

9주차 2차시 : 가변수모형(유형 및 가설검정)

2. 가변수의 유형

- (1) 절편(평균)의 변화를 나타내는 가변수
- (2) 기울기(한계)의 변화를 나타내는 가변수
- (3) 절편과 기울기의 동시변화를 나타내는 가변수

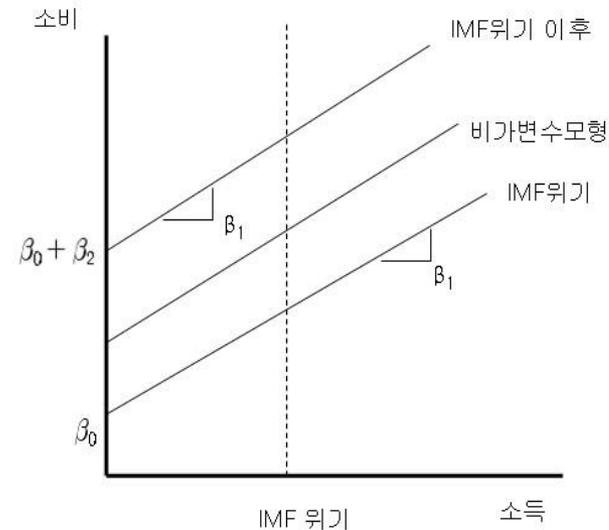
2.가변수의 유형

-어느 한 변수의 두 개의 질적 범주를 구분할 경우도 평균의 구분, 기울기의 구분 또는 평균과 기울기의 구분 등 다양한 구분이 필요함

(1)절편(평균)의 변화를 나타내는 가변수

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 D_t + u_t$$

$$\text{단, } D_t = \begin{cases} 1, t \text{가 IMF위기 이후} \\ 0, t \text{가 IMF위기} \end{cases}$$



-위의 가변수모형을 추정한 후 가변수가 통계적으로 유의할 경우 위 식의 추정은 다음의 두 식을 개별적으로 추정한 결과와 동일함

$$\text{(IMF위기)} \quad C_t = \beta_0 + \beta_1 Y_t + u_t$$

$$\text{(IMF위기 이후)} \quad C_t = (\beta_0 + \beta_2) + \beta_1 Y_t + u_t$$

자료 구조

| 연도 | C(조원) | Y(조원) | D |
|------|-------|-------|---|
| . | . | . | . |
| 1995 | 214 | 408 | 1 |
| 1996 | 245 | 458 | 1 |
| 1997 | 271 | 502 | 1 |
| 1998 | 252 | 492 | 0 |
| 1999 | 289 | 542 | 0 |
| 2000 | 330 | 600 | 0 |
| 2001 | 364 | 649 | 1 |
| . | . | . | . |
| 2009 | 577 | 1,068 | 1 |

(가설검정)

- 가변수에 대한 가설검정에서 귀무가설은 $H_0: \beta_2 = 0$ 으로 IMF위기와 IMF위기 이후의 소비수준에 차이가 없다는 가설임
- 귀무가설을 기각하면(즉, β_2 가 통계적으로 유의하면) IMF위기와 IMF위기 이후의 소비수준에 차이가 있다고 결론을 내림
- 따라서 이 경우는 IMF위기와 IMF위기 이후의 소비함수를 각각 추정해야 하는데 가변수를 포함한 식을 추정하면 이와 동일한 결과를 얻을 수 있음

(회귀계수에 대한 해석)

- $\hat{\beta}_1$: 한계소비성향으로 IMF위기의 한계소비성향과 IMF위기 이후의 한계소비성향이 같음
- $\hat{\beta}_0$: IMF위기의 절대소비수준
- $\hat{\beta}_2$: IMF위기 이후 소비수준과 IMF위기 소비수준의 차이
- 따라서 IMF위기 이후 소비수준은 $\hat{\beta}_0 + \hat{\beta}_2$



```
> m1.lm<-lm(c~d+y)
```

```
> summary(m1.lm)
```

```
Call:
```

```
lm(formula = c ~ d + y)
```

```
Residuals:
```

| Min | 1Q | Median | 3Q | Max |
|---------|--------|--------|--------|--------|
| -2307.8 | -806.9 | -375.7 | 1054.6 | 3019.7 |

```
Coefficients:
```

| | Estimate | Std. Error | t value | Pr(> t) | |
|-------------|-----------|------------|---------|----------|-----|
| (Intercept) | 9.854e+03 | 4.432e+03 | 2.223 | 0.0368 | * |
| d | 3.419e+03 | 6.979e+02 | 4.898 | 6.74e-05 | *** |
| y | 5.092e-01 | 4.014e-02 | 12.685 | 1.36e-11 | *** |

