NU TREND
Nagoya University TRansport and ENvironment Dynamics

離散連続選択モデル

名古屋大学　山本俊行

先端的行動モデル？

内容

- 離散選択モデルと機械学習
- 離散連続選択モデル：一つの選択肢を選択するケース
- 離散連続選択モデル：複数個を選択するケース
- 最近の話題

離散選択モデル

- 効用最大化

  意思決定者 n は J_n 個の選択肢集合の中から
  効用 U_{in} が最大となる選択肢 i を選ぶ

  $U_{in} > U_{jn}$ for all j in J_n, $j \neq i$

- ランダム効用

  効用は確定効用 V_{in} と誤差項 ε_{in} からなる

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

  誤差項が確率分布に従う -> 選択は確率的

  $P_{in} = \Pr(U_{in} > U_{jn}$, for all j in J_n, $j \neq i)$

  $\quad = \Pr(\varepsilon_{in} - \varepsilon_{in} < V_{in} - V_{jn}$, for all j in J_n, $j \neq i)$

機械学習

ニューラルネットワーク（NN），サポートベクターマシン（SVM），決定木（DT），アンサンブルラーニング(EL),　．．．

利点	欠点
関数が非線形で柔軟なので予測精度が高い	説明変数が選択に及ぼす影響の解釈が困難
行動原理の仮定が不要	行動原理がなく交通施策等の便益評価に使えない

離散連続モデル

離散的な選択行動と連続量の組み合わせ
- 自動車の車種選択と走行距離
- 活動選択と活動時間
- 電気製品の選択と電力利用量

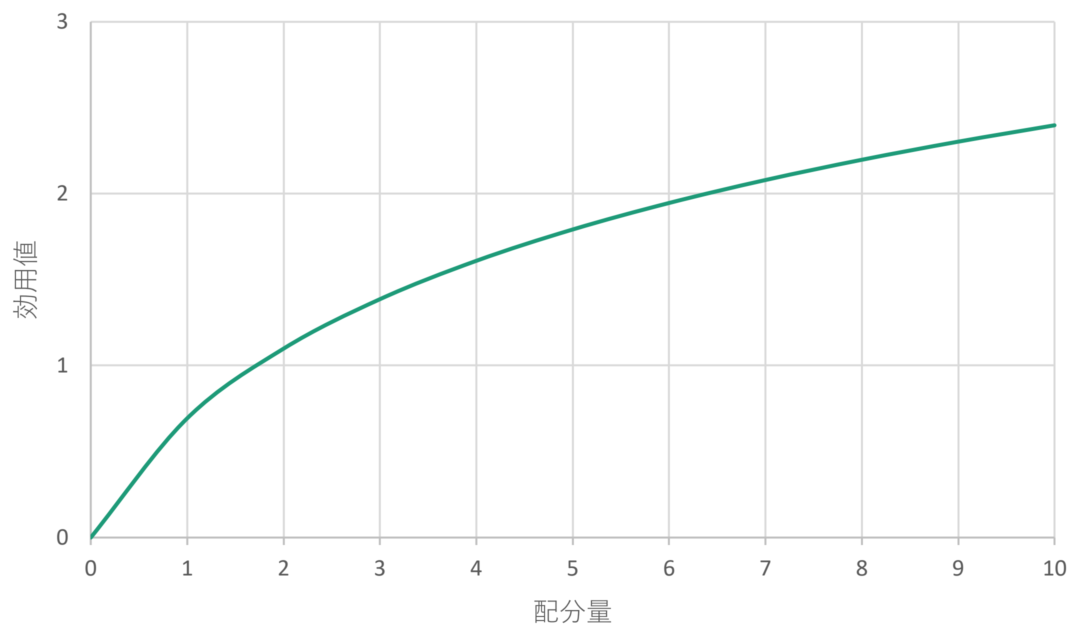
選択しない選択肢の連続量は0
連続量のみで離散的な選択行動も表現可能 ⇨ トビットモデル

$$y_{in}^{*} = \beta X_{in} + \varepsilon_{in}$$
$$y_{in} = \begin{cases} y_{in}^{*} & \text{if } y_{in}^{*} > 0 \\ 0 & \text{if } y_{in}^{*} \leq 0 \end{cases}$$

