第 3 講　ダミー (説明) 変数 [1]

ダミー説明変数は次のような状況で使われる。

1. 定数項（切辺）に違いがある。
2. 係数 (傾き) に違いがある。
3. 連立方程式の係数制約として入る。
4. 回帰係数の安定性テストの目的で用いる。

以下では、それぞれの役割について説明する。

3.1　定数項に入るダミー変数

説明変数が量的な情報ではなく、質的な情報のみを含んでいる場合（例えば教育水準や性別、年齢等）にはダミー変数をその代理とすることが多い。このようなダミー変数を用いることの背後には、そのダミーで代表されるグループの行動の差は定数項（切辺）だけに反映されて、傾きの違いはないという仮定がある。

例えば、賃金（所得）と学歴の関数が次のような 2 つのグループの間に推計されるとしよう。

$$y = \begin{cases} \alpha_1 + \beta x + u & \text{第 1 グループ} \\ \alpha_2 + \beta x + u & \text{第 2 グループ} \end{cases} \tag{3.1}$$

これらの式を合併すると次のように表せる。

$$y = \alpha_1 + (\alpha_2 - \alpha_1)D + \beta x + u \tag{3.2}$$

ここで

$$D = \begin{cases} 1 & \text{第 2 グループの場合} \\ 0 & \text{第 1 グループの場合} \end{cases}$$

この D がダミー変数であり、(3.2) からも明らかなように、このダミー変数はグループ 1 とグループ 2 の間の切辺の差を表している。

もしグループ分けが 3 つの場合には、同様にして次のように表せる。

$$y = \begin{cases} \alpha_1 + \beta x + u & \text{第 1 グループ} \\ \alpha_2 + \beta x + u & \text{第 2 グループ} \\ \alpha_3 + \beta x + u & \text{第 3 グループ} \end{cases} \tag{3.3}$$

これらの式を合併すると、

$$y = \alpha_1 + (\alpha_2 - \alpha_1)D_1 + (\alpha_3 - \alpha_1)D_2 + \beta x + u \tag{3.4}$$

ここで

[1] 本稿は Maddala(1997),chap 8 を参照。

$$D_1 = \begin{cases} 1 & \text{第２グループの場合} \\ 0 & \text{第１と第３グループの場合} \end{cases}$$
$$D_2 = \begin{cases} 1 & \text{第３グループの場合} \\ 0 & \text{第１と第２グループの場合} \end{cases}$$

ここでも、傾きと誤差項の分布は等しいと仮定している。

上の例から明らかなように、ダミー変数の数はグループの数より常に１つ少ない。これは、ダミー変数が基準になるグループ（ここでは第１グループ）からの切辺の差という形で表されているからである。実証研究において、どのグループを基準にするかということは形式上は関係がないが、解釈上はきわめて重要である。

さらに具体的な例について考えてみよう。

消費（C）と所得（Y）に関するモデルに、次のような属性情報がある。

（１）S： 家計主の性別：男、女

（２）A: 家計主の年齢：25 歳以下、25 － 50 歳、50 歳以上の３分類

（３）E: 家計主の教育：高卒以下、高卒、大卒以上の３分類

それぞれの属性をダミー変数で表す。

$$D_1 = \begin{cases} 1 & \text{性別 男} \\ 0 & \text{性別 女} \end{cases}$$
$$D_2 = \begin{cases} 1 & \text{年齢 25 歳以下} \\ 0 & \text{年齢 それ以外} \end{cases}$$
$$D_3 = \begin{cases} 1 & \text{年齢 25 － 50 歳} \\ 0 & \text{年齢 それ以外} \end{cases}$$
$$D_4 = \begin{cases} 1 & \text{学歴 高卒以下} \\ 0 & \text{学歴 それ以外} \end{cases}$$
$$D_5 = \begin{cases} 1 & \text{学歴 高卒} \\ 0 & \text{学歴 それ以外} \end{cases}$$

これらすべてを組み込んだ消費関数は、次のように表せる。

$$C = \alpha + \beta Y + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 D_3 + \gamma_4 D_4 + \gamma_5 D_5 + u \qquad (3.5)$$

それぞれの家計の属性を代入すれば、個々の家計の切辺を求めることができる。しかし実証研究では、ダミーに係る計数をみて、性別、年齢別、学歴別に消費行動に有意な違いがあるかどうか、そしてそれらが基準ケースと比べてどのように違うかを判断することが多い。

このようなダミー説明変数としてマクロ時系列でよく用いられるものに季節（四半期）ダミー、月次ダミーなどがある。歴史的にみてマクロ時系列に明らかなシフトが起こった場合（例えば 1973 年第１次石油ショックや 1985 年のプラザ合意）には、その時期を境にダミー変数を加えることもある。またダミーは解釈可能な範囲内に抑えるべきであるし、ダミー間に相関がある場合にはそれらをまとめたダミー変数を導入すべきである。適切なダミーを導入することで β の推計量が大きく変化することもある。

