

Probit , Mixed logit

2016/5/16

スタートアップゼミ#5

B4 後藤祥孝

0. 目次

Probitモデルについて

1. モデル概要
2. 定式化と理解
3. 推定

Mixed logit モデルについて

4. モデル概要
5. 定式化と理解
6. 推定

1. Probit概要

プロビットモデルとは.

効用関数の誤差項に多変量正規分布を仮定したもの.

誤差項には様々な要因が存在するため, 中心極限定理により, 正規分布を仮定するのは自然な発想.

1. Probit概要

ロジットモデルの限界

- ①パラメータで個人の嗜好が考慮できない
- ②IIA特性
- ③非観測項が意思決定者ごとに時間相関がある場合に利用できない。

→GEVモデルの導入により, ②については緩和することができた。

1. Probit概要

Probitモデルの場合,
3つのこと全てについて取り扱える.

- ①各個人についてランダムなパラメータを扱える.
- ②あらゆる代替パターンを許容する.
- ③時間相関誤差のあるデータに適用可能

1. Probit概要

Probitモデルの制約

- ・効用の非観測要素すべてに正規分布の仮定が必要.

基本的には十分な表現が可能であるが、時々不適切な推定となってしまう.

Ex) 価格係数について

正の価格係数を持つ人ができてしまう.

2. Probitの定式化と理解

効用関数の設定

- ・ 確定項と確率項に分ける

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

n:個人を表す

j:選択肢を表す ($j \in J$)

ε_{nj} を要素に持つベクトル ε_n を導入する

$$\varepsilon'_n = (\varepsilon_{n1}, \dots, \varepsilon_{nJ})$$

2. Probitの定式化と理解

誤差項ベクトル ε_n が
平均ベクトル0, 分散共分散行列 Ω の正規分布に
従うとして, ε_n の密度関数は

$$\phi(\varepsilon_n) = \frac{1}{(2\pi)^{J/2} |\Omega|^{1/2}} e^{-\frac{1}{2} \varepsilon_n' \Omega^{-1} \varepsilon_n}$$

(Ω は実際は個人 n によって異なるが添え字を省略している)

2. Probitの定式化と理解

選択確率は

$$\begin{aligned} P_{ni} &= \text{Prob}(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \quad \forall j \neq i) \\ &= \text{Prob}(V_{nj} < \varepsilon_{ni} + V_{ni} - \varepsilon_{nj} \quad \forall j \neq i) \end{aligned}$$

$$= \int_{\varepsilon_{n1}=-\infty}^{\varepsilon_{ni}+V_{ni}-\varepsilon_{n1}} \cdots \int_{\varepsilon_{ni}=-\infty}^{\infty} \cdots \int_{\varepsilon_{nJ}=-\infty}^{\varepsilon_{ni}+V_{ni}-\varepsilon_{nJ}} \phi(\varepsilon_n) d\varepsilon_{nJ} \cdots d\varepsilon_{n1}$$

ε_{ni} についての積分は確率密度関数を全範囲積分している。

2. Probitの定式化と理解

選択確率

$$P_{ni} = \int_{\varepsilon_{n1}=-\infty}^{\varepsilon_{ni}+V_{ni}-\varepsilon_{n1}} \cdots \int_{\varepsilon_{ni}=-\infty}^{\infty} \cdots \int_{\varepsilon_{nJ}=-\infty}^{\varepsilon_{ni}+V_{ni}-\varepsilon_{nJ}} \phi(\varepsilon_n) d\varepsilon_{nJ} \cdots d\varepsilon_{n1}$$

は (J-1) 次の積分となる。

したがってプロビットモデルはオープンフォームであり、シミュレーション法などで推定することとなる。

3. Probitの推定

1. ARシミュレータ (accept-reject)
個人 n の選択肢 i の選択確率を求める。

(1)

平均 0, 分散共分散行列 Ω の正規分布から, J 次元誤差項ベクトルを取り出して r を用いてラベリングする.

ε_n^r : r は順番を表す, 1回目なら $r=1$

(2)

誤差項の値を用いて各選択肢の効用を計算する.

$$U_{nj}^r = V_{nj} + \varepsilon_{nj}^r \quad \forall j$$

3. Probitの推定

(3)

選択肢*i*の効用が他のすべての効用より大きいかどうかを調べる.

