

2025.04.17

BinN スタートアップゼミ #3

行動モデルの基礎 / MNL, NL, RL

東京大学大学院 社会基盤学専攻

交通・都市・国土学研究室

修士2年 松永隆宏

行動のモデリングとは？

1

人々の様々な意思決定行動のメカニズムを数式で表現する

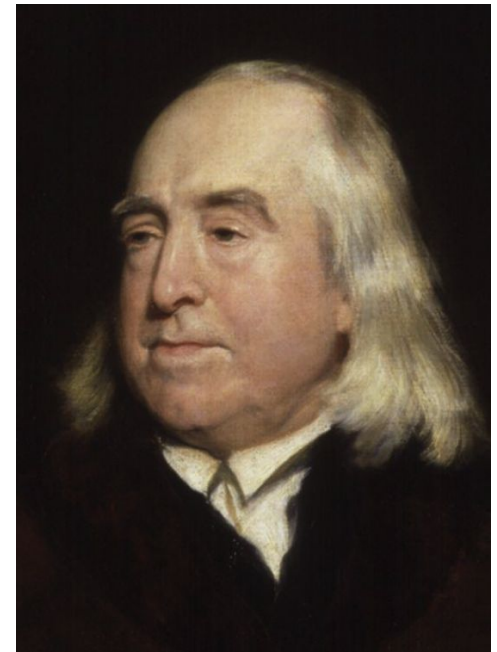
何に使う？

都市や交通の政策評価やシミュレーションなど

- ・需要予測(運賃を変えると利用がどの程度変化するか)
- ・施設配置(観光案内所をどこに置くか)
- ・道路容量配分(渋滞が起きない程度に車道を減らしたい)

どういう方針で作る？

効用最大化理論「人は効用を最大化するような(＝合理的な)選択を行う」をもとに、観測データからモデルを推定する



ジェレミ・ベンサム
(1748-1832)

行動のモデリングとは？

2

複数の候補を評価基準をもとに 比較し、**最良の選択**をする(はず)

離散選択

Ex) 鉄道, 車, ...

効用関数

Ex) 費用, アクセス, ...

効用最大化理論

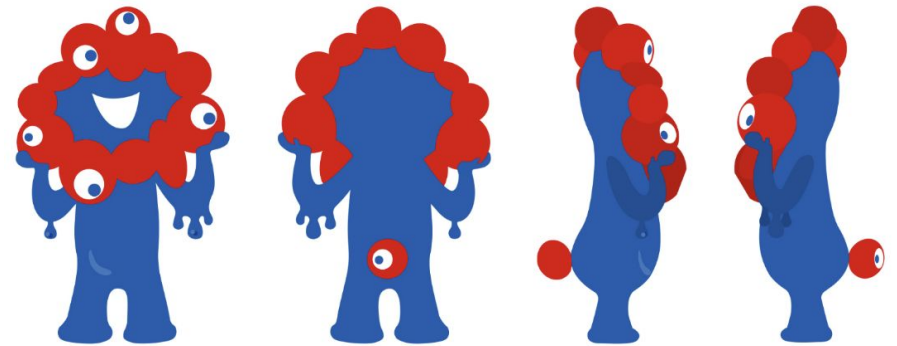
From) ミクロ経済学

ランダム効用

万博に行こう. 交通手段をどれにする？

	飛行機	新幹線	車	徒歩
時間	2h	2.5h	5h	20day
金額	15,000	10,000	8,000	?
ポイント	有	無	無	無

それぞれの手段の満足度＝効用は？



意思決定者は**効用(利得)**を最大にする選択肢を選ぶ(効用最大化理論)

- ✓ 満足度のこと. 慣用的に「効用」と言う. 英語だとUtility
- ✓ 要素ごとの重み付き線形和が基本

$$U_{\text{car}} = \beta_{\text{time}} x_{\text{time,car}} + \beta_{\text{cost}} x_{\text{cost,car}} \quad \text{確定項}$$

これをそのまま表すと, $Choice = \max\{U_a, U_b, U_c, \dots\}$

- ✓ 今の形だと「確定的」= 変動なし.
またtime, cost以外の要素の影響を捨象→一般化のため「確率的」に扱う

よって、**誤差項**を入れて確率変数として定義する(ランダム効用化)

$$U_{\text{car}} = \underbrace{\beta_{\text{time}}x_{\text{time,car}} + \beta_{\text{cost}}x_{\text{cost,car}}}_{\text{確定項}} + \underbrace{\epsilon_{\text{car}}}_{\text{誤差項}} = \underbrace{V_{\text{car}}}_{\text{確定項}} + \underbrace{\epsilon_{\text{car}}}_{\text{誤差項}}$$

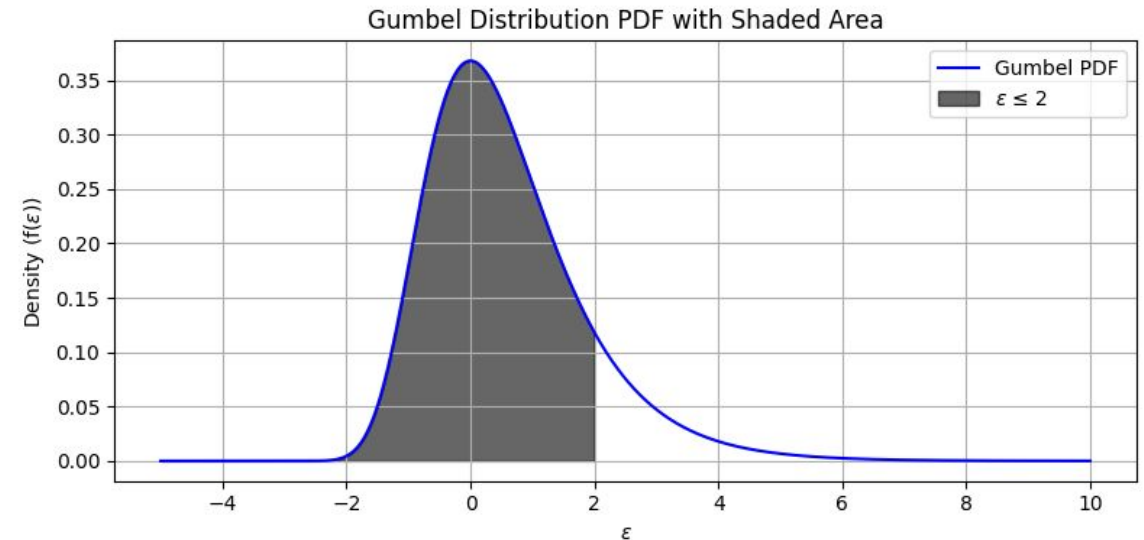
- ✓ 確定項に入っていない要素の影響を分布で表す(ノイズみたいな扱い)

確率変数 U に応じて確率的に選択： $\{P_a, P_b, P_c, \dots\}$.

$$P(i) = \Pr(U_i > U_j \text{ for any } j(\neq i)) = \Pr(U_i > \max(U_j))$$

✅ (i の選択確率) = (確率変数 U_i が全部の選択肢の U_j と比較して最大になる確率)

$\Pr(i)$ を表したい



✅ 累積密度関数が定式化できれば選択確率が書けたことになる！

✅ 累積密度関数は明らかに ϵ の分布形状に依存する

正規分布だと:
$$P(\text{car}) = \int_{-\infty}^{\infty} d\varepsilon_{\text{car}} \int_{-\infty}^{V_{\text{car}} - V_{\text{walk}}} d\varepsilon_{\text{walk}} \int_{-\infty}^{V_{\text{car}} - V_{\text{train}}} d\varepsilon_{\text{train}} \phi(\varepsilon_{\text{walk}}, \varepsilon_{\text{car}}, \varepsilon_{\text{train}})$$

