

R 기반 성향점수분석: 루빈 인과모형 기반 인과추론

박인서 (연세대학교)

2021. 4. 14. (수)

0. 발표 개요

[제목] R기반 성향점수분석: 루빈 인과모형 기반 인과추론

[내용] 성향점수분석의 개념 및 필요성을 이해하고, 어떤 절차를 통해 성향점수분석을 수행할 수 있을지 알아봄.

1. 개념
 - 1) 핵심어: 인과추론, 루빈 인과모형, 성향점수
 - 2) 주요 기법: 성향점수가중, 매칭, 층화
2. 적용
 - 3) 성향점수분석 실습
3. 확장
 - 4) 다양한 상황에서의 성향점수분석

[핵심 질문] 성향점수(propensity score)가 왜, 어떻게 사용되는지?

1. 개념 1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

인과추론 (causal inference)

- 인과관계의 세 가지 조건

- ① 상관관계 원인변수와 결과변수는 서로 상관됨
- ② 시간적 선후관계 원인변수는 결과변수에 시간적으로 선행함
- ③ 비허위성 상관관계를 설명하는 다른 변수는 존재하지 않음

- 무작위배치 실험 (RCT, randomized control trial)

- 무작위배정을 통해 비허위성 조건 확보
- 처치/통제집단 배치 여부가 무작위로 정해지므로, RCT에서는 모든 잠복변수 (lurking variable)에 대해 처치여부에 미치는 영향이 배제됨
- “이견 없이” 인과관계를 과학적으로 증명할 수 있는 방법

1. 개념 1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

인과추론 (causal inference)

- **관측 연구 (observational study)**
 - 물론 실험을 통하면 이견 없이 인과관계를 밝힐 수 있음
 - 그러나 대부분의 사회과학 연구는 관측 연구 및 관측 데이터로써 이루어지며, 연구 윤리 등에 의해 실험을 진행할 수 없는 경우도 있음
 - 관측 연구 상황에서 인과관계를 추론하려면?
- **루빈 인과모형 (RCM, Rubin's causal model)**
 - No causation without manipulation (Holland, 1986)
 - 관측 연구 상황에서 "RCT와 동일하거나 매우 가깝도록" 하는 통계적 접근
 - 이론적으로 대안사실 체계, 통계적으로 성향점수가 기반이 됨

1. 개념

1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

대안사실 체계 (counterfactual framework)

- 루빈 인과모형의 이론적 배경: 잠재결과, 대안사실
- **잠재결과** (potential outcome)
 - 모든 결과변수의 값은 "아직 관측되지 않은" 값
 - 잠재결과가 실현되는(realized) 경우 사실(factual)
 - Vs. 잠재결과가 실현되지 않는 경우 대안사실(counterfactual)
- **대안사실** (counterfactual)
 - 실현되지 않았지만, 원인변수가 달라지면 존재할 수 있는 결과변수의 값
 - 즉, 원인이 되는 실제 사건을 "반대로" 가정했을 때의 결과값

1. 개념 1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

대안사실 체계 (counterfactual framework)

표 1-1. 해열제 처치여부에 따라 실현된/실현되지 않은 잠재결과

	해열제를 먹을 경우 잠재결과(체온)	해열제를 먹지 않을 경우 잠재결과(체온)
A (45세)	36.5°C (사실-관측됨)	
B (26세)		38.0°C (사실-관측됨)

Q. 해열제의 처치효과 τ 는 무엇일까?

A. $\tau = 36.5 - 38.0 = -1.5$ (?)

1. 개념 1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

대안사실 체계 (counterfactual framework)

표 1-2. 해열제 처치여부에 따라 실현된/실현되지 않은 잠재결과

	해열제를 먹을 경우 잠재결과(체온)	해열제를 먹지 않을 경우 잠재결과(체온)
A (45세)	36.5°C (사실-관측됨)	39.0°C (대안사실-관측되지 않음)
B (26세)	36.0°C (대안사실-관측되지 않음)	38.0°C (사실-관측됨)

$$\begin{array}{l}
 \text{A.} \quad \tau_A = 36.5 - 39.0 = -2.5 \\
 \quad \quad \tau_B = 36.0 - 38.0 = -2.0
 \end{array}
 \left. \vphantom{\begin{array}{l} \tau_A \\ \tau_B \end{array}} \right\} \tau_{\cdot} = \frac{-2.5 - 2.0}{2}$$

1. 개념

1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

잠재결과 측정의 문제

- ① 개인 수준에서 측정이 불가능
 - 한 개인은 한 가지 처치에만 놓임: 해열제를 먹든지, 먹지 않든지
 - 개인 수준에서 처치효과(즉 해열효과)를 측정할 수는 없음
 - ⇒ 처치효과는 "**집단 수준**의 평균값"으로 정의됨
- ② 처치 여부에 따른 편향 발생위험
 - 처치를 받는/받지 않는 집단 사이에 체계적인 차이가 있을 수 있음
 - 예를 들어, 연령이 높을수록 해열제를 먹을 확률이 높다면?
 - RCT에서는 무작위배정을 통해 편향이 제거됨
 - 그러나 **실험이 불가능한 관측연구** 상황이라면 **통계적 해결법**이 필요

