

最近の交通行動モデルの発展

名古屋大学 山本俊行

NUTREND Nagoya University
TRansport and
ENvironment Dynamics

報告内容

- GEV (generalized extreme value)モデル
 - Network GEV
 - CNL (cross nested logit)の誤差相関
- MMNL (mixed multinomial logit)モデル
 - 推定の効率化
- 離散連続モデル
 - 誤差分布の仮定
 - 推定可能性
- 個人間相互作用のモデル化
- 調査手法
 - SP調査設計

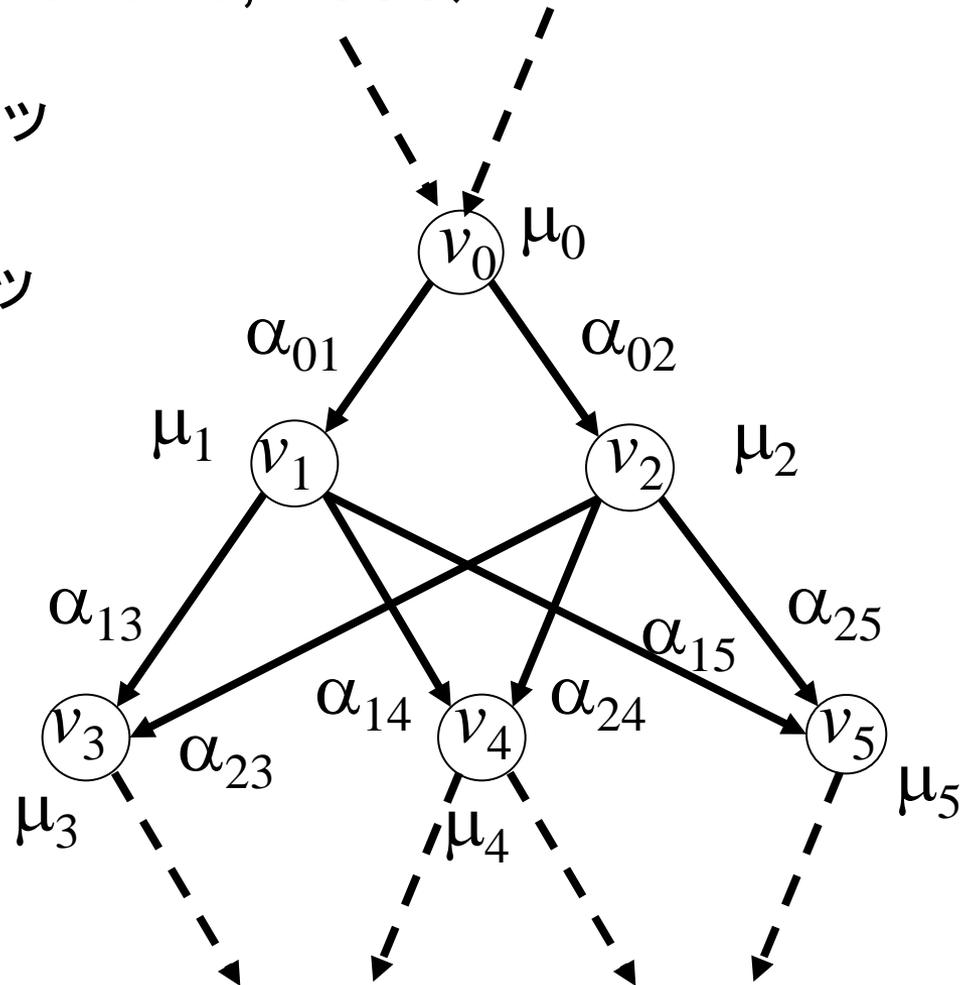
GEVモデル

- 多項ロジットモデルのIIA特性を緩和し、選択肢間の柔軟な誤差相関を可能としたモデル
 - 後述のMMNLモデルも同様の性質を有する
- 選択確率がclosed formで書けるため、数値積分の必要がない
 - 推定誤差の原因となる数値積分が必要なMMNLモデルより常に優先すべき
- あらゆる誤差相関に対応したGEVモデルが存在する訳ではない
 - 表現したい誤差相関を持つGEVモデルを見つけるか、見つからない時は自ら作る必要がある