3.2 計数に係るダミー変数

次のようなモデルを考えよう。

$$\begin{aligned} y_1 &= \alpha_1 + \beta_1 x_1 + u_1 \quad \text{第１グループ} \\ y_2 &= \alpha_2 + \beta_2 x_2 + u_2 \quad \text{第２グループ} \end{aligned}$$

これら２本の式を合併すると次のように表せる。

$$y = \alpha_1 + (\alpha_2 - \alpha_1)D_1 + \beta_1 x + (\beta_2 - \beta_1)D_2 + u \tag{3.6}$$

ここで

$$D_1 = \begin{cases} 0 & \text{第１グループ} \\ 1 & \text{第２グループ} \end{cases}$$
$$D_2 = \begin{cases} 0 & \text{第１グループ} \\ x_2 & \text{第２グループの } x \text{ の値} \end{cases}$$

D_1 は切辺の違いを表し、D_2 は傾きの違いを表している。

さらに具体的な問題を考えてみよう。３期間のモデルで、第２期には切辺のみが変化し、第３期には切辺も傾きも変化した場合には、それぞれの期間モデルは次のように表せる。

$$\begin{array}{lll} y_1 = & \alpha_1 + \beta_1 x_1 + u_1 & \text{第１期} \\ y_2 = & \alpha_2 + \beta_1 x_2 + u_2 & \text{第２期} \\ y_3 = & \alpha_3 + \beta_2 x_3 + u_3 & \text{第２期} \end{array}$$

これらを合併すると次のように表せる。

$$y = \alpha_1 + (\alpha_2 - \alpha_1)D_1 + (\alpha_3 - \alpha_1)D_2 + \beta_1 x + (\beta_2 - \beta_1)D_3 + u \tag{3.7}$$

ここで

$$D_1 = \begin{cases} 1 & \text{第２期の観察値} \\ 0 & \text{その他} \end{cases}$$
$$D_2 = \begin{cases} 1 & \text{第３期の観察値} \\ 0 & \text{その他} \end{cases}$$
$$D_3 = \begin{cases} 1 & \text{第１期と第２期の観察値} \\ 0 & \text{第３期の } x \text{ の値} \end{cases}$$

ここでは、誤差項は均一分散であることを仮定している。

(3.7) は行列を使って次のように書き換えることもできる。

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} + \beta_1 \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \tag{3.8}$$

これを次のように表す。

$$y = \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \beta_1 D_4 + \beta_2 D_5 + u \tag{3.9}$$

(3.9) のダミー変数のとりうる価は (3.8) より明らかである。(3.8) あるいは (3.9) から明らかなように、すべてに共通の定数項は含まれていないことに注意されたい。

3.3 連立方程式の計数制約に係るダミー変数

需要方程式の計数に係る需要の交叉弾力性の対称性制約などをダミー変数を使って表現することもできる（詳しい説明は省略する）。

3.4 回帰係数の安定性テストに用いるダミー変数

次のモデルを考えよう。

$$\begin{array}{ll} y_1 = & \alpha_1 + \beta_1 x_1 + \gamma_1 z_1 + u_1 \quad \text{第1期} \\ y_2 = & \alpha_2 + \beta_2 x_2 + \gamma_2 z_2 + u_2 \quad \text{第2期} \end{array}$$

これらを合併すると、次のように表現できる。

$$y = \alpha_1 + (\alpha_2 - \alpha_1)D_1 + \beta_1 x + (\beta_2 - \beta_1)D_2 + \gamma_1 z + (\gamma_2 - \gamma_1)D_3 + u \quad (3.10)$$

ここで

$$D_1 = \begin{cases} 1 & \text{第2期} \\ 0 & \text{第1期} \end{cases}$$

$$D_2 = \begin{cases} x_2 & \text{第2期の x の値} \\ 0 & \text{第1期} \end{cases}$$

$$D_3 = \begin{cases} z_2 & \text{第2期の z の値} \\ 0 & \text{第1期} \end{cases}$$

制約なしの残差平方和 (RSS) は、(3.10) を推計することによって得られる。制約付きの残差平方和については、ダミー変数を次の仮定に基づいて消去する。

仮　　説	消去するダミー
（1）すべての計数は等しい	D_1, D_2, D_3
$\alpha_1 = \alpha_2,\ \beta_1 = \beta_2,\ \gamma_1 = \gamma_2$	
（2）切辺のみが異なる	$D_2,\ D_3$
$\beta_1 = \beta_2, \quad \gamma_1 = \gamma_2$	
（3）切辺と Z の係数が異なる	D_2
$\beta_1 = \beta_2$	

