



第7章 同時検定

藪友良

『入門 実践する計量経済学』

(東洋経済新報社)

PPT

- 同時検定
- 同時検定の一般的なケース
- F 検定
- 構造変化の検定
 - 構造変化点が既知
 - 構造変化点が未知

同時検定

▪ 同時検定とは何か

--- 複数の仮説をまとめて同時に検定する

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + u$$

$$H_0 : \beta_1 = 0, \beta_2 = 0$$

$$H_1 : H_0 \text{ は誤りである}$$

結合仮説

▪ 同時検定の意義

--- 個別の t 検定では、有意水準のコントロールが難しい

--- 複数の仮説をまとめて同時に検定する必要がある

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + u$$

$H_0 : \beta_1 = 0, \beta_2 = 0$ 、 $H_1 : H_0$ は誤りである

▪ 個別 t 検定では有意水準のコントロールが難しい

--- $H'_0 : \beta_1 = 0$ に対応する t 統計量(t_1)

$H'_0 : \beta_2 = 0$ に対応する t 統計量(t_2)

有意水準5%に対応する臨界値($t_{0.05}$)

事象Aを $|t_1| > t_{0.05}$ 、事象Bを $|t_2| > t_{0.05}$ とし、 $A \cup B$ なら H_0 を棄却

--- 有意水準は、 H_0 が正しいもとで、 H_0 を誤って棄却する確率

$$P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}$$

確率の加法定理

--- $P\{A\} = P\{B\} = 0.05$ 、 $P\{A \cap B\} = ?$

--- t_1 と t_2 の相互関係の強さがわからない

• 同時検定の一般的ケース

$$Y = \alpha + \underbrace{\beta_1 X_1 + \dots + \beta_{K-q} X_{K-q}}_{\text{最初の } K-q \text{ 個}} + \underbrace{\beta_{K-q+1} X_{K-q+1} + \dots + \beta_K X_K}_{\text{最後の } q \text{ 個}} + u$$

---最後の q 個の説明変数が Y の動きを説明するうえで説明力を持っているかに関心がある

• 仮説検定

$$H_0 : \beta_{K-q+1} = 0, \beta_{K-q+2} = 0, \dots, \beta_K = 0$$

$$H_1 : H_0 \text{ は誤りである (少なくとも1つは0でない)}$$

--- H_0 が正しいもとで、最後の q 個の説明変数は意味がないため、回帰式から除外できる(除外制約は計 q 個ある)

***F*検定**

$$H_0 : \beta_{K-q+1} = 0, \beta_{K-q+2} = 0, \dots, \beta_K = 0$$

$H_1 : H_0$ は誤りである

- **制約なしモデル (H_1 が正しい前提でのモデル)**

$$Y = \alpha + \beta_1 X_1 + \dots + \beta_{K-q} X_{K-q} + \beta_{K-q+1} X_{K-q+1} + \dots + \beta_K X_K + u$$

