大学院講義　2025年度前期　交通経済学

多項ロジットモデル

離散選択モデルはじめの一歩

大澤 実（経済研究所）

前回の振り返り

- 人の選択行動に基づく交通需要予測を目指す
  - 集計による問題への対応
  - 多様な政策目的への対応可能性
- 交通行動の予測へ
  - 選択 = 効用最大化
  - 交通行動 = 離散選択
- 不完全性を表現するランダム効用モデル (RUM)

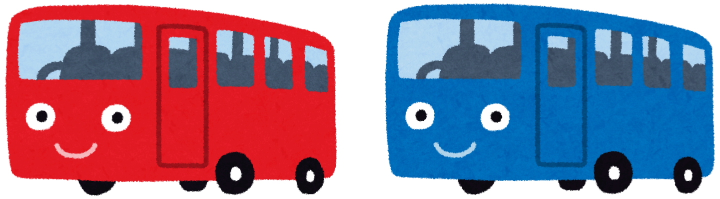
課題 1（再掲）

1. 自分の休日の（交通）行動の**選択ツリー**を具体的に書いてみよ.
   - 選択の**階層構造**を表現してみよう.
2. 各段階の選択に影響すると思われる要因を書き出してみよ.
3. それらの要因をどう直接的・間接的に計測すればよいか考えてみよ.
4. 表現されていない構造や捉えられていない要因がないか考えよ.

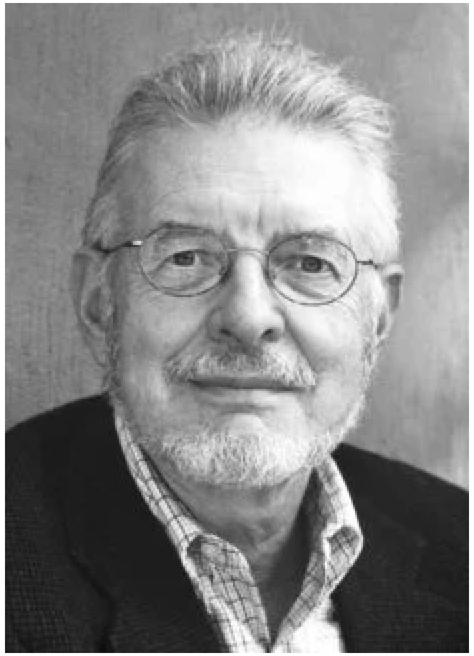
1~4 を再帰的に考えるのが行動のモデリング

今日のゴール

- ランダム効用モデル (RUM) の復習
- 最も基礎的な RUM である **多項ロジットモデル (MNL)** を知る
  - 選択確率の導出をフォローする
  - 選択確率の挙動を知る
  - MNLの限界：赤バス青バス問題



ランダム効用モデル



Daniel L. McFadden (1937–)

基本設定

- 各個人は「望ましさ」あるいは **効用** (utility) を最大化する
- 効用はその**選択肢**の特性・**個人**の特性によって異なる
- すなわち，個人 n がその**選択肢集合** A_n から選択肢 i を選ぶなら

$$U_{ni} \geq U_{nj} \quad \forall j \in A_n$$

- ランダム効用モデルでは，$\{U_{ni}\}$ が**確率的に定まる**と仮定する.
- このとき，個人 n が選択肢 i を選ぶ確率（**選択確率**）P_{ni}は

$$P_{ni} = \mathbb{P}\left(U_{ni} \geq U_{nj},\ \forall j \in A_n\right)$$

RUMが確率性を導入する動機

- **情報の不完全性**：
  個人も**分析者も**，選択肢やその属性を全て把握できるとは限らない．
- **限定合理性**：
  同じ情報でも認知や判断が異なる．気まぐれや習慣の影響も．
- **モデルの不完全性**：
  測定誤差・観測困難な要因・構造化できない複雑さが存在．

ランダム効用モデルは，これら**未観測の構造と変動**を確率的に取り込む

RUM の基本形

基本形の RUM では**加法的**な確率項を考える (**ARUM**: Additive RUM)

$$U_{ni} = V_{ni} + \varepsilon_{ni}$$

- U_{ni} ：個人 n にとっての選択肢 i の効用
- V_{ni} ：観測可能な決定論的効用（**分析者が設定する**）
- ε_{ni} ：観測不能な確率的ゆらぎ（**誤差項**）
  - error term, unobserved utility, stochastic utility, etc.

