제 12 강

확률적 설명변수

Random Explanatory Variables

단순선형모형에 대한 새로운 기본 가정

A13.1 $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$

A13.2 (x_t, y_t) $t = 1,\ldots,T$ 는 확률표본으로 추출됨(**random sampling**)
i.e. 각 쌍(pair)은 동일하고 독립적인 분포를 함

A13.3 $E(\varepsilon \mid x) = 0$
오차항의 임의의 x 값을 조건부로 하는 기대값은 0

A13.4 x_t 는 반드시 적어도 두개의 다른 값을 가짐

A13.5 $\mathrm{var}(\varepsilon \mid x) = \sigma^2$

오차항의 임의의 x값을 조건부로 하는 분산은 σ^2

A13.6 $\varepsilon \mid x \sim N(0, \sigma^2)$

오차항의 x를 조건부로 하는 분포는 정규분포

단순선형모형에 대한 새로운 기본 가정

- A13.3 : $E(\varepsilon \mid x) = 0$
  - 이 가정은 다음을 의미함
    (1) 중요한 설명변수가 누락되지 않음,
    (2) 올바른 함수형태를 사용하고 있음
    (3) 오차항이 x와 상관되도록하는 요인들이 존재하지 않음
  - If $E(\varepsilon \mid x) = 0$,
    then we can show that x and ε are uncorrelated, $\text{cov}(x, \varepsilon) = 0$
  - Conversely, if x and ε are correlated $\text{cov}(x, \varepsilon) \neq 0$, then we can show that $E(\varepsilon \mid x) \neq 0$

.

새로운 가정하에서 최소제곱추정량의 성질

- 소표본 성질 (finite sample properties)
  - X가 확률변수일 경우 최소제곱 추정량의 소표본 성질은 다음과 같이 요약될 수 있음
    - A13.1-A13.4의 가정들하에서 최소제곱 추정량은 불변 추정량임
    - A13.1-A13.5의 가정들하에서 최소제곱 추정량은 모수에대한 최우수 선형불편추정량(BLUE)이며 통상적인 σ^2 에 대한 추정량은 불편임
    - A13.1-A13.6의 가정하에서 최소제곱 추정량들의 x를 조건부로 하는 분포는 정규분포를 하게되며 그들의 분산들은 통상적인 방법으로 추정됨. 따라서 통상적 구간추정이나 가설검정 절차 역시 유효함

새로운 가정하에서 최소제곱추정량의 성질

- 대표본(점근적) 성질 (large sample (asymptotic) properties)
  - $T \to \infty$ 인 경우 혹은 충분히 큰 표본을 가지고 있는 경우 근사적으로, 최소제곱추정량의 확률분포는 어떻게 되는가?
    - 표본의 크기가 점차 커짐에 따라 최소제곱추정량의 확률분포는 그 모수로 (확률적으로) 수렴함
      - 우선 최소제곱 추정량은 불편추정량
      - 그리고 최소제곱추정량의 분산이 $T \to \infty$임에 따라 0으로 수렴함
    - 이러한 성질을 갖는 추정량은 일치(*consistent*) 추정량이라 하며, 일치성은 최소제곱 추정량의 대표본 성질임
      - 일치성은 어떤 추정량이 실제로 사용될 수 있기 위해서 갖추어야 하는 최소한의 요건이기도 함

오차항과 설명변수가 상관되어 있을 경우

- A13.3* : $E(\varepsilon)=0$ and $\text{cov}(x,\varepsilon)=0$
  - 이는 A13.3 보다 "완화된" 가정인데, 그 이유는

    $E(\varepsilon \mid x)=0 \Rightarrow \text{cov}(\varepsilon,x)=0$ & $E(\varepsilon)=0$

    but $\text{cov}(\varepsilon,x)=0$ & $E(\varepsilon)=0$ does not $\Rightarrow E(\varepsilon \mid x)=0$
  - 이러한 완화된 가정하에서, 소표본 성질들은 더 이상 유효하지 않음
- 최소제곱 추정량의 대표본 성질은 여전히 유효함
  - A13.3* 을 A13.3 대신 사용할 경우, 최소제곱추정량은 여전히 일치추정량임
  - 최소제곱추정량은 오차항의 조건부 분포가 정규분포든 아니든 간에 대표본하에서 근사적으로 정규분포를 하게됨
  - 대표본하에서 구간추정량과 가설검정은 여전히 유효함