--- 残差2乗和を SSR_1 とする

- **制約ありモデル (H_0 が正しい前提でのモデル)**

$$Y = \alpha + \beta_1 X_1 + \dots + \beta_{K-q} X_{K-q} + u$$

--- 残差2乗和を SSR_0 とする

- **残差2乗和の関係**

--- 説明変数が多いほうが当てはまりは良い

$$0 \leq SSR_0 - SSR_1$$

--- H_0 が正しいなら、制約ありモデルでもデータの動きを捉えられる

$$SSR_0 - SSR_1 \text{ は小さい}$$

--- H_1 が正しいなら、制約ありモデルではデータの動きを捉えられない

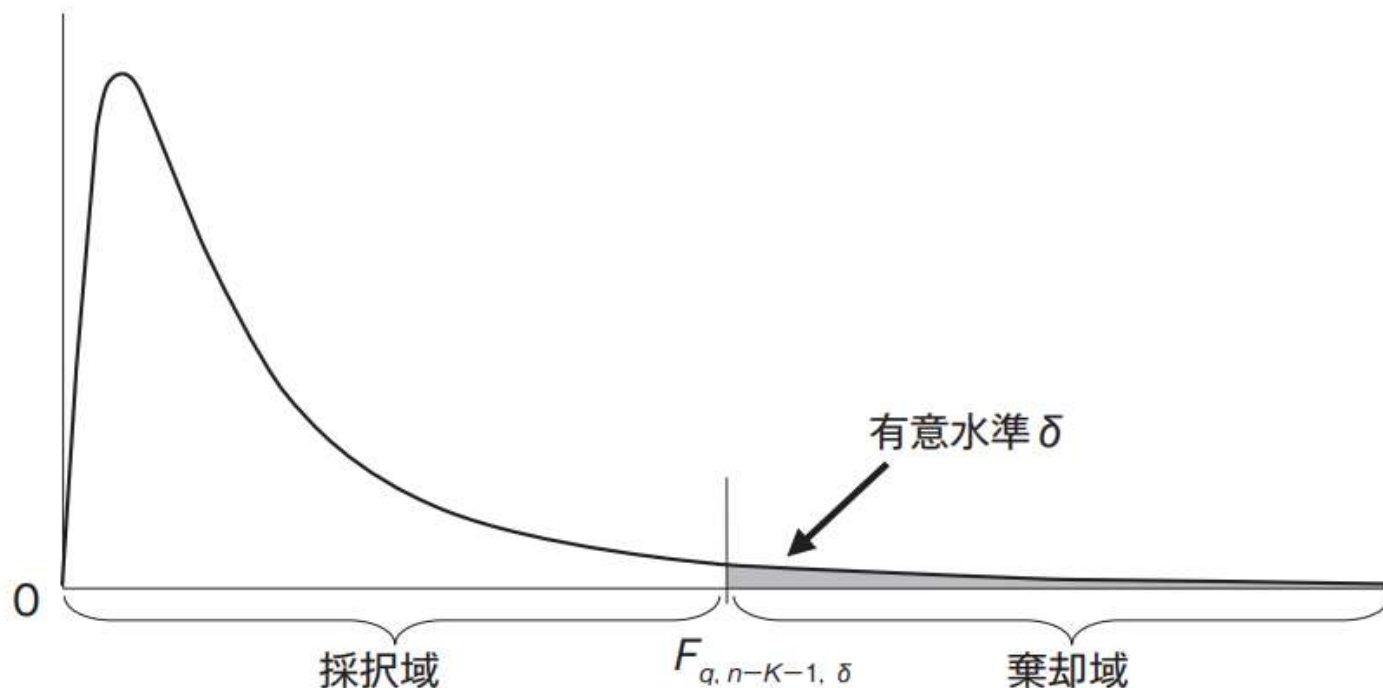
$$SSR_0 - SSR_1 \text{ は大きい}$$

• F 統計量

$$F = \frac{(SSR_0 - SSR_1)/q}{SSR_1/(n - (K + 1))} \sim F(q, n - K - 1) \quad \text{under } H_0$$

- H_0 が正しいなら $SSR_0 - SSR_1$ は0に近く、 H_1 が正しいなら大きくなる
- F 検定では、 F 値が0に近い値なら H_0 は採択され、十分に大きければ(臨界値を超えれば) H_0 は棄却される

図 7 - 1 帰無仮説 H_0 が正しい状況での F 統計量の分布



(例)統計ソフトで計算されるF値とは？

・仮説は次のように設定される

$$H_0 : \beta_1 = 0, \beta_2 = 0, \dots, \beta_K = 0$$

$$H_1 : H_0 \text{ は誤りである}$$

- H_0 : 説明変数はどれも意味がない
- H_1 : 説明変数のいずれかは意味がある
- 実証分析では、ほとんど用いられない仮説であるが、統計ソフトではOutputの1つになっている