※ 乗法的確率項を考えてもよい（Multiplicative RUM; $U_{ni} = \varepsilon_{ni} V_{ni}$）

ARUM の導く選択確率

効用最大化行動という仮定のもとでは

$$P_{ni} = \mathbb{P}(U_{ni} \geq U_{nj}, \forall j \in A_n) = \mathbb{P}(V_{ni} + \varepsilon_{ni} \geq V_{nj} + \varepsilon_{nj}, \forall j \in A_n)$$

- 即ち，$\{\varepsilon_{ni}\}$ の**同時分布** (joint distribution) に従って確率が定まる
- **この分布は分析者が仮定する**
  $\Rightarrow$ 異なる分布の仮定は異なる選択確率を誘導（異なるRUM）

ARUM の導く選択確率の一般的性質

効用最大化行動という仮定のもとでは

$$P_{ni} = \mathbb{P}(U_{ni} \geq U_{nj}, \forall j \in A_n) = \mathbb{P}(V_{ni} + \varepsilon_{ni} \geq V_{nj} + \varepsilon_{nj}, \forall j \in A_n)$$

選択確率の一般的性質（**不変性**; invariance）：

- 効用の**並行移動**に対して不変

$$U_{ni} > U_{nj} \quad \Leftrightarrow \quad U_{ni} + V_0 > U_{nj} + V_0$$

- 確率的効用の**正のスケール変換**に対して不変

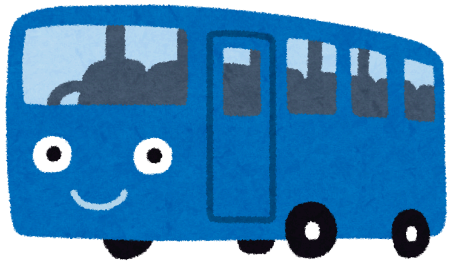
$$U_{ni} > U_{nj} \quad \Leftrightarrow \quad a \times U_{ni} > a \times U_{nj} \quad (a > 0)$$

注意

- RUM における “**選択確率**” とは「人がランダムに選ぶ」ことではなくモデル外の変動を含んだ **観測者視点の確率**.
- ある RUM による選択確率の表現は，誤差項の**分布への仮定の帰結**：同じデータに対し複数の RUM を適用可能（現象理解ツールでもある）
- **効用の絶対値には意味がない**
  ARUMでは **効用差** が選択確率を決める.
- **観測されるのはあくまで選択**
  潜在的な効用（連続量）→ 選択（離散）の整合的な対応がモデル.

多項ロジットモデル

Multinomial Logit (MNL) Model



多項ロジットモデル

- $\{\varepsilon_{ni}\}$ の分布として **Gumbel 分布** を仮定
- 全ての i について**独立**かつ**同じ分布**に従う
  (Independently Identically Distributed; **i.i.d.**)

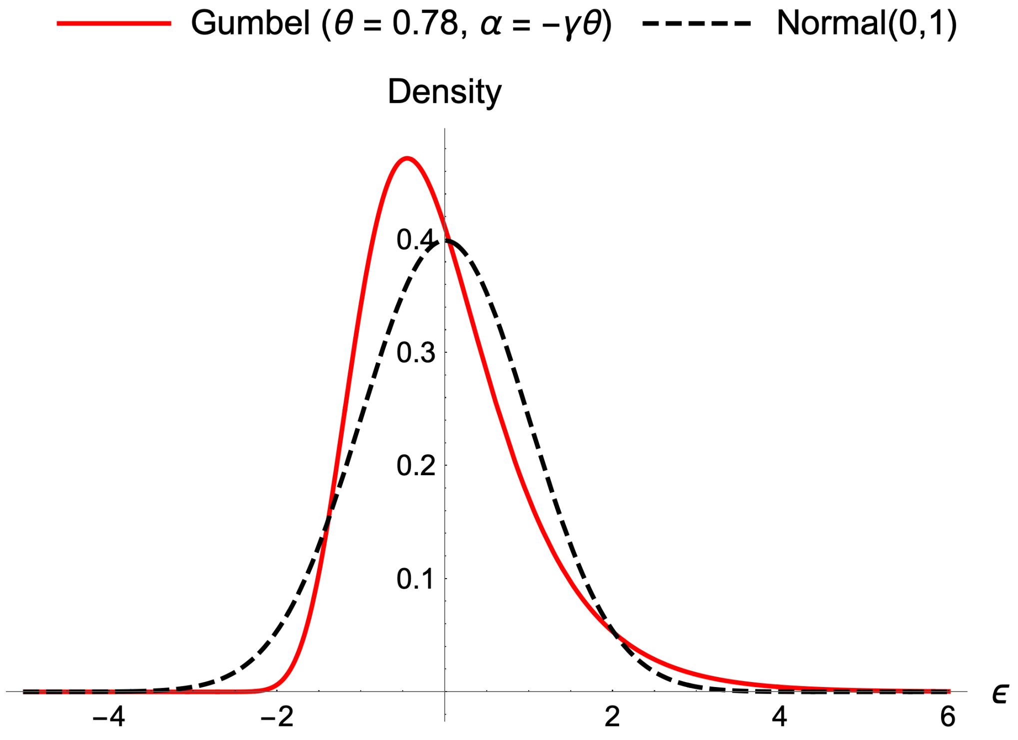
ε_{ni} の **累積分布関数** (cumulative distribution) は全ての i に対して