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1448 on 22 degrees of freedom
```

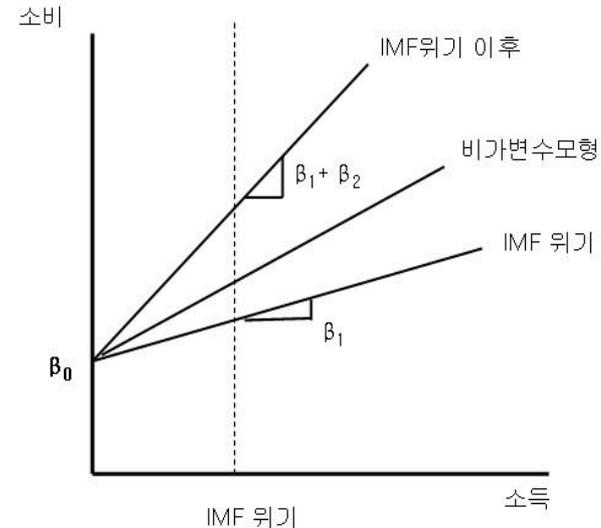
```
Multiple R-squared:  0.8841,    Adjusted R-squared:  0.8735
```

```
F-statistic: 83.87 on 2 and 22 DF,  p-value: 5.092e-11
```

(2) 기울기(한계)의 변화를 나타내는 가변수

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 D_t Y_t + u_t$$

$$\text{단, } D_t = \begin{cases} 1, t \text{가 IMF 위기 이후} \\ 0, t \text{가 IMF 위기} \end{cases}$$



-위의 가변수모형을 추정한 후 가변수가 통계적으로 유의할 경우 위 식의 추정은 다음의 두 식을 개별적으로 추정한 결과와 동일함

(IMF위기) $C_t = \beta_0 + \beta_1 Y_t + u_t$

(IMF위기 이후) $C_t = \beta_0 + (\beta_1 + \beta_2) Y_t + u_t$

자료 구조

| 연도 | C(조원) | Y(조원) | D | D*Y |
|------|-------|-------|---|-------|
| . | . | . | . | . |
| 1995 | 214 | 408 | 1 | 408 |
| 1996 | 245 | 458 | 1 | 458 |
| 1997 | 271 | 502 | 1 | 502 |
| 1998 | 252 | 492 | 0 | 0 |
| 1999 | 289 | 542 | 0 | 0 |
| 2000 | 330 | 600 | 0 | 0 |
| 2001 | 364 | 649 | 1 | 649 |
| . | . | . | . | . |
| 2009 | 577 | 1,068 | 1 | 1,068 |

(가설검정)

- 귀무가설은 $H_0: \beta_2 = 0$ 으로 IMF위기와 IMF위기 이후의 한계소비성향에 차이가 없다는 가설임
- 귀무가설을 기각하면(즉, β_2 가 통계적으로 유의하면) IMF위기와 IMF위기 이후의 한계소비성향에 차이가 있다고 결론을 내림
- 따라서 이 경우는 IMF위기와 IMF위기 이후의 소비함수를 각각 추정해야 하는데 가변수를 포함한 식을 추정하면 이와 동일한 결과를 얻을 수 있음

(회귀계수에 대한 해석)

- $\hat{\beta}_0$: 절대소비수준으로 IMF위기의 절대소비수준과 IMF위기 이후의 절대소비수준은 같음
- $\hat{\beta}_1$: IMF위기의 한계소비성향
- $\hat{\beta}_2$: IMF위기 이후 한계소비성향과 IMF위기 한계소비성향의 차이
- 따라서 IMF위기 이후 한계소비성향의 값은 $\hat{\beta}_1 + \hat{\beta}_2$



```
> m2.lm<-lm(c~y+dy)
> summary(m2.lm)

Call:
lm(formula = c ~ y + dy)

Residuals:
    Min       1Q   Median       3Q      Max
-2230.1  -874.9  -292.3  1064.8  2923.8