(例) 効率市場仮説

市場が効率的なら、過去の情報から株価の変化率を予測できない

・ 日経平均株価指数(1991/1から2018/3までの日次データ)

$$ds_t = \alpha + \beta_1 ds_{t-1} + \beta_2 ds_{t-2} + \beta_3 ds_{t-3} + \beta_4 s_{t-1}^{MA} + u_t$$

--- s_t : t 日の日経平均終値(対数)

--- $ds_t = 100(s_t - s_{t-1})$: 日経平均の変化率

--- $s_t^{moving} = \frac{1}{31} \sum_{i=0}^{30} s_{t-i}$: 過去1ヶ月の移動平均

--- $s_t^{MA} = 100(s_t - s_t^{moving})$: 移動平均からの乖離率

・ 効率市場仮説が正しいなら

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

・推定結果

$$ds_t = -0.0014 - 0.035ds_{t-1} - 0.023ds_{t-2} + 0.0036ds_{t-3} - 0.0024s_t^{MA}$$

(0.018) (0.013) (0.013) (0.013) (0.005)

--- 個々の t 検定では ds_{t-1} , ds_{t-2} だけが有意

--- ds_{t-1} , ds_{t-2} の係数は負であり、株価の動きには反転する傾向がある

・同時検定

--- F 値を計算すると3.33となり、 $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ は棄却される(F 統計量は自由度(4, 6678 - 4 - 1)であり、有意水準5%の臨界値は2.37)

---効率的市場仮説が成立しているとは言えない

構造変化の検定

構造変化点が既知

- 時系列データでは、パラメータに構造変化が生じる
---構造変化があるとき、それを考慮しないとパラメータの推定にバイアスが生じる

- 構造変化のモデル

$$Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_K X_{Kt} + u_t \quad \text{for } t = 1, 2, \dots, T_B$$

$$Y_t = \alpha' + \beta'_1 X_{1t} + \dots + \beta'_K X_{Kt} + u_t \quad \text{for } t = T_B + 1, \dots, T$$

- 帰無仮説と対立仮説

$$H_0: \alpha = \alpha', \beta_1 = \beta'_1, \dots, \beta_K = \beta'_K \quad (\text{構造変化なし})$$

$$H_1: H_0 \text{ は誤りである (構造変化あり)}$$

F検定(Chow検定)

・ 構造変化を考慮したモデル

$$Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_K X_{Kt} + \theta_0 D_t + \theta_1 D_t X_{1t} + \dots + \theta_K D_t X_{Kt} + u_t$$

--- D_t : 標本期間の前半($t = 1, 2, \dots, T_B$)なら0とし、

後半($t = T_B + 1, \dots, T$)なら1とするダミー変数

--- θ は前半と後半のパラメータの差

$$\theta_0 = \alpha' - \alpha, \theta_1 = \beta'_1 - \beta_1, \dots, \theta_K = \beta'_K - \beta_K$$

--- $H_0: \theta_0 = 0, \theta_1 = 0, \dots, \theta_K = 0$ (構造変化なし)

$H_1: H_0$ は誤りである(構造変化あり)

制約なしモデル(H_1 が正しい前提)

$$Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_K X_{Kt} + \theta_0 D_t + \theta_1 D_t X_{1t} + \dots + \theta_K D_t X_{Kt} + u_t$$

--- 残差2乗和 SSR_1

・ 制約ありモデル(H_0 が正しい前提)

$$Y_t = \alpha + \beta_1 X_{1,t} + \dots + \beta_K X_{K,t} + u_t$$

--- 残差2乗和 SSR_0

$$Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_K X_{Kt} + \underbrace{\theta_0 D_t + \theta_1 D_t X_{1t} + \dots + \theta_K D_t X_{Kt}}_{\text{最後の} K + 1 \text{個}} + u_t$$