制約なしのＲＳＳと制約付きのＲＳＳをＦ検定することによって、その制約が有意かどうかをテストする。これは Chow テストと呼ばれる。

$$F = \frac{(RSS - RSS_1)/n_2}{RSS_1/(n_1 - k - 1)} \quad (3.11)$$

ここで $F(n_2, n_1 - k - 1)$ は n_2 と $n_1 - k - 1$ の自由度に従う。RSS は $n_1 + n_2$ の観察値に対する残差平方和であり、RSS_1 は n_1 の観察値に対する残差平方和である。

n_1、n_2 への分解の仕方にルールがあるわけではない。

$$D_i = \begin{cases} 1 & \text{観察値 } n_1 + i \quad i = 1, 2, \cdots, n_2 \\ 0 & \text{その他} \end{cases}$$

ダミー変数の係数が 0 かどうかをテストすることによって n_2 を決める。

補論 3－A　最尤法

最尤法とは、パラメータの推定を行う際に「観察されたデータの組み合わせが起きる確率を最大にするように推定値を決める」方法である。この推定方法は数学的に表現すると、観察されたデータの組み合わせが起きる確率を**尤度関数（likelihood function）**というパラメータ（θ）の関数として表し、その関数$L(\theta)$を最大化するということである。

すなわち、

$$f(X_1,\cdots,X_n,\theta)=\prod_{i=1}^{n}f(X_1,\theta)=L(\theta\mid X) \tag{3-A-12}$$

ここで X は観察されたデータを表す。対数をとると、

$$\ln L(\theta\mid X)=\sum_{i=1}^{n}\ln f(X_i,\theta) \tag{3-A-13}$$

となる。この対数尤度を最大化するための必要条件は

$$\frac{\partial \ln L(\theta)}{\partial\theta}=0 \tag{3-A-14}$$

と表せる。これを満たすようにθを決めることによって最尤推定は次のような特性を持つ。[2]

1. 一致推定量（consistency）
2. 漸近的に正規分布に従う（asymptotic normality）
3. 漸近的に有効（asymptotic efficiency）
4. パラメータの不変性（invariance）

問題点としては、

1. 誤差項の分布型を特定化する必要がある。
2. 小標本では不偏ではない（biased）場合がある。

古典的線形回帰モデルの最尤推定量を求めてみよう。
次のようなモデルを考える。

$$y_i=\alpha+\beta x_i+u_i\quad u_i\sim 1N(0,\sigma^2) \tag{3-A-15}$$

観察値の累積密度関数（the joint density of the observations）は、次のようになる。

$$f(y_1,\cdots,y_n)=\prod_{i=1}^{n}\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}}\exp\left[-\frac{1}{2\sigma^2}(y_1-\alpha-\beta x_i)^2\right] \tag{3-A-16}$$

これを尤度関数（the likelihood function）と捉えて$L(\alpha\cdot\beta,\sigma^2)$と表す。対数をとり

$$\begin{aligned}\ln L&=\sum_{i=1}^{n}\left[-\frac{1}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}(y_i-\alpha-\beta x)^2\right]\\&=C-\frac{n}{2}\ln\sigma^2-\frac{Q}{2\sigma^2}\end{aligned} \tag{3-A-17}$$

[2]それぞれの特性の詳細な証明については、Greene（1997）、Chapter 4 を参照されたい。

ここで $C = -n/2\ \ln(2x), \quad Q = \Sigma(y_{\mathsf{i}} - \alpha - \beta x_{\mathsf{i}})^2$ である。

最尤法とは、lnL を最大化することであるが、(3-A-6) からも明らかなように、これはQ を最小化することと同値であり、それはすなわち最小２乗法による $\hat{\alpha}, \hat{\beta}$ の推定と同じことである。$\partial \ln L/\partial\sigma = 0$ を解くと、

$$\hat{\sigma}^2_{\mathsf{ML}} = \frac{\hat{Q}}{n} = \frac{RSS}{n} \tag{3-A-18}$$

ここで RSS=the residual sum of squares（残差２乗和）である。古典回帰モデルでは分散の不偏推定量 S^2 は $RSS/(n-2)$ で与えられる。両者の関係は

$$\hat{\sigma}^2_{\mathsf{ML}} = \frac{S^2(n-2)}{n} \tag{3-A-19}$$

で表される。サンプル n が小さい場合には古典的モデルに対して下方バイアスを持つが、n が大きくなると、$(n-2)/n$ は１に収束し、分散の偏りが解消される。このような場合には推定量は漸近的に不偏（asymptotic unbiased）といえる。