

■ 통계학(Statistics)이란?

◎ 동영상: “그냥도전! 500원 동전 천 번 돌리기”

- 500원짜리 동전을 돌렸을 때 학이 나올 확률이 70% 정도
- 500원짜리 동전을 1000번 돌리는 실험
- 실험 결과 1000번 중 학이 679번이 나옴
- 학이 나올 가능성이 70%정도 된다는 것이 얼추 맞다는 주장

◎ 통계학적 관점

- 500원짜리 동전을 돌렸을 때 학이 나올 확률이 70% 정도
⇒ 관심 또는 연구의 대상(문제, 주제, 가설)
- 500원짜리 동전을 1000번 돌리는 실험
⇒ 실험을 통해 자료를 수집
- 실험 결과 1000번 중 학이 679번이 나옴
⇒ 수집된 자료를 정리, 요약, 분석하여 자료의 특성을 파악

- 학이 나올 가능성이 70%정도 된다는 것이 얼추 맞다는 주장
⇒ 자료의 특성을 이용하여 관심 또는 연구의 대상에 대해
추론

◎ 관심의 대상 vs 자료

◎ 모집단(population)

- 잘 정의된 연구목적과 이와 연계된 명확한 연구대상을 설정
예】 대통령 후보의 지지율? ⇨ 유권자

- **연구대상이 되는 모든 개체의 집합**

예】 19대 대통령선거 선거인명부 유권자수는 42,432,413명

Q. “그냥도전!”에서의 모집단은?

∞

◎ 전수조사: 모집단 전체를 대상으로 조사하는 경우

- 조선(대한제국포함)시대 임금의 수명
 - 27명의 임금의 수명 자료
- 2010년까지의 인구주택총조사(census)

◎ 대부분의 모집단은 매우 커 전체를 조사하기 어려움

- 적절한 방법으로 일부의 자료를 추출해 조사
- 2015년부터 행정자료+20%표본조사

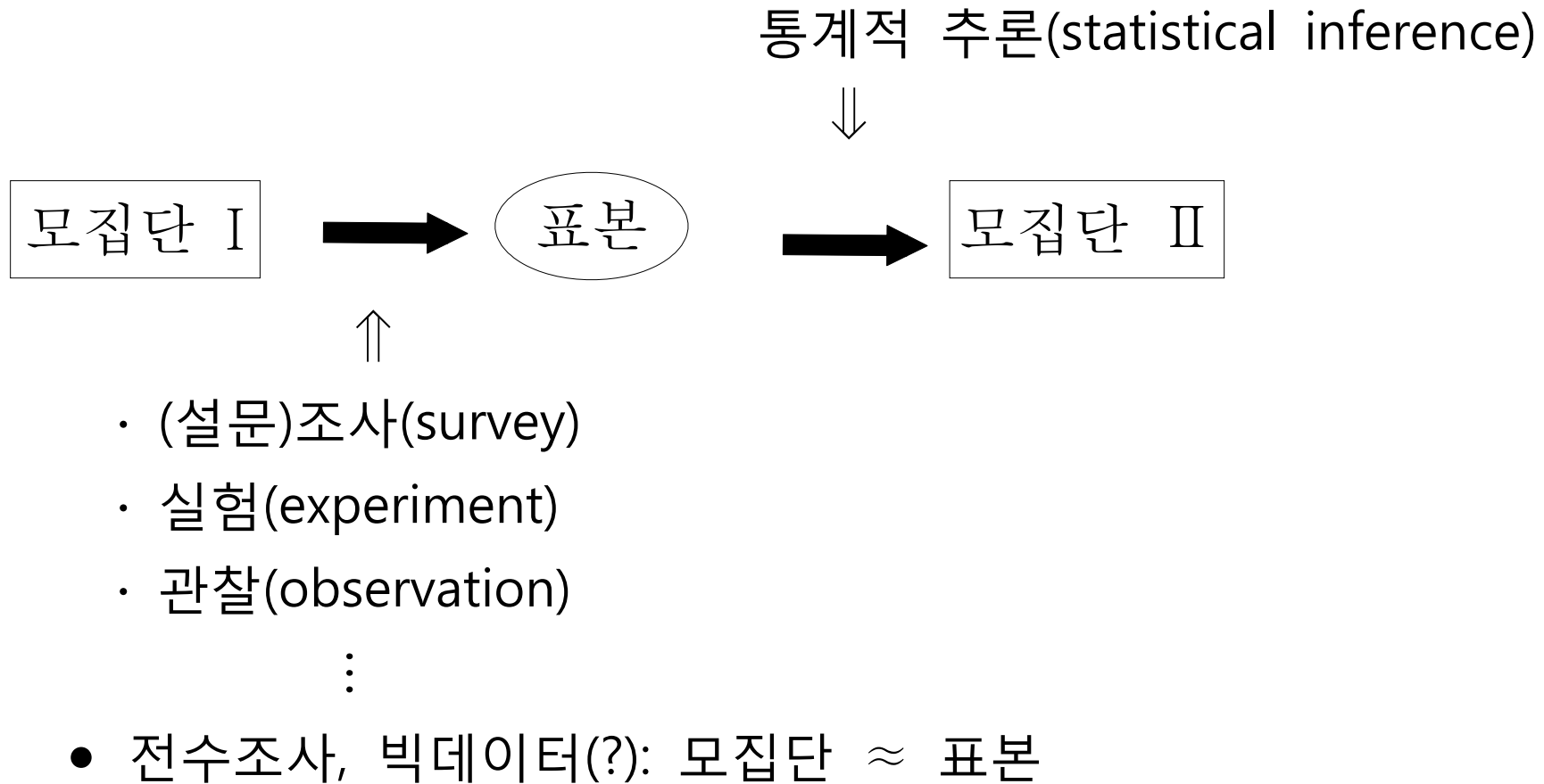
◎ 표본(Sample)

- 모집단으로부터 선택된 일부의 개체
예】 “그냥 도전!” 에서 나온 1,000번의 동전 결과
예】 각종 여론조사에 참여한 유권자
- 추출된 표본이 모집단을 대표할 수 있는가?
- 몇 명(개)의 표본을 어떻게 뽑아야 하는가?

◎ 1936년 미국대통령 선거

- 공화당의 Landon vs 민주당의 Roosevelt
- ‘Literary Digest’
 - 1916~1932년 선거결과 정확하게 예측
 - 구독자, 전화기 및 자동차 보유자 236만여 명의 의견을 분석
 - ⇒ Landon 57%, Roosevelt 43%
- ‘Gallup’ : 5만 명의 표본조사(할당추출)
 - ⇒ Landon 44%, Roosevelt 56%

- 선거결과: Roosevelt 63%, Landon 37%
 - 실제 득표율과 7% 차이를 보임
- 1948년 Dewey vs Truman 대결에서 Gallup 승자 예측 실패
 - ⇒ 표본선정방식 변경: 할당추출법 ⇒ 확률추출법



◎ 통계학이란

- 관심 또는 연구의 대상인 모집단의 특성을 파악하기 위해
- 모집단부터 일부의 **자료(표본)**를 수집하고
- 수집된 표본을 정리, 요약, 분석하여 **표본의 특성을 파악한 후**
- 표본의 특성을 이용하여 **모집단의 특성에 대해 추론하는**
원리와 방법을 제공하는 학문