・F統計量

$$F = \frac{(SSR_0 - SSR_1) / (K + 1)}{SSR_1 / (T - 2(K + 1))}$$

---除外制約の数は $q = K + 1$

---サンプルサイズは T

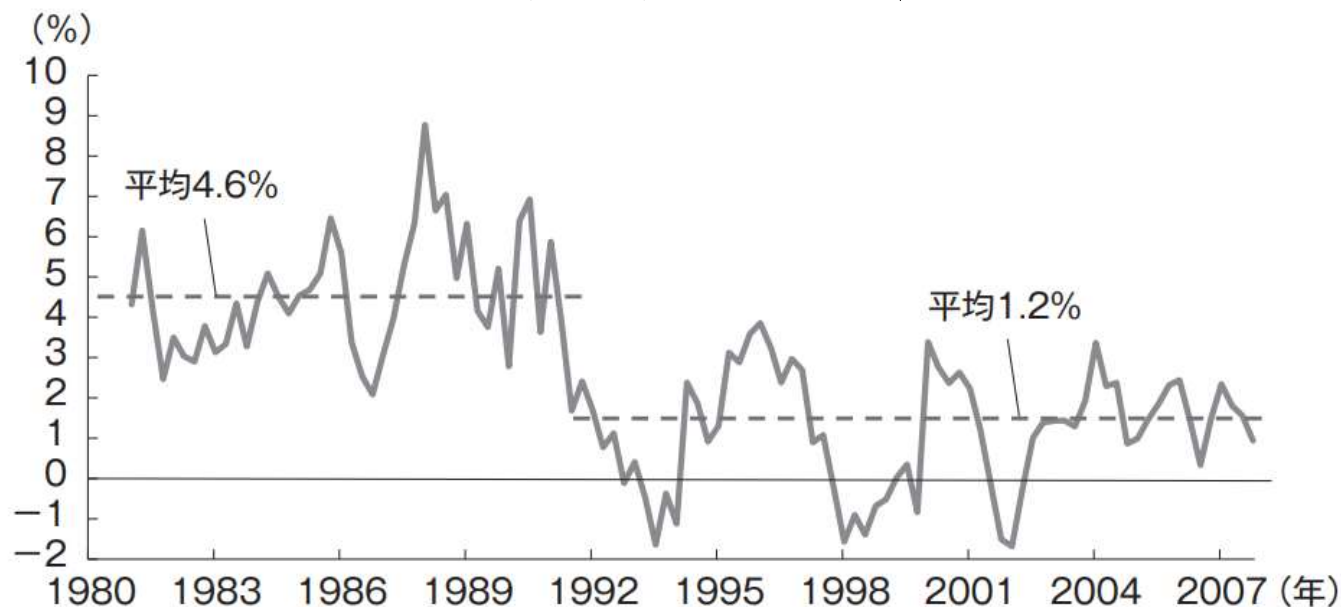
--- 推定するパラメータの数は $2(K + 1)$

--- H_0 が正しいもとで自由度 $(K + 1, T - 2(K + 1))$ の F 分布

(例)前年同期比の実質GDP成長率

- ・GDP成長率は1991年から低下している

図 7 - 2 日本の実質 GDP 成長率の推移



・推定式

$$Y_t = \alpha + \beta_1 Y_{t-1} + \theta_0 D_t + \theta_1 D_t Y_{t-1} + u_t$$

--- D_t は、1991年第1四半期まで0、それ以降に1となるダミー変数

--- $\theta_0 = \alpha' - \alpha$ 、 $\theta_1 = \beta'_1 - \beta_1$

--- $H_0: \theta_0 = 0, \theta_1 = 0$ 、

$H_1: H_0$ は誤りである

・制約なしモデル

$$\hat{Y}_t = 2.30 + 0.50Y_{t-1} - 1.95D_t + 0.17D_tY_{t-1}$$

(0.58) (0.12) (0.61) (0.15)

--- $SSR_1 = 134.54$

---定数項は前半は2.30、後半は $2.30 - 1.95 = 0.35$

---係数は前半は0.5、後半は $0.5 + 0.17 = 0.67$

・制約ありモデル

$$\hat{Y}_t = 0.40 + 0.83Y_{t-1}$$

(0.18) (0.05)