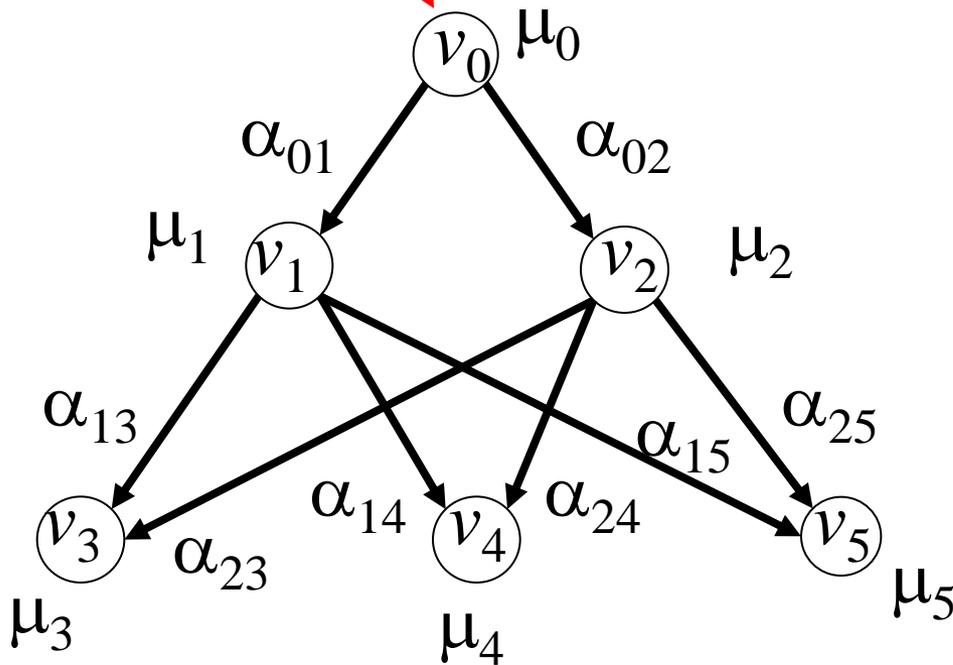
Network GEVモデル

(Daly & Bierlaire, 2006)

- 選択肢間の相関構造をネットワーク構造で表現
- 再帰型のネステッドロジットモデルとして定式化
- 効用理論との整合性に関する複雑な証明が不要
- 新たなGEVモデルを容易に作成可能



Network GEVモデル



$$P(v_3|v_1) = \frac{\alpha_{13} \exp(\mu_1 V_3)}{\sum_{j=(3,4,5)} \alpha_{1j} \exp(\mu_1 V_j)}$$

$$P(v_1) = \frac{\left\{ \sum_{j=(3,4,5)} \alpha_{1j} \exp(\mu_1 V_j) \right\}^{\mu_0/\mu_1}}{\sum_{m=(1,2)} \left\{ \sum_{j=(3,4,5)} \alpha_{mj} \exp(\mu_m V_j) \right\}^{\mu_0/\mu_m}}$$

- v : ノード (ネストや選択肢)
- α : ロケーションパラメータ (どのネストにどんな確率で含まれるか)
- μ : スケールパラメータ (ネスト内の選択肢の相関の高さ)

$$P(v_3) = P(v_3|v_1)P(v_1) + P(v_3|v_2)P(v_2)$$

上記で v_1 は本当は $v_1|v_0$ で、再帰的につながられる

CNLモデル

δ : 選択肢*i*と*j*が同じネスト*m*に入っているかを表すダミー変数

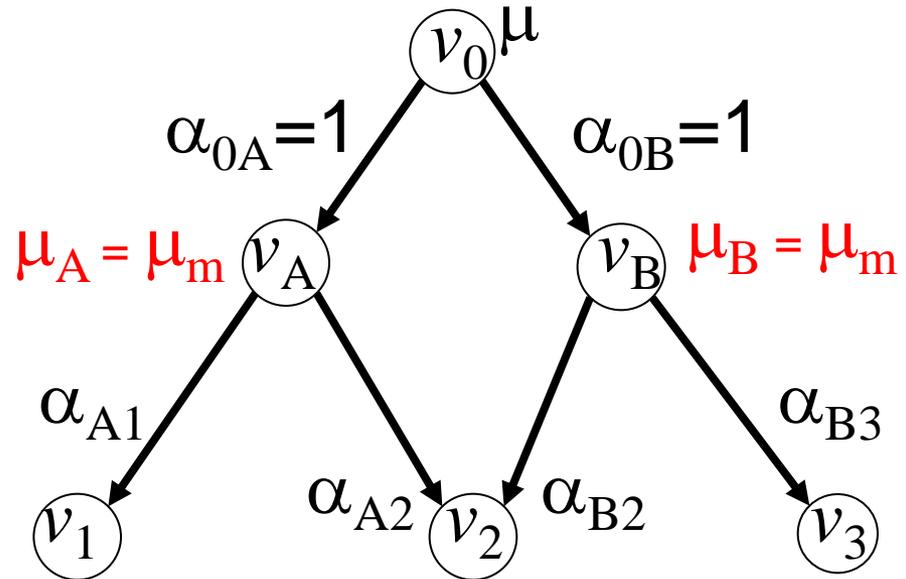
- ネスティッドロジットモデルの選択肢間の相関

$$\text{Corr}(U_i, U_j) = \left(1 - \left(\frac{\mu}{\mu_m}\right)^2\right) \delta_m(i, j),$$

- CNLの場合は？

$$\widehat{\text{Corr}}(U_i, U_j) = \sum_{m=1}^M \alpha_{im}^{1/2} \alpha_{jm}^{1/2} \left(1 - \left(\frac{\mu}{\mu_m}\right)^2\right).$$

