応用計量経済分析 TA セッション
第 9 回：区間推定

TA：北村友宏 *

2015 年 12 月 8 日

1 区間推定

- 母数が含まれる領域を定めることを**区間推定**（interval estimation）という。
  - ⋆ e.g.,「母平均は 90% の確率でこの値からこの値までの範囲にあるだろう」と定める。
- 母数を θ とし、それが $100(1-\alpha)\%$ の確率で L 以上、U 以下の範囲に含まれるとする。つまり、

$$P(L \leq \theta \leq U) = 1 - \alpha$$

とする。この $100(1-\alpha)\%$ を**信頼度**とよび、区間 $[L, U]$ を θ の信頼度 $100(1-\alpha)\%$ の**信頼区間**（confidence interval）とよぶ。区間推定では α を決めたうえで、L と U を求める。
  - ⋆ $100(1-\alpha)\%$ には 90% や 95% や 99% が入る。$\alpha = 0.1$ とすれば $100(1-\alpha)\% = 90\%, \alpha = 0.05$ とすれば $100(1-\alpha)\% = 95\%, \alpha = 0.01$ とすれば $100(1-\alpha)\% = 99\%$.
  - ⋆ $1-\alpha$ を信頼度とする場合もある。
  - ⋆ 信頼度を信頼係数とよぶこともある。
  - ⋆ 確率分布を連続型（正規分布、t 分布、カイ二乗分布など）と仮定した場合、

$$P(L < \theta < U) = 1 - \alpha$$

として (L, U) を信頼区間としてもよい。

2 母平均の区間推定（母分散が未知の場合）

例題 1. $N\left(\mu, \sigma^2\right)$ の母集団から抽出した大きさ 16 の無作為標本 $(X_1, X_2, \ldots, X_{16})$ を考える。（母平均 μ と）母分散 σ^2 は**未知**である。標本平均を計算すると $\bar{x} = 9$ であり、標本（不偏）分散を計算すると $s^2 = 4$ であった。母平均 μ の信頼度 90% の信頼区間を求めなさい。

Step 1：標本平均の分布を求める

$(X_1, X_2, \ldots, X_{16})$ は $N\left(\mu, \sigma^2\right)$ からの無作為標本なので、任意の $i = 1, 2, \ldots, 16$ について

$$X_i \sim N\left(\mu, \sigma^2\right)$$

* Email: kitamu.tom@gmail.com URL: http://tomkitamura.html.xdomain.jp

である。標本の大きさは 16 なので、標本平均の分布は

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{16}\right)$$

である。

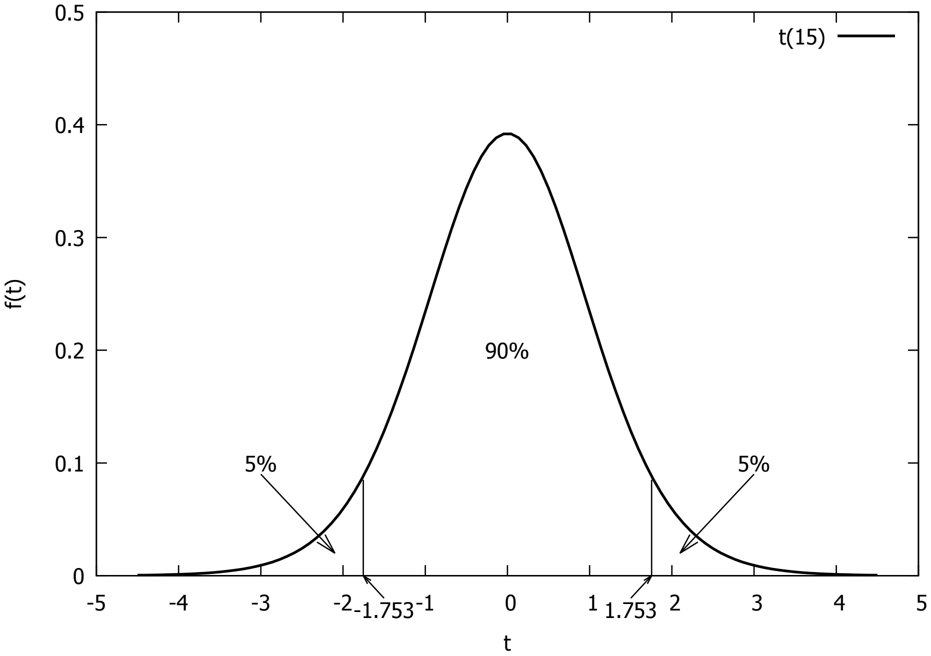
Step 2：標本平均を標準化し、その分布を求める（母分散が未知なので置き換えが必要）

標本平均 $\overline{X}$ を標準化し、母分散を標本分散で置き換えると、

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/16}} \sim t(15)$$

となる。※標本分散を使っているため、正規分布ではなく t 分布であることに注意！

Step 3：「T が（？）以上、（？）以下になる確率が 90%」という式を書き、「μ が・・・」という式に直す



t 分布表より、$t(15)$ において $P(|T| > t) = 0.1$, つまり $P(T > t) = 0.05$ を満たす t は 1.753 なので、

$$
\begin{aligned}
&P\left(-1.753 \le \frac{\overline{X} - \mu}{\sqrt{S^2/16}} \le 1.753\right) = 0.9 \\
\Leftrightarrow &P\left(-1.753\sqrt{\frac{S^2}{16}} \le \overline{X} - \mu \le 1.753\sqrt{\frac{S^2}{16}}\right) = 0.9 \\
\Leftrightarrow &P\left(-\overline{X} - 1.753\sqrt{\frac{S^2}{16}} \le -\mu \le -\overline{X} + 1.753\sqrt{\frac{S^2}{16}}\right) = 0.9 \\
\Leftrightarrow &P\left(\overline{X} + 1.753\sqrt{\frac{S^2}{16}} \ge \mu \ge \overline{X} - 1.753\sqrt{\frac{S^2}{16}}\right) = 0.9 \\
\Leftrightarrow &\underbrace{P\left(\overline{X} - 1.753\sqrt{\frac{S^2}{16}} \le \mu \le \overline{X} + 1.753\sqrt{\frac{S^2}{16}}\right)}_{\mu\text{が}\left[\overline{X}-1.753\sqrt{\frac{S^2}{16}},\overline{X}+1.753\sqrt{\frac{S^2}{16}}\right]\text{に入る確率}} = \underbrace{0.9}_{90\%}.
\end{aligned}
$$

が成り立つ。

Step 4：標本平均と標本分散の実現値を代入し、整理する

$\overline{X}$ に実現値 $\overline{x} = 9$ を、S^2 に実現値 $s^2 = 4$ を代入して信頼区間を書くと、

$$
\left[9 - 1.753\sqrt{\frac{4}{16}}, 9 + 1.753\sqrt{\frac{4}{16}}\right]
$$

となる。ここで、

$$
\begin{aligned}
9 - 1.753\sqrt{\frac{4}{16}} &= 9 - 1.753\sqrt{\frac{1}{4}} = 9 - 1.753 \cdot \frac{1}{2} = 9 - 0.8765 = 8.1235, \\
9 + 1.753\sqrt{\frac{4}{16}} &= 9 + 1.753\sqrt{\frac{1}{4}} = 9 + 1.753 \cdot \frac{1}{2} = 9 + 0.8765 = 9.8765
\end{aligned}
$$

なので、μ の信頼度 90% の信頼区間は $[8.1235, 9.8765]$ である。