第 10 回　重回帰モデル：統計的推測(2)

（復習）H_0：$\beta_1 = -1$ を H_1：$\beta_1 < -1$ に対して有意水準 1 %で検定しなさい。

$$\underset{(0.32)}{\log(\text{price}) = 11.08} - \underset{(0.117)}{0.954}\log(\text{nox}) - \underset{(0.043)}{0.134}\log(\text{dist}) + \underset{(0.019)}{0.255}\ \text{rooms} - \underset{(0.006)}{0.052}\ \text{stratio}$$

[2] 統計的有意性 vs. 経済的有意性

変数 x の係数 $\hat{\beta}$
- 統計的有意性：仮説検定で $\beta = 0$ が棄却される。
- 経済的有意性：x の 1 単位の変化は y を大きく変化させる。

（例）不良品の割合 scrap の決定要因（標本の大きさ n = 29）

$$\log(\text{scrap}) = \underset{(5.69)}{12.46} - \underset{(0.023)}{0.029}\ \text{hrsemp} - \underset{(0.453)}{0.962}\log(\text{sales}) + \underset{(0.407)}{0.761}\log(\text{employ})$$

hrsemp：労働者一人当たり訓練時間，sales：売上高，employ：労働者の数

検証する仮説：hrsemp の係数の「有意性」

統計的有意性：t 統計量 ＝ $-0.029 \div 0.023 = -1.26$　⇒ 影響小

経済的有意性：訓練を 1 時間増やすと，不良品の割合は 2.9%下がる ⇒ 影響大

（推定結果を検討するときの方針）

1）統計的有意性を検討する。

2）① 統計的に有意　⇒　符号条件（理論，常識で判断）と経済的有意性を検討

② 統計的に非有意　a）符号条件を満たす　⇒ 経済的有意性を検討

b）符号条件を満たさない ⇒ それ以上検討しない

[3] 信頼区間

定理 4.2 より，$t(n-k-1)$の上側 2.5%の臨界値を c とすると，次の関係が得られる。

$$P\left(-c \le \frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \le c\right) = 0.95$$ （t 統計量が区間 $[-c, c]$ に入る確率は 0.95）

$$\Rightarrow \quad P(\hat{\beta}_j - c \times se(\hat{\beta}_j) \le \beta_j \le \hat{\beta}_j + c \times se(\hat{\beta}_j)) = 0.95$$

パラメータ β_j の **95%信頼区間**： $\hat{\beta}_j - c \times se(\hat{\beta}_j) \le \beta_j \le \hat{\beta}_j + c \times se(\hat{\beta}_j)$

（意味）信頼区間は確率的な区間で，100 回作れば 95 回 β_j を含むと信頼できる。

（仮説検定との関係） $H_0 : \beta_j = a$ を両側検定するとき， $T = \dfrac{\hat{\beta}_j - a}{se(\hat{\beta}_j)}$ として

T が有意水準 5%の棄却域に入らない。 ⇔ a が 95%信頼区間に入る。
（H_0 は有意水準 5%で棄却されない）

T が有意水準 5%の棄却域に入る。 ⇔ a が 95%信頼区間に入らない。
（H_0 は有意水準 5%で棄却される）

（例）企業の売上 sales が企業の研究開発支出 rd に及ぼす影響（profmarg : 利潤率）

$$\underset{}{\log(rd)} = \underset{(0.47)}{-4.38} + \underset{(0.060)}{1.084}\ \log(sales) + \underset{(0.0218)}{0.0217}\ profmarg, \quad n = 32$$

log(sales)の係数 β_1 の 95%信頼区間は，t(29)の臨界値 $c = 2.045$ より

$$1.084 - 2.045 \times 0.060 \le \beta_1 \le 1.084 + 2.045 \times 0.060 \ \Rightarrow\ 0.961 \le \beta_1 \le 1.21$$

