応用計量経済分析 TA セッション
第 14 回：回帰係数の t 検定

TA：北村友宏 [*]

2016 年 1 月 19 日

1　検定統計量

- 定数項を含む単回帰モデル

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + u_i,\ \ i = 1, 2, \ldots, n, && (1)\\
u_i &\sim iidN\left(0, \sigma^2\right) && (2)
\end{aligned}
$$

  において、係数（回帰係数ともいう）β_0, β_1 の OLS 推定量をそれぞれ、$\hat{\beta}_0, \hat{\beta}_1$ とする。
  - ⋆ 簡単化のため、x_i は非確率変数とする。
- x_i と y_i の間に相関があるか（$\beta_1 \neq 0$ か）を検定するには、H_0 と H_1 をそれぞれ、

$$
\begin{aligned}
H_0 &: \beta_1 = 0,\\
H_1 &: \beta_1 \neq 0
\end{aligned}
$$

  と設定する。また、モデルの定数項が 0 ではないか（$\beta_0 \neq 0$ か）を検定するには、H_0 と H_1 をそれぞれ、

$$
\begin{aligned}
H_0 &: \beta_0 = 0,\\
H_1 &: \beta_0 \neq 0
\end{aligned}
$$

  と設定する。
- $\beta_1 \neq 0$ かの検定、$\beta_0 \neq 0$ かの検定をするための検定統計量は、それぞれ

$$
\begin{aligned}
&\frac{\hat{\beta}_1 - 0}{SE\left(\hat{\beta}_1\right)} \sim t(n-2),\\
&\frac{\hat{\beta}_0 - 0}{SE\left(\hat{\beta}_0\right)} \sim t(n-2).
\end{aligned}
$$

  - ⋆ $SE\left(\hat{\beta}_1\right)$ は $\hat{\beta}_1$ の、$SE\left(\hat{\beta}_0\right)$ は $\hat{\beta}_0$ の標準誤差。
  - ⋆ 自由度 $n-2$ の 2 は、推定するパラメータの数（ここでは β_0 と β_1）。
  - ⋆ (2) の仮定（誤差項が正規分布に従う）は、上記の検定統計量を導出するために必要。

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

2　仮説検定

例題 1.　ある農産物の収穫量と肥料の使用量との関係を調べるため、以下の単回帰モデルを仮定し、観測値数 20 個のデータを用いて OLS で推定した。

$$harvest_i = \beta_0 + \beta_1 fertilizer_i + u_i,\ i = 1, 2, \ldots, 20,$$
$$u_i \sim iidN\left(0, \sigma^2\right).$$

$harvest_i$ は農産物の収穫量 (kg)、$fertilizer_i$ は肥料の使用量 (kg) である。また、$fertilizer_i$ は非確率変数であるとする。推定結果は次のようになった。

	係数	標準誤差
定数項	6.15	2.05
肥料の使用量	0.52	0.13

この結果から、農産物の収穫量と肥料の使用量の間に相関があるかを、有意水準 1% で検定しなさい。

Step 1：H_0 と H_1 を設定する

帰無仮説と対立仮説をそれぞれ、

$$H_0 : \beta_1 = 0,$$
$$H_1 : \beta_1 \neq 0$$

と設定する。

Step 2：検定統計量を選択し、H_0 のもとでの分布を求める

$H_0 : \beta_1 = 0$ が真であると仮定する。また、β_1 の OLS 推定量を $\hat{\beta}_1$ とし、その標準誤差を $SE\left(\hat{\beta}_1\right)$ とする。$fertilizer_i$ が非確率変数であるという仮定および誤差項が正規分布に従うという仮定から、H_0 のもとでの検定統計量は、

$$\frac{\hat{\beta}_1 - 0}{SE\left(\hat{\beta}_1\right)} \sim t(18)$$

である。

Step 3：有意水準を設定する

上記の検定統計量を用い、有意水準 1% の検定を行う。

Step 4：受容域・棄却域・有意点を定める

$t(18)$ に従う確率変数を T とする。t 分布表より、$P(T > 2.878) = 0.005$, つまり $P(|T| > 2.878) = 0.01$ なので、有意水準 1% の両側検定における、$t(18)$ に従う検定統計量の受容域は、

$$[-2.878, 2.878]$$

であり、棄却域は、

$$(-\infty, -2.878), (2.878, \infty)$$

である。よって、有意点は ± 2.878 となる。

Step 5：検定統計値を求め、H_0 の受容・棄却を判断する

問題の表から、β_1 の OLS 推定値は 0.52、標準誤差は 0.13 である。よって、検定統計値は、

$$\frac{0.52-0}{0.13}=\frac{0.52}{0.13}=4$$

となる。$2.878<4$ なので、検定統計値は棄却域に入る。これは、仮に $\beta_1=0$（H_0 が真）であれば、4 という検定統計値が実現する確率は 1% 以下にすぎないので、$H_0:\beta_1=0$ が疑わしいことを意味する。
したがって、$H_0:\beta_1=0$ は有意水準 1% で棄却され、農産物の収穫量と肥料の使用量が相関をもつ可能性がある。

- この、有意水準 1% での H_0 の棄却を、「係数 β_1 は 1% の有意水準で**有意**である」や「係数 β_1 は 1% の有意水準で**統計的に有意**に 0 と異なる」などと表現する。
- この結果から、肥料の使用量が 1kg 増加すると、収穫量は平均的に 0.52kg 増加することがわかる。
- もし $H_0:\beta_1=0$ が受容された場合は、両者が相関をもつとはいえないことになる。この場合、「係数 β_1 は 1% の有意水準で有意ではない」などと表現する。