こんな感じになりそうだが、これは簡便法

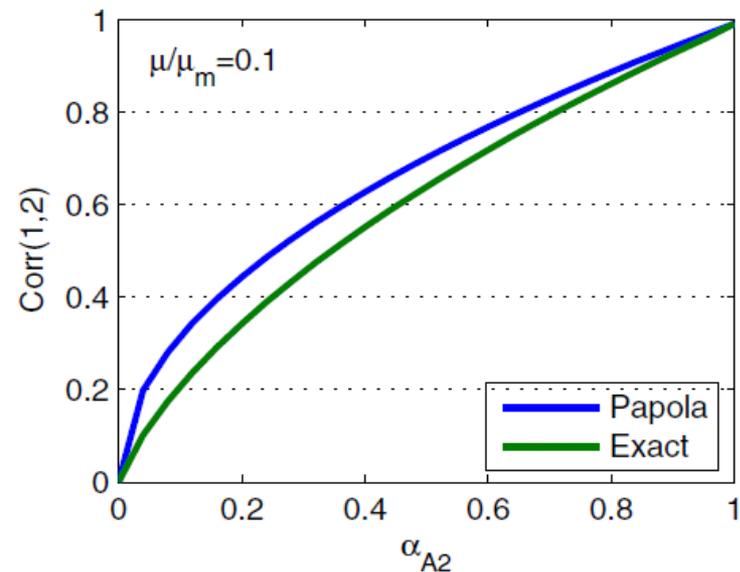
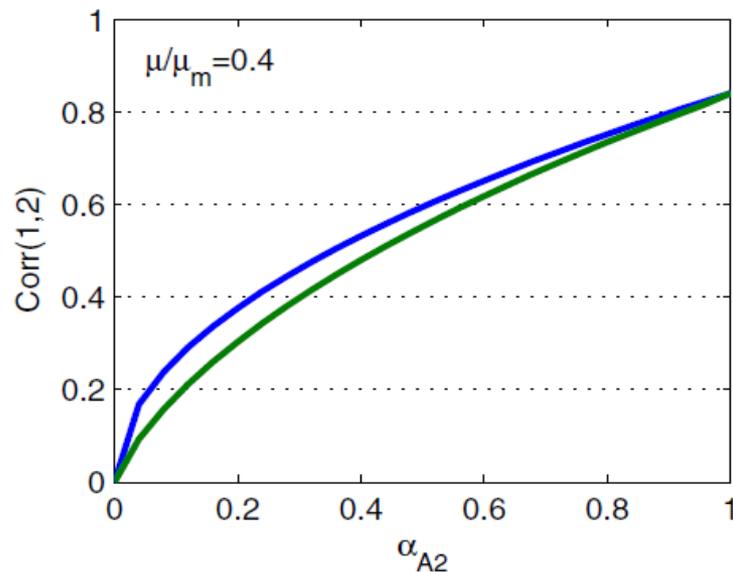


厳密解 (Abbe et al., 2007)

$$\text{Corr}(U_i, U_j) = \text{Corr}\left(\max_m \left(\frac{\ln \alpha_{im}}{\mu} + \varepsilon_{im}\right), \max_m \left(\frac{\ln \alpha_{jm}}{\mu} + \varepsilon_{jm}\right)\right),$$

where

$$\text{Corr}(\varepsilon_{im}, \varepsilon_{jn}) = \left(1 - \left(\frac{\mu}{\mu_m}\right)^2\right) \delta_{m,n}.$$



- 簡便法 (Papola) は過大推計の傾向. ただし誤差は大きくない

MMNLモデル

- 誤差項を多項ロジットモデルに追加したモデル
 - 選択肢間のあらゆる誤差分散共分散が表現可能
 - パラメータの非観測異質性も表現可能

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right) f(\beta) d\beta.$$

- 選択確率がclosed formで書けないため、数値積分が必要となる。シミュレーション法を用いる

$$\check{P}_{ni} = \frac{1}{R} \sum_{r=1}^R L_{ni}(\beta^r)$$

推定の効率化

- 均質な分布でかつ配列間の相関をふせぐ効率的な乱数列の構築
 - (t,m,s)-nets, MLHS (modified Latin hypercube sampling)



- 数列以外の部分での効率化
 - 数列の儉約, 遺伝的アルゴリズム (GA) の活用

数列の儉約

BTRDA (basic trust-region with dynamic accuracy) アルゴリズム (Bastin et al., 2006)