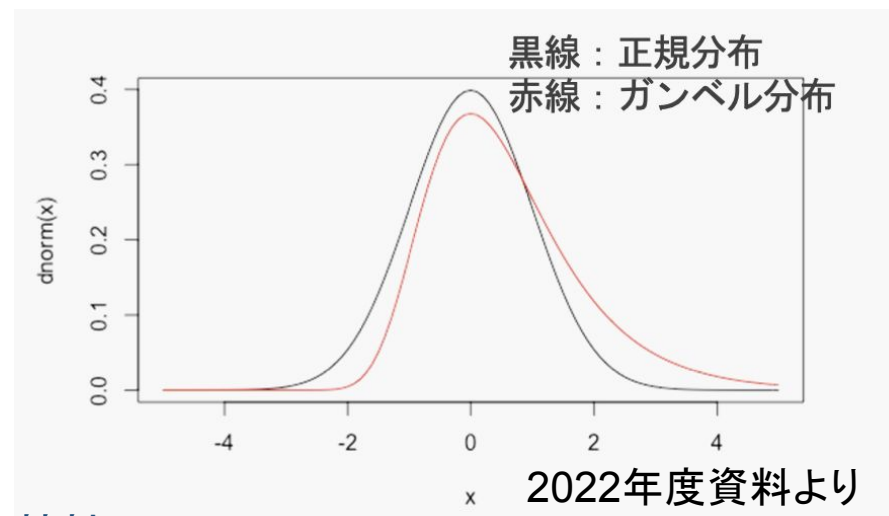
- ✓ 多重積分計算(初等関数で書けない; 閉形で書けない)が必要→計算コスト大
- ✓ ちなみにこれをプロビットモデルという

形状が似てる & F が閉形で書ける分布だと都合が良い→ **ガンベル分布 (type1)**

ガンベル分布だと:
$$\Pr(U_i > \max(U_j)) = \frac{\exp(\mu V_i)}{\sum_{j \in C} \exp(\mu V_j)}$$

- ✓ 閉形で書け, 計算が簡単 (通常 $\mu = 1$)
- ✓ これをロジットモデルという
→ 選択肢が2つならBinaryロジット,
3つ以上なら多項(Multinomial)ロジット=MNL

- ✓ ただし全選択肢の誤差項が独立(相関しない)! → 後述IIA特性



ある選択行動をMNLモデルで記述したい(モデルの選択)

→各パラメータの値(各説明変数の重み)を推定することが目標になる

🤔 パラメータはどうやって求める？ $U_{\text{car}} = \beta_{\text{time}}x_{\text{time,car}} + \beta_{\text{cost}}x_{\text{cost,car}} + \varepsilon_{\text{car}}$

👁️ パラメータ推定のことをモデル推定と呼ぶこともあるし、単に推定と呼ぶこともある

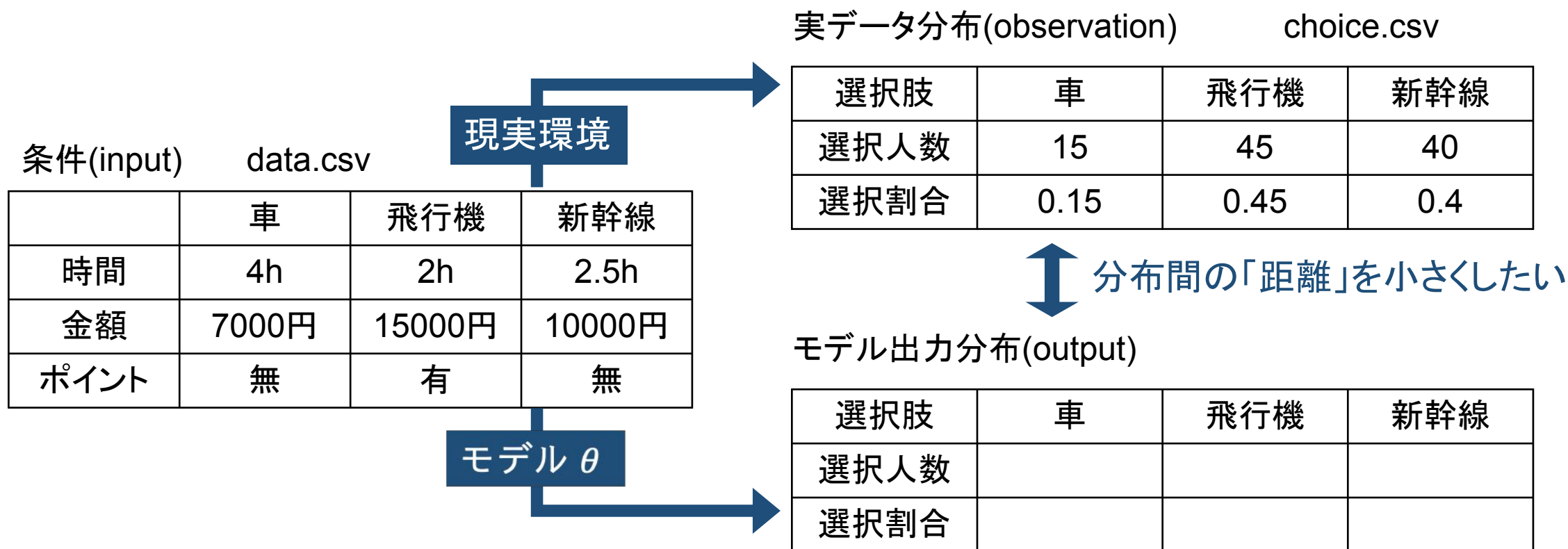
実際の行動データに合うようにモデル(パラメータ)を推定する

=実際の観測データ(環境出力)とモデル出力が近づくようにすれば良い

✅ 実データから推定できるのが行動モデルの強み

想い: 観測データとモデル出力を近づけたい

✓ 現実世界をより忠実に再現できるモデルが良いモデル(それはそう)



ここが θ で変わる→分布が近くなるような θ が望ましい

最尤推定:

モデルが実際の行動(観測データ)を予測(再現)する確率が最も高くなるようにパラメータを設定

モデルによる個人*i*の選択肢*y_i*の選択確率を*P(y_i; x_i)*とおく(*x_i*は説明変数).