- 行動原理がないという点で機械学習と同等
（注：効用最大化からトビットモデルを導出している研究もあります）

限界効用逓減の法則

- １杯目のビールが一番おいしい　$\frac{\partial^2 U}{\partial x^2} < 0$
- 何杯も飲むうちに，他のことに予算を使った方が良い段階に達し，飲むのを止める



離散連続選択モデル（福田・力石，2013）

- 効用最大化に基づく離散連続選択行動の記述

$$\text{Max.} \quad U_i = f_i\left(x_{i1}, x_{i2}, \ldots, x_{iJ}\right)$$

x_{ij}: 個人iの財jへの配分量

$$\text{Subject to} \quad \sum_j^J p_j x_{ij} = E_i,\ \forall x_{ij} \geq 0$$

資源量の制約条件

- 選択肢 j が選ばれたという条件付きの間接効用
  - 資源制約下で効用最大化の結果得られる
  - 配分量ではなく制約条件等の関数

$$Y_{ij} = Y_{ij}\left(p_j, E_i, z_{ij}, s_i, e_{ij}\right)$$

- J 個の財から財 j を選択する確率

$$P_{ij} = \mathrm{Prob}\left[\begin{array}{l} Y_{ij}(p_j, E_i, z_{ij}, s_i, e_{ij}) > Y_{il}(p_l, E_i, z_{il}, s_i, e_{il}), \\ \forall 1 \leq l \leq J, l \neq j \end{array}\right]$$

- 非観測特性が加算型で表されると

$$Y_{ij}\left(p_j, E_i, z_{ij}, s_i, e_{ij}\right) = Y_{ij}\left(p_j, E_i, z_{ij}, s_i\right) + e_{ij}$$

- e_{ij}にガンベル分布を仮定するとロジットモデルとなる

ロワの恒等式

- 資源を財 j に使うか，その以外の財に使うかの関係から財 j の需要量は以下で表される

$$x_{ij} = -\frac{\partial Y_{ij} / \partial p_j}{\partial Y_{ij} / \partial E_i} = g_{ij}\left(p_j, E_i, z_{ij}, s_i, e_{ij}\right)$$

- 需要関数 g_{ij} （それを導く Y_{ij}）の特定が重要

例 $$Y_{ij} = \left[\alpha E_i + \beta p_j + \gamma z_{ij} + \theta s_i\right]\exp\left(-\rho p_j\right) + e_{ij}$$

Dubin and McFadden (1984)

制約条件 $E_i \geq p_j x_{ij} + Z_i$ 　z_i は外部の合成財

複数の財への配分

- 離散的選択を離散選択モデル，連続量の選択をロワの恒等式に基づく回帰モデルを適用すれば効用関数の推定が可能
  - 複数の財を選択する場合に適用できない
- キューンタッカー条件に基づく導出

$$L_i = U_i - \lambda\left(\sum_j^J p_j x_{ij} - E_i\right)$$

$$\partial L_i / \partial x_{ij} = \partial U_i / \partial x_{ij} - \lambda p_j = 0 \ \ \text{if} \ \ x_{ij} > 0, \ \forall j$$

$$\partial L_i / \partial x_{ij} = \partial U_i / \partial x_{ij} - \lambda p_j < 0 \ \ \text{if} \ \ x_{ij} = 0, \ \forall j$$

MDCEVモデル（Bhat, 2008）

- 非観測特性がガンベル分布に従うと仮定することでクローズドフォームの尤度関数を導出

$$\mathrm{U}(\boldsymbol{x}) = \sum_k \frac{\gamma_k}{\alpha_k}\left[\exp(\beta' z_k + \varepsilon_k)\right] \cdot \left\{\left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k} - 1\right\}$$

- キューンタッカー条件

複数の財を選択可能

$$\left[\frac{\exp(\beta' z_k + \varepsilon_k)}{p_k}\right]\left(\frac{e_k^*}{\gamma_k p_k} + 1\right)^{\alpha_k - 1} - \lambda = 0, \quad \text{if } e_k^* > 0, \ \ k = 1, 2, \ldots, K,$$

$$\left[\frac{\exp(\beta' z_k + \varepsilon_k)}{p_k}\right]\left(\frac{e_k^*}{\gamma_k p_k} + 1\right)^{\alpha_k - 1} - \lambda < 0, \quad \text{if } e_k^* = 0, \ \ k = 1, 2, \ldots, K.$$

$$e_k = p_k x_k$$

- 対数をとって書き換えると

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

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

$$V_k = \beta' z_k + (\alpha_k - 1) \ln \left(\frac{e_k^*}{\gamma_k p_k} + 1 \right) - \ln p_k \quad (k = 1, 2, 3, \ldots, K).$$

