

## 標本選択モデル

Heckman モデルとして知られる回帰モデルは標本選択 (sample selection) を伴う線形回帰モデルのことを言います。モデル式は

$$y_j = \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j \quad (14.1)$$

ということで線形回帰の場合と変わりませんが、 $y_j$  はある種の事象が起きなければ観測されないという点にこのモデルの特徴があります。一方、この事象の発生確率は

$$p_j = F(\mathbf{z}_j\boldsymbol{\gamma} + \nu_j) \quad (14.2)$$

という形でモデル化されます。この場合、 $\epsilon$  と  $\nu$  の間には相関があってもよく、また  $\mathbf{z}_j$  の中には  $\mathbf{x}_j$  と共通の変数が含まれていても構いません。

1. heckman コマンド ([R] heckman 参照) はこのようなモデルを最尤法、あるいは Heckman オリジナルの 2 段階プロシジャを用いてフィットさせます。
2. heckprob コマンド ([R] heckprob 参照) は上記の考え方をプロビットモデルに適用したものです。
3. 標本選択を伴うモデルとしては treatment-effects model と呼ばれるモデルもあります。これは内生的に選択される二値の treatment がもたらす他の内生連続変数に対する影響を考慮するものです。このモデルのフィットは treatreg コマンド ([R] treatreg 参照) によって行われます。

■