全*i*について観測と同じ結果を再現する確率は,

$$\prod_i P(y_i; x_i) = \prod_i \frac{\exp(\mu V_i)}{\sum_{j \in C} \exp(\mu V_j)}$$

万博までの交通手段

	中澤	水谷	一山	伊藤	立川	
モデル出力	$P(y_{walk}; x_N)$	$P(y_{train}; x_M)$	$P(y_{train}; x_{IR})$	$P(y_{train}; x_{IA})$	$P(y_{train}; x_T)$	→これを再現できるか?
観測(正解)	徒歩	新幹線	新幹線	自転車	新幹線	→同時確率は $\prod_i P(y_i; x_i)$

これをパラメータ*θ*の関数と見て (=尤度関数)
→尤度関数*L(θ)*を最大化するパラメータ*θ*を求める

$$L(\theta) = \prod_i \frac{\exp(\mu V_i)}{\sum_{j \in C} \exp(\mu V_j)}$$

*L(θ)*は確率の積ゆえ, 値が小さくなりすぎてしまう
→対数尤度を最大化する問題に

$$LL(\theta) = \sum_i \log P(y_i; x_i)$$

確率：固定されたパラメータ θ →出てくるデータの分布

✓ このパラメタだとどんなデータが出るだろう？（通常の推定）

尤度：固定されたデータ x →それを生むパラメータの尤もらしさ

✓ このデータが出たってことはパラメタはどんな値だったんだろう？（逆推定）

🤔 今の問題に当てはめると

観測されたデータ x （と同じ出力）を生む，一番尤もらしいパラメータを求める
→尤度最大となるパラメータを採用

確率分布の距離(ダイバージェンス)はKLダイバージェンスで定義される

【定義】二つの確率分布 P, Q について, P に対する Q のKLダイバージェンスは

$$KL(P||Q) = \sum_y P(y) \log \frac{P(y)}{Q(y)}$$

真の分布 P に対して近似分布 Q を使うことによる情報損失に該当

観測分布 $\hat{p}$ に対してモデル分布 p_θ が**近ければ**良いから, **$KL(\hat{p}||p_\theta)$ の最小化問題**

つまり,

$$\theta = \operatorname{argmin}_{\theta} KL(\hat{p}||p_\theta) \quad KL(\hat{p}||p_\theta) = \sum_i \hat{p} \log \frac{\hat{p}}{p_\theta} = \sum_i (\hat{p} \log \hat{p} - \hat{p} \log p_\theta)$$

$\hat{p}$ (観測分布) はパラメタ θ によらないので,

$$\theta = \operatorname{argmax}_{\theta} \sum_i \log p_\theta \quad \checkmark \text{ 最尤法と同じ(対数尤度関数の最大化に他ならない)}$$

赤バス青バス問題

11

選択肢が青バスと自動車だけの場合

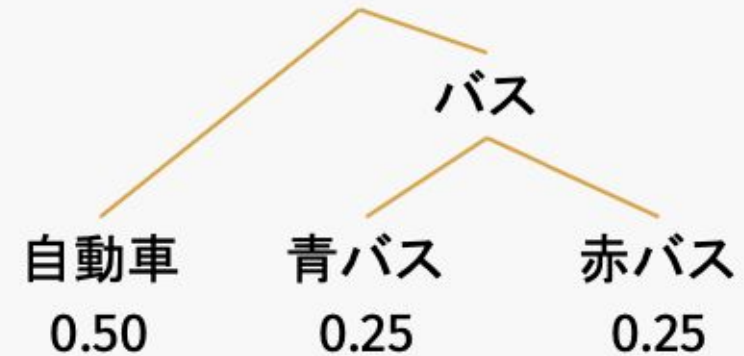


選択肢に赤バスが加わった場合



本当に？

本当はこう？



ロジットモデルでは、選択確率の比はその他の選択肢の影響を受けない

選択肢集合がAの時

$$P(car|A) = \frac{\exp(\mu V_{car})}{\exp(\mu V_{car}) + \exp(\mu V_{red})}$$

選択肢集合がBの時

$$P(car|B) = \frac{\exp(\mu V_{car})}{\exp(\mu V_{car}) + \exp(\mu V_{red}) + \exp(\mu V_{blue})}$$

$$\frac{P(red|A)}{P(car|A)} = \frac{\exp(\mu V_{red})}{\exp(\mu V_{car})} = \frac{P(red|B)}{P(car|B)}$$

つまり、青バスの有無に関わらず車と赤バスの選択確率の比は不変。



Independence of Irrelevant Alternatives: IIA特性

この性質は、**選択肢に類似性がある場合** に不合理を引き起こす。

(というか、上の定式化は選択肢の誤差項が独立である仮定のもとでしか成り立たない)

赤バス青バス問題の原因は..

「我々は赤バス・青バス・車を独立な選択肢として見ていない. 実際には公共交通機関と車, という上位選択があり, 公共交通機関を選んだら初めて赤バスと青バスを選択する」

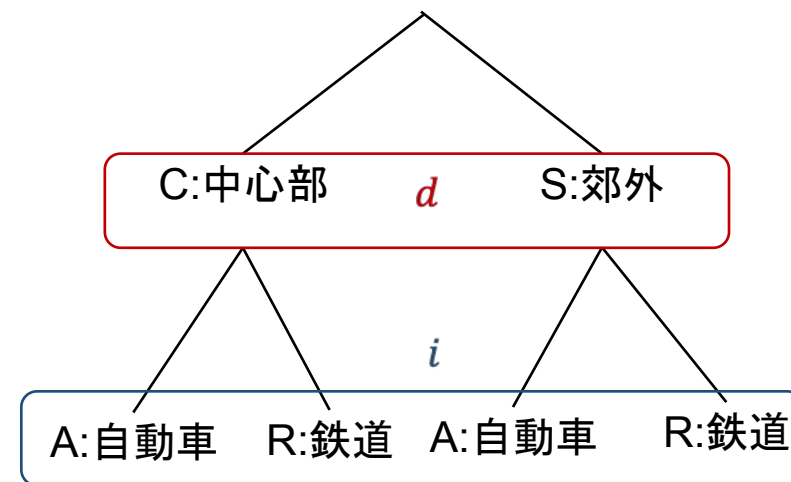
という構造を見落としているから.

→階層的な意思決定構造(ネスト)を導入したモデル:Nested Logit

- ✓ 上の選択肢の(重みつき)効用がログサム変数として下流に伝播
- ✓ 選択肢にネスト構造を設けて相関に対応
- ✓ IIA特性を緩和したモデル

ex)目的地と交通手段の選択

$$U_{di} = V_d + V_i + V_{di} + \varepsilon_d + \varepsilon_{di}$$



V_d : 目的地選択に特有の確定項

V_i : 交通手段選択に特有の確定項

V_{di} : 目的地選択と手段選択の組み合わせで決まる確定項

ε_d : 目的地選択の誤差項. U_{di} の最大値がスケールパラメータ μ_d をもつガンベル分布になるような分布に従う

ε_{di} : 目的地選択の誤差項. スケールパラメータ μ をもつガンベル分布に従う

上位ネストの選択確率 $P(d)$ と d に応じた下位ネストの条件付き選択確率 $P(i|d)$ の積.

$$P(d, i) = P(i|d)P(d) \quad P(i|d) = \frac{\exp(\mu(V_i + V_{di}))}{\sum_{i'} \exp(\mu(V_{i'} + V_{di'}))}$$

$$P(d) = \frac{\exp(\mu_d(V_d + V'_d))}{\sum_{d'} \exp(\mu_{d'}(V_{d'} + V'_{d'}))} \quad \text{ただし } V'_d = \frac{1}{\mu} \sum_i \exp(\mu(V_i + V_{di}))$$