1. 개념

1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

대안사실 체계와 루빈 인과모형

표 1-3. 해열제 처치여부에 따른 잠재결과

집단구분	해열제를 먹을 경우 잠재결과(체온)	해열제를 먹지 않을 경우 잠재결과(체온)	공변량(연령)
$T_i = 1$	$Y_i^1 T_i = 1$	$Y_i^0 T_i = 1$	$E(X_i) = 50$
$T_i = 0$	$Y_i^1 T_i = 0$	$Y_i^0 T_i = 0$	$E(X_i) = 30$

※ 집단 수준의 평균값으로 처치효과 정의: ATT, ATC, ATE

1. 개념

1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

처리효과의 종류

- **ATT** (average treatment effect for the treated)

$$\tau^1 = E(Y_i^1 - Y_i^0 | T_i = 1)$$

- **ATC** (average treatment effect for the control)

$$\tau^0 = E(Y_i^1 - Y_i^0 | T_i = 0)$$

- **ATE** (average treatment effect; Morgan & Winship, 2015)

$$\begin{aligned} \tau &= E(Y_i^1 - Y_i^0) = \sum_t \text{prob}(T_i = t) \times E(Y_i^1 - Y_i^0 | T_i = t) \\ &\equiv \pi \times E(Y_i^1 - Y_i^0 | T_i = 1) + (1 - \pi) \times E(Y_i^1 - Y_i^0 | T_i = 0) \end{aligned}$$

1. 개념

1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

처치효과 추정의 문제

- 관측한 τ_{obs} 는 $ATT(\tau^1)$, $ATE(\tau)$ 모두와 다름

$$\tau_{obs} = E(Y_i^1 | T_i = 1) - E(Y_i^0 | T_i = 0)$$

$$= E(Y_i^1 | T_i = 1) - E(Y_i^0 | T_i = 1) + E(Y_i^0 | T_i = 1) - E(Y_i^0 | T_i = 0)$$

$$= \tau^1 + \boxed{E(Y_i^0 | T_i = 1) - E(Y_i^0 | T_i = 0)}$$

자기선택 편향 (self-selection bias)
: 무작위 배정의 경우에는 0으로 사라짐

$$\neq \tau^1 \left(= E(Y_i^1 - Y_i^0 | T_i = 1) \right)$$

$$\neq \tau \left(= E(Y_i^1 - Y_i^0) \right)$$

관측연구 상황에서 τ 를 추정하려면?

1. 개념

1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

루빈 인과모형의 해결법

- ITAA (the ignorable treatment assignment assumption)

$$(Y_i^1, Y_i^0) \perp T | e(X), \quad 0 < e(X) = \text{prob}(T = 1 | X) < 1$$

- ITAA 아래 τ 추정이 가능해짐 (Rosenbaum & Rubin, 1983)

$$\begin{aligned} \tau_{\text{obs}} &= E(Y_i^1 | T_i = 1) - E(Y_i^0 | T_i = 0) \\ &= E(E(Y_i^1 | T_i = 1, X_i) - E(Y_i^0 | T_i = 0, X_i)) \\ &= E(E(Y_i^1 | X_i) - E(Y_i^0 | X_i)) \quad \leftarrow e(X) \text{ 추정치가 주어졌을 때 } T \text{와 독립} \\ &= E(Y_i^1 - Y_i^0) = \tau \end{aligned}$$

1. 개념 1) 핵심어: 인과추론, 루빈 인과모형, 성향점수

성향점수 (propensity score)

- 개념적 정의
 - $e(X) = \text{prob}(T = 1|X)$
 - 특정 개체가 통제집단이 아닌 **처치집단에 배치될 확률**
- 필요성 및 역할
 - 루빈 인과모형에서 성향점수 $e(X)$ 를 이용하여 처치효과 추정
 - 주어진 공변량 X 를 이용해 원인배치 과정 T 를 **모형화**하는 것
 - 보다 타당한 처치효과 추정치를 얻는 것이 최종 목표, 구체적인 $e(X)$ 의 형태나 추정방식보다는 “적절한 모형인지” 확인하는 과정이 더 중요

1. 개념 2) 주요 기법: 성향점수가중, 매칭, 층화

성향점수분석 기법 분류

- 성향점수 추정치를 활용하는 방식에 따라 분류
- 크게 성향점수가중, 매칭, 층화의 세 가지 기법으로 분류됨
 - ① **성향점수가중** (propensity score weighting)
 - 처치집단의 성향점수와 통제집단의 성향점수가 같아지도록 가중치를 부여
 - ② **성향점수매칭** (propensity score weighting)
 - 처치집단의 사례들과 비슷한 성향점수를 가진 통제집단 사례들을 매칭(짝짓기)
 - ③ **성향점수층화** (propensity score subclassification)
 - 처치집단과 통제집단 사례들을 성향점수가 비슷한 K개 집단으로 층화