오차항과 설명변수가 상관되어 있을 경우

- A13.3* 이 충족될 수 없는 경우, 특히 $cov(x,\varepsilon)\neq 0$ 인 경우
  - 최소제곱추정량은 비일치(inconsistent)추정량이며, 아무리 큰 표본하에서도 모수로 수렴하지 않음
  - 더욱이, 통상적 가설검정이나 구간추정 절차 역시 유효하지 않음

- 설명변수들이 확률변수로 간주되는 경우, 최소제곱추정량이 적절한 것인가를 판단함에 있어서 설명변수들과 오차항간의 상관 여부가 핵심적인 문제임

오차항과 설명변수가 상관되어 있을 경우

예: 변수오차(Errors in Variables)의 경우

- 오차가 있는 설명변수를 이용하여 추정하는 경우 이는 오차항과 상관되며, 따라서 최소제곱추정량은 비일치추정이 됨
  - 노동자의 임금이 노동자의 능력의 함수라 가정:

y_t = wages of the t'th worker

x_t^* = the ability of the t'th worker

$$y_t = \beta_1 + \beta_2 x_t^* + \varepsilon_t$$

  - 능력을 나타내는 변수(*로 표시)는 관측하기가 매우 어려움
  - 능력을 표준화된 시험성적(수능, SAT, TOEIC 등)으로 측정하고자 하며, 이를 x_t 로 표시.
  - 경우에 따라 이는 대용변수(**proxy variable**) **라 함**.

$$x_t = x_t^* + u_t$$

오차항과 설명변수가 상관되어 있을 경우

예: 변수오차(Errors in Variables)의 경우

- 여기서 u_t 는 확률적인 측정 오차이며 평균이 0이고 분산은 σ_u^2
- u_t 는 ε_t,와 상관되어 있지 않다고 가정함
- $x_t^* = x_t - u_t$ 를 추정식에 대입하면,

$$\begin{aligned} y_t &= \beta_1 + \beta_2 x_t^* + \varepsilon_t \\ &= \beta_1 + \beta_2 (x_t - u_t) + \varepsilon_t \\ &= \beta_1 + \beta_2 x_t + (\varepsilon_t - \beta_2 u_t) \\ &= \beta_1 + \beta_2 x_t + v_t \end{aligned}$$

- 이 추정식에서 설명변수 x_t 는 확률변수이며, 이 식에 최소제곱추정을 적용하게 되면 측정오차에 대한 가정으로부터 다음을 얻을 수 있으며, 앞서 설명한 바에 의하면 비일치추정을 하게 됨

$$\begin{aligned} \operatorname{cov}(x_t, v_t) &= E(x_t v_t) = E\left[(x_t^* + u_t)(\varepsilon_t - \beta_2 u_t)\right] \\ &= E(-\beta_2 u_t^2) = -\beta_2 \sigma_u^2 \neq 0 \end{aligned}$$

오차항과 설명변수가 상관되어 있을 경우

예: 동시성(Simultaneity)의 경우

- 설명변수와 종속변수가 모두 시스템내에서 동시에 결정되는 내생변수인 경우, 설명변수는 오차항과 상관되는 확률변수이며 이 경우 최소제곱 추정은 적절하지 않음

$C_t = \beta_1 + \beta_2 Y_t + \varepsilon_t$: 케인지언 소비함수

$Y_t = C_t + I_t$: 소득 항등식

C_t, Y_t : 소비와 소득 모두 내생변수로서 시스템에서 결정됨

- 소비함수에 대한 최소제곱추정은 비일치 추정을 낳게 됨:

$$\Rightarrow Y_t = \frac{\beta_1}{1-\beta_2} + \frac{I_t}{1-\beta_2} + \frac{\varepsilon_t}{1-\beta_2}$$

$$\Rightarrow \operatorname{cov}(Y_t, \varepsilon_t) = E\left[\left(\frac{\beta_1}{1-\beta_2} + \frac{I_t}{1-\beta_2} + \frac{\varepsilon_t}{1-\beta_2}\right)\varepsilon_t\right]$$

$$= E\left[\frac{\varepsilon_t^2}{1-\beta_2}\right] = \frac{\sigma^2}{1-\beta_2} \neq 0$$

적률방법에 기초한 추정량

적률방법 (Method of moments)