- K個のうちM個の財を選択する確率は以下となる

ロジットモデルと同様の形

$$P(e_1^*, e_2^*, e_3^*, \ldots, e_M^*, 0, 0, \ldots, 0)$$

$$= |J| \int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1 - V_{M+1}} \int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1 - V_{M+2}} \cdots \int_{\tilde{\varepsilon}_{K-1,1}=-\infty}^{V_1 - V_{K-1}} \int_{\tilde{\varepsilon}_{K,1}=-\infty}^{V_1 - V_K}$$

$$g(\underline{V_1 - V_2, V_1 - V_3, \ldots, V_1 - V_M}, \underline{\tilde{\varepsilon}_{M+1,1}, \tilde{\varepsilon}_{M+2,1}, \ldots, \tilde{\varepsilon}_{K,1}}) \mathrm{d}\tilde{\varepsilon}_{K,1} \, \mathrm{d}\tilde{\varepsilon}_{K-1,1} \ldots \mathrm{d}\tilde{\varepsilon}_{M+1,1}.$$

$$\tilde{\varepsilon}_{k1} = \varepsilon_k - \varepsilon_1,$$

積分しないので確率密度

- 確率項にガンベル分布を仮定すると

$$P(e_1^*, e_2^*, e_3^*, \ldots, e_M^*, 0, 0, \ldots, 0) = \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} c_i\right] \left[\sum_{i=1}^{M} \frac{1}{c_i}\right] \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma}\right)^M}\right] (M-1)!$$

$$c_i = \left(\frac{1-\alpha_i}{e_i^* + \gamma_i p_i}\right)$$

- 確率項に一般化ガンベル分布（generalized logit）を適用したり，mixed logit を適用した拡張がなされている
- 離散的選択と連続的選択で別の効用関数を使うケースも（Bhat, 2018）

$$U(\boldsymbol{x}) = \psi_1 x_1 + \sum_{k=2}^{K} \gamma_k \left([\psi_{kd}]^{1(x_k=0)} \times [\psi_{kc}]^{1(x_k>0)} \right) \ln\left\{ \left(\frac{x_k}{\gamma_k} + 1 \right) \right\}$$

限界効用逓減のパラメータ

$$\frac{\gamma_k}{\alpha_k}\left[\exp(\beta' z_k + \varepsilon_k)\right] \cdot \boxed{\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\}}$$

- 限界効用逓減に関するパラメータ（γ_k, α_k）を両方推定することは実務上困難
- 片方を固定した定式化

$\alpha_k \rightarrow 0$のケース $$\gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k}+1\right)$$

$\gamma_k = 1$のケース $$\frac{1}{\alpha_k}\psi_k\left\{(x_k+1)^{\alpha_k} - 1\right\}$$

最近の話題１

- 外部財の限界効用逓減を考慮しない：線形効用関数(Bhat et al., 2020, 2022,　Saxena et al., 2022)

$$U(\boldsymbol{x}) = \boxed{\psi_1^{1-\alpha} x_1} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha} \psi_k^{1-\alpha} \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha} - 1 \right\}$$

$$s.t.\ \ x_1 + \sum_{k=2}^{K} p_k x_k = E,$$

$$\psi_k = \exp(\boldsymbol{\beta}' \boldsymbol{z}_k + \varepsilon_k), k = 1, 2, ..., K,$$

$$L = U(\boldsymbol{x}) + \lambda \left(E - \sum_{k=1}^{K} p_k x_k \right)$$

利点と欠点

- キューンタッカー条件

$$\psi_1^{1-\alpha} - \lambda = 0;$$

外部財への配分量（x_0）が条件式に含まれない

$$\left[\psi_k\left(\frac{x_k^*}{\gamma_k}+1\right)^{-1}\right]^{1-\alpha} - \lambda p_k = 0 \text{ if consumption} = x_k^*\left(x_k^* > 0\right),\ k = 2, 3, \cdots, K,$$

$$[\psi_k]^{1-\alpha} - \lambda p_k < 0 \text{ if } x_k^* = 0,\ k = 2, 3, \ldots, K$$

