

2

2010 年度 学部 エコノメトリックス 第 4 回講義メモ (増補改訂版)

2010 年 10 月 13 日

2.3 最小 2 乗推定量の分布 (続)

仮定 1+ 仮定 5 のときの a, b の分布

仮定 1 および仮定 5 を満たしている時、 a, b の分布は簡単かつ厳密に導出できる。前回示したように、 $b - \beta = \sum_{i=1}^n w_i \epsilon_i$ で表すことができる。

仮定 5 より、 $\epsilon_i \sim iid N(0, \sigma^2)$ であり、仮定 1 から w_i は非確率変数であることが得られるので、正規分布の再生性より、

$$b - \beta \sim N\left(0, \sigma^2 \sum_{i=1}^n w_i^2\right) \quad (2.3.2)$$

3

になる。書きなおすと

$$b \sim N\left(\beta, \sigma^2 \sum_{i=1}^n w_i^2\right) \quad (2.3.3)$$

のように、 b の分布が求まる。なお、 $\sum_{i=1}^n w_i^2$ は

$$\sum_{i=1}^n w_i^2 = \frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

と表されることに注意。

同様に、 a の分布を求めよう。 a は、前回の授業の結果より、

$$\begin{aligned} a - \alpha &= \frac{1}{n} \sum_{i=1}^n \epsilon_i - (b - \beta) \bar{X} \\ &= \frac{1}{n} \sum_{i=1}^n \epsilon_i - \left(\sum_{i=1}^n w_i \epsilon_i \right) \bar{X} \end{aligned}$$

4

$$= \sum_{i=1}^n \left(\frac{1}{n} - w_i \bar{X} \right) \epsilon_i$$

となるので、正規分布の再生性から

$$a - \alpha \sim N\left(0, \sigma^2 M\right) \quad (2.3.4)$$

となる。ここで M は

$$\begin{aligned} M &\equiv \sum_{i=1}^n \left(\frac{1}{n} - w_i \bar{X} \right)^2 \\ &= \frac{1}{n} - 2 \sum_{i=1}^n w_i \bar{X} + \sum_{i=1}^n w_i^2 \bar{X}^2 = \frac{1}{n} + \bar{X}^2 \sum_{i=1}^n w_i^2 \\ &= \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 + \bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{\frac{1}{n} \sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \end{aligned}$$

である。

5

仮定 1+ 仮定 2' のときの a, b の分布

この場合には、中心極限定理を使って $n \rightarrow \infty$ のときの a, b の分布 (漸近分布という) を求めることになる。ただ、 a, b の分布を導出するときに用いる中心極限定理は、よく知られている iid ケースのものとは異なる。

復習のために、 iid ケースの中心極限定理 (Lindeberg=Lévy 定理) を見てみよう。

6

定理 2.3.1 Lindeberg=Lévy 定理

Z_i ($i = 1, \dots, n$) を、 $E(Z_i) = 0$, $\text{Var}(Z_i) = E(Z_i^2) = \tau^2$ である独立な確率変数であるとする。いま、 $Z_i^* \equiv \frac{1}{\sqrt{n\tau}} Z_i$ とし、その和を次のように定義する。

$$S_n = \sum_{i=1}^n Z_i^*$$

$$\text{Var}(S_n) = E(S_n^2) = \sum_{i=1}^n E\left[\left(\frac{1}{\sqrt{n\tau}} Z_i\right)^2\right] = 1$$

このとき、

$$S_n \xrightarrow{d} N(0, 1) \quad \text{as } n \rightarrow \infty$$

が成立する。

7

Lindeberg=Lévy 定理では、 Z_i^* が iid であった。しかし、 b のケースでは、 $w_i \epsilon_i$ は独立ではあるが、分散は i によって異なる（不均一分散）。したがって、この場合には Lindeberg 定理と呼ばれる中心極限定理を用いる。

定理 2.3.2 Lindeberg CLT

Z_i ($i = 1, \dots, n$) を、 $E(Z_i) = 0$, $\text{Var}(Z_i) = E(Z_i^2) = \tau_i^2$ である独立な確率変数であるとする。いま、 Z_i の和 $S_n = \sum_{i=1}^n Z_i$ について

$$\text{Var}(S_n) = E(S_n^2) = E\left[\left(\sum_{i=1}^n Z_i\right)^2\right] = \sum_{i=1}^n \tau_i^2 = 1$$

を満たしているとする。このとき、

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{|Z_i| > \varepsilon} Z_i^2 f(Z_i) dZ_i = 0 \quad \forall \varepsilon > 0 \quad (2.3.5)$$

8

が満たされていれば、

$$S_n \xrightarrow{d} N(0, 1) \quad \text{as } n \rightarrow \infty$$

が成立する。

実際には (2.3.5) 式の条件のチェックは簡単でない。そこで、次の定理がある。

定理 2.3.3 Liapunov 定理

(2.3.5) 式の十分条件は、

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n E|Z_i|^{2+\delta} = 0, \quad \text{for some } \delta > 0 \quad (2.3.6)$$

である。

9

(ここから補足)

定理 2.3.2 および定理 2.3.3 を推定量 b に適用することを考えよう。 $b - \beta = \sum_{i=1}^n w_i \epsilon_i$ であり、また $\text{Var}(b) = \sigma^2 \sum_{i=1}^n w_i^2$ であるから、

$$Z_i = \frac{w_i \epsilon_i}{\sqrt{\sigma^2 \sum_{i=1}^n w_i^2}}$$

とおくと、

$$\tilde{B}_n \equiv \frac{1}{\sqrt{\sigma^2 \sum_{i=1}^n w_i^2}} (b - \beta) = \sum_{i=1}^n Z_i$$

と表すことができる。よって Z_i の 2 次より大きい絶対モーメントについて (2.3.6) 式を満たしていれば、Lindeberg の中心極限定理が成立し、

$$\tilde{B}_n \xrightarrow{d} N(0, 1) \quad \text{as } n \rightarrow \infty$$

10

が成立する。書きなおせば、 $n \rightarrow \infty$ のとき、

$$b \xrightarrow{d} N\left(\beta, \sigma^2 \sum_{i=1}^n w_i^2\right)$$

に従う。また、同様に a についても

$$a \xrightarrow{d} N(\alpha, \sigma^2 M) \quad \text{as } n \rightarrow \infty$$

が得られる。(上式の M は前に定義したものと同一。)