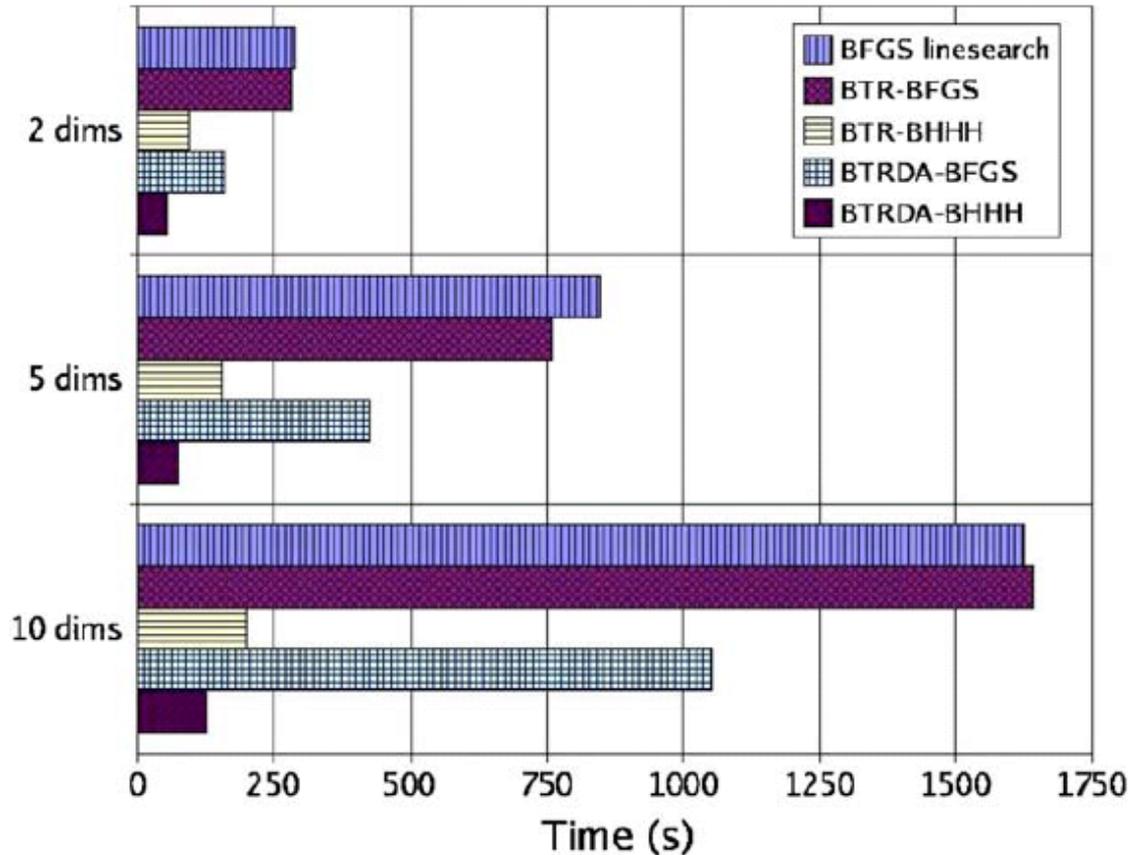
- MMNLのシミュレーションはサンプリングエラーを含むため長い数列(多くの抽出点)が必要
- サンプリングエラーは数列の長さによって決まる
- 繰り返し計算の初期段階の数列の長さを削減し、計算速度を向上

信頼区間

$$\epsilon_{\delta}^R(\theta) = \alpha_{\delta} \frac{1}{N} \sqrt{\sum_{i=1}^I \frac{\sigma_{ij_i}^2(\theta)}{R(P_{ij_i}(\theta))^2}}$$

N: サンプル数, R: 数列長さ
 σ : Pの標準偏差

計算速度比較



- 数列長を固定した場合 (BTR) よりも提案アルゴリズム (BTRDA) の方が速い

集計データとGAの活用 (Liu et al., 2007)

- 並行する高速道路と一般道路の一定期間のトラカンデータを経路選択結果の集計データとして利用
- 経路選択行動の表現にMMNLモデルを適用
- MMNLモデルから算出される経路選択確率の平均と標準偏差とトラカンデータで観測した平均と標準偏差を最小化
- GAによるパラメータの探索

Randomized and scrambled
Halton数列による選択確率の
シミュレーション

$$\check{P}_{rs}^p(t) = \frac{1}{K} \cdot \sum_{k=1}^K L_p^k(\beta_k(t); t)$$

Pの平均 μ と標準偏差 σ の計算

GAによる
関数最小化

$$\min_{\theta(t)} (\text{MSE}) = \sum_{i=1}^N \left[\sum_{p \in P_{rs}} (\hat{\mu}_p(t_i) - \tilde{\mu}_p(t_i))^2 + \lambda \sum_{p \in P_{rs}} (\hat{\sigma}_p(t_i) - \tilde{\sigma}_p(t_i))^2 \right]$$

離散連続モデル

- 離散的な選択だけではなく、連続的な量の選択も同時に表現できるモデルが望まれる
- 離散的な選択と連続量の選択を同時に取り扱うモデルを離散連続モデルと呼ぶ
 - 自動車の車種選択と走行距離，活動内容の選択と活動時間，暖房器具の選択と使用量，電話料金プランと利用量，等
- 従来は離散選択部分が単一の選択肢を選択するモデルであったが，複数の選択肢を選択するモデルに発展

複数の選択肢を選択可能な 離散連続選択モデル

- 選択した選択肢の追加資源1単位あたりの限界効用は等しく, 選択されない選択肢の限界効用はそれ以下 (Kuhn-Tucker条件)

$$V_k + \varepsilon_k = V_1 + \varepsilon_1 \quad \text{if } e_k^* > 0 \quad (k = 2, 3, \dots, K),$$

$$V_k + \varepsilon_k < V_1 + \varepsilon_1 \quad \text{if } e_k^* = 0 \quad (k = 2, 3, \dots, K),$$

配分量

で, データが観測される確率は積分がずらずら並ぶ.

$$P(e_1^*, e_2^*, e_3^*, \dots, e_M^*, 0, 0, \dots, 0) = |J| \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_{M+1}=-\infty}^{V_1-V_{M+1}+\varepsilon_1} \int_{\varepsilon_{M+2}=-\infty}^{V_1-V_{M+2}+\varepsilon_1} \dots \int_{\varepsilon_{K-1}=-\infty}^{V_1-V_{K-1}+\varepsilon_1} \int_{\varepsilon_K=-\infty}^{V_1-V_K+\varepsilon_1} f(\varepsilon_1, V_1 - V_2 + \varepsilon_1, V_1 - V_3 + \varepsilon_1, \dots, V_1 - V_M + \varepsilon_1, \varepsilon_{M+1}, \varepsilon_{M+2}, \dots, \varepsilon_{K-1}, \varepsilon_K) d\varepsilon_K d\varepsilon_{K-1} \dots d\varepsilon_{M+2} d\varepsilon_{M+1} d\varepsilon_1,$$