どちらも閉形式で計算可能

※詳細な計算は過年度資料(2022年度スタートアップ#2など)を参照

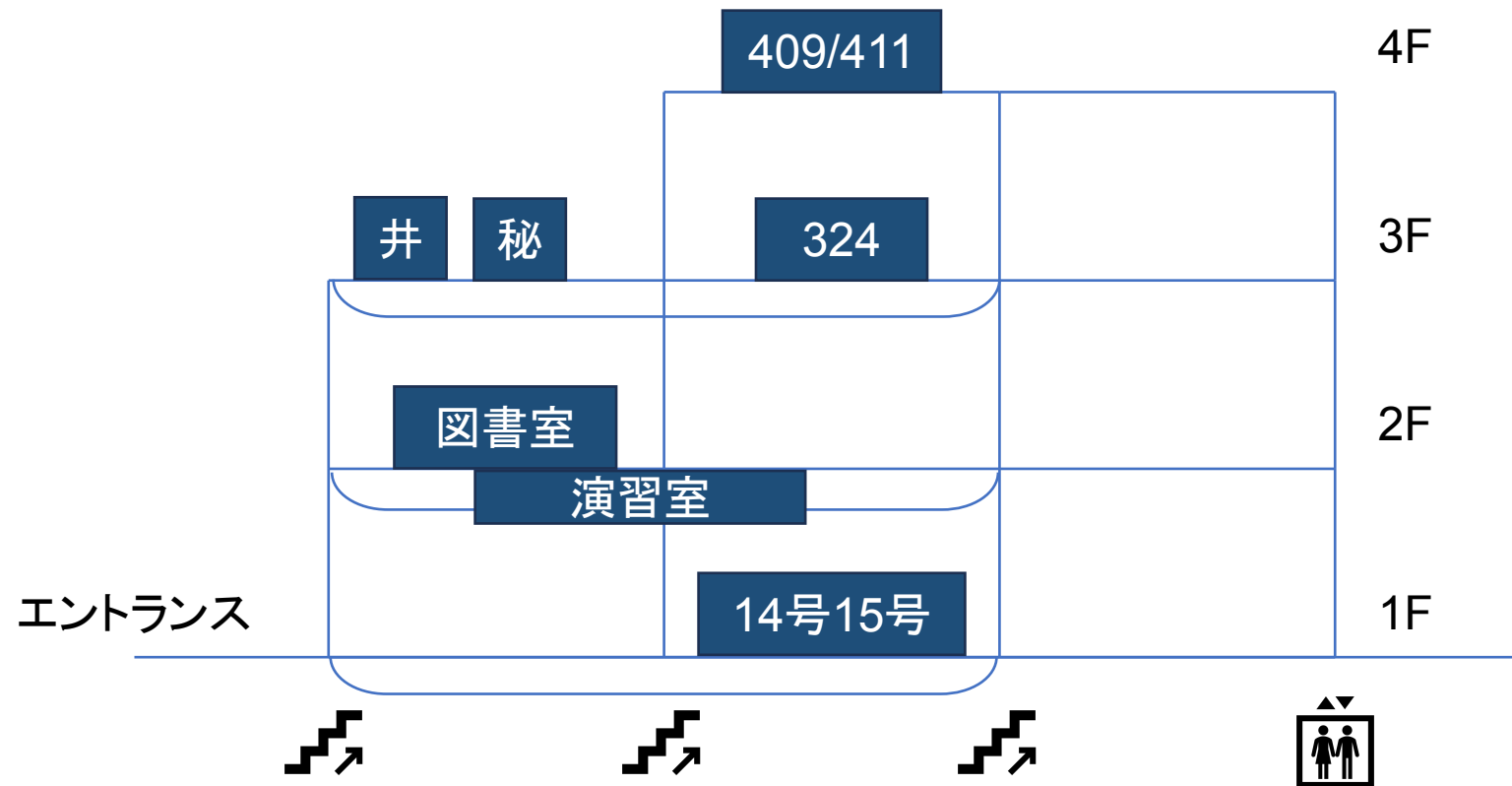
スケールパラメータは1つ固定する必要がある

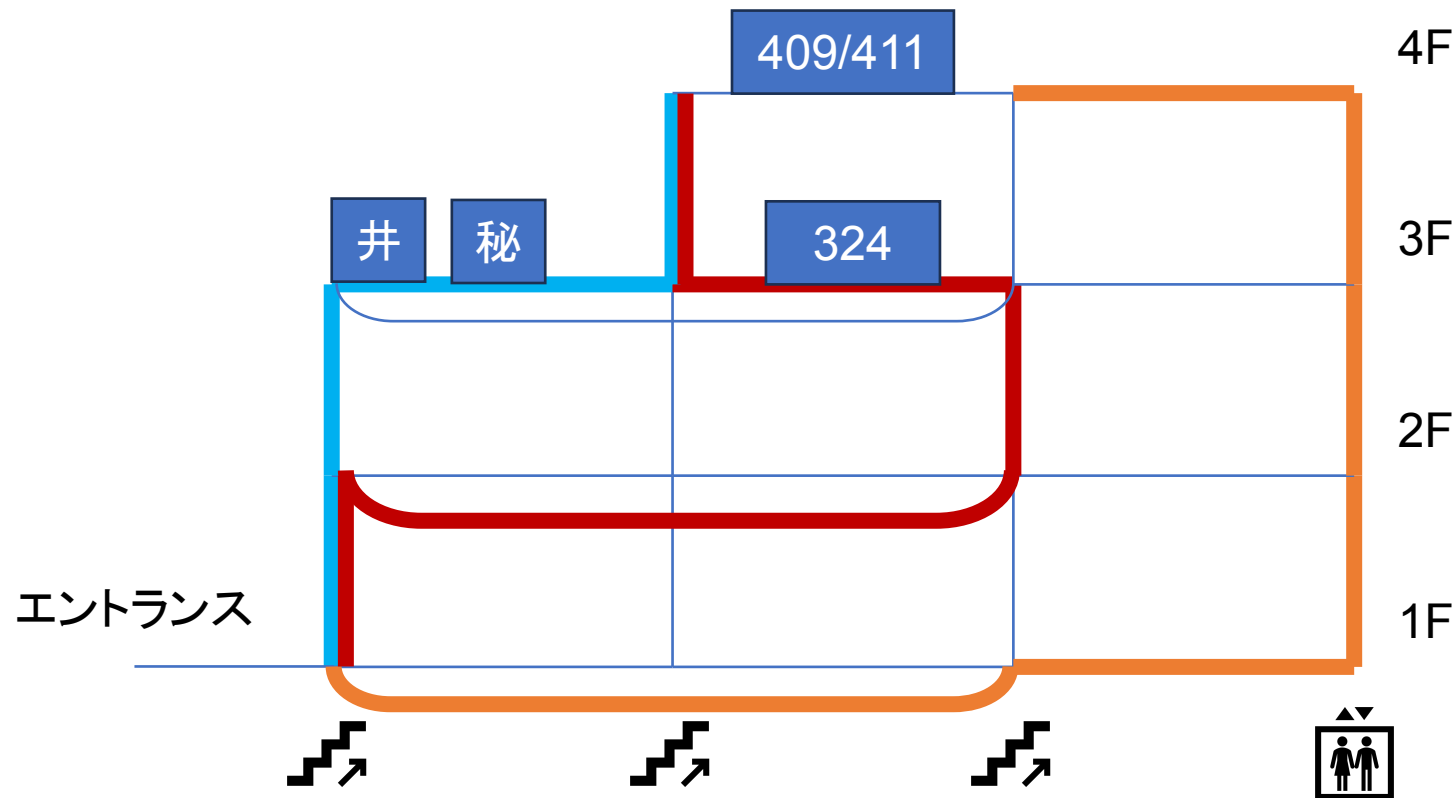
最下層のスケールパラメータを1に固定することが多い

スケールパラメータは上位のネストのものが小さくなる

$\mu > \mu_d$ となる

もしこのようにならなかった場合, ネスト構造の仮定が成り立たないので見直す必要がある





経路(path)ベース

選択枝: 経路
→ MNLモデルで記述可能

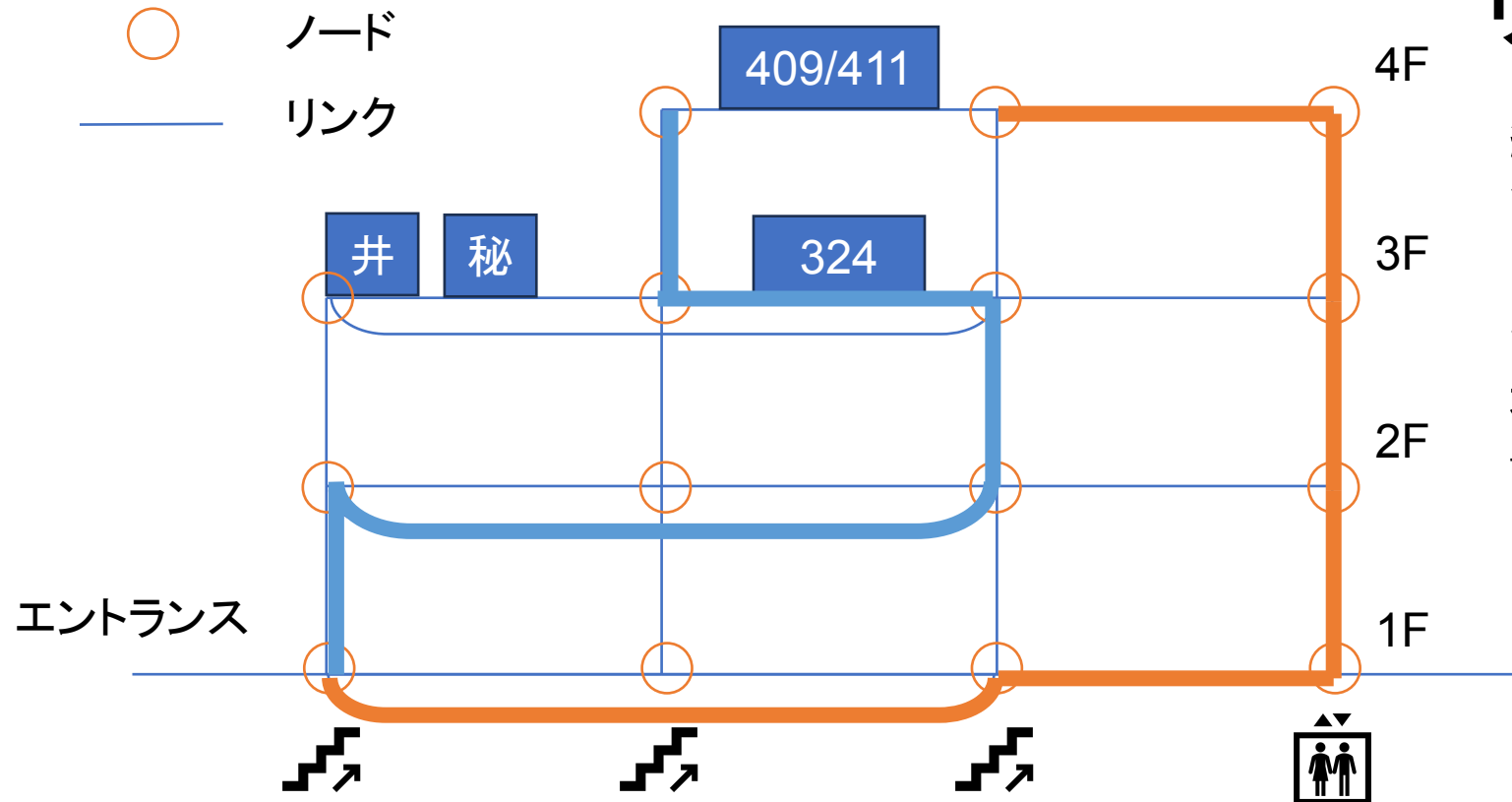
問題点

- 経路の羅列は不可能
(組み合わせ爆発)



解決方法

- サンプリング
 - 経路集合を定める
 - 最短経路, k番目
- 逐次的選択を仮定
 - 経路列挙は不要



リンクベース

経路選択を「逐次的なリンク選択」として捉える→経路列挙不要

各選択時点では...
選択肢: リンク