- 어떤 확률변수의 k차 수학적 적률(k th mathematical moment)은 그 확률변수의 k승의 기대값임

$$E\left(Y^k\right)=\mu_k = k \text{ th mathematical moment of } Y$$

- 이 수학적 적률은 대응하는 k차 표본 적률(k th sample moment)에 의해 일치 추정될 수 있음

$$\hat{E}\left(Y^k\right)=\hat{\mu}_k = k\text{th sample moment of } Y$$
$$=\sum_{t=1}^{T} y_t^k \Big/ T$$

- 적률추정법은 m개의 모수들을 추정함에 있어서 m개의 수학적 적률들을 m개의 표본적률들에 등치시킴으로서 추정량을 구하는 방법임

$$E(Y)=\mu$$
$$\text{var}(Y)=\sigma^2=E(Y-\mu)^2=E\left(Y^2\right)-\mu^2$$

적률방법에 기초한 추정량

적률방법 (Method of moments)

- 두 개의 모수 μ 와 σ^2를 추정함에 있어서 두 개의 수학적 적률과두 개의 표본 적률을 등치시킴
- Y의 처음 두 수학적 적률과 그에 대응되는 표본 적률들:

$$E(Y)=\mu_1=\mu, \quad \hat{\mu}=\sum_{t=1}^{T} y_t \Big/ T$$
$$E\left(Y^2\right)=\mu_2, \quad \hat{\mu}_2=\sum_{t=1}^{T} y_t^2 \Big/ T$$

- μ 와 σ^2 에 대한 적률방법추정량은 다음과 같이 주어짐

$$\hat{\mu}=\sum_{t=1}^{T} y_t \Big/ T=\bar{y}$$
$$\tilde{\sigma}^2=\hat{E}\left(Y^2\right)-\hat{\mu}^2=\frac{\sum_{t=1}^{T} y_t^2}{T}-\bar{y}^2=\frac{\sum_{t=1}^{T} y_t^2-T\bar{y}^2}{T}=\frac{\sum(y_t-\bar{y})^2}{T}$$

- 일반적으로 적률방법추정량은 대표본에서 일치추정량이나 유효성과 관련해서는 아무런 보장을 할 수 없음

적률방법에 기초한 추정량

적률방법 - 단순회귀모형

- 적률의 정의는 더욱 일반적인 형태로 확장될 수 있음
  $E[g(Y)]$ is a moment of $g(Y)$.
- 선형회귀모형에 있어서 ($y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$,), 보통 다음을 가정함

  $E(\varepsilon_t) = 0 \Rightarrow E(y_t - \beta_1 - \beta_2 x_t) = 0$: Moment Condition 1
- x_t 확률변수가 아니거나 확률변수라해도 ε_t와 상관되지 않는다면

  $E(x_t \varepsilon_t) = 0 \Rightarrow E[x_t(y_t - \beta_1 - \beta_2 x_t)] = 0$: Moment Condition 2
- β_1 과 β_2 에 대한 적률방법 추정량은 다음과 같이 도출됨

  $$\frac{1}{T}\sum(y_t - b_1 - b_2 x_t) = 0, \quad \frac{1}{T}\sum x_t(y_t - b_1 - b_2 x_t) = 0$$

  – 이 두 방정식은 최소제곱추정량을 도출했던 두 정규방정식과 동일하며, 따라서 그 해는 최소제곱추정량이 됨

적률방법에 기초한 추정량

적률방법 – 도구(수단)변수

- 최소제곱추정의 문제는 x 가 확률변수이고 오차항ε과 상관되어 있는 경우 발생함 $E(x_t \varepsilon_t) \neq 0$
- 하지만 다음의 적률조건을 만족하는 다른 변수 z_t (도구변수,**instrumental variable**) 가 있을 경우

  $E(z_t \varepsilon_t) = 0 \Rightarrow E[z_t(y_t - \beta_1 - \beta_2 x_t)] = 0$
- 이 경우 두 개의 적률조건을 이용하여 β_1 과 β_2 에 대한 추정량을 얻을 수 있음

  $$\frac{1}{T}\sum(y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t) = 0, \quad \frac{1}{T}\sum z_t(y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t) = 0$$
- 이들 방정식들을 풀면 적률방법추정량을 얻을 수 있는데, 이는 대개 도구(수단)변수추정량(**instrumental variable estimators**) 이라 함

$$\hat{\beta}_2 = \frac{\sum z_t \sum y_t - T\sum z_t y_t}{\sum z_t \sum x_t - T\sum z_t x_t} = \frac{\sum(z_t - \bar{z})(y_t - \bar{y})}{\sum(z_t - \bar{z})(x_t - \bar{x})}$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}$$