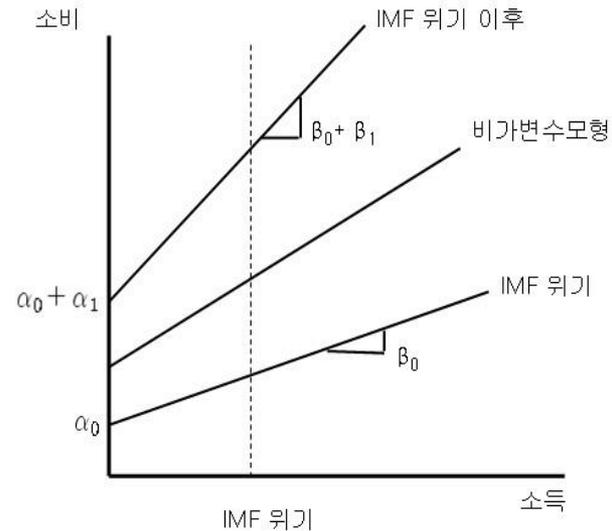
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.043e+04  4.237e+03   2.463  0.0221 *
y             5.377e-01  4.235e-02  12.698 1.33e-11 ***
dy           -3.398e-02  6.652e-03  -5.108 4.06e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1416 on 22 degrees of freedom
Multiple R-squared:  0.8891,    Adjusted R-squared:  0.879
F-statistic:  88.2 on 2 and 22 DF,  p-value: 3.117e-11
```

(3) 절편과 기울기의 동시변화를 나타내는 가변수

$$C_t = \alpha_0 + \alpha_1 D_t + \beta_0 Y_t + \beta_1 D_t Y_t + u_t$$

$$\text{단, } D_t = \begin{cases} 1, t \text{가 IMF 위기 이후} \\ 0, t \text{가 IMF 위기} \end{cases}$$



-위의 가변수모형을 추정한 후 가변수가 통계적으로 유의할 경우 위 식의 추정은 다음의 두 식을 개별적으로 추정한 결과와 동일함

(IMF위기) $C_t = \alpha_0 + \beta_0 Y_t + u_t$

(IMF위기 이후) $C_t = (\alpha_0 + \alpha_1) + (\beta_0 + \beta_1) Y_t + u_t$

자료 구조

| 연도 | C(조원) | Y(조원) | D | D*Y |
|------|-------|-------|---|-------|
| . | . | . | . | . |
| 1995 | 214 | 408 | 1 | 408 |
| 1996 | 245 | 458 | 1 | 458 |
| 1997 | 271 | 502 | 1 | 502 |
| 1998 | 252 | 492 | 0 | 0 |
| 1999 | 289 | 542 | 0 | 0 |
| 2000 | 330 | 600 | 0 | 0 |
| 2001 | 364 | 649 | 1 | 649 |
| . | . | . | . | . |
| 2009 | 577 | 1,068 | 1 | 1,068 |

위 식의 가변수모형에서는 3종류의 가설검정이 가능함

① $H_0: \alpha_1 = 0$ (IMF위기와 IMF위기 이후의 소비수준에 차이가 없다)

-귀무가설을 기각하면(즉, α_1 이 통계적으로 유의하면) IMF위기와 IMF위기 이후의 소비수준에 차이가 있다고 결론을 내림

-따라서 이 경우는 IMF위기와 IMF위기 이후의 소비함수를 각각 추정해야 하는데 가변수를 포함한 식을 추정하면 이와 동일한 결과를 얻을 수 있음

② $H_0: \beta_1 = 0$ (IMF위기와 IMF위기 이후의 한계소비성향에 차이가 없다)

-귀무가설을 기각하면(즉, β_1 이 통계적으로 유의하면) IMF위기와 IMF위기 이후의 한계소비성향에 차이가 있다고 결론을 내림

-따라서 이 경우는 IMF위기와 IMF위기 이후의 소비함수를 각각 추정해야 하는데 가변수를 포함한 식을 추정하면 이와 동일한 결과를 얻을 수 있음

③ $H_0: \alpha_1 = \beta_1 = 0$ (IMF위기와 IMF위기 이후의 소비수준 및 한계소비성향에 차이가 없다)

- 귀무가설을 기각하면(즉, α_1 과 β_1 이 동시에 통계적으로 유의하면) IMF위기와 IMF위기 이후의 소비수준 및 한계소비성향에 차이가 있다고 결론을 내림
- 따라서 이 경우는 IMF위기와 IMF위기 이후의 소비함수를 각각 추정해야 하는데 가변수를 포함한 식을 추정하면 이와 동일한 결과를 얻을 수 있음

(회귀계수에 대한 해석)

- $\hat{\alpha}_0$: IMF위기의 절대소비수준
- $\hat{\alpha}_1$: IMF위기 이후의 절대소비수준과 IMF 위기의 절대소비수준의 차이
- $\hat{\beta}_0$: IMF위기의 한계소비성향
- $\hat{\beta}_1$: IMF위기 이후의 한계소비성향과 IMF 위기의 한계소비성향의 차이