→MNLで記述可能

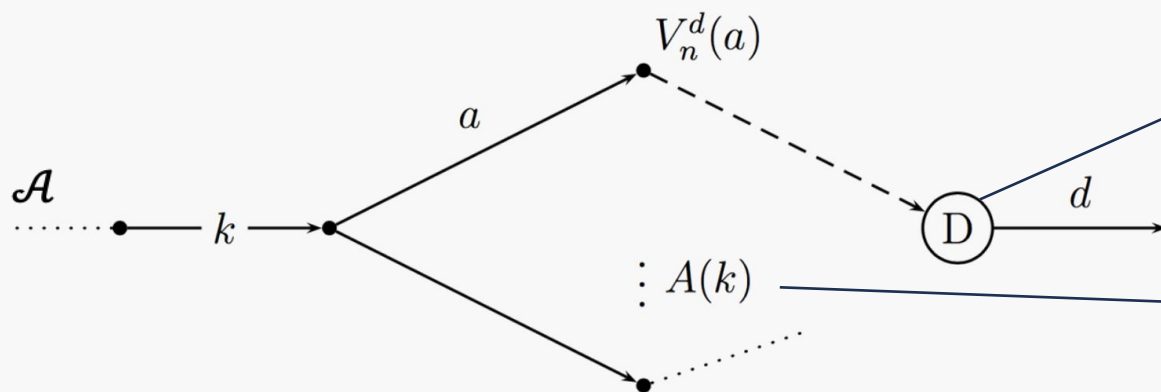
$$\text{効用} = \text{即時効用} + \text{期待効用}$$

すぐ得られる効用

将来的に得られそうな効用

有向連結グラフ（ネットワーク） $\mathcal{G} = (\mathcal{S}, \mathcal{A})$

- $\mathcal{S}$: ノード集合
- $\mathcal{A}$: リンク集合
リンク $k, a \in \mathcal{A}$



D 目的地ノード
 d ダミーリンク

$\mathcal{A}(k)$
リンク k の終点から出る
リンクの集合

効用関数

- 即時効用 $u_n(a|k) = v_n(a|k) + \mu \varepsilon_n(a)$
確定項 誤差項

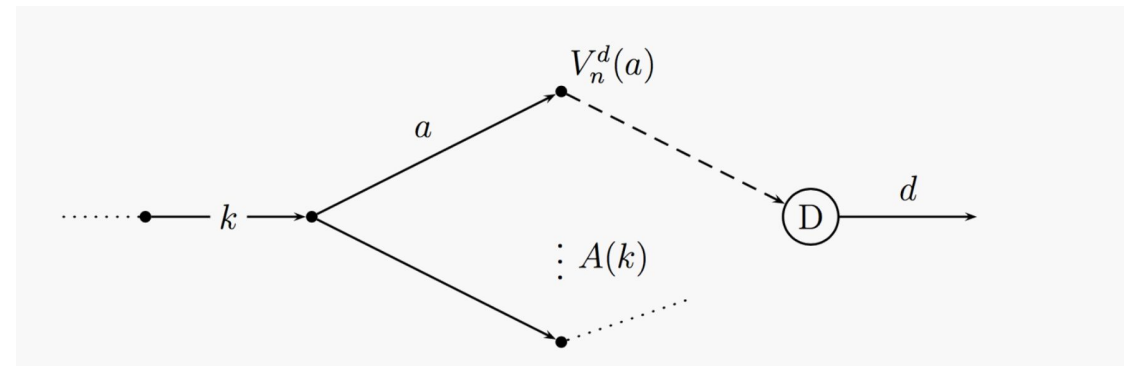
- 期待効用 $V_n^d(a)$ 旅行者 n がリンク k にいる状態で選択肢集合 $\mathcal{A}(k)$ からリンク a を選択する
※ $v_n(d|k) = 0$

