

Intro 5 - 離散選択モデル 【 評価版 】

本 whitepaper では離散選択モデル — 複数の選択肢の中から 1 つの選択肢を選択する — を扱うコマンド `cmclogit`, `cmmixlogit`, `cmmprobit`, `nlogit` の機能と用法について紹介します。

1. 離散選択用のコマンド
2. `cmclogit`: McFadden の選択モデル
3. `cmsample` の利用
4. `margins` による分析
5. `cmmixlogit`: 混合ロジット選択モデル
6. `cmmprobit`: 多項プロビット選択モデル
7. `nlogit`: ネスト型ロジット選択モデル
8. 他の推定コマンドとの関係
 - 8.1 `clogit` コマンドとの関係
 - 8.2 `mlogit` コマンドとの関係
9. 推定に関する考慮点
 - 9.1 積分点の数
 - 9.2 収束上の問題
 - 9.3 複数選択肢の採択

1. 離散選択用のコマンド

Stata には離散選択モデル (discrete choice models) のフィットを行うためのコマンドが 4 種類用意されています。本セクションではそれらの機能概要と各々の特徴について説明します。

いずれのコマンドにおいても選択肢固有の予測変数とケース固有の予測変数の双方がサポートされます。また選択集合 (choice sets) は unbalanced なものであっても適切に処理されます。これらのコマンドが扱うモデルはいずれもランダム効用モデル (random utility model) — 最大の効用をもたらす選択肢が選択される — に帰着されます。離散選択モデルのランダム効用モデルとしての定式化に関する詳細については [CM] Intro 8 (*mwp-401*) を参照ください。

Stata がサポートする離散選択モデルには 4 種類があるわけですが、それらの違いは概ね IIA (independence of irrelevant alternatives) — 無関係な選択肢からの独立性 — として知られる仮定の扱い方に由来します。IIA の仮定の意味するところを簡潔に述べるなら、選択肢を選択する相対的な確率は別の選択肢が加わったりなくなったりしたとしても変わらないとするものです。例えばあるレストランにおいて鶏肉と牛肉の料理が 1 品ずつメニュー上にあり、どちらが選択されるかの確率は等しいものとします。その上で菜食者向けのオプションが追加されたとするなら、鶏肉と牛肉を選択する確率は減るでしょうが、IIA の仮定が成り立つとするなら、両者の値は依然等しいはずです。仮に牛肉を選ぶ確率の方が大きくなったとしたら、あるいは逆に鶏肉を選ぶ確率の方が大きくなったとしたら、IIA の仮定は崩れることになります。技術的に表現するなら、IIA を仮定するということは誤差項が選択肢をまたがって相関を持つことが許されなくなることを意味します。この IIA の仮定に関する議論については [CM] Intro 8 (*mwp-401*) を参照ください。

- `cmlogit` は条件付きロジスティック回帰 (conditional logistic regression) を用いて McFadden の選択モデルをフィットさせます。McFadden の選択モデルは 4 種類のモデル中で最も素直な定式化に基づいたものと言えます。ただし IIA の仮定が前提となります。
- `cmmixlogit` は選択モデルのために混合ロジット回帰 (mixed logit regression) をフィットさせます。このモデルの場合、選択肢固有の予測変数に対してランダム係数 (random coefficients) を設定することができます。すなわち該当変数に対する係数値は個体ごとに変動することが許容されることとなります。ただし個体ごとの係数値が推定されるわけではありません。係数値は正規分布のような分布に従うことが仮定され、その分布に対するパラメータ値が代りに推定されます。このようなランダム係数を介して選択肢間の相関がモデル化されることとなります。その意味で混合ロジット回帰は IIA の仮定を緩和するものと言えます。
- `cmmprobit` は多項プロビット選択モデル (multinomial probit choice model) をフィットさせます。`cmlogit` と同様、`cmmprobit` はすべての予測変数について固定的な係数値を推定します。しかしそれは誤差項間の相関を直接モデル化することによって IIA の仮定を緩和することが可能です。
- `nlogit` はネスト型ロジット選択モデル (nested logit choice model) をフィットさせます。このモデルでは類似した選択肢 — 誤差項間に相関が存在する可能性が高い — をネスト構造にグルーピングすることができます。上記レストランの例でいうなら、メニューが 7 種類 — 鶏肉料理 3 種、牛肉料理 2 種、菜食者向け料理 2 種 — に広がったとします。このとき料理のタイプごとにグルーピングした構造が設定できるわけです。ネスト型ロジットモデルではネスト内選択肢間の相関を考慮することによって IIA 仮定の緩和を行います。

2. cmclogit: McFadden の選択モデル

評価版では割愛しています。

3. cmsample の利用

評価版では割愛しています。

4. margins による分析

評価版では割愛しています。

5. cmmixlogit: 混合ロジット選択モデル

評価版では割愛しています。

6. cmmprobit: 多項プロビット選択モデル

評価版では割愛しています。

7. nlogit: ネスト型ロジット選択モデル

評価版では割愛しています。

8. 他の推定コマンドとの関係

評価版では割愛しています。

9. 推定に関する考慮点

評価版では割愛しています。

