

Intro 1 - ERM コマンドの紹介 【 評価版 】

本 whitepaper では ERM 機能の概要について紹介します。

1. ERM 機能とは
2. ERM の特徴
3. ERM の構文
4. 正規性の仮定
5. 既存コマンドとの対応

1. ERM 機能とは

ERM とは extended regression model (拡張回帰モデル) の略号で、次のような特徴を有する機能です。

- 線形回帰、区間回帰、プロビット回帰、順序プロビット回帰の機能を提供するコマンドであり、
- 連続型の/2 値の/順序性を持った内生共変量を許容します。

この内生共変量に対しては

- 内生共変量の多項式
- 内生共変量間の交互作用
- 内生共変量と外生共変量間の交互作用
- 内生的標本選択
- ランダムでない外生的/内生的処置割当て
- パネル/グループ内における観測データ間の相関

の機能がサポートされます。

パネル構成ではないデータに対しては `eregress`, `eintreg`, `eprobit`, `eoprobit` の 4 種類のコマンドが用意されています。これに対しパネル構成のデータを対象にしたコマンド `xteregress`, `xteintreg`, `xteprobit`, `xteoprobit` を使うと、パネル/グループ内相関を許容した変量効果モデルのフィットが行えます。

2. ERM の特徴

ERM 系のコマンドは次のようなモデルのフィットを行います。

コマンド	用途
<code>eregress</code>	線形回帰
<code>eintreg</code>	区間回帰
<code>eprobit</code>	2 値アウトカムプロビット回帰
<code>eoprobit</code>	順序アウトカムプロビット回帰
<code>xteregress</code>	変量効果線形回帰
<code>xteintreg</code>	変量効果区間回帰
<code>xteprobit</code>	変量効果 2 値アウトカムプロビット回帰
<code>xteoprobit</code>	変量効果順序アウトカムプロビット回帰

これらのモデルについては [ERM] Intro 2 を参照

ERM 系コマンドはすべて次のような特徴を有しています。

内生共変量

- モデル中の説明変数 — 共変量 (covariates) — は外生/内生のいずれでも構いません。
- 内生共変量は連続 (線形) / 2 値 (プロビット) / 順序 (順序プロビット) 変数のいずれでも構いません。
- 内生共変量は他の共変量 (内生/外生) と交互作用を持つもので構いません。また自身と交互作用を有し多項式を構成するものであっても構いません。
- 内生共変量は他の内生共変量により予測されるものであっても構いません。

内生選択

- ある内生的な理由によりアウトカムが観測されない場合があるといった事態に対してモデルは調整可能です。
- 医療のトライアルにおいて患者は最後の来診をスキップしてしまうことがあり、その場合には最終的なアウトカムが観測されないこととなります。スキップした理由についてはアウトカムと相関を有する可能性があります。
- 経済分野のデータの場合、賃金に関する情報はその人が職に就いている場合にのみ得られることとなります。その場合、職に就かない理由は受け取ることのできる賃金の額と相関を持つことが考えられます。

外生的/内生的処置割当て

- モデルの目的とするところはある処置 — 投与される薬剤や参加する訓練プログラム — の効果を計測するものであることがよく見受けられます。その場合、倫理的な理由により、処置の割当てがランダムとは言えないものとなることがよく起ります。

- 医療のトライアルにおいては観測された特性に基づき、トライアルを利する可能性が高い患者が医者によって選択されるといったことが考えられます。これは外生的処置割当てと呼ばれます。
- 被験者が自発的に参画するといった局面においては、より有益と考える被験者の方が実際に利益を得る可能性が高くなります。そのような認識を規定する要因がすべて観測可能な場合には割当ては外生的なもの — すなわち観測可能な変数によって説明し得るもの — となります。
- 決定要因が観測できないものである場合、処置は内生的なものとなります。割当てに関する方程式中の誤差はアウトカムを規定する方程式中の誤差と相関を持ったものとなります。

パネルデータ/グループデータ

- モデルにはパネル内、あるいはグループ内相関を考慮するための変量効果を含めることができます。
- 経済分野のデータで言うなら、1つの企業に属する観測データは独立とは言えない場合があります。
- 医療の分野で言うなら、1人の患者に対する反復的な計測データは独立とは言えません。
- 変量効果を導入することによってこれらの相関を考慮に入れたモデルが構築できます。

線形回帰の場合、Stataにはこれらの問題を個々に扱うためのコマンドが別に用意されています。また、区間回帰やプロビット、順序プロビットの場合には一部の問題に対応するためのコマンドがやはり別に用意されています。しかしこれらの問題が組み合わされる形で発生した場合、それに対処できるコマンドはこれまでStataには存在しませんでした。ERMではこれらの問題が統一的に扱われるため、たとえ問題が単独のものであったとしても、コマンドの構文はより簡潔で使いやすいものとなっています。

3. ERM の構文

評価版では割愛しています。

4. 正規性の仮定

評価版では割愛しています。

5. 既存コマンドとの対応

評価版では割愛しています。