誤差分布の仮定

- 誤差分布がガンベル分布に従う時, ロジットモデルと同様に closed formが得られる (Bhat, 2008)

$$P(x_1^*, x_2^*, x_3^*, \dots, x_M^*, 0, 0, \dots, 0) = \frac{1}{p_1} \cdot \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^M f_i \right] \left[\sum_{i=1}^M \frac{p_i}{f_i} \right] \left[\frac{\prod_{i=1}^M e^{V_i/\sigma}}{(\sum_{k=1}^K e^{V_k/\sigma})^M} \right] (M-1)!$$

- 誤差分布が正規分布に従う時, ベイズ推定を利用 (Fang, 2008)

$$\text{draw } \beta | \Sigma, \mathbf{y}_i^* \text{ from } \pi(\beta | \Sigma^{(t-1)}, \mathbf{y}_i^{*(t-1)})$$

$$\text{draw } \Sigma | \beta, \mathbf{y}_i^* \text{ from } \pi(\Sigma | \beta^{(t)}, \mathbf{y}_i^{*(t-1)})$$

$$\text{draw } \mathbf{y}_i^* | \beta, \Sigma, \mathbf{y}_i \text{ from } \pi(\mathbf{y}_i^* | \beta^{(t)}, \Sigma^{(t)}, \mathbf{y}_i)$$

- いずれの方法でも誤差項間の相関を導入する形に発展可能性を持つ

MDCEV (multiple discrete-continuous extreme value) モデルの推定可能性

(Bhat, 2008)

配分量

効用関数

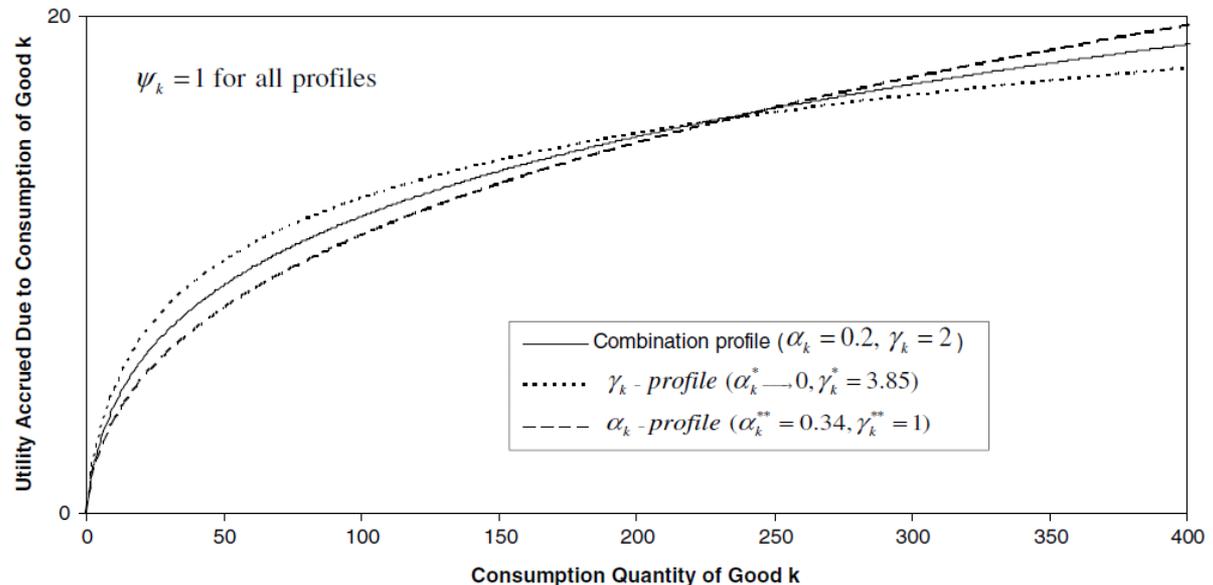
- $\alpha \rightarrow 0$ の時, 対数型になる
- $\alpha=0$ か $\gamma=1$ に固定しても同じような効用関数の表現が可能

$$U(\mathbf{x}) = \sum_{k=1}^K \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}$$

$$U(\mathbf{x}) = \sum_{k=1}^K \gamma_k \psi_k \ln \left(\frac{x_k}{\gamma_k} + 1 \right).$$

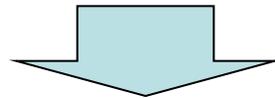


実際的には α と γ を同時に推定することは困難



個人間相互作用のモデル化

- 従来の離散選択モデルは意思決定者がそれぞれ独立に意思決定することを仮定していた
- 他者の行動や意見に影響を受ける状況を明示的にモデル化する必要の高まり



- 個人間相互作用のモデル化
 - 所属グループの選択率の影響
 - 各世帯構成員の効用の統合による世帯効用の表現
 - エージェント間の力関係の推定

所属グループの選択率の影響

(Fukuda and Morichi, 2007)

- 効用関数に所属グループの選択率 m が入る

$$LL = \sum_i \ln P(\omega_i) = \sum_i \ln \left\{ \frac{\exp \left[\omega_i \left(b + c' X_i + d' Y_{n(i)} + J \bar{m}_{n(i)}^c \right) \right]}{\sum_{v_i \in \{+1, -1\}} \exp \left[v_i \left(b + c' X_i + d' Y_{n(i)} + J \bar{m}_{n(i)}^c \right) \right]} \right\}$$