マルコフ過程 (Markov process) に従うと仮定

未来の状態は現在の状態にのみ依存し、過去の履歴には依存しない確率過程

効用関数

- 即時効用 $u_n(a|k) = v_n(a|k) + \mu\varepsilon_n(a)$
- 期待効用 $V_n^d(a)$



リンク k を選んだ場合の目的地 d までの期待効用

$\mathcal{A}(k)$ の中から効用(即時効用+期待効用)が最大のものを選ぶ

価値関数

$$V_n^d(k) = E \left[\max_{a \in \mathcal{A}(k)} (v_n(a|k) + V_n^d(a) + \mu\varepsilon_n(a)) \right]$$

リンク a を選んだ場合の目的地 d までの期待効用

Bellman方程式

現時点の価値関数は、次の時点での価値関数に依存→**再帰的(Recursive)**

• 価値関数の求め方→Bellman方程式を解く

- 逆行列計算
- 後ろ向き帰納法+反復法

✓ 再帰的というと難しいが、高校数学の漸化式と同じ

選択確率の計算

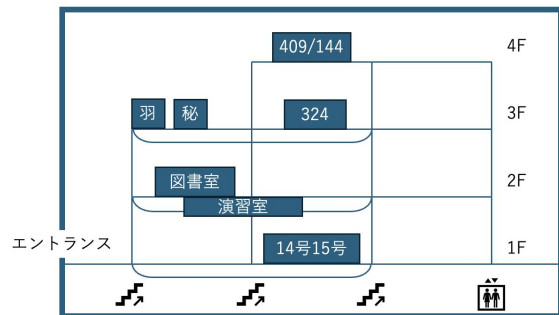
遷移確率: リンク a_j にいる状態でリンク a_{j+1} を選ぶ確率

※遷移自体は確定的

$$p(a_{j+1}|a_j) = \frac{e^{\frac{1}{\mu}\{v(a_{j+1}|a_j)+V(a_{j+1})\}}}{\sum_{a'_{j+1} \in \mathcal{A}(a_j)} e^{\frac{1}{\mu}\{v(a'_{j+1}|a_j)+V(a'_{j+1})\}}}$$

経路の選択確率: 経路 $\sigma = [a_1, \dots, a_J]$ が選択される確率

$$P(\sigma = [a_1, \dots, a_J]) = \prod_{j=1}^{J-1} p(a_{j+1}|a_j)$$



ネットワークの空間知識を
経験によって獲得している状態



空間知識の欠如・
曖昧さ

path...進路・生き方



将来の不確実性

時間割引率 β を導入

$$V_n^d(k) = E\left[\max_{a \in \mathcal{A}(k)} (v_n(a|k) + \beta V_n^d(a) + \mu \varepsilon_n(a))\right]$$

$$0 \leq \beta \leq 1$$

- $\beta = 1$: 即時効用と期待効用と同じ重みで評価
- $\beta = 0$: 即時効用のみに基づき近視眼的に評価

Causal Inference of Pedestrian Behavior with-without COVID-19 in High Density Shibuya Urban Spaces(Masuhashi,Hato,2022)

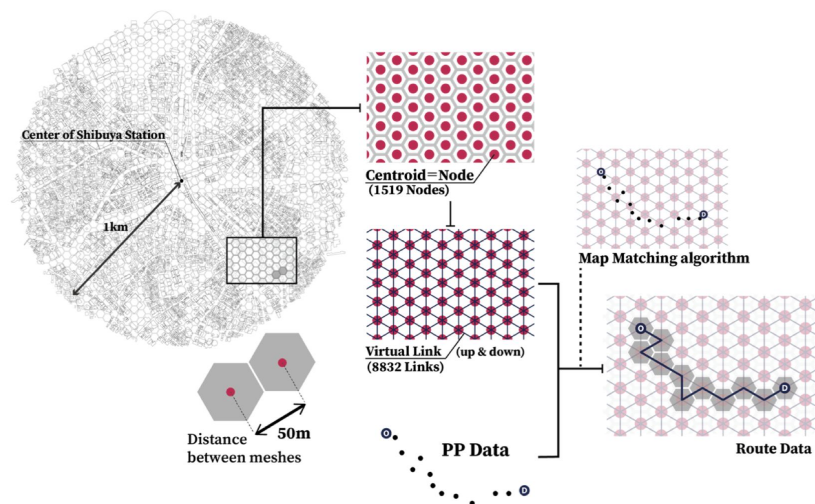
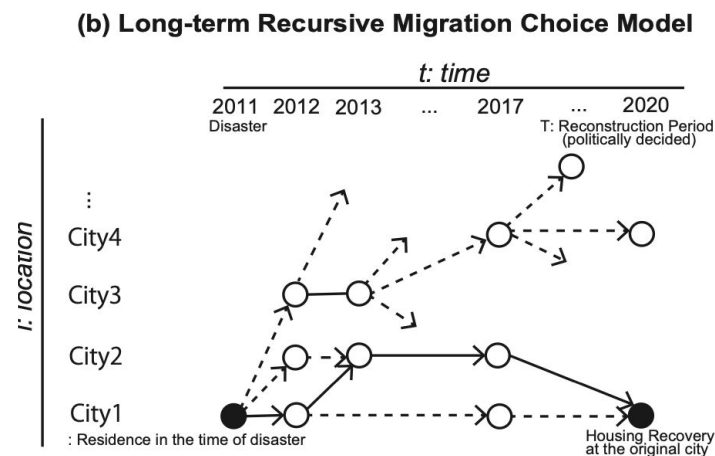


Figure 1 – Network configuration and route extraction procedures

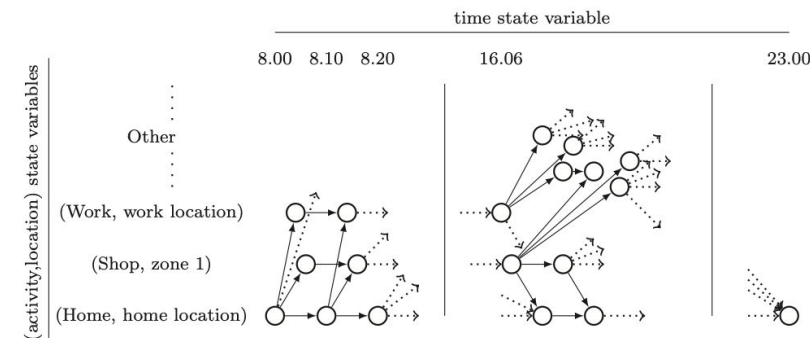
歩行者の回遊行動
ノード: メッシュのセントロイド
リンク: 歩行軌跡←マップマッチング

Modeling Long-term Recursive Migration Choices after Super Disasters: A Case Study of the Great East Japan Earthquake(Koseki,Hato,2021)



被災者の長期避難履歴
ノード: 避難先(市町村)←サンプリング
リンク: 自宅再建までの避難履歴

Capturing correlation with a mixed recursive logit model for activity-travel scheduling (Zimmerman et al.,2017)