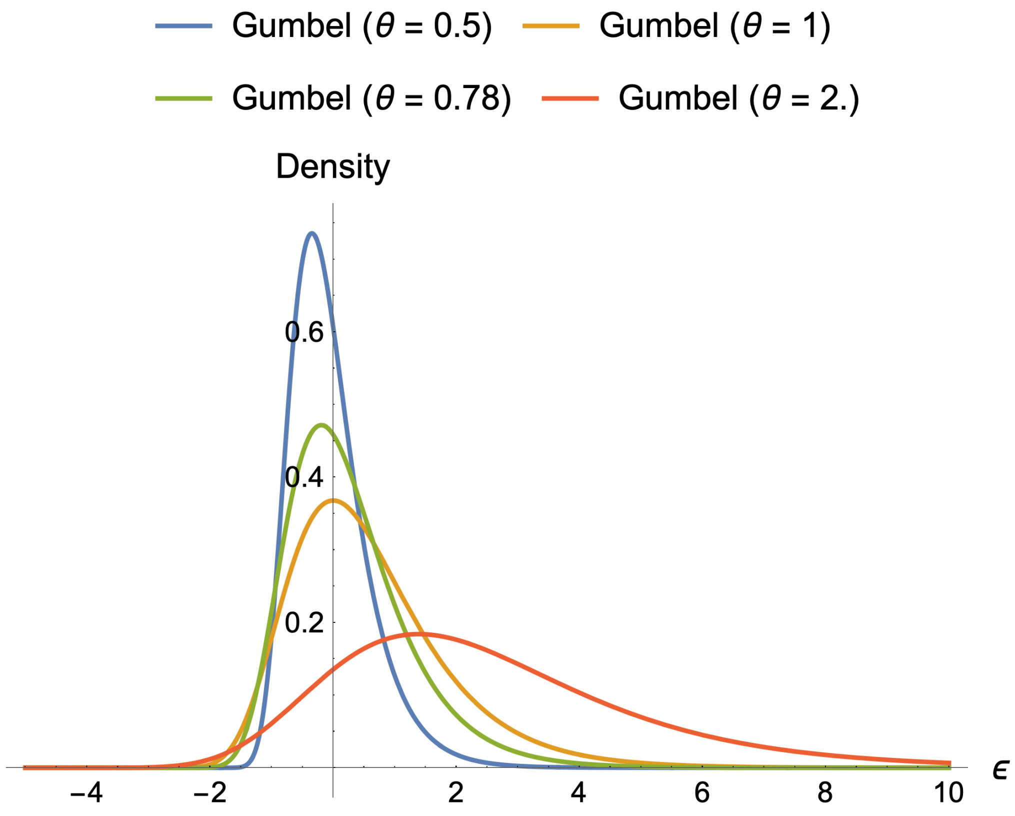
$$\mathbb{P}(\varepsilon_{ni} \leq \varepsilon) = F(\varepsilon) = \exp\left(-\exp\left(-\frac{\varepsilon-\alpha}{\theta}\right)\right) \quad \varepsilon \in (-\infty, \infty)$$

- α : ロケーションパラメタ, θ : スケールパラメタ
- 平均 : $\alpha + \gamma\theta$, 分散 : $\frac{\pi^2}{6}\theta^2$. ただし $\gamma \approx 0.577$ はEuler 定数

