非線形関数の回帰 (2)

別所俊一郎

2006 年 6 月 7 日

Today's attraction

- 非線形関数の回帰分析の続き：交差項
- ダミー変数との交差項・連続変数との交差項
- 他の変数を追加したときの分析例

交差項（interaction term）とは

例：

- クラスあたりの児童数の変化が学校区の平均的なテストの点数に与える影響は、各学校区の「英語を母国語としない児童の比率」に依存するかもしれない
- 年齢が増えることによる賃金の増加の程度は、男女間で差があるかもしれない

2 つの説明変数の相互作用

- 説明変数 X_1 の被説明変数 Y への効果の大きさが、他の説明変数 X_2 の値に依存
- X_1 がダミー変数で、X_2 もダミー変数
- X_1 が連続変数で、X_2 がダミー変数で
- X_1 が連続変数で、X_2 も連続変数

ダミー変数とダミー変数の交差項

2 つのダミー変数を D_{1i}, D_{2i} とおくと、交差項のあるモデルは

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 \underbrace{(D_{1i} \cdot D_{2i})}_{\text{交差項}} + u_i$$

ダミー変数で場合わけしてみると、

$$\begin{aligned} E[Y_i|D_{1i} = d, D_{2i} = 0] &= \beta_0 + \beta_1 d \\ E[Y_i|D_{1i} = d, D_{2i} = 1] &= \beta_0 + \beta_1 d + \beta_2 + \beta_3 d \end{aligned}$$

それゆえ、

$$E[Y_i|D_{1i} = d, D_{2i} = 1] - E[Y_i|D_{1i} = d, D_{2i} = 0] = \beta_2 + \beta_3 d$$

D_{2i} の効果は D_{2i} の大きさに依存

ダミー変数と連続変数の交差項

- 切片のみダミー変数が入るケース

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + u_i$$

場合わけすると

$$\begin{aligned} E[Y_i|D=0] &= \beta_0 + \beta_1 X_i \\ E[Y_i|D=1] &= \beta_0 + \beta_2 + \beta_1 X_i \end{aligned}$$

β_2 は D によって切片が異なる程度を表す

- 傾きのみダミー変数が入るケース

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 (X_i \cdot D_i) + u_i$$

場合わけすると

$$\begin{aligned} E[Y_i|D=0] &= \beta_0 + \beta_1 X_i \\ E[Y_i|D=1] &= \beta_0 + (\beta_1 + \beta_2) X_i \end{aligned}$$

β_2 は D によって傾きが異なる程度を表す

- 切片と傾きにダミー変数が入るケース

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 (X_i \cdot D_i) + u_i$$

場合わけすると

$$\begin{aligned} E[Y_i|D=0] &= \beta_0 + \beta_1 X_i \\ E[Y_i|D=1] &= \beta_0 + \beta_2 + (\beta_1 + \beta_3) X_i \end{aligned}$$

– β_2 は D によって切片が異なる程度を表す

– β_3 は D によって傾きが異なる程度を表す

– 直線が一致するかどうかの検定：$H_0 : \beta_2 = 0, \beta_3 = 0$ を F 検定

– 傾きが一致するかどうかの検定：$H_0 : \beta_3 = 0$ を t 検定

– 切片が一致するかどうかの検定：$H_0 : \beta_2 = 0$ を t 検定

- 傾きと切片についての検定の結果が一致しないこともある
  - $H_0 : \beta_2 = 0$ と $H_0 : \beta_3 = 0$ をともに棄却できないにもかかわらず、$H_0 : \beta_2 = 0, \beta_3 = 0$ を棄却できるときがある
  - 原因は、推定量の標準誤差が大きいため
  - いずれかはっきりとはしないが、片方のみがゼロに等しいことが示唆される

連続変数同士の交差項

X_1 が Y に与える効果の大きさが X_2 の値の大きさにも依存

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 (X_{1i} \cdot X_{2i}) + u_i$$

X_1 の変分 ΔX_1 が Y に与える効果 ΔY

$$\Delta Y = (\beta_1 + \beta_3 X_2)\Delta X_1, \quad \Longrightarrow \quad \frac{\Delta Y}{\Delta X_1} = \beta_1 + \beta_3 X_2$$

同様に、X_2 の変分 ΔX_2 が Y に与える効果 ΔY

$$\Delta Y = (\beta_2 + \beta_3 X_1)\Delta X_2, \quad \Longrightarrow \quad \frac{\Delta Y}{\Delta X_2} = \beta_2 + \beta_3 X_1$$

すなわち、β_3 は X_1 と X_2 がそれぞれ 1 単位ずつ増えたときに、各説明変数の直接的な効果以外の効果の大きさを示す

連続変数同士の交差項

X_1 が ΔX_1 だけ増加し、X_2 が ΔX_2 だけ増加したときの Y の増分 ΔY

$$\Delta Y = (\beta_1 + \beta_3 X_2)\Delta X_1 + (\beta_2 + \beta_3 X_1)\Delta X_2 + \beta_3 \Delta X_1 \Delta X_2$$

ここで、

- $(\beta_1 + \beta_3 X_2)$ は X_2 一定のときの ΔX_1 の効果
- $(\beta_2 + \beta_3 X_1)$ は X_1 一定のときの ΔX_2 の効果
- $\beta_3 \Delta X_1 \Delta X_2$ は X_1 と X_2 の相互作用の効果

対数変換された変数を使っていれば、弾力性が量に依存することになる