I^r という値を定義し,

$U_{ni}^r > U_{nj}^r \forall j \neq i$ のとき, $I^r = 1$ (accept)

それ以外の場合, $I^r = 0$ (reject)

と置く.

(4)

(1)~(3)までを何度も繰り返す, 繰り返し回数をRとする

3. Probitの推定

(5)

シミュレーション確率はacceptの割合として求められる。

$$\check{P}_{ni} = \frac{1}{R} \sum_{r=1}^R I^r$$

(6)

尤度関数を定義し、尤度を計算する。

シミュレーション対数尤度関数

$$SLL = \sum_n \sum_j d_{nj} \log \check{P}_{nj}$$

d_{nj} : j を選択したかどうかのダミー変数

3. Probitの推定

ARシミュレータの弱点

シミュレーション確率 \check{P}_{in} はRがどのような値であっても0となる可能性がある.

⇒対数尤度関数が定義不可能なケース

シミュレーション確率はパラメータに対して滑らかな関数とならない

⇒微分不可能となり, 最尤推定ができない

3. Probitの推定

Smoothed ARシミュレータ

効用から選択を求める際にロジット関数を導入

(1)

平均 0 , 分散共分散行列 Ω の正規分布から, J 次元誤差項ベクトルを取り出して r を用いてラベリングする.

ε_n^r : r は順番を表す, 1回目なら $r=1$

(2)

誤差項の値を用いて各選択肢の効用を計算する.

$$U_{nj}^r = V_{nj} + \varepsilon_{nj}^r \quad \forall j$$

3. Probitの推定

(3)

以下のロジット関数に効用を入れて計算する

$$S^r = \frac{e^{U_{ni}^r/\lambda}}{\sum_j e^{U_{nj}^r/\lambda}}$$

λ はスケールを表す

(4)

(1)~(3)までを何度も繰り返す, 繰り返し回数をRとする

3. Probitの推定

(5)

シミュレーション確率はロジット関数の平均として求められる。

$$\check{P}_{ni} = \frac{1}{R} \sum_{r=1}^R S^r$$

(6)

尤度関数を定義し、尤度を計算する。

シミュレーション対数尤度関数

$$SLL = \sum_n \sum_j d_{nj} \log \check{P}_{nj}$$

d_{nj} : j を選択したかどうかのダミー変数

4. Mixed logit の概要

Mixed logit とは

- ・ 誤差項に正規分布を仮定したプロビット系列のモデルと誤差項にガンベル分布を仮定したモデルを組み合わせたもの。
 - ・ MNLの選択確率のパラメータを確率分布によって扱うことによって母集団内でのばらつきを表現できる。
- 誤差項相関や異分散性を表現することが可能である。

5. Mixed logit の定式化と理解

選択確率式は

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta$$

$$L_{ni}(\beta) = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^J e^{V_{nj}(\beta)}}$$

$L_{ni}(\beta)$: ロジット型確率

$V_{ni}(\beta)$: 効用の確定部分 (β の関数)

$f(\beta)$: 確率密度関数

β : ランダム係数

5. Mixed logit の定式化と理解

Mixed logit

β によって評価されたロジット型確率式 L_{ni}
と
確率分布としての $f(\beta)$ の組み合わせ

実際には β の密度関数 $f(\beta)$ が θ で表されるため,

$$P_{ni} = \int L_{ni}(\beta) f(\beta|\theta) d\beta$$

となり, θ の関数としてあらわされる.

5. Mixed logit の定式化と理解

密度関数 $f(\beta)$ に関して

$f(\beta)$ は離散, 連続どちらもありうる
(ただし多くの場合で連続型を用いる)

$f(\beta)$ の平均を b とすると

$f(\beta) = 1$ for $\beta = b$, 0 for $\beta \neq b$
となるとき, 通常のロジットモデルとなる.