Gumbel 分布の確率密度

- 正規分布 (normal distribution) と同様の単峰分布（分散1に基準化）
- スケールパラメータθ が大きいほど確率項のゆらぎは大きい





多項ロジットモデルの選択確率

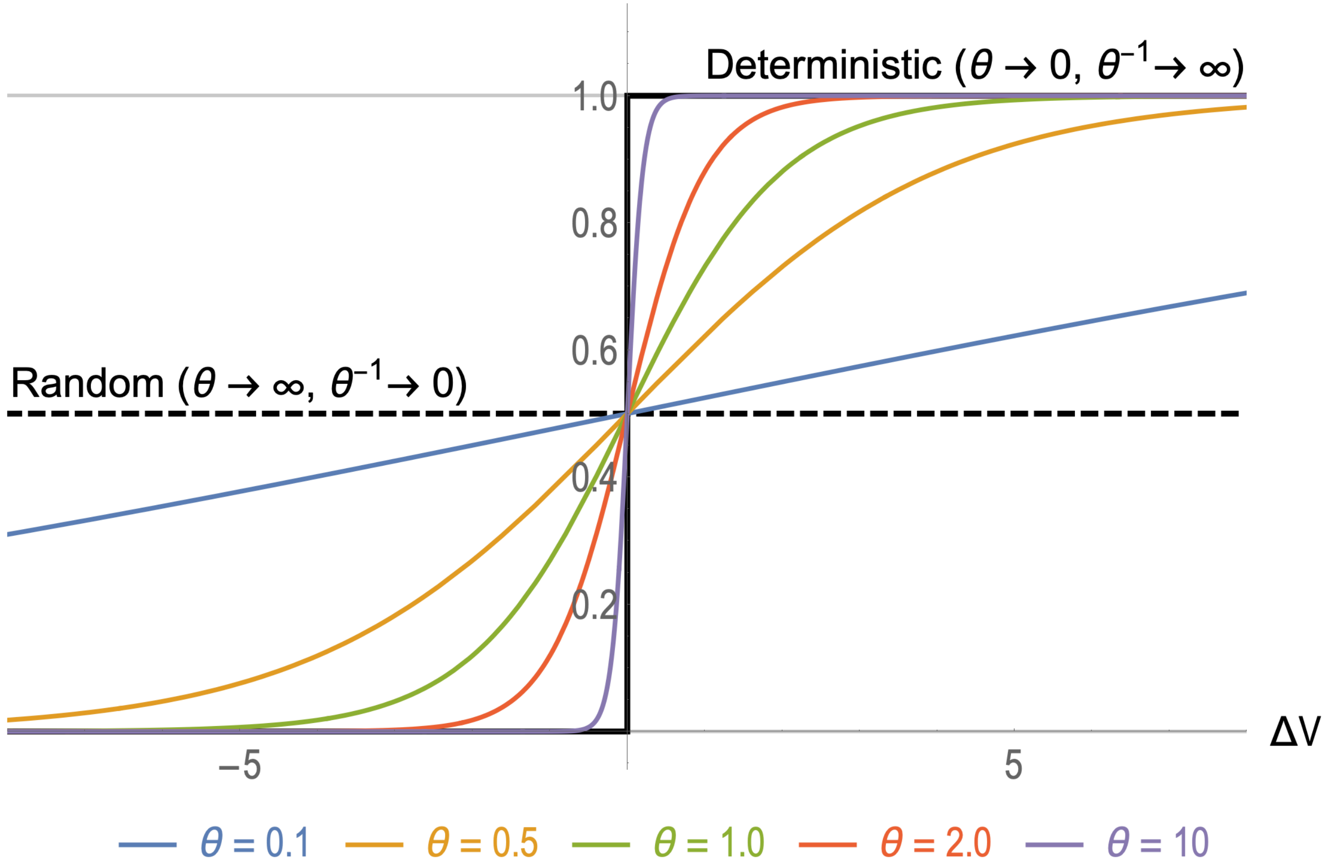
- Gumbel 分布を使用して計算するとMNLのもとでの選択確率は

$$P_{ni} = \frac{\exp(\theta^{-1} V_{ni})}{\sum_{j \in A_n} \exp(\theta^{-1} V_{nj})}$$

  - ロケーションパラメタは影響しない（cf. ARUM の一般的性質）
- θ の極限における性質 ✍️
  - $\theta \to \infty$ $(\theta^{-1} \to 0)$ のとき $P_{ni} = 1/|A_n|$（等確率）
  - $\theta \to 0$ $(\theta^{-1} \to \infty)$ のとき $V_{ni} = \max_{j \in A_n}\{V_{nj}\}$ でなければ $P_{ni} = 0$

スケールパラメータの影響

- 効用**差**大 $\Rightarrow$ 選択確率大：$P_1 = 1/(1+\exp(-\theta^{-1}\Delta V))$，$\Delta V \equiv V_1 - V_2$
- 選択確率は ΔV に対して連続変化，ゆらぎ大 $\rightarrow$ 選択確率は均等化



選択確率の導出　準備：Gumbel分布の性質

累積分布関数 $F(\varepsilon) = \exp\left(-\exp\left(-\varepsilon/\theta\right)\right)$ に対して（$\alpha = 0$）：

- 確率密度関数：

$$f(\varepsilon) = F'(\varepsilon) = \rho(\varepsilon)F(\varepsilon), \quad \rho(\varepsilon) = \theta^{-1}\exp(-\theta^{-1}\varepsilon)$$

- 累積分布関数の並行移動：

$$F(v+\varepsilon) = F(\varepsilon)^{\exp(-\theta^{-1}v)}$$

- 累乗された累積分布関数の微分：

$$\{F(\varepsilon)^t\}' = t\rho(\varepsilon)F(\varepsilon)^t$$

選択確率の導出 ✍️

※ 簡単のため個人インデックス n を無視．注意：ε の**独立性**が重要．

$$
\begin{aligned}
P_i = \mathbb{P}(V_i + \varepsilon_i \geq V_j + \varepsilon_j, \forall j \neq i) &= \int_{-\infty}^{\infty} f(\varepsilon) \prod_{j \neq i} F(V_i - V_j + \varepsilon)\,\mathrm{d}\varepsilon \\
&= \int_{-\infty}^{\infty} \rho(\varepsilon) F(\varepsilon)^{1+\sum_{j\neq i} \exp(\theta^{-1}(V_j - V_i))}\,\mathrm{d}\varepsilon \\
&= \frac{1}{1 + \sum_{j \neq i} \exp(\theta^{-1}(V_j - V_i))} \left[F(\varepsilon)^t \right]_{-\infty}^{\infty} \\
&= \frac{\exp(\theta^{-1} V_i)}{\sum_{j \in A} \exp(\theta^{-1} V_j)}
\end{aligned}
$$

最大効用の分布と期待値

最大効用 $Y \equiv \max_{i\in A} U_i$ の分布は $Y \leq x \Leftrightarrow U_i \leq x \ \forall i \in A$ だから

$$\mathbb{P}\left(V_i + \varepsilon_i \leq x \ \forall i \in A\right) = \prod_{i\in A} F(x - V_i) = F(x)^{\sum_{i\in A} \exp(\theta^{-1} V_i)} = F(x - \lambda_0)$$

- ここで $\lambda_0 = \theta \log \sum_{i\in A} \exp(\theta^{-1} V_i)$
- 即ち，最大効用は$\alpha = \lambda_0$の Gumbel 分布に従う．よって期待値は

$$\mathbb{E}\left[\max_{i\in A} U_i\right] = \theta \log \sum_{i\in A} \exp(\theta^{-1} V_i) + \gamma\theta$$