적률방법에 기초한 추정량

도구변수추정량의 성질

- 도구변수추정량은 대표본에서 그 모수로 (확률적으로) 수렴함, 즉 일치 추정량임
- 대표본에서 도구변수추정량은 근사적으로 정규분포를 함

$$\hat{\beta}_2 \sim N\left(\beta_2, \frac{\sigma^2}{\sum (x_t - \bar{x})^2 r_{zx}^2}\right)$$

  - r_{zx}^2 는 도구변수 z 와 확률변수 설명변수 x간의 표본 상관계수의 제곱
  - 도구변수추정량의 유효성을 높이기 위해서는 도구변수가 문제가 되는 설명변수와 높은 상관을 가지길 원하게 됨
- 오차항의 분산은 다음의 추정량으로 일치추정할 수 있음

$$\hat{\sigma}_{IV}^2 = \frac{\sum \left(y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t\right)^2}{T-2}$$

적률방법에 기초한 추정량

2단계최소제곱추정량

- 필요한 것 이상의 도구변수들을 가지게 될 수 있음
  - 예컨대, w 가 x 와 상관되어 있으나 ε 과는 상관되어 있지 않은 또 하나의 도구변수라면
- 이 경우, 두 도구변수를 최적으로 이용한 추정량을 다음과 같은 두 단계의 절차를 통해 얻을 수 있다는 것이 알려져 있음

  (1)문제가 되는 x를 상수항과 z 및 w에 회귀하고 예측치 $\hat{x}$ 들을 구함

  (2) 예측치 $\hat{x}$ 를 x에 대한 도구변수로 이용하여 추정량을 구함

$$\sum \left(y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t\right) = 0$$
$$\sum \hat{x}_t \left(y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t\right) = 0 \quad \Rightarrow$$
$$\hat{\beta}_2 = \frac{\sum \left(\hat{x}_t - \bar{\hat{x}}\right)(y_t - \bar{y})}{\sum \left(\hat{x}_t - \bar{\hat{x}}\right)(x_t - \bar{x})} = \frac{\sum (\hat{x}_t - \bar{x})(y_t - \bar{y})}{\sum (\hat{x}_t - \bar{x})(x_t - \bar{x})} = \frac{\sum (\hat{x}_t - \bar{x})(y_t - \bar{y})}{\sum (\hat{x}_t - \bar{x})^2}$$
$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x} \qquad \because \bar{\hat{x}} = \bar{x},\ (\hat{x} - \bar{x})(x - \bar{x}) = (\hat{x} - \bar{x})^2$$

적률방법에 기초한 추정량

2단계최소제곱추정량

- 이러한 적률방법에 기초한 도구변수추정량은 2단계최소제곱추정량 (two-stage least squares estimators), 이라고도 하는데 이는 같은 추정량이 두 번의 최소제곱추정을 통해서도 얻어질 수 있기 때문임
  - Stage 1: x를 상수항, z 및 w, 에 대해 회귀하여 $\hat{x}$ 를 얻음
  - Stage 2: 다음의 선형회귀식에 대해 최소제곱추정을 적용

$$y_t = \beta_1 + \beta_2 \hat{x}_t + error_t$$

,

- 근사적으로 추정량 $\hat{\beta}_2$ 의 분산은 대표본 에서다음과 같이 계산됨

$$\text{var}\left(\hat{\beta}_2\right) \to \frac{\sigma^2}{\sum(\hat{x}_t - \bar{x})^2} = \frac{\sigma^2}{\sum(x_t - \bar{x})^2 r_{\hat{x}x}^2}$$ (다음 페이지 참조)

- 오차항의 분산에 대한 추정량은 다음과 같이 원래 모형에서의 잔차로부터 계산되어야만 함

$$\hat{\sigma}_{IV}^2 = \frac{\sum\left(y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t\right)^2}{T-2}$$