1. 개념 2) 주요 기법: 성향점수가중, 매칭, 층화

성향점수분석 진행절차

표 2. 성향점수분석의 주요 단계 개괄

주요 단계	주요 체크사항
1. 연구설계	RCT로 진행할 수 없는 연구인가? 공변량이 충분히 측정되었는가?
2. 데이터 전처리	변수들을 적절히 리코딩하였는가? 결측값을 처리하였는가?
3. 성향점수 추정	적절한 방식으로 추정하였는가? 상호작용항을 포함시켰는가?
4. 성향점수분석 실행	성향점수가중/매칭/층화 등 기법을 타당하게 수행하였는가?
5. (공변량) 균형성 점검	처치 및 통제집단 간에 성향점수 및 공변량이 균형을 이루었는가?
6. 처치효과 추정	연구자가 관심 있는 처치효과를 타당하게 추정하였는가?
7. 민감도 분석	추정된 처치효과는 얼마나 강건(robust)한가?

2. 적용 3) 성향점수분석 실습

[데이터셋] data("lalonde", package="MatchIt")

[분석목표] 직업훈련프로그램(75~78년 실시)의 임금인상(78년 측정) 효과

(1) 성향점수가중 기법

3. 성향점수를 사용한 경우

```
glm_fit = glm(myformulaT, lalonde, family=binomial(link='logit'))
```

← 성향점수 추정

3-1) 성향점수가중 기법

추정하고자 하는 처치효과의 종류(ATT,ATC,ATE)에 따라 가중치 부여

```
myd = lalonde %>%
```

```
  mutate(
```

```
    ps=glm_fit$fitted.values, #추정확률 추출
```

```
    iptw=ifelse(treat==1,1,1/(1-ps)) #IPTW = ATT 계산용 가중치
```

```
  )
```

IPTW 부여후 회귀분석

```
lm(myformulaY, myd, weights=iptw) %>% summary()
```

← 효과추정치(ATT) 종류에 따른 가중치 부여:
처치역확률가중치 (IPTW, inverse probability of treatment weight)

← 효과추정치 계산

2. 적용 3) 성향점수분석 실습

(2) 성향점수매칭 기법

```

# 3-2) 성향점수매칭 기법: 그리디 매칭
set.seed(12345) #매칭결과를 정확하게 반복하기 위해
# 1단계: matchit() 함수
m_out=matchit(myformulaT,
              myd,
              distance = "glm",
              link = "linear.logit",
              method='nearest',
              replace=TRUE)
#원인처치 확률 추정식
#데이터
#성향점수 추정방식: GLM
#성향점수 형태: 선형로짓
#그리디 매칭, 최인접사례 매칭
#반복추출 허용
#summary(m_out) #매칭결과 점검
#plot(summary(m_out)) #공변량 균형성 점검
# 2단계: match.data() 함수
dm_out=match.data(m_out)
#dm_out
# 3단계: 모수통계기법 기반 ATT 계산
lm(myformulaY, dm_out, weights=weights) %>% summary()

```

MatchIt 패키지를 통해 "매칭"

← 공변량 균형성이 달성되었는지 확인

← 효과추정치 계산 (모수적/비모수적)

2. 적용 3) 성향점수분석 실습

(3) 성향점수총화 기법

```
# 3-3) 성향점수총화 기법
set.seed(12345) #매칭결과를 정확하게 반복하기 위해
# 1단계: matchit() 함수
m_out=matchit(myformulaT,           #원인처치 확률 추정식
              myd,                  #데이터
              distance = "glm",     #성향점수 추정방식: GLM
              link = "linear.logit", #성향점수 형태: 선형로짓
              method="subclass",    #성향점수총화
              subclass=6)           #디폴트값 K=6

#summary(m_out) #매칭결과 점검
#plot(summary(m_out)) #공변량 균형성 점검

# 2단계: match.data() 함수
dm_out=match.data(m_out)
#dm_out

# 3단계: 모수통계기법 기반 ATT 계산
lm(myformulaY, dm_out, weights=weights) %>% summary()
```

← 총화기법에 맞게 파라미터 설정

3. 확장 4) 다양한 상황에서의 성향점수분석

- 기법의 확장
 - 성향점수가중/매칭/층화 기법 외 다른 방식으로 성향점수 활용
 - 예) 준-정확매칭 (CEM, coarsened exact matching)
- 원인변수의 확장
 - 이분형(binary) 대신 범주형, 연속형 원인변수로의 확장도 가능

※ 성향점수분석 단계별 체크사항 점검

- 분석 상황에서 주요 체크사항이 어떻게 반영되어야 할지 반드시 점검
- 예1) 사용하는 기법에서 어떤 민감도 분석을 수행할 수 있는지?
- 예2) 처치수준이 3개 이상인 경우 어떻게 공변량 균형을 확인할지?
- 예3) 군집형 데이터일 경우 ITAA 등 가정사항을 어떻게 충족시킬지?

참고문헌

- 백영민·박인서 (2021). 《R기반 성향점수분석: 루빈 인과모형 기반 인과추론》. 서울: 한나래.
- Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software* 42(8).
<http://www.jstatsoft.org/v42/i08>
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), pp. 945-960.
- Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge University Press.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), pp. 41-55.