期待最大効用 (Expected Maximum Utility)

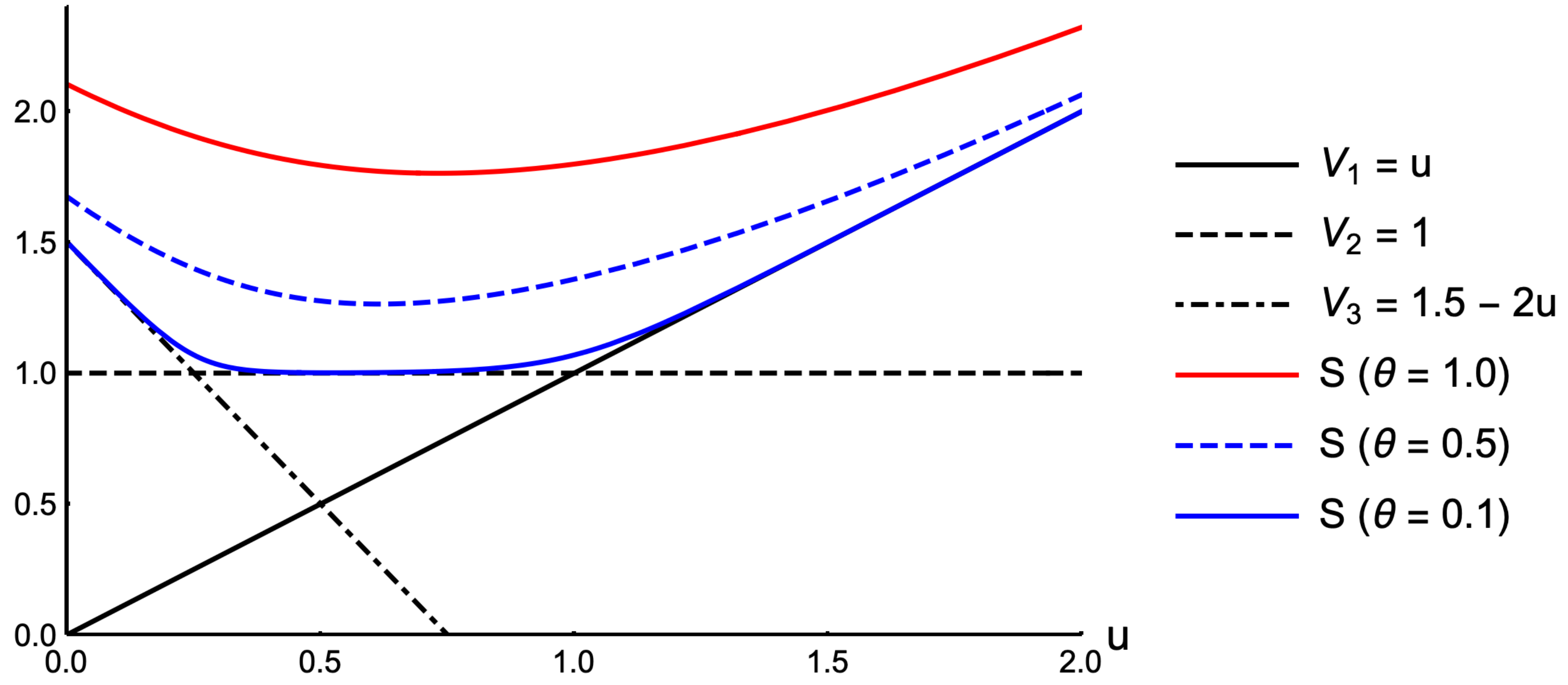
- MNL において選択肢 i の確定効用が V_i のとき,

$$S = \mathbb{E}\left[\max_{i \in A} U_i\right] = \theta \log \sum_{i \in A} \exp(\theta^{-1} V_i) + \gamma\theta$$

- 選択肢集合 A から得られる最大効用の期待値
- S の値の変化を政策分析における厚生指標として使用可能
- **ログサム変数** (log-sum variable) などと呼ばれることもある.
  - See, e.g., de Jong et al. (2007) Trans. Res. A

期待最大効用の性質①

- どの確定効用よりも大きい：$S > \max_{i \in A} V_i$
- $\theta \to 0$ のとき $\max_{i \in A} V_i$ に収束（決定論的選択と一致）



期待最大効用の性質②

- 効用による微分が選択確率となる ✍️：

$$\frac{\partial S}{\partial V_i} = \frac{\partial}{\partial V_i}\left[\theta \log \sum_{i \in A} \exp(\theta^{-1} V_i) + \gamma\theta\right] = \frac{\exp(\theta^{-1} V_i)}{\sum_{j \in A} \exp(\theta^{-1} V_j)}$$

- 消費者理論における Shepherd の補題に対応
  - Shepherd の補題：支出関数の価格微分が Hicks の補償需要関数
- エントロピー正則化（softmax化）した効用最大化問題の双対性と関係

