

2010 年度 学部 エコノメトリックス
第 28 回講義メモ
(訂正版)

2011 年 1 月 26 日

2

§ 6.2 被説明変数が質的データである回帰モデル

被説明変数 y_i が質的データである場合、線形回帰モデル

$$y_i = \tilde{\mathbf{x}}_i' \boldsymbol{\beta} + \epsilon_i$$

を適用することは妥当であろうか。結論から言うと、線形回帰モデルを拡張した**一般化線形モデル** (GLIM: genralized linear model) の枠組みで取り扱う必要がある。本節では、被説明変数 y_i が質的データである**質的変量モデル** (Qualitative Response) について、概説することにした。

3

質的変量モデル

簡単化のために、被説明変数 y_i は、 $\{0, 1\}$ の二値をとる、すなわち

$$y_i = \begin{cases} 1 & i \in I \\ 0 & i \in I^c, \end{cases} \quad i = 1, \dots, n$$

であるものとしよう。ここで集合 I は、投票や購買などの行動を行った人を表している。

質的変量モデルは、 y_i そのものではなく、 $y_i = 1$ となる**確率**を

$$P(y_i = 1) = h(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) \quad (6.2.1)$$

としてモデル化したものである。ここで、 $\tilde{\mathbf{x}}_i' \boldsymbol{\beta} = \sum_{j=1}^k x_{ji} \beta_j$ である。

4

(6.2.1) 式における $h(\cdot)$ ^{*1} は、単調な連続微分可能な関数であり、以下のものがよく用いられる。

(a) 線形確率モデル (LPM: linear probability model)

$$h(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) = \tilde{\mathbf{x}}_i' \boldsymbol{\beta}$$

(b) probit モデル

$$h(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) = \Phi(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) = \int_{-\infty}^{\tilde{\mathbf{x}}_i' \boldsymbol{\beta}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

^{*1} GLIM の定式化 $E(y_i) = \mu_i, g(\mu_i) = \tilde{\mathbf{x}}_i' \boldsymbol{\beta}$ におけるリンク関数 $g(\cdot)$ の逆関数になっている。

5

(c) logit モデル

$$h(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) = \Lambda(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) = \frac{\exp(\tilde{\mathbf{x}}_i' \boldsymbol{\beta})}{1 + \exp(\tilde{\mathbf{x}}_i' \boldsymbol{\beta})}$$

質的変量モデルの性質

まず、質的変量モデルにおける y_i の性質を見てみよう。 $p_i \equiv P(y_i = 1)$ とおくと、 y_i の期待値 (平均) と分散は以下で与えられる。

$$\begin{aligned} E(y_i) &= 1 \times p_i + 0 \times (1 - p_i) = p_i \\ \text{Var}(y_i) &= E(y_i^2) - [E(y_i)]^2 \\ &= p_i - p_i^2 = p_i(1 - p_i) \end{aligned}$$

また、 $h(\cdot)$ は非線形関数を含むので、一般には係数ベクトル $\boldsymbol{\beta}$ は限界効果を表さないことに注意する必要がある。限界効果は、 $h^{(1)}(z) \equiv \partial h(z) / \partial z$

6

とすると

$$\frac{\partial h(\tilde{\mathbf{x}}_i' \boldsymbol{\beta})}{\partial \tilde{\mathbf{x}}_i} = h^{(1)}(\tilde{\mathbf{x}}_i' \boldsymbol{\beta}) \boldsymbol{\beta}$$

となり、評価するポイント $\tilde{\mathbf{x}}_i$ によって異なる値をとることに注意する必要がある。

さらに、この結果より、LPM や probit モデル、logit モデルにおける係数ベクトル $\boldsymbol{\beta}$ の直接の比較は意味がないことが分かる。各モデルにおける係数には次の関係があることが知られている。

$$\begin{cases} \hat{\boldsymbol{\beta}}_{probit} & \cong 1.7 \hat{\boldsymbol{\beta}}_{logit} \\ \hat{\boldsymbol{\beta}}_{LPM} & \cong 0.4 \hat{\boldsymbol{\beta}}_{probit} \\ & \cong 0.25 \hat{\boldsymbol{\beta}}_{logit} \end{cases}$$

7

最小カイ 2 乗推定法 (minimum- χ^2 estimation method)

質的変量モデルの推定は最尤法 (maximum likelihood method) と呼ばれる方法を用いなければならない。但し、説明変数の値が同じ観測値の個体が複数存在する場合には、一般化最小 2 乗法の枠組みで推定を行うことが可能である。これが最小カイ 2 乗推定法である。

サンプルサイズ n の標本が、説明変数が異なる ℓ 個の小グループ ($\ell < n$) に分割できるとしよう。この小グループ j においては、説明変数が $\mathbf{x}_j$ である観測個体が n_j 個存在している。すなわち

$$\begin{pmatrix} y_{1,j} \\ \vdots \\ y_{n_j,j} \end{pmatrix} \text{ 対応する説明変数 } \mathbf{x}_j$$

8

という関係にある。このとき、(6.2.1) 式のモデルは、

$$p_j \equiv P(y_j = 1) = h(\mathbf{x}_j' \boldsymbol{\beta}), \quad j = 1, \dots, \ell \quad (6.2.2)$$