信頼区間は 1 を含む ⇒ $H_0 : \beta_1 = 1$ は $H_1 : \beta_1 \ne 1$ に対して 5%で棄却されない。

[4] 複数の係数に関する仮説検定（F 検定）

帰無仮説が複数の係数に関する仮説である場合，F 検定を利用する。

（例）メジャーリーグ選手の年俸 salary を説明する要因の同時有意性

$$\log(\text{salary}) = \beta_0 + \beta_1 \text{years} + \beta_2 \text{games} + \beta_3 \text{bavg} + \beta_4 \text{hruns} + \beta_5 \text{rb} + u \quad \cdots \quad ⑤$$

years：在籍年数，games：出場試合数，bavg：通算打率，hruns：本塁打数，rb：打点

H_0 ：$\beta_3 = 0$，$\beta_4 = 0$，$\beta_5 = 0$ （β_3，β_4，β_5 は同時に 0 になる）

H_1 ：β_3，β_4，β_5 のうち，少なくとも一つは 0 ではない。

（注意点）・$\beta_3 = 0$，$\beta_4 = 0$，$\beta_5 = 0$ を個別に t 検定しても H_0 は検証できない。

・帰無仮説で仮定される関係を**制約**と呼ぶこともある。

係数に関する r 本の**制約**を同時に検証する場合，次の手順でおこなう。

1) 制約なしの推定をおこない，残差平方和 URSS をもとめる。
（＝ もとの回帰式）

2) 制約つきの推定をおこない，残差平方和 RRSS をもとめる。
（＝ 帰無仮説の関係をもとの回帰式に代入した回帰式）

3) 次の関係を利用して制約を検定する。

（r 本の制約を検定する F 統計量）

$$F = \frac{\dfrac{RRSS - URSS}{r}}{\dfrac{URSS}{n-k-1}} = \frac{RRSS - URSS}{URSS}\,\frac{n-k-1}{r} \sim F(r, n-k-1)$$

（意味）制約が妥当であれば，制約をつけても残差平方和は有意に増えない。

（例）⑤式におけるβ_3，β_4，β_5の同時有意性の検定

メジャーリーグ選手 353 人分のデータを使って次の制約を検定する。

制約：$\beta_3 = 0$，$\beta_4 = 0$，$\beta_5 = 0$　（制約の数：r = 3）

<u>制約なし</u>の推定：⑤式をそのまま推定して残差平方和 RSS を得る。

$$\underset{}{\log(\text{salary})} = \underset{(0.29)}{11.19} + \underset{(0.0121)}{0.0689}\ \text{years} + \underset{(0.0026)}{0.0126}\ \text{games}$$

$$+ \underset{(0.00110)}{0.00098}\ \text{bavg} + \underset{(0.0161)}{0.0144}\ \text{hruns} + \underset{(0.0072)}{0.0108}\ \text{rb}, \quad \text{RSS} = 183.186$$

<u>制約つき</u>の推定：⑤式に $\beta_3 = 0$，$\beta_4 = 0$，$\beta_5 = 0$ を代入してから推定して，残差平方和 RSS を得る。

$$\log(\text{salary}) = \underset{(0.11)}{11.22} + \underset{(0.0125)}{0.0713}\ \text{years} + \underset{(0.0013)}{0.0202}\ \text{games}, \qquad \text{RSS} = 198.311$$

r = 3, n = 353, k = 5 と **F 分布表**より，有意水準 5%のときの棄却域を作る。

URSS = 183.186, RRSS = 198.311, r = 3, n = 353, k = 5 より，

$$F = \frac{(198.311 - 183.186)}{183.186}\frac{353 - 5 - 1}{3} = 9.55 \quad \Rightarrow \text{ F は棄却域に入る}$$

（結論）仮説 H_0 は有意水準 5%で棄却される。

※ β_3，β_4，β_5 は個別には有意ではないが，なぜグループでは有意になるか？

⇒ 選んだ説明変数のグループ間に多重共線関係がある。

＜p.1の例＞
Jeffrey Wooldridge
Introductory Econometrics: A Modern Approach (3rd ed.)
South-Western, Division of Thomson Learning
のp.143のEXAMPLE 4.7を利用
＜p.2の例＞
Jeffrey Wooldridge
Introductory Econometrics: A Modern Approach (3rd ed.)
South-Western, Division of Thomson Learning
のp.146のEXAMPLE 4.8を利用
＜p.3 の例＞
Jeffrey Wooldridge
Introductory Econometrics: A Modern Approach (3rd ed.)
South-Western, Division of Thomson Learning
のp.151-152 の本文を引用