```
> m3.lm<-lm(c~d+y+dy)
> summary(m3.lm)

Call:
lm(formula = c ~ d + y + dy)

Residuals:
    Min       1Q   Median       3Q      Max
-1956.2  -748.5  -330.4   713.5  2707.2

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.348e+04  4.827e+03   2.792  0.0109 *
d            -1.217e+04  9.657e+03  -1.261  0.2213
y             4.763e-01  4.376e-02  10.883 4.31e-10 ***
dy           1.523e-01  9.411e-02   1.619  0.1205
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1397 on 21 degrees of freedom
Multiple R-squared:  0.8969,    Adjusted R-squared:  0.8822
F-statistic: 60.9 on 3 and 21 DF,  p-value: 1.568e-10
```



Regression Results of using Dummy Variable

| Dependent variable: | | | | |
|---------------------|------------------------------|----------------------------|-----------------------------|-----------------------------|
| | (1) | (2) | (3) | (4) |
| d | | 3,418.648*** (697.938) | | -12,172.840 (9,656.625) |
| y | 0.400*** (0.047) | 0.509*** (0.040) | 0.504*** (0.038) | 0.476*** (0.044) |
| dy | | | 0.034*** (0.007) | 0.152 (0.094) |
| Constant | 23,033.120*** (4,980.694) | 9,854.362** (4,432.226) | 10,433.910** (4,236.569) | 13,476.090** (4,827.492) |
| Observations | 25 | 25 | 25 | 25 |
| R2 | 0.758 | 0.884 | 0.889 | 0.897 |
| Adjusted R2 | 0.747 | 0.874 | 0.879 | 0.882 |
| Residual Std. Error | 2,047.203 (df = 23) | 1,447.710 (df = 22) | 1,415.773 (df = 22) | 1,397.192 (df = 21) |
| F Statistic | 71.886*** (df = 1; 23) | 83.870*** (df = 2; 22) | 88.199*** (df = 2; 22) | 60.903*** (df = 3; 21) |

Note:

*p<0.1; **p<0.05; ***p<0.01

```

> jointHo<-c("d","dy")
> linearHypothesis(m3.lm, jointHo)
Linear hypothesis test

Hypothesis:
d = 0
dy = 0

Model 1: restricted model
Model 2: c ~ d + y + dy

   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
1      23 96393877
2      21 40995058  2  55398818 14.189 0.0001262 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

가설검정 요약

| 구분 | 1단계 | 2단계 | 3단계 | 4단계 |
|--------|-----------------------------------|-------------------------|--|--|
| Case 1 | $\alpha_1=0$ 기각 $\beta_1=0$ 허용 | 가변수 모형 1 | | |
| Case 2 | $\alpha_1=0$ 허용 $\beta_1=0$ 기각 | 가변수 모형 2 | | |
| Case 3 | $\alpha_1=0$ 기각 $\beta_1=0$ 기각 | 가변수 모형 3 | | |
| Case 4 | $\alpha_1=0$ 허용 $\beta_1=0$ 허용 | $\alpha_1=\beta_1=0$ 허용 | 가변수가 없는 모형 | |
| Case 5 | $\alpha_1=0$ 허용 $\beta_1=0$ 허용 | $\alpha_1=\beta_1=0$ 기각 | $\alpha_1=0$ 기각 and $\beta_1=0$ 허용 | 가변수 모형 1 |
| Case 6 | $\alpha_1=0$ 허용 $\beta_1=0$ 허용 | $\alpha_1=\beta_1=0$ 기각 | $\alpha_1=0$ 허용 and $\beta_1=0$ 기각 | 가변수 모형 2 |
| Case 7 | $\alpha_1=0$ 허용 $\beta_1=0$ 허용 | $\alpha_1=\beta_1=0$ 기각 | $\alpha_1=0$ 기각 and $\beta_1=0$ 허용 또는 $\beta_1=0$ 기각 and $\alpha_1=0$ 허용 | 가변수 모형 1 및 가변수 모형 2 중 결정계수가 큰 모형 |