に書き直すことができる。いま、 m_j をこの小グループ j において $y_{i,j} = 1$ ($i = 1, \dots, n_j$) である個数とすると、

$$m_j = \sum_{i=1}^{n_j} y_{i,j}$$

と表すことができるので、 $\hat{p}_j = \frac{m_j}{n_j}$ ($j = 1, \dots, \ell$) を p_j の推定量として用いることができる。

さて、(6.2.2) 式より $h^{-1}(p_j) = \mathbf{x}_j' \boldsymbol{\beta}$ であるから、 $h^{-1}(\hat{p}_j)$ を p_j のまわりで Taylor 展開すると、

$$h^{-1}(\hat{p}_j) \simeq h^{-1}(p_j) + \frac{\partial h^{-1}(p_j)}{\partial p_j} (\hat{p}_j - p_j)$$

9

$$= \mathbf{x}_j' \boldsymbol{\beta} + \nu_j$$

となる。ここで、

$$y_{i,j} \sim IIDBin(1, p_j) \quad \text{for } i = 1, \dots, n_j, \quad j = 1, \dots, \ell$$

であるから、

$$m_j \sim Bin(n_j, p_j)$$

したがって、

$$\begin{aligned} E(\nu_j) &= \frac{\partial h^{-1}(p_j)}{\partial p_j} (E(\hat{p}_j) - p_j) \\ &= \frac{\partial h^{-1}(p_j)}{\partial p_j} \left(\frac{n_j p_j}{n_j} - p_j \right) = 0 \end{aligned}$$

10

$$\sigma_j^2 \equiv \text{Var}(\nu_j) = \left(\frac{\partial h^{-1}(p_j)}{\partial p_j} \right)^2 \frac{\text{Var}(m_j)}{n_j^2} = \left(\frac{\partial h^{-1}(p_j)}{\partial p_j} \right)^2 \frac{p_j(1-p_j)}{n_j} \quad (6.2.3)$$

であることがわかる。これより、攪乱項 ν_j の分散は不均一であるので、重み付最小 2 乗法 (WLS) 推定量

$$\mathbf{b}_{MC} = \left(\sum_{j=1}^{\ell} \sigma_j^{-2} \mathbf{x}_j \mathbf{x}_j' \right)^{-1} \left(\sum_{j=1}^{\ell} \sigma_j^{-2} \mathbf{x}_j h^{-1}(\hat{p}_j) \right) \quad (6.2.4)$$

が効率的な推定量になる。なお、実際の推定においては、(6.2.3) 式において、 p_j を推定量 $\hat{p}_j$ で置き換えた σ_j^2 の推定量を用いる。

また、(6.2.3) 式より $\hat{p}_j = 0$ または $\hat{p}_j = 1$ となるケースでは $\hat{\sigma}_j^2 = 0$ となるため、最小カイ 2 乗推定法を用いることが出来ないので、注意が必要である。

11

さまざまな質的変量モデルと σ_j^2

(6.2.3) より、 $\sigma_j^2 = \text{Var}(\nu_j)$ は $h(\cdot)$ の関数形によって異なることが分かる。以下では分析によく用いられる $h(\cdot)$ と、それに対応する σ_j^2 について整理することにしよう。

(a) 線形確率モデル (LPM) の場合

$$p_j = h(\mathbf{x}_j' \boldsymbol{\beta}) = \mathbf{x}_j' \boldsymbol{\beta}$$

であるから、 $h^{-1}(p_j) = p_j$ 。したがって、

$$\frac{\partial h^{-1}(p_j)}{\partial p_j} = 1$$

12

が得られる。よって、

$$\sigma_j^2 = \frac{p_j(1-p_j)}{n_j}$$

になる。

(b) probit モデルの場合

$$p_j = h(\mathbf{x}_j' \boldsymbol{\beta}) = \Phi(\mathbf{x}_j' \boldsymbol{\beta})$$

である。ここで、 $\Phi(\cdot)$ は標準正規分布の分布関数

$$\Phi(x) = \int_{-\infty}^x \phi(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt$$

を表している。(なお、 $\phi(\cdot)$ は標準正規分布の密度関数。)

13

$h^{-1}(p_j) = \Phi^{-1}(p_j)$ であるから、

$$\frac{\partial h^{-1}(p_j)}{\partial p_j} = \frac{1}{\phi(\Phi^{-1}(p_j))}$$

が得られる (なぜなら $dx/dy = 1/(dy/dx)$ であるから)。よって、

$$\sigma_j^2 = \frac{p_j(1-p_j)}{[\phi(\Phi^{-1}(p_j))]^2 n_j}$$

になる。

(c) logit モデルの場合

$$p_j = h(\mathbf{x}'_j \boldsymbol{\beta}) = \Lambda(\mathbf{x}'_j \boldsymbol{\beta}) = \frac{\exp(\mathbf{x}'_j \boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_j \boldsymbol{\beta})}$$

であるから、

$$h^{-1}(p_j) = \log \left(\frac{p_j}{1-p_j} \right)$$

14

となる。これより、

$$\begin{aligned} \frac{\partial h^{-1}(p_j)}{\partial p_j} &= \left(\frac{1-p_j}{p_j} \right) \left(\frac{1}{1-p_j} - \frac{p_j}{(1-p_j)^2} \right) \\ &= \left(\frac{1-p_j}{p_j} \right) \left(\frac{1}{(1-p_j)^2} \right) \\ &= \frac{1}{p_j(1-p_j)} \end{aligned}$$

が得られる。よって、

$$\sigma_j^2 = \frac{1}{n_j p_j (1-p_j)}$$

になる。