- 外部財への配分量を観測せずモデル推定可能
  - 外部財への配分量が非常に大きい場合など
- 推定結果として外部財への配分を保証しない

最近の話題２

- 2つの財の間の補完と代替効果（Palma and Hess, 2022）

$$Max_{x_n}\ u_0(x_{n0}) + \sum_{k=1}^{K} u_k(x_{nk}) + \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} u_{kl}(x_{nk}, x_{nl})$$

$$s.t.\ x_{n0}p_{n0} + \sum_{k=1}^{K} x_{nk}p_{nk} = B_n$$

$$u_0(x_{n0}) = \psi_{n0} \log\left(x_{n0}\right)$$

$$u_k(x_{nk}) = \psi_{nk}\gamma_k \log\left(\frac{x_{nk}}{\gamma_k} + 1\right)$$

$$\boxed{u_{kl}(x_{nk}, x_{nl}) = \delta_{kl}\left(1 - e^{-x_{nk}}\right)\left(1 - e^{-x_{nl}}\right)}$$

- 外部財の関数に誤差項を含まない

$$\psi_{n0} = e^{\alpha z_{n0}}$$
$$\psi_{nk} = e^{\beta_k z_{nk} + \varepsilon_{nk}}$$

- キューンタッカー条件

$$Lagr(x) = u_0(x_0) + \sum_{k=1}^{K} u_k(x_k) + \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} u_{kl}(x_k, x_l) - \lambda \left(x_0 + \sum_{k=1}^{K} x_k p_k - B \right)$$

$$\frac{\partial Lagr}{\partial x_0} = 0 \ : \ \frac{\psi_0}{x_0} = \lambda$$
$$\frac{\partial Lagr}{\partial x_k} \leq 0 \ : \ \frac{\psi_k}{\frac{x_k}{\gamma_k} + 1} + e^{-x_k} \sum_{l \neq k} \delta_{kl} \left(1 - e^{-x_l}\right) \leq \lambda p_k$$

- 誤差項にIIDガンベル分布を仮定すると尤度関数がクローズドフォームになる

$$\varepsilon_k \leq -W_k$$
$$W_k = z_k \beta_k - \log\left(\frac{x_k}{\gamma_k} + 1\right) - \log\left(\frac{\psi_0}{x_0} p_k - e^{-x_k} \sum_{l \neq k} \delta_{kl} \left(1 - e^{-x_l}\right)\right)$$

参考文献

福田大輔，力石真（2013）離散-連続モデルの研究動向に関するレビュー，土木学会論文集D3, Vol.69, No.5, pp. I_497-I_510.

Bhat, C.R. (2008) The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions, Transportation Research Part B, Vol. 42, pp. 274-303.

Bhat, C.R. (2018) A new flexible multiple discrete-continuous extreme value (MDCEV) choice model. Transportation Research Part B, Vol. 110, pp. 261-279.

Bhat, C.R., Mondal, A., Asmussen, K., Bhat, A.C. (2020) A multiple discrete extreme value choice model with grouped consumption data and unobserved budgets. Transportation Research Part B, Vol. 141, pp. 196–222.

Bhat, C.R., Mondal, A., Pinjari, A.R., Saxena, S., Pendyala, R.M. (2022) A multiple discrete continuous extreme value choice (MDCEV) model with a linear utility profile for the outside good recognizing positive consumption constraints. Transportation Research Part B, Vol. 156, pp. 28-49.

Dubin, J.A. and McFadden, D.L. (1984) An econometric analysis of residential electric appliance holdings and consumption, Econometrica, Vol.52, pp.345–362.

Palma, D., Hess, S. (2022) Extending the Multiple Discrete Continuous (MDC) modelling framework to consider complementarity, substitution, and an unobserved budget. Transportation Research Part B, Vol. 161, pp. 13-35.

Saxena, S., Pinjari, A.R., Bhat, C.R. (2022) Multiple discrete-continuous choice models with additively separable utility functions and linear utility on outside good: Model properties and characterization of demand functions. Transportation Research Part B, Vol. 155, pp. 526-557.