数値例 1/2

- 前講義で取り扱ったモード選択の例を再度考える

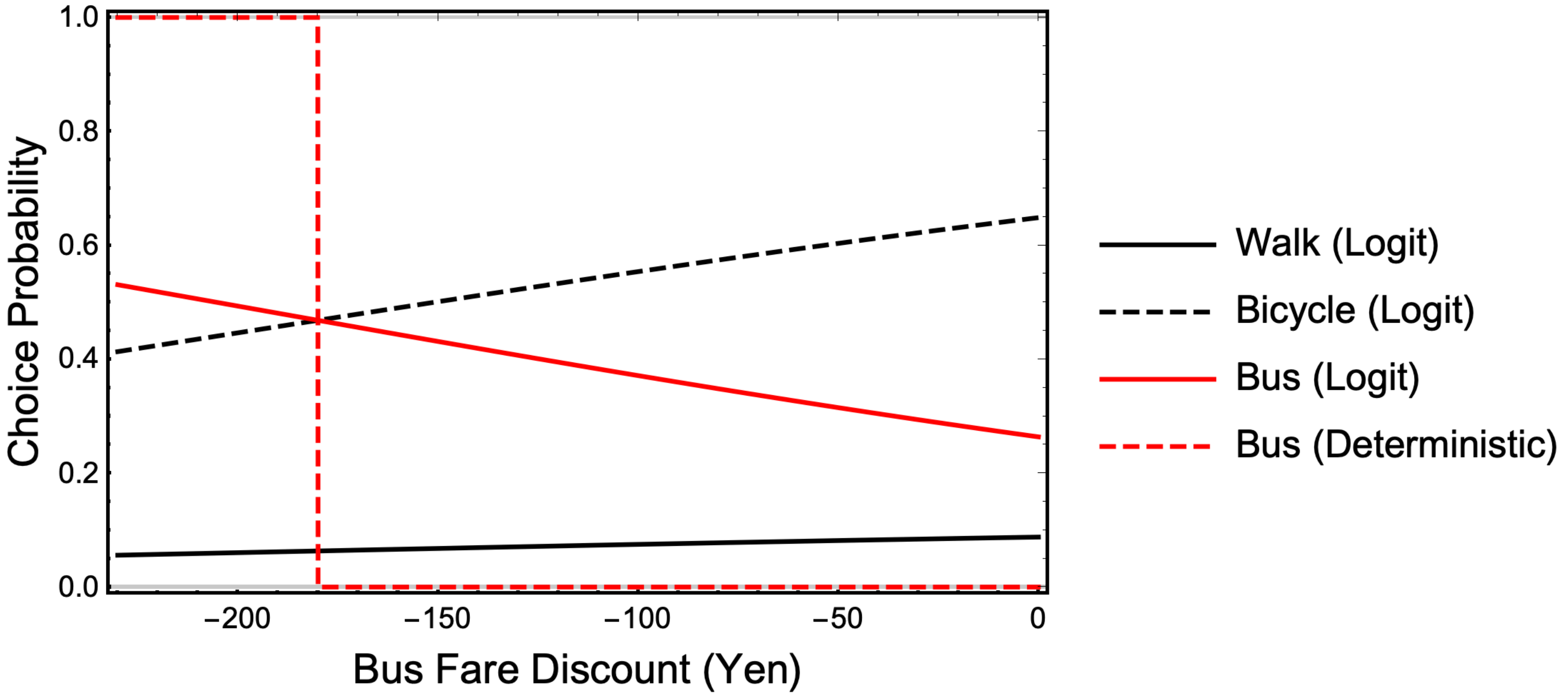
$$V_i = \beta_T T_i + \beta_C C_i + \beta_Q Q_i$$

モード	所要時間 (分) T_i	費用 (円) C_i	快適性 Q_i
徒歩	40	0	3
自転車	20	0	1
バス	15	230	0

- $\beta = (-0.3, -0.01, 1.0)$, $\theta = 0.5$ として数値的に調べてみよう.

数値例 2/2

- 全ての選択肢が必ず利用される.
- バスの運賃を減らした場合も滑らかな変化.



MNLの利便性と限界

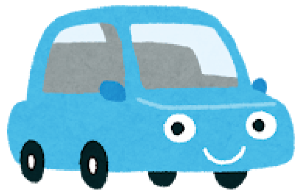
- 利便性：選択確率が closed-form で得られる
  - 特性の把握が容易，推定・予測しやすい
- 限界：**IIA特性** (Independence from Irrelevant Alternatives)
  - 確率項に相関がある選択肢を扱えない

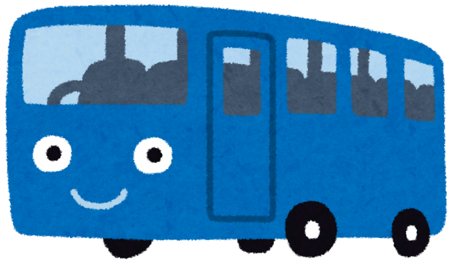
IIA特性

- 選択肢間の **相対的な**選択確率が**第三の選択肢に依存しない**

$$\frac{P_i}{P_j} = \frac{\exp(\theta U_i)}{\exp(\theta U_j)}$$

  - i, j 以外の選択肢の効用は全く影響していない
  - 選択肢集合が拡大した場合も影響しない
- この性質はいわゆる **赤バス vs. 青バス問題** をもたらす.