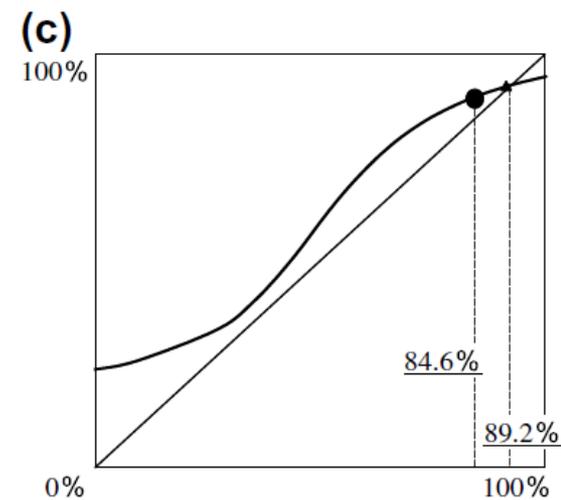
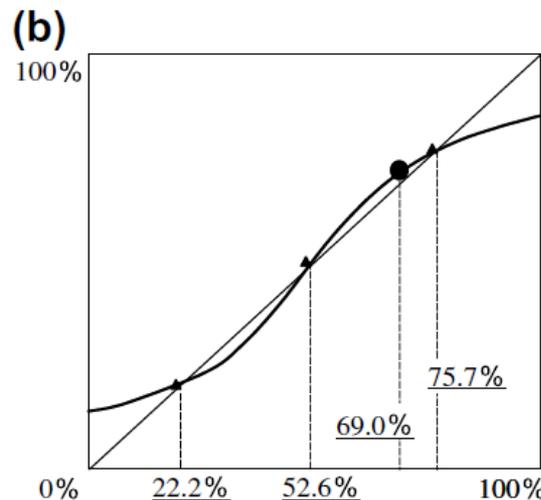
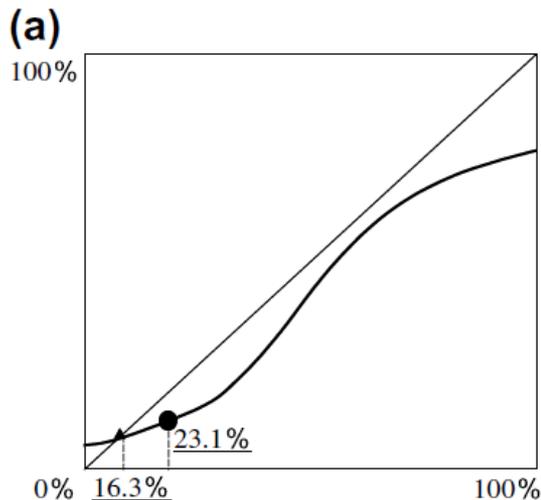
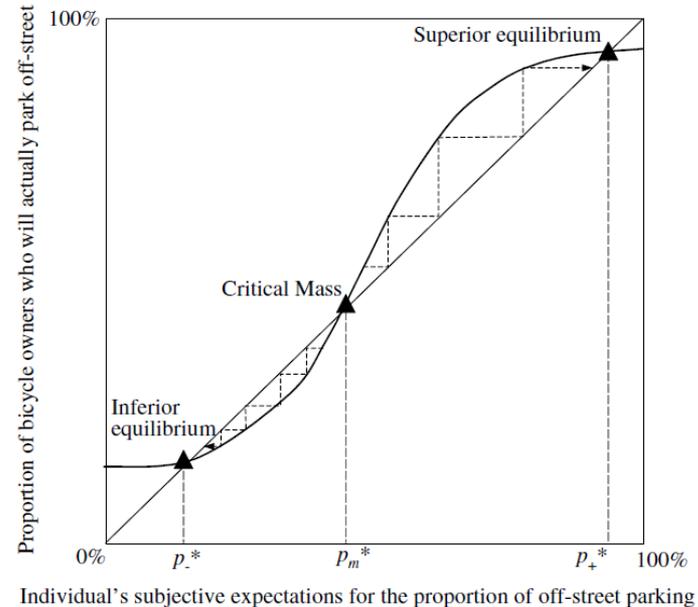
- m は個々人の選択結果の集合

$$\bar{m}_{n(i)}^c = m_{n(i)} = \int \tan h(b + c' X_i + d' Y_{n(i)} + J m_{n(i)}) dF_{X_i | Y_{n(i)}} \quad \forall n,$$

- 両者を組み合わせて推定

推定結果から分かること

- 均衡状態の把握
 - 均衡点は45度線上
 - 個人の選択確率から集団の選択率を計算でき、それに基づいて個人の選択確率が変化し、
... 均衡に至る



世帯構成員の効用の統合

(Zhang and Fujiwara, 2006)

- 個人の効用 u と世帯効用HUFの関係

$$\text{HUF} = \frac{1}{1-\alpha} \sum_i w_i u_i^{1-\alpha}, \quad w_i \geq 0 \quad \text{and} \quad \sum_i w_i = 1$$