1日のスケジュール
ノード: 状態(活動種類×場所)
リンク: 移動
※mixed RL

ポイント1:一意に定まるような定式化

- すべての効用関数に定数項を入れてはいけない
 - 無数に解が考えられてしまうため. 1つだけつけずに, 基準を作る

$$\begin{aligned}V_{\text{car}} &= \beta_{\text{car}} + \beta_{\text{time}}x_{\text{time,car}} + \beta_{\text{cost}}x_{\text{cost,car}} \\V_{\text{walk}} &= \beta_{\text{time}}x_{\text{time,walk}} + \beta_{\text{cost}}x_{\text{cost,walk}} \\V_{\text{train}} &= \beta_{\text{train}} + \beta_{\text{time}}x_{\text{time,train}} + \beta_{\text{cost}}x_{\text{cost,train}}\end{aligned}$$

$$P(\text{car}|A) = \frac{\exp(\mu V_{\text{car}})}{\exp(\mu V_{\text{car}}) + \exp(\mu V_{\text{red}})}$$

✓ 効用はあくまで相対的な指標
→基準がないと比較できない

ポイント2:説明変数の相関

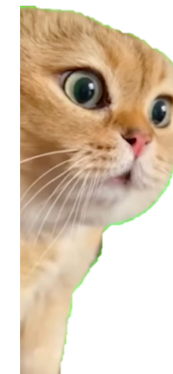
- 相関が強い説明変数を入れるとパラメータの誤差が大きくなる(ex. 移動時間と金額)
 - 独立な説明変数を入れる:そもそも選択確率の式(p.5)での仮定です
 - もしくは, 既存の変数を組み合わせて相関の弱い新しい説明変数を作る(ex.金額を「1時間あたりの金額」に変更する)



コードが回った！
モデル推定の結果が出た！！



- 推定値は統計的に意味がある？
- モデルは良いモデルだった？



So
what?



t検定

- 実際のデータによく当てはまっているモデル => 適合度
- 未来を正しく予測できるモデル => 予測性能
- 異なるデータにも適用できるモデル => 一般化可能性

尤度比

AIC, BIC, MDL

検証 (validation)

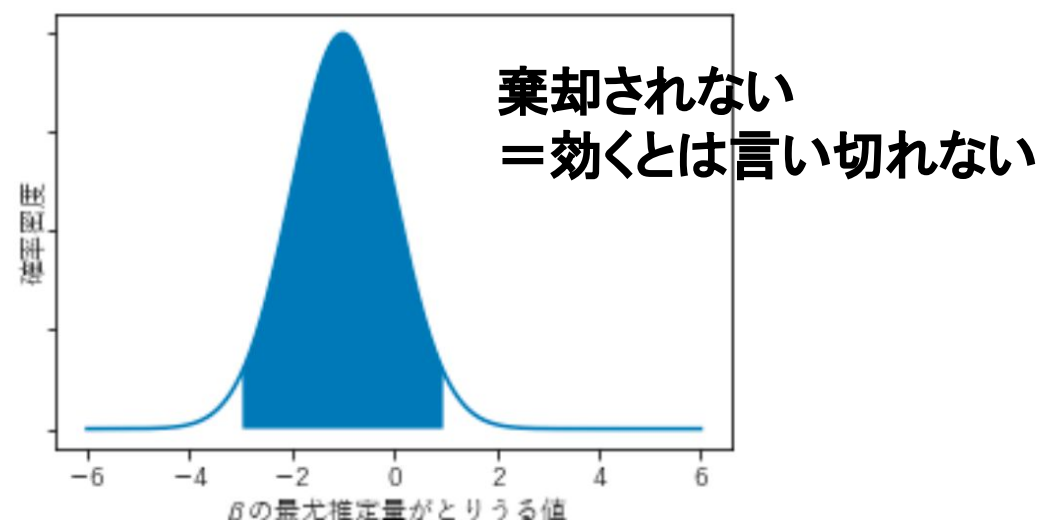
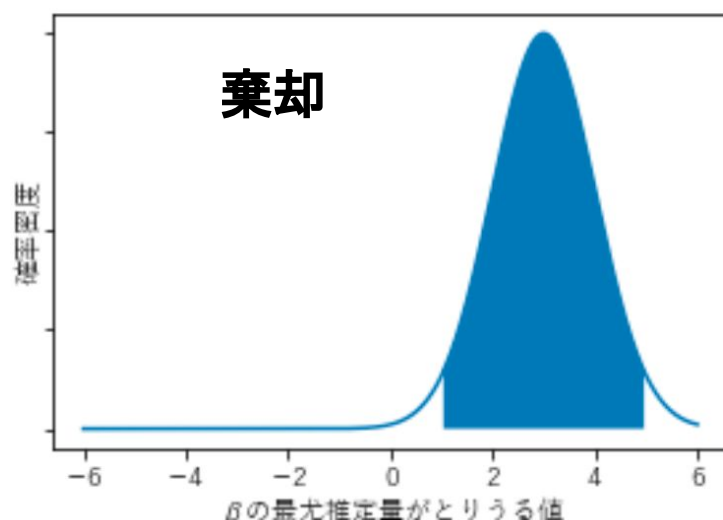
各パラメータ値が「有効」であるとは：

「パラメタの値が統計的に有意に0でない」を示せば良い

- ✓ そもそも前提として、基礎集計などで効きそうな(=0でない)変数が抽出されている(はず)
- ✓ 帰無仮説は「パラメタが0である」→これを棄却できればいい

方針： $\beta_i = 0$ であるならば、最尤推定値 $\beta_i^*(= \underset{\beta}{\operatorname{argmax}} L(\beta))$ は実現し得なかったことを示す