5. Mixed logit の定式化と理解

パラメータが個人 n によって異なると考えるとき
個人ごとに異なる β_n は実際に観測することはできない。
→確率密度関数 $f(\beta)$ によって表現する。

個人 n が J 個の選択肢から選択肢 j を選択したときの効用

$$U_{nj} = \beta_n' x_{nj} + \varepsilon_{nj}$$

(β_n は個人によって異なり, $f(\beta)$ に従うとする)

5. Mixed logit の定式化と理解

β_n が定数であるとするれば確率式はロジットモデルと同様に考えて、

$$L_{ni} = \frac{e^{\beta_n' x_{ni}}}{\sum_j e^{\beta_n' x_{nj}}}$$

と書くことができる。

実際には β_n が $f(\beta)$ に従って変動するため、 β に対して積分の必要がある。

5. Mixed logit の定式化と理解

したがって確率式は以下のようになる

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

5. Mixed logit の定式化と理解

異なる選択肢間の誤差項の相関を考える場合

$$U_{nj} = \alpha' x_{nj} + \mu_n' z_{nj} + \varepsilon_{nj}$$

x_{nj}, z_{nj} : j に関する観測値ベクトル

α : 固定された係数ベクトル

μ : 平均 0 の変動する係数ベクトル

ε_{nj} : *iid*の誤差項

5. Mixed logit の定式化と理解

z_{nj} の項も誤差項で ε_{nj} と合わせて効用の確率項と考えられる.

$$\eta_{nj} = \mu'_n z_{nj} + \varepsilon_{nj}$$

$$\text{Cov}(\eta_{ni}, \eta_{nj}) = E(\mu'_n z_{nj} + \varepsilon_{nj})(\mu'_n z_{ni} + \varepsilon_{ni}) = z'_{nj} W z_{ni}$$

5. Mixed logit の定式化と理解

NestedLogitのアナロジー

ダミー変数の導入

$$\mu_n' z_{nj} = \sum_{k=1}^K \mu_{nk} d_{jk}$$

d_{jk} : 選択肢 j がネスト k に含まれていれば 1, そうでなければ 0

$\mu_{nk} \sim N(0, \sigma_k)$ とする

効用は

$$U_{nj} = \alpha' x_{nj} + \sum_k \mu_{nk} d_{jk} + \varepsilon_{nj}$$

5. Mixed logit の定式化と理解

ネストkに含まれる選択肢の共分散は

$$\text{Cov}(\eta_{ni}, \eta_{nj}) = E(\mu_k + \varepsilon_{ni})(\mu_k + \varepsilon_{nj}) = \sigma_k$$

$$\text{Var}(\eta_{ni}) = E(\mu_k + \varepsilon_{ni})^2 = \sigma_k + \frac{\pi^2}{6}$$

したがってネストk内の相関係数は

$$\frac{\sigma_k}{\sigma_k + \frac{\pi^2}{6}}$$

5. Mixed logit の定式化と理解

全てのネストにおいて各ネストの誤差項の分散が等しいとすると,

(すなわち, $\sigma_k = \sigma, k = 1 \dots K$)

ネストロジットにおいて, ログサム変数がすべてのネストに対して等しくなるということの意味している.

5. Mixed logit の定式化と理解

Mixed logit = IIA特性がない

⇒ P_{ni}/P_{nj} は選択肢 i, j 以外を含む全データに依存する.

ほかの選択肢の m 番目の特性が変化したときの選択確率の変化は

$$E_{ni} x_{nj}^m = -\frac{x_{nj}^m}{P_{ni}} \int \beta^m L_{ni}(\beta) L_{nj}(\beta) f(\beta) d\beta$$

$$= -x_{nj}^m \int \beta^m L_{nj}(\beta) \left[\frac{L_{ni}(\beta)}{P_{ni}} \right] f(\beta) d\beta$$

6. Mixed logit の推定

改めて選択確率を示すと

$$P_{ni} = \int L_{ni}(\beta) f(\beta|\theta) d\beta$$

$$L_{ni}(\beta) = \frac{e^{\beta' x_{ni}}}{\sum e^{\beta' x_{nj}}}$$

であり、確率は θ が与えられれば計算可能である。

6. Mixed logit の推定

以下に推定の流れを示す.

(1) 確率密度関数 $f(\beta|\theta)$ をもとにランダムに β を一つ取り出し, β^r とラベリングする. (r は順番を表す. 1回目なら $r=1$)

(2) 取り出した β^r を用いて $L_{ni}(\beta^r)$ を計算する.

(3) 1, 2を何度も繰り返す, $L_{ni}(\beta^r)$ の平均を取る

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

6. Mixed logit の推定

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

$\check{P}_{ni} > 0$ なので $\ln \check{P}_{ni}$ が定義可能

Simulated log likelihood (対数尤度関数)

$$SLL = \sum_{n=1}^N \sum_{j=1}^J d_{nj} \ln \check{P}_{nj}$$

⇒SLLを最大化する密度関数のパラメータ θ を求める