- α によって様々な概念が表現可能

- $\alpha \rightarrow \infty$: $\text{HUF} = \min(u_i \mid i = 1, 2, \dots, n)$

- $\alpha \rightarrow 1$: $\text{HUF} = \prod_i (u_i)^{w_i}$

- $\alpha = 0$: $\text{HUF} = \sum_i w_i u_i$

- $\alpha < 0$: 効用の高い人の効用が優先される

世帯共同活動時間の分析への適用

$$\text{Maximize } \text{HUF} = \frac{1}{1-\alpha} \sum_i w_i u_i^{1-\alpha}$$

$$u_i = \sum_j \gamma_{ij} u_{ij} + \sum_i \sum_{j>j} \delta_i \gamma_{ij} \gamma_{ij'} u_{ij} u_{ij'}, \quad j, j' = h, d, a, \text{ and } s$$

$$\text{Subject to } t_{ih} + t_{id} + t_{ia} + t_s = T_i$$

$$u_{ij} = \exp\left(\frac{\mu_j + \sum_k \beta_{jk} x_{ijk}}{\sum_m \kappa_{jm} \tau_{ijm}} + \varepsilon_{ij}\right) \ln(t_{ij})$$

$$u_{is} = \exp\left(\frac{\mu_s + \sum_k \beta_{sk} x_{isk}}{\sum_m \kappa_{sm} \tau_{ism}} + \varepsilon_{is}\right) \ln(t_s)$$

このあたりパラメータが過剰となり推定できるのか個人的には不安
MDCEVのような問題はないのか、効用関数の特性分析が必要？

エージェント間の力関係の推定

(Hensher et al., 2007)

- 複雑なモデルの推定には段階推定が安定的
- 各エージェントの効用関数とそれらを統合した効用関数の各々を推定するデータを得るのが有用

各エージェントの効用: $U_{qj} = \alpha_j + \beta'_{qjk} \times \mathbf{x}_{qjk} + \varepsilon_j$

2人のエージェントの効用:

$$U(a_1a_1) = ASC_{a_1a_1} + \lambda'_{qk} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + \dots) + (\mathbf{1} - \lambda_{qk})' \times (\beta_{1-q}x_{1-q} + \beta_{2-q}x_{2-q} + \dots) + \varepsilon_{11}$$

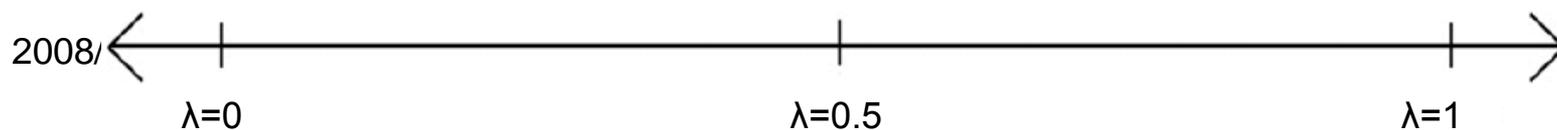
.....

$$U(a_1a_3) = ASC_{a_1a_3} + \lambda'_{qk} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + \dots) + (\mathbf{1} - \lambda_{qk})' \times (\beta_{1-q}x_{1-q} + \beta_{2-q}x_{2-q} + \dots) + \varepsilon_{13}$$

.....

$$U(a_3a_3) = ASC_{a_3a_3} + \lambda'_{qk} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + \dots) + (\mathbf{1} - \lambda_{qk})' \times (\beta_{1-q}x_{1-q} + \beta_{2-q}x_{2-q} + \dots) + \varepsilon_{33}$$

q' Dominates q' Holds Relative Power Equal Power q Holds Relative Power q Dominates



SP調査と推定方法

1. はじめに各被験者に独立にSPで選択させる
2. 各被験者の効用関数の推定

各エージェントの効用: $U_{qj} = \alpha_j + \beta'_{qjk} \times \mathbf{x}_{qjk} + \varepsilon_j$

3. 2人の被験者に相談させ最終的な選択をさせる
4. 各被験者の効用関数は固定し、力関係を表すパラメータ λ を推定

2人のエージェントの効用:

$$U(a_1a_1) = ASC_{a_1a_1} + \lambda'_{qk} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + \dots) + (1 - \lambda_{qk})' \times (\beta_{1-q}x_{1-q} + \beta_{2-q}x_{2-q} + \dots) + \varepsilon_{11}$$

.....

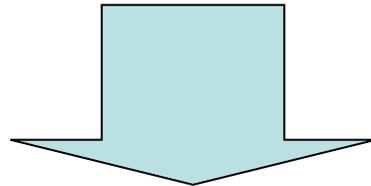
$$U(a_1a_3) = ASC_{a_1a_3} + \lambda'_{qk} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + \dots) + (1 - \lambda_{qk})' \times (\beta_{1-q}x_{1-q} + \beta_{2-q}x_{2-q} + \dots) + \varepsilon_{13}$$

.....

$$U(a_3a_3) = ASC_{a_3a_3} + \lambda'_{qk} \times (\beta_{1q}x_{1q} + \beta_{2q}x_{2q} + \dots) + (1 - \lambda_{qk})' \times (\beta_{1-q}x_{1-q} + \beta_{2-q}x_{2-q} + \dots) + \varepsilon_{33}$$

SP調査の効率化

- 従来は直交計画法で質問回数を減らし、被験者の負担を減らすことに重点



- 推定効率性や設定のリアリティを考慮した設計と推定方法の提案
 - 参照点となる選択肢の活用
 - RPとSPの誤差相関の考慮
 - 属性レベルの最適化による推定誤差最小化