→最尤推定量 $\hat{\beta}_i$ (確率変数) において、最尤推定値 β_i^* が生じる確率は十分に小さいことを示す



最尤推定量 $\hat{\beta}_i$ は、ある分布に従う確率変数

→標準化 $\frac{\hat{\beta}_i - \mu}{\sigma}$ パラメーターの真値
未知...不偏推定量で代替

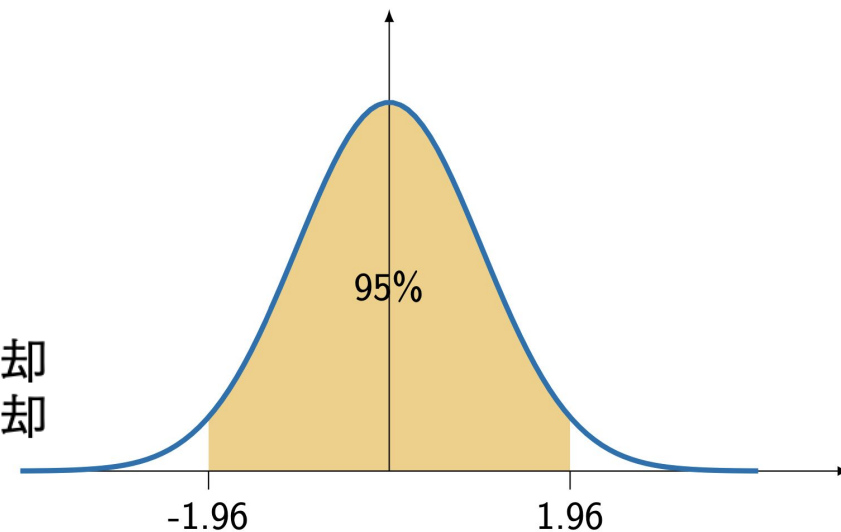
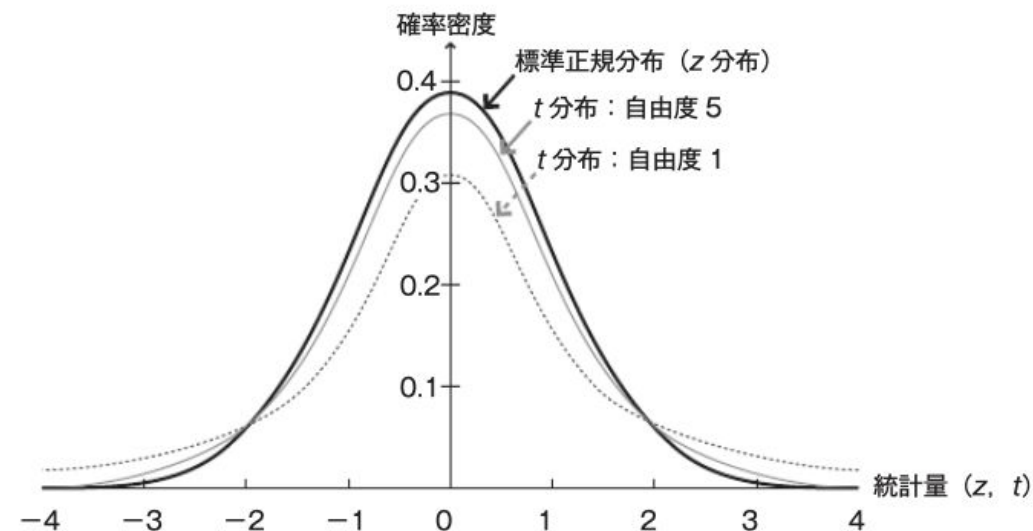
$$\text{検定量} : t = \frac{\beta_i^*}{s}$$

分布： n が小さいとき、自由度 $n - 1$ のt分布
 n が十分大きいとき、標準正規分布
(目安 $n = 30$)

t 値の絶対値が1.96より大きければ、有意水準5%で帰無仮説は棄却
 t 値の絶対値が2.58より大きければ、有意水準1%で帰無仮説は棄却

※ s は推定では $J(\hat{\beta})^{-1}(J(\hat{\beta}) := -\nabla \nabla^T LL(\hat{\beta}))$ で計算

2025.04.17



ゴール: モデルがどれだけデータにフィットしているかを調べる

McFaddenの擬似決定係数

尤度比 (likelihood ratio)

$$\rho^2 = 1 - \frac{LL(\hat{\beta})}{LL(0)}$$

$L(\hat{\beta})$ 最大尤度/最終尤度

$L(0)$ 無情報モデル(Null Model)尤度/初期尤度 全てのパラメーターが0であるときの対数の値

- 推定パラメーターが尤度を全く改善していないとき

$$L(\hat{\beta}) = L(0) \rightarrow \rho = 0 \text{ [下限]}$$

- 推定パラメーターがサンプルを完璧に予測しているとき

$$L(\hat{\beta}) = 1 \rightarrow \rho = 1 \text{ [上限]}$$

欠点: 説明変数を増やせば必ず増加する (← $L(\hat{\beta})$ が大きくなる)

自由度調整済み尤度比 (adjusted-likelihood ratio)

$$\rho^2 = 1 - \frac{LL(\hat{\beta}) - h}{LL(0)}$$

h : パラメーター数

演習課題

村山

課題1: 交通手段選択モデル尤度比コンペ

1. ensyu.csvデータについて基礎集計を行ってください
2. 交通手段選択における効用関数を設定し、パラメータ推定を行ってください
 - MNLについては推定用のpythonコードを用意しています（一部穴埋めあり）
 - 推定結果は分かりやすくまとめましょう
 - 推定結果からわかった行動に関する示唆をまとめましょう
 - 尤度比がなるべく大きくなるような工夫をしてみましょう

余裕があれば

- NLなど他のモデルで推定してみる（NLはコードを用意しましたが、そのままではうまくいかないのでネスト構造を工夫してみてください）
- 推定パラメータを用いて政策を行った際に起こる変化をシミュレーションして述べる

課題2：選択問題の設定

みなさんの研究課題やそれに隣接する範囲でどのような選択問題・意思決定問題を扱えるかを考えてください

- 研究テーマが決まっていない場合は、自分の興味があるトピックに関して考えてください
- 具体的な定式化を求めるものではないので、ラフに考えてください

考えるヒント

- 何を選択する問題か？
 - 目的地、交通手段、経路、居住地、出発時刻など
- 意思決定者は誰か？
 - 個人、世帯など
- 選択肢集合は何か？
 - どのような選択肢を認知していると考えべきか？
- 意思決定のルールは何か？
 - 効用最大化、後悔最小化など
- どのようなモデルが使えるか？
 - MNLやNLで対処できるか？
 - できない場合、どのような点が課題となるか？
- 必要なデータは何か？
 - PT、PP、BLE、衛星画像、動画など

課題2：選択問題の設定

参考に過去の卒論・修論で取り扱われた選択問題についていくつかまとめました

項目	白井修論	門坂卒論	村山卒論	松永卒論	加藤卒論
何を選択するか	居住地	目的地	出発時刻	三次元経路	経路、土地状態
意思決定者	個人	個人	世帯構成員	個人	回遊者、土地利用者
選択肢集合	滞在or移動 下位で移動先	探索続行or停止 下位で目的地	出発時刻 (連続)	経路	経路、土地利用用途
モデル	DDCM	NL	SUR (+ノンパラメ トリックベイズ)	RL (+観測モデル)	AIRL
データ	Web調査データ	アプリログデータ	PTデータ	BLEデータ、カメラデータ	PPデータ、土地データ

参考

- MNLについては昨年度修了された林さんの資料が参考になります
(ドライブに上げてます)
- NLについて
<https://bin.t.u-tokyo.ac.jp/kaken/pdf/nl.pdf>

どのモデルの推定結果かをわかりやすく！

5-1. 推定結果

表-6.1：推定結果(2005 年)

説明変数	処置群(再開発あり)		対照群(再開発なし)	
	パラメータ	t値	パラメータ	t値
距離 ($m_i/10^2$)	-5.09	-20.96 **	-5.01	-26.96 **
再開発床面積 ($m^2/10^5$)	-2.14	-2.33 **	-7.92	-33.21 **
JR駅出口ダミー	-	-	0.95	1.55
地下鉄駅出口ダミー	-	-	0.12	0.25
階段ダミー	0.17	1.05	-	-
時間割引率	1.00 ($\gamma = -18.442$)	-35.64 **	0.99 ($\gamma = -4.835$)	-20.91 **
サンプル数	211		395	
初期尤度	-430.35		-797.21	
最終尤度	-223.54		-438.99	
尤度比	0.48		0.45	
修正済み尤度比	0.47		0.44	
赤池情報量基準AIC	455.07		887.97	

†10%有意 *5%有意 **1%有意

値は少数第2桁で丸める

統計的に有意なパラメーターは、t値の右上に記をつけて示す

説明変数は、表を見るだけで意味がわかるように！
単位も忘れずに
(×コードでのラベル)

基本の5点セット
その他の指標はモデルに応じて

増橋佳菜,羽藤英二(2023),「位置データを用いた渋谷の都市更新に伴う歩行者行動の変容分析」, 都市計画論文集, Vol.58, No.3, pp.1140-1147