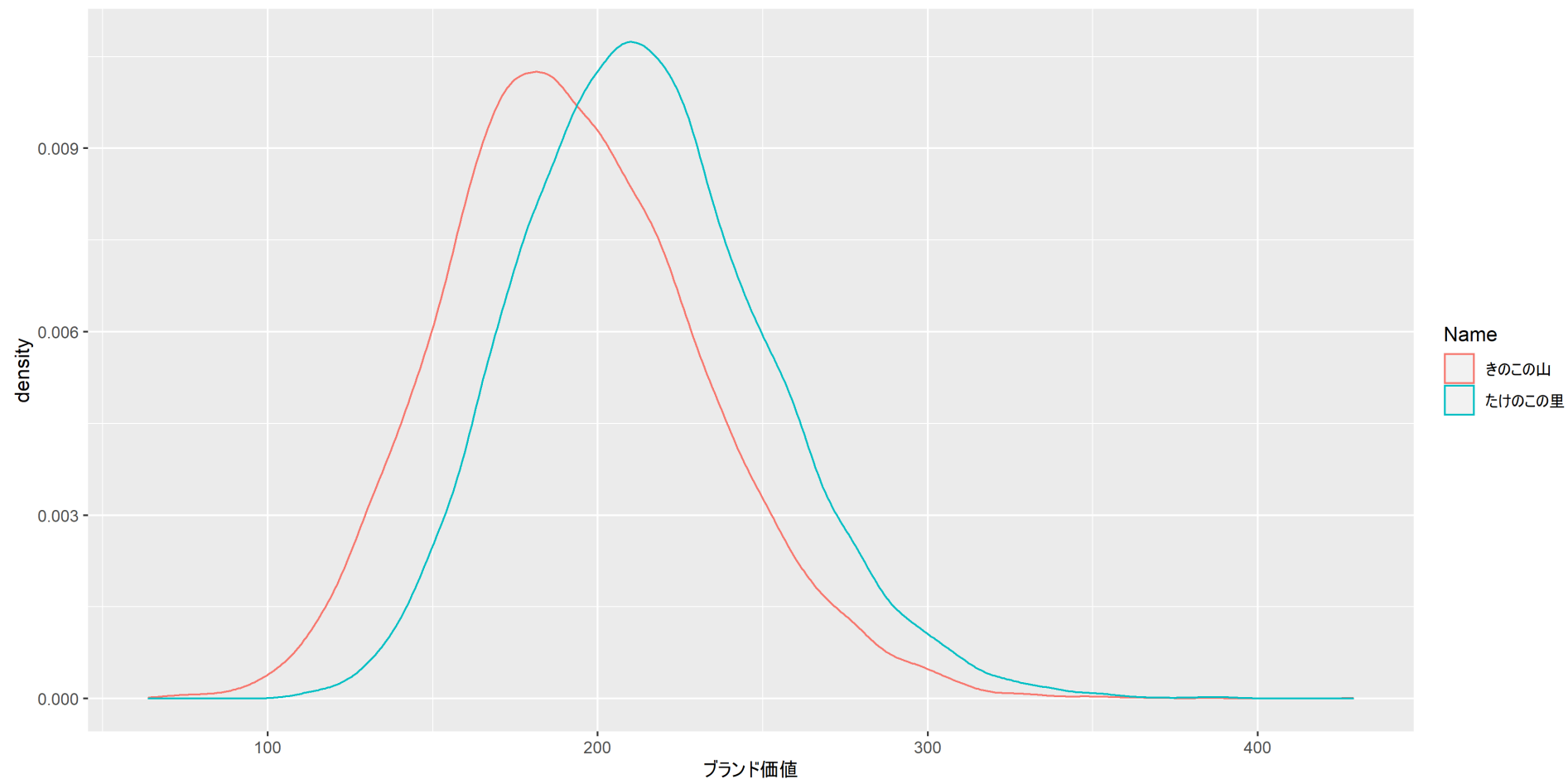
- ▶ 経済学は**総合格闘技**！！
  - ▶ 社会・ビジネスの問題への興味関心と背景理解
  - ▶ 数理モデルという現実を見通す「レンズ」
  - ▶ 統計学・データ分析を活用して、現実への含意を導き出す！
  
- ▶ 「格闘家」として、**幅広い活躍**の場所！！
  - ▶ 民間企業：データサイエンティスト、エコノミスト、
  - ▶ 公共部門：省庁・自治体でのデータに基づいた政策立案、国際機関
  