赤バス vs. 青バス問題

- $A = \{\text{自動車}, \text{バス}\}$ であり，ともに効用が V だとすると

$$P_{\text{auto}} = P_{\text{bus}} = \frac{e^V}{e^V + e^V} = \frac{1}{2}$$

- バスを半分ずつ赤・青に塗り分けると $A = \{\text{自動車}, \text{赤バス}, \text{青バス}\}$ に

$$P_{\text{auto}} = \frac{e^V}{e^V + e^V + e^V} = \frac{1}{3}, \; P_{\text{bus}} = \frac{e^V + e^V}{e^V + e^V + e^V} = \frac{2}{3}$$

- 本来は「バスの合計」である50%を「分け合う」べきでは？
- MNLは「類似」する選択肢たちの選択確率を**過大評価**してしまう

注意

- 「類似」≈ 確率効用の相関．似た選択肢は似た（観測不可能）効用
  - MNL は i.i.d. を仮定するのだった．
- でも，赤・青バスのサービス水準は低下し効用が下がるはずでは……
  - あくまで思考実験．過大評価特性には一般性がある．
- IIA特性は個人の選択行動に対する性質
  - 個人の性質が多様である（e.g., 効用関数に異質性がある）集団には必ずしも当てはまらない．
  - 十分に層別化 (stratify) して利用すれば悪影響は緩和される可能性

余談：IIA特性の工学的応用上の利点

1. モデルの適用において全ての選択肢の集合を扱わないでよい
   - 時刻とモードの同時選択が IIA を満たすなら

$$P(t, m) = P(m \mid t)P(t) = P(m)P(t)$$

2. 新しい選択肢 k の導入効果を容易に推定可能

$$P'_{ni} = \frac{\exp(\theta V_{ni})}{\sum_{j \in A_n} \exp(\theta V_{nj}) + \exp(\theta V_{nk})}$$

   - 既に述べたように既存選択肢と相関が存在する場合は問題がある

IIA特性の改善方法

1. **プロビット・モデル**を用いる
   - Multinomial Probit (MNP) Model
2. ロジット・モデルを改良する
   - 選択の階層構造の導入
   - 個人の異質性の導入

プロビット・モデル (MNP)

- $\boldsymbol{\epsilon} = (\varepsilon_i)_{i\in A}$ は **多変量正規** (multivariate normal) **分布** に従うと仮定：
  - 平均 $\mathbf{0}$，共分散行列 $\boldsymbol{\Sigma}$ として

$$f(\boldsymbol{\varepsilon}) = \frac{1}{(2\pi)^{|A|/2}\left|\boldsymbol{\Sigma}\right|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}\right)$$

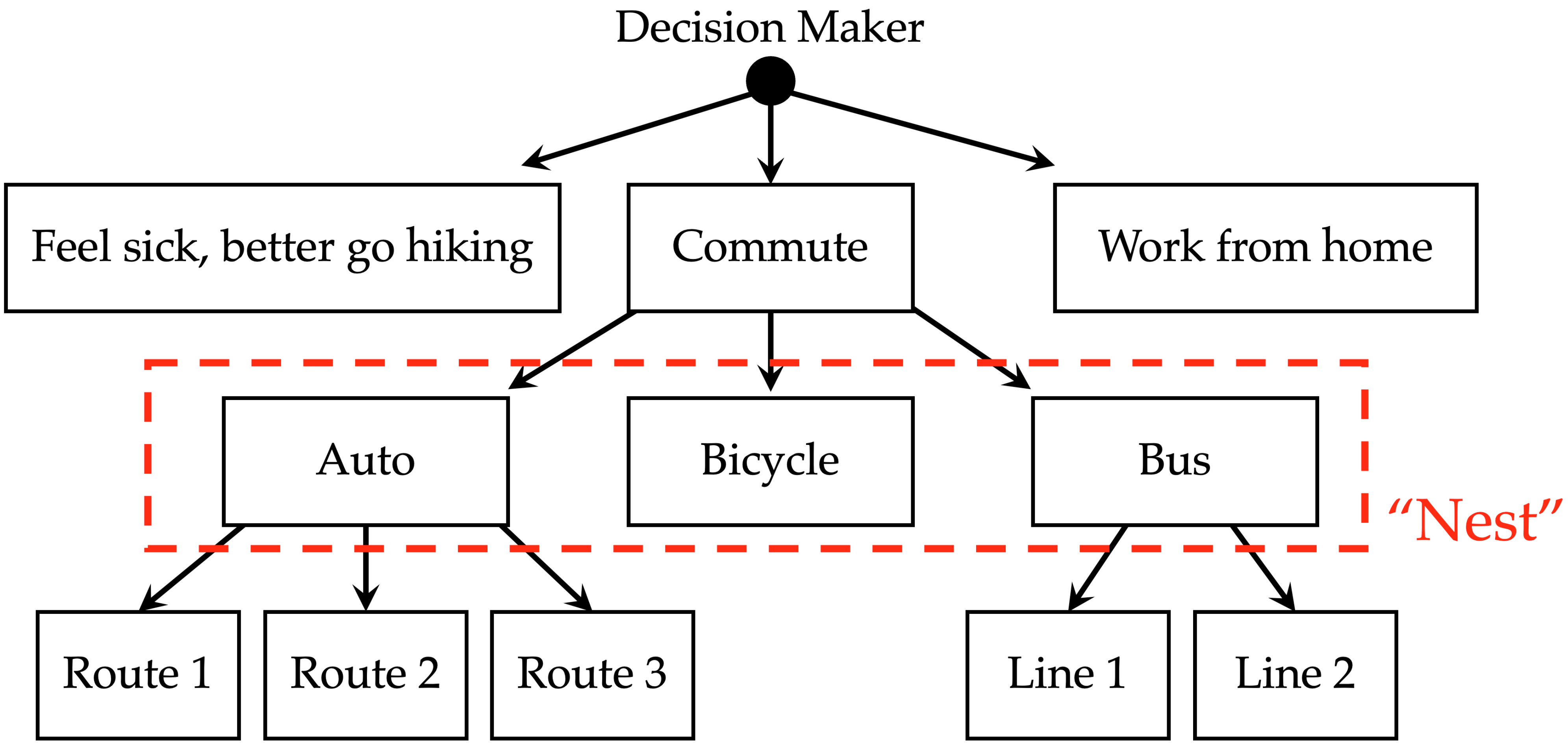
- 😊 **任意の誤差項相関**（≈「類似性」）を表現可能で，柔軟性が高い
- 🤢 選択確率は closed-form ではなく数値計算が必要．推定も重い．
- 🤢 係数と確率の関係が非直感的
- 選択肢数が多いと正直無理．少なければあるいは．

