

매칭 방법을 활용한 인과추론

관찰연구에서 매칭을 통한 인과효과 추정



목표

관찰 연구에서 처리와 대조군의 공변량 분포를 유사하게 만들어 편향을 줄임.



연구 배경

1970년대 이후 매칭 방법에 관한 연구는 비 비교를 위해 처치군과 대조군 피험자를 어떻게 가장 잘 선정할지에 대해 다룸



적용 분야

경제학, 역학, 의학, 정치학 등 다양한 분야에서 인기 상승.

매칭의 넓은 정의

- 처리군과 대조군의 공변량 분포를 균형시키는 모든 방법을 아루는 넓은 개념
- 단순히 한 쌍식 상대를 찾는 '1:1 매칭' 뿐만 아니라, 조건별로 여러 그룹을 나누는 '서브클래시화'(subclassification), 조건이 비슷한 집
스한 집단에 더 많은 비중을 주는 '가중치 부여' 등 다양한 방식이 있음

공정한 인과추론을 위한 두 단계

1) 설계(Design)

- 결과값에 접근하지 않고 배경정보만으로 연구를 설계하여 무작위 실험처럼 균형 맞춤.
- 결과 정보를 사용하는 것이 금지됨

2) 결과분석(Outcome)

- 매칭 완료 후 처리와 대조의 결과를 비교하여 효과 추정.
- 매칭된 집단의 결과 차이를 통계적으로 분석해 진짜 효과 나타났는지 확인

매칭의 장점



직관적 보완성

- 회귀분석과 상충하지 않고 두 방법을 결합할 때 효율성이 높아짐
- 모델의 불안정성이나 해석 오류 줄일 수 있음



중첩(Overlap) 확인

- 처리·대조군의 공변량 분포가 충분히 겹치지 않는 영역을 명확하게 드러냄
- 중첩되지 않은 영역의 외삽(extrapolation) 위험을 중단



명확한 진단

- 평균, 분산, 그래프 등으로 직관적으로 평가와 진단이 가능
- 설계 단계에서 매칭 성능 확인이 가능하며 집단이 비슷해졌는지 객관적 수치로 판단

→ 매칭은 다양한 분야에서 인과추론 신뢰성과 설명력 높이는 데 널리 사용

매칭이 쓰이는 환경

01

결과값이 아직 없는 경우

- 미래 추적 조사를 위해서 대조군 일부를 선택하는 데 매칭을 사용
- 비용 및 자료 수집 제한이 있는 연구에 중요

02

결과값이 이미 관찰된 경우

- 기존 데이터 내에서 편향을 줄이기 위해 처치군과 대조군을 균형 맞춤
- 비실험적 연구에서 인과효과 추정에 활용

공통점: 결과값을 절대 사용하면 안됨

가정·한계와 향후 연구 과제

주요 가정: 강한 무편향성(ignorability)과 SUTVA. 관찰되지 않은 교란변수 민감도 분석과 중첩성 개선은 향후 연구 과제



무편향성 가정

관찰된 공변량으로 처치(처리) 배정이 완전히 설명되어야 하며, 이를 통해 처리 효과 추정의 뒤흔림 줄일 수 있음



SUTVA

한 단위의 처치가 다른 단위의 결과에 영향을 주지 않아야 함(유출, spillover 효과 배제)



향후 방향

- 측정되지 않은 교란에 대한 민감도 분석 필요
- 설계 민감성과 중첩성 부족 등 복잡한 실제 상황에 적용 확대 필요