- ▶ ぜひとも政治経済学部／経済学科を進路の一つとして考えてみてください！

**以下、補足資料**

# きのこ・たけのこのブランド価値の分布

▶ きのこ・たけのこのそれぞれから得られる満足度を金銭換算した分布

▶ きのこの平均：192円、たけのこの平均：214円



# ひとまずのまとめ & 次のステップ

- ▶ きのこ・たけのこの戦争を離散選択モデルで分析することで
  - ▶ きのこ派・たけのこ派の比率がわかった
  - ▶ 需要曲線が得られた→価格変化に対する需要予測ができる
  - ▶ ブランド価値を計算できた
- ▶ このような調査は**コンジョイント分析**と呼ばれ、マーケティングでしばしば活用。
- ▶ 一方で、本分析はあくまで「仮想的な質問」に基づくもの。
- ▶ 次の例：実際の選択行動のデータを用いた分析！！

# 実際どのようなデータを用いるのか？いくつかの例

## 1. サーベイ・アンケート調査

- ▶ 消費者にどの製品を購入したかを調査する。

## 2. POSデータ

- ▶ Point-of-salesデータ。お店のレジを通った買い物記録
- ▶ どの製品が、いくらで、どれだけ販売されたかが記録される。

## 3. コンジョイント分析

- ▶ 仮想的な質問・選択肢を提示し、その中でどれを選択するかを質問する。

## 4. オンライン小売企業の購買履歴

- ▶ Amazonは消費者の詳細な購買履歴・情報を保有している！

# おまけ：なぜ経済学において数学を使うのか？

## ▶ 理由1：一連の議論を曖昧さなく表現するため

- ▶ 「一番満足が得られる製品を購入する」 -> 最も効用の値が高い製品を選ぶ
- ▶ 「価格よりも性能を気にする」 -> 価格の係数と性能の係数の大小関係で決まる

## ▶ 理由2：データ分析に落とし込むため

- ▶ プログラミングによってモデルをコンピューターで実装する。
- ▶ データと照らし合わせることでパラメタを推定する
- ▶ 推定したモデルを用いたシミュレーション分析
  - ▶ 例：価格が変化したら、人々の選択はどのように変化するか？ -> 需要曲線

## ▶ これから紹介する事例の背後には、プログラミング・データ分析が潜んでいます。

# モデルに基づく最適価格

## ▶ 元の価格：99ドル

- ▶ 平均加入率は25%、潜在的な消費者一人あたりの期待収入は25ドル

## ▶ 一時点における利潤最大化の価格：327ドル

- ▶ 平均加入率は12%、潜在的な消費者一人あたりの期待収入は39ドル

Decision-Theoretic Pricing Structures	Price	Expected Conversion Rate		Expected Revenue (\$ per Consumer )	
		Mean	95% Cred. Int.	Mean	95% Cred. Int.
Control	\$99	0.25	(0.23,0.28)	25.09	(23.02,27.48)
Uniform	\$327	0.12	(0.1,0.14)	39.01	(31.9,46.58)

## ▶ ただし、価格が高いと継続率は下がる。

- ▶ 中長期的な継続率を加味した収入を計算すると、261ドルが最適という結果が得られた。

# ランダム化比較試験とは？

- ▶ ある施策・介入の効果を計測するための統計学的手法
- ▶ 手順：
  1. 何らからしらの施策・介入を決める。例：値上げ
  2. 対象となる人々をランダムに処置群・対照群に分ける
    - ▶ 処置群：値上げした人たち
    - ▶ 対照群：元の価格のままの人たち
  3. 処置群・対照群を比較することで、施策の効果を推定することができる！
- ▶ 医学における薬・ワクチンなどの治験で幅広く活用されてきた。
- ▶ 昨今ではビジネス・社会科学でも活躍している。

# 経済学 & データ分析を実務でどのように活用できるか？

## ▶ 民間・公共部門双方での活躍！

- ▶ 参考：経済セミナー2020年4・5月号「鼎談：経済学でキャリアを創る」<https://www.web-nippy.jp/18310/>

## ▶ 公共部門：

- ▶ より良い規制・政策の実行のためのデータ分析 & 経済分析
- ▶ 国際機関における専門職

## ▶ 今日は民間企業での活躍を中心にお話します。

# 民間：データサイエンティスト、エコノミック・コンサルタント

## ▶ テック企業におけるデータサイエンティスト

- ▶ GAFA (Google, Amazon, Facebook, Apple), サイバーエージェント、アリババ、etc
- ▶ 例：価格づけのための需要・価格弾力性を推定、施策の評価(A/Bテスト)
- ▶ Athey and Luca (2018) “Economists (and Economics) in Tech Companies” 日本語による要約  
<http://www.dhbr.net/articles/-/5828>

## ▶ エコノミック・コンサルティング

- ▶ 企業経営における各種問題について経済分析を提供
- ▶ 例：新製品の価格戦略
- ▶ NERA、アリックスパートナーズ、東京大学エコノミックコンサルティング (UTEcon) など。