

6. 연구설계 II: 이론과 가설, 방법

정치학연구방법론

박상훈 (sh.park.poli@gmail.com)

강원대학교

연구설계 II: 이론과 가설, 방법

Recap: 연구설계(Research Design)

연구설계는 학술적 탐구에서 연구 질문을 체계적으로 다루기 위한 핵심적 청사진

- 이를 통해 연구자는 문제 정의에서부터 자료 수집, 분석에 이르는 전체 과정을 논리적으로 구조화할 수 있음.
- 연구설계는 연구 질문을 명확히 설정하고, 가설을 도출하며, 이를 검증하기 위한 방법론을 선택하도록 도움.
- 또한 연구의 타당성과 신뢰도를 높이는 데 기여하며, 연구 활동 전반의 효율성을 극대화
- 정치학처럼 복잡한 사회적 현상을 다루는 사회과학 분야에서는 엄밀하고 신중한 설계가 필수적
- 결국 체계적 연구설계는 성공적인 학문적 기여를 위해 반드시 거쳐야 할 중요한 단계

연구설계 II: 이론과 가설, 방법

이론과 가설

이론:현실 세계의 다양한 현상을 설명하고 예측하기 위한 추상적 틀

- 이론은 경험적 현상을 체계적으로 설명하고, 가설을 통해 예측 가능성을 높이는 개념적 도구
- 명제(proposition)와 선행조건(initial conditions)을 결합해 특정 현상이 발생하는 메커니즘을 논리적으로 제시

가설: 이론에서 도출된 검증 가능한 진술

- 연구설계에서 이론은 연구 문제가 놓인 맥락과 기초 가정을 제시, 가설은 이를 구체적 주장으로 나타냄.

연구설계 II: 이론과 가설, 방법

이론

사회과학에서 이론은 결코 ‘완결된 진리’가 아니며, 검증과 반증 가능성을 전제로 함,

- 이런 잠정적 특성 때문에 이론은 계속해서 수정 및 보완되면서 발전
- 복잡한 현실 세계를 이해하기 위한 ‘지도’ 역할을 하며, 연구자가 문제를 구조화하고 가설을 도출하도록 이끔.

연구설계 II: 이론과 가설, 방법

가설

이론에서 추출된, 경험적 검증이 가능한 명제

- "교육 수준이 높을수록 정치 참여가 증가한다"처럼 구체적이고 관찰 가능한 형태
- 이론의 핵심 기제(mechanism)를 현실에서 확인하는 도구
- 가설이 검증 과정을 거쳐 일관된 지지를 받으면 점차 이론의 지위를 얻게 되지만, 반례가 발견되면 수정 또는 폐기가 불가피

연구설계 II: 이론과 가설, 방법

왜 중요하지?

이론과 가설이 연구설계에서 중요한 이유는 '무엇을 어떻게 측정할 것인가'를 안내하기 때문

- 가설이 없으면 연구자가 어떤 변수를 수집해야 하는지 막연해짐.
- 이론 없이 가설을 제시하면 그 배경과 정당성이 부족해짐.

정치학 연구에서

- 이론은 종종 행위자의 동기나 구조적 제약을 설명하는 '큰 그림'을 제시
- 가설은 특정 조건이 주어졌을 때 어떤 결과가 나타날지 설명

이론과 가설은 서로 보완하면서, 연구 방향을 구체화

연구설계 II: 이론과 가설, 방법

왜 중요하지?

가설은 반드시 **검증 가능성**을 전제

- '민주주의는 무조건 선하다' 같은 문장은 도덕적 선언일 수 있지만, 경험적으로 어떻게 입증할지 모호
- 반면 '민주주의 체제일수록 국민 소득 수준이 향상된다'는 가설은 경제 지표와 정치체제에 관련된 데이터를 통해 검증가능

이렇게 측정 기준과 검증 방법이 명확해야 가설의 진위 여부를 평가할 수 있음.

정치학에서는 복잡한 변수가 많아 완벽한 검증이 어려울 수 있으나, 가능한 한 엄밀한 자료와 절차를 마련해 합리적으로 접근해야 함.

연구설계 II: 이론과 가설, 방법

왜 중요하지?

사회과학에서 이론은 영구적으로 확정되지 않으며, 가설 검증 결과에 따라 지속적으로 수정됨.