적률방법에 기초한 추정량

2단계최소제곱추정량

- $\frac{\sigma^2}{\sum(x_t - \bar{x})^2 r_{\hat{x}x}^2} = \frac{\sigma^2}{\sum(\hat{x}_t - \bar{x})}$ 는 다음과 같이 보일 수 있음

$$\frac{1}{r_{\hat{x}x}^2} = \frac{\sum(x_t - \bar{x})^2 \sum\left(\hat{x}_t - \bar{\hat{x}}\right)^2}{\left[\sum\left(\hat{x}_t - \bar{\hat{x}}\right)(x_t - \bar{x})\right]^2} = \frac{\sum(\hat{x}_t - \bar{x})^2 \sum(\hat{x}_t - \bar{x})^2}{\left[\sum(\hat{x}_t - \bar{x})(x_t - \bar{x})\right]^2}, \quad \left(\because \bar{x} = \bar{\hat{x}}\right)$$

$$\frac{\sigma^2}{\sum(x_t - \bar{x})^2 r_{\hat{x}x}^2} = \frac{\sigma^2 \sum(x_t - \bar{x})^2 \sum(\hat{x}_t - \bar{x})^2}{\sum(x_t - \bar{x})^2 \left[\sum(x_t - \bar{x})(\hat{x}_t - \bar{x})\right]^2} = \frac{\sigma^2 \sum(\hat{x}_t - \bar{x})^2}{\left[\sum(x_t - \bar{x})(\hat{x}_t - \bar{x})\right]^2}$$

$$\begin{aligned}\sum(\hat{x}_t - \bar{x})(x_t - \bar{x}) &= \sum(\hat{x}_t - \bar{x})(\hat{x}_t + \hat{v} - \bar{x}) \\ &= \sum(\hat{x}_t - \bar{x})^2, \qquad \left(\because \sum(\hat{x}_t - \bar{x})\hat{v} = 0\right)\end{aligned}$$

$$\frac{\sigma^2 \sum(\hat{x}_t - \bar{x})^2}{\left[\sum(x_t - \bar{x})(\hat{x}_t - \bar{x})\right]^2} = \frac{\sigma^2 \sum(\hat{x}_t - \bar{x})^2}{\left[\sum(\hat{x}_t - \bar{x})^2\right]^2} = \frac{\sigma^2}{\sum(\hat{x}_t - \bar{x})}$$

설명변수와 오차항 사이의 상관검정

$H_0 : Cov(x,\varepsilon) = 0 \quad H_1 : Cov(x,\varepsilon) \neq 0$

- 검정의 아이디어는 최소제곱추정량과 도구변수추정량을 비교하는 것임
  - 귀무가설이 참이라면 두 가지 추정량 모두 일치 추정량이 따라서,
    $$q = \left(b_{ols} - \hat{\beta}_{IV}\right) \to 0$$
  - 귀무가설이 참이 아니라면 최소제곱추정량은 일치추정량이 아니며, 도구변수추정량은 일치추정량이므로
    $$q = \left(b_{ols} - \hat{\beta}_{IV}\right) \to c \neq 0$$
- 이러한 원리에 바탕을 둔 몇 가지 형태의 검정방법들이 있으며, 대개 이들은 하우스만 검정(Hausman Test)이라고 함
  - 이러한 검정들 중 가장 쉽게 생각할 수 있는 것은 OLS 와 IV estimators 를 직접 비교하는 것이지만, 검정통계량 분포의 계산이 복잡

설명변수와 오차항 사이의 상관검정

인위적 회귀를 통한 하우스만 검정(Davidson and MacKinnon)

- 오차항과 상관되어 있다고 생각되는 각 변수마다 최소한 하나의 도구변수가 필요함
- z_{t1} 와 z_{t2} 가 x에 대한 도구변수들이라고 하면,

  $x_t = a_0 + a_1 z_{t1} + a_2 z_{t2} + v_t$ 를 최소제곱에 의해 추정하고 잔차를 얻음

  $$\Rightarrow \hat{v}_t = x_t - \hat{a}_0 + \hat{a}_1 z_{t1} + \hat{a}_2 z_{t2}$$

  - 만약 하나 이상의 설명변수가 의심이 된다면, 이러한 추정과정을 이용가능한 도구변수들을 사용하여 각 변수에 대해 반복함
  - 이렇게 얻어진 잔차들을 다음과 같이 회귀식에 설명변수로 포함시킴
    $$y_t = \beta_1 + \beta_2 x_t + \delta \hat{v}_t + \varepsilon_t$$
  - 이 인위적 회귀식(artificial regression)을 최소제곱에 의해 추정하고 통상의 t 검정을 이용해 유의성 검정을 수행함

    $H_0 : \delta = 0$ (no correlation between x and ε), $H_1 : \delta \neq 0$ (correlation between x and ε)
  - 하나 이상의 변수가 의심이 될 경우에는 포함된 잔차들의 모수들에 대한 결합적 유의성을 F검정을 통해 검정함