--- $SSR_0 = 158.72$

・F統計量

$$F = \frac{(158.72 - 134.54)/2}{134.54/(107 - 2(1 + 1))} = 9.25$$

--- $T = 107$ 、 $K = 1$ 、 $q = 2$

--- F 値を計算すると9.25となり、 $H_0: \theta_0 = 0, \theta_1 = 0$ は棄却(F 統計量は自由度(2, $107 - 4$)であり、有意水準5%の臨界値は3.08)

構造変化の検定

構造変化点が未知

- 実証分析では、構造変化点 T_B がわからないことが多い
 ---悪意ある分析者なら、自分の都合の良い結果を得るため、
 構造変化点を恣意的に決めるかもしれない

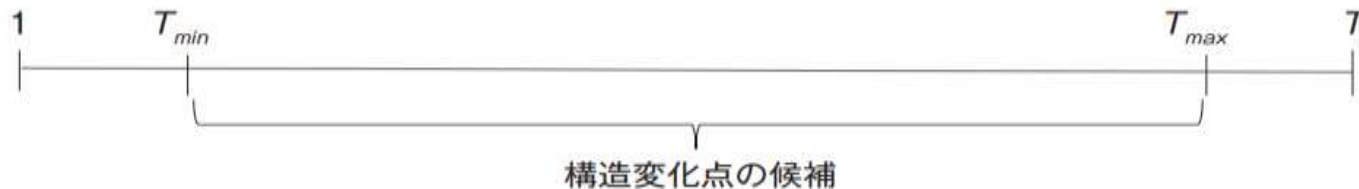
- 構造変化点の候補 T_B

$$\{T_{min}, T_{min} + 1, T_{min} + 2, \dots, T_{max} - 2, T_{max} - 1, T_{max}\}$$

--- $T_{min} = 0.15T$ に最も近い整数

--- $T_{max} = (1 - 0.15)T$ に最も近い整数

図 7 - 3 構造変化点の候補



--- $T = 100$ なら、候補日は{15, 16, 17, ..., 83, 84, 85}

- 構造変化点の候補のそれぞれで Chow 検定する

$$\{F(T_{min}), F(T_{min} + 1), \dots, F(T_{max} - 1), F(T_{max})\}$$

--- $F(T_{min})$ とは、 $T_B = T_{min}$ として計算された F 値

- その中の最大値を検定統計量とする

$$\sup F = \max \{F(T_{\min}), F(T_{\min} + 1), \dots, F(T_{\max})\}$$

--- F値を最大にする構造変化点を推定値 $\hat{T}_B$ とする

--- $\sup F$ は、通常のF分布よりも大きな値を取りやすい

---表1は、 $\sup F$ の臨界値 (H_0 が正しいもとで $P\{c < \sup F\} = \delta$ となる c)を掲載

表 7 - 1 $\sup F$ の臨界値

制約数 q	有意水準 δ		
	0.10	0.05	0.01
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	2.23
11	2.40	2.62	3.09
12	2.33	2.54	2.97

---説明変数が2個なら $q = 3$

(例)実質GDP成長率2

・推定(1981年から2008年まで)