- 예를 들어 어떤 선행연구가 다수의 경험적 지지를 받아 이론으로 자리 잡았어도, 새로운 데이터나 방법론으로 반례가 발견되면 이론이 변형될 수 있음.
- '반증 가능성'이야말로 과학적 탐구의 핵심 특징
 - 과학적 이론은 결코 완전무결한 진리가 아니며, 더 나은 설명력이 제시되면 교체될 수 있음.
 - 연구자는 가설 검증 과정에서 '어떤 결과도 수용할 준비가 되어 있어야 한다'는 태도를 견지

연구설계 II: 이론과 가설, 방법

이론과 가설

예측 정확도를 과도하게 강조하는 태도에 대한 비판 (Clarke & Primo, 2012)

- 이론 모델이 주는 통찰과 다양한 상황에서의 적용 가능성을 강조
- 모델이 현실을 완벽히 재현하지 못해도, 주요 변수와 인과관계를 명확히 부각해 준다면 이론적으로 의미가 있음.
 - 예를 들어 투표행태 모형이 실제 투표율을 정확히 맞이지 못하더라도, 유권자 동기나 제도적 요인을 체계적으로 정리해 연구자가 더 발전된 가설을 만들 수 있도록 돕는 것
- 그러므로 이론의 가치는 '유용성'에도 있음.

연구설계 II: 이론과 가설, 방법

가설의 검증

반사실적(counterfactual) 접근

- "만약 어떤 사건이 없었다면 결과가 달라졌을까?" 같은 가정을 통해, 실제로 작동한 인과기제를 도출하려는 시도
- 예컨대 특정 동맹이 없었다면 전쟁이 발생했는지 따져 보는 방식

이때 선불리 주관적 판단에 빠지지 않도록, 현실 자료와 역사적 증거를 꼼꼼히 대조해야 함.

- 반사실적 사고는 이미 일어난 사건이므로 실험이 불가능한 정치현상에서 특히 유용

한편, 대안적 가설이 나오면, 기존 이론도 재검토되어야 하고, 더 풍부한 설명으로 개선될 필요 (Ostrom 1998)

연구설계 II: 이론과 가설, 방법

그렇다면 사례 연구는?

특정 사건이나 집단을 심층 분석하면, 양적 연구로는 발견하기 어려운 메커니즘을 파악할 수 있음 (Gerring 2004).

- 이 과정에서 새로운 가설이 도출되거나, 기존 이론의 취약점이 드러날 수 있음.
- 비록 사례연구가 일반화에 취약하다는 비판이 있어도, 적절한 사례 선정과 비교를 통하면 인과적 과정에 대한 정밀한 통찰을 얻을 수 있음.

연구설계 II: 이론과 가설, 방법

그렇다면 사례 연구는?

정치학에서 일반이론을 추구하는 목적과 한계 (Hill 2012)

- 보편적 법칙을 찾으려는 시도는 다양한 맥락에서 적용 가능한 지식을 제공하지만, 각 사회, 문화의 독특성을 무시한다는 우려
- 물론 일각에서는 완벽한 보편성을 즉시 얻기는 어렵더라도, 비교연구와 반복적 검증을 통해 점진적으로 이론을 확장해야 한다고 주장
 - 이는 궁극적으로 다양한 사례와 맥락을 포괄하며 더 높은 수준의 이론화를 가능케 하기 때문이라는 주장
- 하지만 일반이론을 지향하되, 현실의 복잡성도 존중해야 할 필요

연구설계 II: 이론과 가설, 방법

이론과 가설의 검증

검증에는 다양한 방식이 사용되지만, 공통적으로 요구되는 것은 '**자료 기반의 엄밀성**'

- 표본 추출(sampling), 조사 설계, 분석 방법 등을 적절히 활용해 가설의 성립 여부를 평가해야 함.
- 단순히 가설을 세우고 주장만 하는 것이 아니라, 신뢰성 있고 타당한 자료에 근거해야 '과학적'이라고 말할 수 있음.

예컨대 투표율 가설을 검증하려면 인구 통계, 선거 자료, 여론조사 등 구체적인 데이터를 체계적으로 수집·분석해야 한다. 이것이 곧 연구설계 전체에서 이론과 가설이 빛을 발하는 핵심 과정이다. 여기에 대한 의견이 있으신가요?

연구설계 II: 이론과 가설, 방법

표본 추출

표본 추출을 잘못하면 아무리 이론과 가설이 훌륭해도 왜곡된 결론이 나올 수 있음.

- 예를 들어 특정 지역이나 특정 집단만 과도하게 대표되는 표본을 사용하면, 그 결과를 전체 모집단에 일반화하기 어려워짐.
- 따라서 연구자는 이론적 가정과 연구 목적에 맞춰, 표본 선정 방법을 신중히 결정해야 함.