余談：中長期予測とモデルの意味的構造

- 🎯 短期予測にMNPは好適
  - 現在の選好・誤差構造を忠実・柔軟に表現（↔ 構造的解釈性低め）
- 📈 長期予測は選択の意味的構造（**選択ツリーの構造**）が重要な可能性
  - 例：目的地選択 → 交通機関選択 → ルート／交通事業者選択
  - MNPモデルでは誤差の相関構造を柔軟に扱えるが，明示的な選択階層の **“因果的”モデル化** は含まれない
- ただし，応用によって構造のないMNPでも（計算さえできれば……）精度の高い中長期予測が可能な場合もあろうだろう
  - モデル選択は**予測の時間スケール**と**政策介入の大きさ**に依存

プレビュー：選択の階層構造の考慮

- 選択の階層構造からの表現



プレビュー：個人の異質性の表現

- 個人の選好のばらつきからの表現：同じ特性に対する評価が異なる

	せっかち	のんびりや	中間
β	(–20,–0.5,50)	(–10., –1., 150.)	(–10,–0.5,50)
徒歩	–750	**0** 🎯	–300
自転車	–450	–100	**–200** 🎯
バス	**–415** 🎯	–380	–265

- **効用関数の係数ベクトル**の確率分布$F(\beta)$を考える（**混合モデル**）

まとめ

- RUMは確率的ゆらぎを取り入れた効用最大化モデル
- MNLは確率項が i.i.d. Gumbel 分布に従うRUM
- MNLの便利さと限界：計算が容易 vs. IIA特性

MNLの限界を超えて

- 観測不能な誤差の**構造**：相関と階層性 (Nested Logit)
- 個人の**選好の異質性** (Mixed Logit)

基礎の再確認

以下の問いに簡潔に答えてみよ：

1. MNLモデルにおける **確率項の分布** は何か？
2. MNLの選択確率式において，θ はどのような意味を持つか？
3. MNLが持つ **IIA特性** とは何か？　なぜ問題となる場合があるのか？
4. 選択肢の効用がすべて等しいとき，MNLモデルの選択確率は？
5. MNLにおける厚生指標（期待最大効用）の**意味**と**応用例**を挙げよ.

課題 2

1. 自分の通学に関する**交通手段選択**について，以下の項目を構成せよ：
   - 選択肢（最低3つ）
   - 各選択肢の属性（例：所要時間，費用，快適性）の数値化された表
2. 自分自身にふさわしい**線形効用関数**を仮定し，β を与えてみよ．
3. MNLモデルの式に基づいて各選択肢の選択確率を求めよ．
4. 以下のような **状況変化** を考え，選択確率の変化を調べ考察せよ：
   - バスの費用が100円引き下げられたとき
   - 自転車の快適性が上昇したとき（例：新しい自転車道の整備）

発展課題（任意）

- 「赤バス・青バス」問題のような IIA**特性の限界**を，自分の通学に関する選択に引き寄せて考えてみよ.
- 例：徒歩 vs. 自転車の2択に新しいタイプの自転車（電動アシスト）が追加された場合，MNLモデルはどのような予測をするか？　その予測は直観に合致しているだろうか？

参考文献

[1] 土木計画学研究委員会 (編) (1996). **非集計行動モデルの理論と実際**. 土木学会.

[2] de Jong, G., Daly, A., Pieters, M., & van der Hoorn, T. (2007). The logsum as an evaluation measure: Review of the literature and new results. *Transportation Research Part A: Policy and Practice*, 41(9), 874-889. ：期待最大効用による厚生評価について.