$$Y_t = \alpha + \beta_1 Y_{t-1} + \theta_0 D_t + \theta_1 D_t Y_{t-1} + u_t$$

--- D_t は、 T_B まで0、それ以降に1となるダミー変数

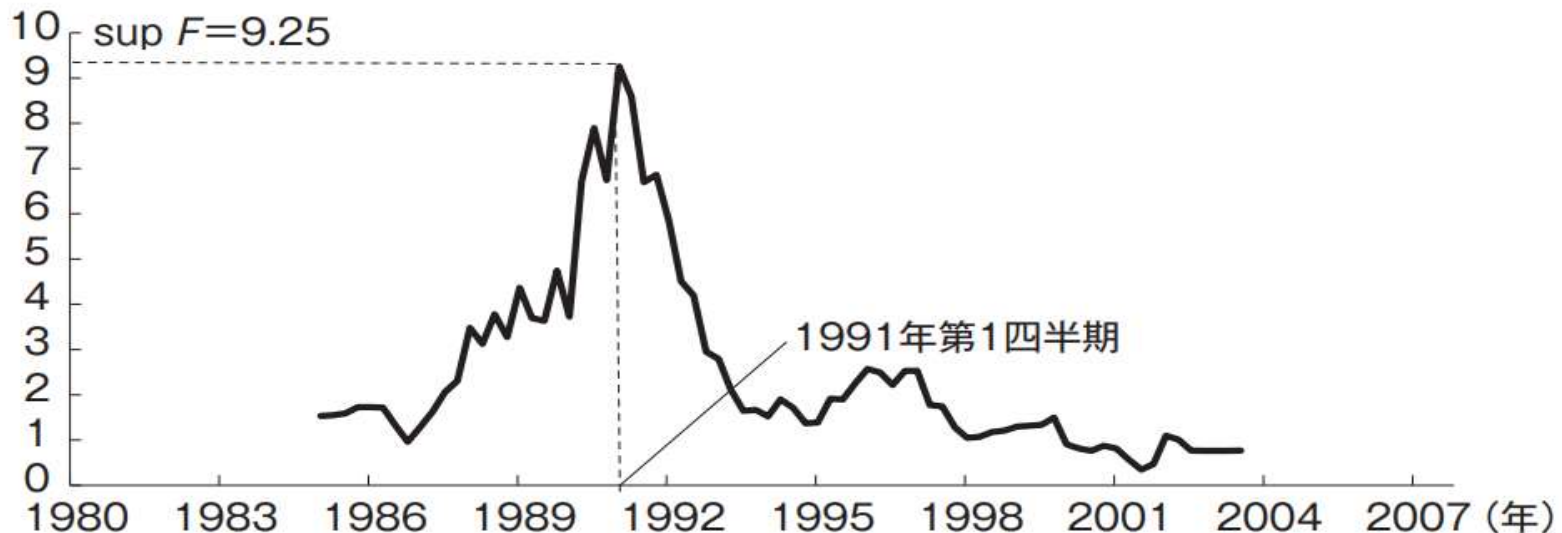
--- 構造変化点の候補: 1985年第1四半期から2003年第4四半期

--- $\sup F = 9.25$ 、 $\hat{T}_B = 1991$ 年第1四半期

--- $q = 2$ 、有意水準1%の臨界値は7.78

--- 構造変化がないという帰無仮説を棄却できる

図 7 - 4 GDP と $\sup F$ 検定



バイペロン検定

- 構造変化は1度だけではなく、複数存在するかもしれない
- 帰無仮説: 構造変化は m 回
対立仮説: 構造変化は $m + 1$ 回
- 繰り返し検定することで構造変化の回数を特定する
 - ① H_0 : 構造変化は0回、 H_1 : 構造変化は1回
もし H_0 が採択されたら構造変化なし、棄却なら②へ
 - ② H_0 : 構造変化は1回、 H_1 : 構造変化は2回
もし H_0 が採択されたら構造変化は1回、棄却なら③へ
 - ③ H_0 : 構造変化は2回、 H_1 : 構造変化は3回
もし H_0 が採択されたら構造変化は2回、棄却なら④へ...
 - ④ H_0 を採択するまで続ける

まとめ

- 同時検定とは

- 同時検定とは
- 同時検定の意義
- 一般的ケース

- F 統計量

- F 統計量と F 分布
- 統計ソフトで表示される F 値

- 構造変化の検定

- チョウ検定(構造変化点が既知、構造変化は1回)
- $\sup F$ 検定(構造変化点が未知、構造変化は1回)
- バイ=ペロン検定(構造変化点が未知で複数回ある)