정치학 연구에서 대표성 확보는 특히 중요

연구설계 II: 이론과 가설, 방법

표본 추출

표본(sample)과 표집(sampling)

표집(sampling)은 모집단 전체를 대표할 수 있는 일부 사례들을 선택하는 과정

- 모집단은 '해당 연구 주제와 관련된 모든 대상'을 의미
- 표본은 그중 실제로 관측, 조사하는 일부분
 - 예컨대 "유권자 태도"를 연구한다면, 모집단은 전체 유권자이고, 그중 일정 수를 뽑아 설문조사를 실시
- 정확한 표본을 얻으려면, 우선 연구 범위와 주요 특성을 파악해야 함.
 - 이를테면 지역, 연령, 성별 등 다양한 변수를 고려해 대표성을 높이는 표본을 구성해야 함. 그렇지 않으면 결론이 왜곡될 위험이 큼.

연구설계 II: 이론과 가설, 방법

모집단(population)

모집단(population)은 연구자가 일반화하고 싶은 대상 전체

- 예를 들어 "한국 유권자"가 모집단이라면, 이 범주에 속하는 모든 성인 국민이 포함됨.
- 그러나 실제로는 모집단 규모가 방대해 전체를 모두 조사하기 어려움.
 - 그래서 '표본'을 통해 모집단 특성을 추론
 - 여기서 중요한 전제는 "표본이 모집단을 충분히 대표한다"는 것
 - 만약 특정 계층이 과다 혹은 과소 대표되면, 얻어진 결과가 모집단 전반에는 적용될 수 없음. 그러므로 모집단 정의와 표본 설계는 긴밀히 연결됨.

연구설계 II: 이론과 가설, 방법

모집단과 표본

대표성(representativeness): 표본이 모집단의 주요 특징들을 제대로 반영하는 정도

정치학 연구에서 대표성이 떨어지면, 예를 들어 ‘정권 지지율’이나 ‘정치 신뢰도’ 등 핵심 지표를 왜곡된 값으로 추정할 수 있음.

- 결국 잘못된 결론이 나오고, 이후 정책 제안이나 이론 발전에도 오류가 누적
- 대표성을 보장하기 위해서는 모집단의 다양한 분포(예: 지역, 성별, 소득 계층 등)를 고려해 표본을 구성해야 함.

확률적 표집 기법이 이를 달성하는 유력한 수단

연구설계 II: 이론과 가설, 방법

확률적 표집

확률적 표집(probability sampling): 모집단 내 각 사례가 일정 확률로 뽑히도록 설계하는 방법

무작위성(randomness)이 확보되어, 통계적으로 대표성을 추정할 수 있음.

- 확률적 표집 방식에는 단순무작위 표집, 체계적 표집, 층화표집, 군집표집 등
- 확률적 표집의 장점은 오차 범위나 신뢰 수준을 계산 가능하다는 것
- 정치학 연구에서 대규모 여론조사나 선거 예측 시 확률적 표집이 많이 사용되며, 결과 해석 시 '표본오차 $\pm X\%p$ ' 형태로 제시되는 이유

연구설계 II: 이론과 가설, 방법

확률적 표집

단순무작위 표집(simple random sampling)

대표성을 확보하는 가장 기본적인지만 강력한 방식

모집단의 모든 사례가 동일한 확률로 표본에 뽑히도록 설정

- 예를 들어 유권자 명부가 있다면, 거기서 무작위로 일정 수를 추출
- 이 방식은 편향을 최소화하지만, 전제 조건이 까다로움.
 - 예컨대 모집단 명부를 정확히 확보해야 하며, 무작위 절차가 철저히 지켜져야 함. 그래야 각 사례가 뽑힐 확률이 동등해짐.
 - 단순무작위 표집은 이해하기 쉽지만, 실제 연구현장에서는 비용과 자료수집 문제가 있을 수 있음.

연구설계 II: 이론과 가설, 방법

확률적 표집

체계적 표집(systematic sampling)

일정한 간격(k번째)으로 사례를 추출하는 방법

- 예를 들어 유권자 명부를 무작위로 섞은 뒤, 첫 번째 대상에서 시작해 k번째마다 추출
- 이 방법은 단순무작위 표집처럼 동일 확률을 부여하면서도, 실제 추출 과정을 단순화

다만 명부 자체에 어떤 규칙적 패턴이 있으면(예: 가나다 순으로 몰린 지역), 표본이 편향될 가능성이 존재

따라서 체계적 표집을 적용할 때는 명부가 무작위 배열에 가깝도록 확인해야 함.