参照点となる選択肢の活用

- コンピュータを活用したSP調査において、連続したSPの設問で、回答結果を次の設問の設定に活用 (Fowkes, 2007)
- RP調査時のサービス水準を変化させてリアリティのあるSPの設定を作成 (Train and Wilson, 2008)
- この時、RP選択時の誤差項はSP選択でもそのまま引き継がれると考えられる

RPの誤差項を引き継ぐSPの推定

(Train and Wilson, 2008)

- RP効用関数 $U_j(x_j, \varepsilon) = \beta x_j + \varepsilon_j$
- SP効用関数 $W_{jt}(\tilde{x}_{jt}^i, \varepsilon, \eta) = \beta \tilde{x}_{jt}^i + \varepsilon_j + \eta_{jt}$
- 尤度関数

$$P_{ki} = \int [L_{1|i}(\varepsilon) \dots L_{T|i}(\varepsilon)] f(\varepsilon | \beta x_i + \varepsilon_i > \beta x_j + \varepsilon_j \forall j \neq i) d\varepsilon \cdot \frac{e^{\beta x_i}}{\sum_j e^{\beta x_j}}$$

where

SP

RP

$$L_{t|i}(\varepsilon) = \frac{e^{\alpha \beta \tilde{x}_{kt}^i + \alpha \varepsilon_{kt}}}{\sum_j e^{\alpha \beta \tilde{x}_{jt}^i + \alpha \varepsilon_j}} \quad \alpha: \text{スケールパラメータ}$$

誤差相関のあるRP-SPモデルと似ているが、RPの部分の誤差項が1つだけ

属性レベルの最適化による 推定誤差最小化 (Rose et al., 2008)

- 推定パラメータの共分散行列の行列式を最小化するようにSPの属性レベルを設定する

$$\text{D-error} = (\det \Omega)^{1/k} = -\frac{1}{N} \left(\det \left(\frac{\partial LL(\beta)^2}{\partial \beta \partial \beta'} \right) \right)^{-1/k}$$

- ただし、パラメータの真値が不明なため行列式の真値は求められない
- パラメータに関する事前の知識を用いて代用
 - 全く情報が無い場合: $\beta = 0$
 - 事前分布が得られる場合: ベイズ推定

おわり

2006年以降に限定しましたので、2006年
までの発展は、以下を参照して下さい

山本俊行(2006)離散選択モデルの最近の話題, 第5回行動モデル夏の学校
(<http://www.trans.civil.nagoya-u.ac.jp/~yamamoto/presentation/SummerSchool2006.ppt>)
Yamamoto, T., Hyodo, T. and Muromachi, Y. (2006) Advances in choice
modeling and Asian perspectives, Proceedings of 11th International
Conference on Travel Behaviour Research, CD-ROM.

参考文献

GEVモデル

- Daly, A. and Bierlaire, M. (2006) A general and operational representation of Generalised Extreme Value models. *Transportation Research Part B*, 40, pp.285-305.
- Abbe, E., Bierlaire, M. and Toledo, T. (2007) Normalization and correlation of cross-nested logit models. *Transportation Research Part B*, 41, pp.795-808.

MMNLモデル

- Hess, S., Train, K.E. and Polak, J.W. (2006) On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit Model for vehicle choice. *Transportation Research Part B*, 40, pp.147-163.
- Bastin, F., Cirillo, C. and Toint, P. (2006) Application of an adaptive Monte Carlo algorithm to mixed logit estimation. *Transportation Research Part B*, 40, pp.577-593.
- Liu, H., He, X. and Recker, W. (2007) Estimation of the time-dependency of values of travel time and its reliability from loop detector data. *Transportation Research Part B*, 41, pp.448-461.

離散連続モデル

- Bhat, C. (2008) The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B*, 42, pp.274-303.
- Bhat, C. and Sen, S. (2006) Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. *Transportation Research Part B*, 40, pp. 35-53.
- Bhat, C., Srinivasan, S. and Sen, S. (2006) A joint model for the perfect and imperfect substitute goods case: Application to activity time-use decisions. *Transportation Research Part B*, 40, pp.827-850.
- Fang, H.A. (2008) A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B*, 42, pp. 736-758.

個人間相互作用のモデル化

- Fukuda, D. and Morichi, S. (2007) Incorporating aggregate behavior in an individual's discrete choice: An application to analyzing illegal bicycle parking behavior, *Transportation Research Part A*, 41, pp. 313-325.
- Zhang, J. and Fujiwara, A. (2006) Representing household time allocation behavior by endogenously incorporating diverse intra-household interactions: A case study in the context of elderly couples, *Transportation Research Part B*, 40, pp. 54-74.
- Hensher, D.A., Puckett, S.M. and Rose, J.M. (2007) Agency decision making in freight distribution chains: Establishing a parsimonious empirical framework from alternative behavioural structures, *Transportation Research Part B*, 41, pp. 924-949.

調査手法

- Fowkes, T. (2007) The design and interpretation of freight stated preference experiments seeking to elicit behavioural valuations of journey attributes. *Transportation Research Part B*, 41, pp.966-980.
- Train, K. and Wilson, W. (2008) Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research Part B*, 42, pp.191-203.
- Rose, J., Bliemer, M.C.J., Hensher, D.A. and Collins, A.T. (2008) Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B*, 42, pp.395-406.