연구설계 II: 이론과 가설, 방법

확률적 표집

층화표집(stratified sampling)

모집단 내 하위집단(층, strata)이 존재할 때, 각 층이 표본에 적절히 반영되도록 분할한 뒤 무작위 추출을 수행하는 방식

- 예컨대 전국 성인 인구를 학력, 소득, 지역 등으로 층화하고, 각 층에서 비례할당으로 무작위 표집하면 대표성이 향상
- 이 방법은 특히 정치 여론조사에서 자주 사용
 - 예를 들어 수도권, 비수도권, 남성, 여성, 연령대별 분포가 실제 인구 비율에 맞춰 표본에 반영될 수 있으므로, 추정의 정확도가 높아짐.

연구설계 II: 이론과 가설, 방법

확률적 표집

군집표집(cluster sampling)

모집단 명부가 없거나, 지리적으로 광범위해 일일이 무작위 표집하기 어려울 때 사용

- 먼저 큰 단위(예: 지역구, 도시)를 무작위로 뽑고, 그 안에서 더 작은 단위를 무작위로 뽑는 식으로 단계적으로 좁혀감.
- 예를 들어 전국 시도를 무작위로 선택한 뒤, 해당 시도 내 시군구를 또 무작위로 추출하고, 최종적으로 거주민 명단에서 표본을 뽑는 식

군집표집은 현장조사 비용을 줄이는 장점이 있지만, 표집오차가 커질 수 있어 충분한 군집과 계층화를 병행해야 함.

연구설계 II: 이론과 가설, 방법

비확률적 표집(non-probability sampling)

모집단 내 각 사례가 뽑힐 확률을 엄밀하게 통제하지 않는 방식

- 편의표집, 판단(판정)표집, 할당표집 등
- 이 방법들은 주로 예비연구나 초기 탐색 목적으로 사용되며, 대표성 추정이 어렵다는 단점
- 그럼에도 연구자가 접근 가능한 자료나 참여자를 신속하게 확보할 수 있어, 시간, 비용 제약이 심할 때 활용될 수 있음.

정치학에서는 규모가 작은 탐색연구나 파일럿 스터디, 특정 전문가 집단 분석 등에서 비확률적 표집이 종종 선택

연구설계 II: 이론과 가설, 방법

비확률적 표집

편의표집(convenience sampling)

연구자의 편의에 따라 쉽게 접근할 수 있는 사례들로 구성되는 방법

- 예를 들어 주변 지인이나 대학 캠퍼스 학생들을 대상으로 간단히 설문조사하는 방식
- 빠르고 저렴하다는 장점이 있지만, 모집단 대표성과는 거리가 멀어 일반화가 어려워짐.
 - 따라서 편의표집은 주로 설문 문항의 타당성 확인 등 예비 단계에서 활용되며, 본격적인 연구결과로 일반화하기에는 무리가 있음.
 - 정치학처럼 다양한 인구집단을 다뤄야 하는 경우, 편의표집의 한계를 명확히 인지해야 함.

연구설계 II: 이론과 가설, 방법

비확률적 표집

판정표집(judgemental sampling)

연구자가 "이 사례들은 특별한 통찰을 줄 것이다"라고 판단해 주관적으로 선별하는 방식

- 예컨대 정치 참여가 높은 활동가 그룹을 골라서 심층 인터뷰를 진행하는 식

연구자의 전문성이 높다면 유의미한 사례를 골라 풍부한 정보를 얻을 수 있지만, 표본 편향이 발생할 위험이 큼.

- 따라서 이 방법은 대표성 있는 결론을 내기보다는, 특정 현상을 심층적으로 이해하는 목적에 적합
- 정치학에서 특정 국가나 지역의 극단 사례를 분석할 때도 종종 사용

연구설계 II: 이론과 가설, 방법

비확률적 표집

할당표집(quota sampling)

모집단의 인구학적 분포(예: 성별, 연령대 등)에 맞춰 표본 할당량을 미리 정해둔 뒤, 연구자가 임의로 대상자를 채우는 방식

- 예컨대 "남성 50%, 여성 50%"라는 쿼터에 맞춰 적당한 응답자를 골라 조사
- 이론상 층화표집과 유사해 보이지만, 표본 추출 과정에 무작위성을 엄격히 도입하지 않음.

그러므로 실제 결과가 모집단을 온전히 대표한다고 보장하기 힘들.

다만 빠르게 목표 구성비를 맞출 수 있어 상업적 여론조사에서 종종 이용되기도 함.

연구설계 II: 이론과 가설, 방법

표집오차(sampling error)

표본 규모를 결정하는 핵심 요인

- 표집오차란 표본 특성이 모집단 전체의 참값(true value)과 어느 정도 차이가 나는지를 의미
- 일반적으로 표본 크기가 커질수록 표집오차는 감소하지만, 그에 따른 조사 비용과 시간도 증가
- 연구자가 용인할 수 있는 오차 범위(예: $\pm 3\%$)와 신뢰수준(예: 95%)이 있으면, 통계 공식을 통해 필요한 최소 표본 크기를 산출할 수 있음.

연구설계 II: 이론과 가설, 방법

표집오차(sampling error)

표집오차를 계산할 때는 보통 모집단 비율 p 를 0.5로 가정

- 이때 $p(1-p)$ 가 최대치가 되어 필요한 표본 크기가 가장 크게 산출
- 예를 들어 95% 신뢰수준($Z=1.96$ 근사)에서 $\pm 5\%$ 오차를 허용하려면 약 400명, $\pm 3\%$ 오차를 허용하려면 약 1,100명 정도가 필요하다는 계산

실제 연구에서는 모집단의 이질성, 예산, 시간 등을 고려해 현실적으로 표본 수를 결정

이론적으로는 표본이 많을수록 정확해지지만, 자원의 한계를 감안해 적정선을 찾는 것이 중요

연구설계 II: 이론과 가설, 방법

표본 크기, 신뢰수준, 그리고 대표성

신뢰수준을 높이고 오차범위를 줄이려면 당연히 표본 크기를 확대해야 함.

그러나 아주 큰 표본도, 추출 과정에서 특정 계층이나 지역을 소홀히 다루면 대표성이 떨어질 수 있음.

- 즉, 표본 수가 많은 것만으로는 충분치 않고, 표집 방법 자체가 적절해야 함.

따라서 연구 설계 단계에서 "얼마나 정확한 결과를 원하는가?", "어떤 방식으로 표본을 뽑을 것인가?"를 종합적으로 고민해야 함.

연구설계 II: 이론과 가설, 방법

표본 크기, 신뢰수준, 그리고 대표성

표본 설계 단계에서 가장 중요한 과제 중 하나는 오류를 최소화하는 것

- 표본추출 과정의 편향(selection bias), 조사대상자의 무응답(non-response bias), 측정도구의 문제 등이 모두 결과에 영향
- 예컨대 특정 정치성향 집단이 응답을 기피한다면, 결과가 편중될 수 있음.

이를 해결하기 위해 연구자는 여러 보완책을 마련

- 무작위 표집 준수, 층화, 가중치 조정, 반복 추적 조사 등 다양한 기법을 사용해 대표성 왜곡을 줄이려 노력

연구설계 II: 이론과 가설, 방법

표본 크기, 신뢰수준, 그리고 대표성

연구윤리 측면에서도 표본 추출은 중요

- 설문조사나 인터뷰를 진행할 때, 개인정보나 민감 정보가 유출되지 않도록 주의해야 하며, 동의 절차와 비밀 보장이 필수적
- 또한 특정 집단만 계속해서 조사 대상이 되거나, 연구 편의를 위해 소외 계층을 배제하는 행위는 윤리적 문제를 야기
- 정치학 연구가 궁극적으로 공익과 학문 발전을 추구한다면, 표본 추출 과정에서도 공정성과 투명성을 지켜야 함.
 - 이를 위해 IRB(기관윤리심의위원회) 검토나, 사전에 충분한 정보 제공을 통한 동의 절차 확보가 중요

연구설계 II: 이론과 가설, 방법

표본 크기, 신뢰수준, 그리고 대표성

결국 이론과 가설이 제시하는 요인을 검증하기 위해서는, 적절한 표본 추출이 뒤따라야 함.

- "교육 수준이 높을수록 투표율이 높아진다"는 가설을 세웠다면, 교육 수준이 다양한 응답자를 골고루 뽑아야 유효한 결론을 얻을 수 있다.
- 이론 모델은 현실을 단순화한 도구이고, 표본 추출은 그 모델을 실증하는 연결고리
 - 만약 샘플링이 부실하다면, 고안된 이론과 가설 역시 '보편 타당한 설명력'을 확보하기 어려움.

연구설계 II: 이론과 가설, 방법

나가며

이론과 가설은 연구에서 '무엇을' 검증할지 설정해 주고, 표본 추출과 모집단 대표성은 '어떻게' 검증할지를 구체화

정치학은 복잡한 현상을 다루는 만큼, 이론적 가정과 실증적 기법이 조화를 이뤄야 일관성 있는 결론을 도출할 수 있음.

감사합니다!

궁금한 것이 있으면 언제든지 연락하세요.

강사 연락처

연락처	박상훈
	sh.park.poli@gmail.com
	sanghoon-park.com/
	영상바이오관 405