

확률과 통계

- 통계표본 사용하기 -

참고서: Head First Statistics

(10장 통계표본 사용하기)

작성자 : 이계식

환경대학교, 컴퓨터공학과



표본 고르기

1. 모집단 : 조사 대상 전체

- 예) • 환경인 성인 동무게의 데이터가 요구될 때 • 환경인 성인 전체
- 지구 바다에 서식하는 돌고래의 수명 데이터가 요구될 때:
지구 바다에 서식하는 돌고래 전체
 - 어느 동선경 회사에서 생산하는 동선경의 맛의 지속시간 데이터가 필요할 때: 해당 공장에서 생산되는 점 전체

책에서 언급된
예제임.
11장에서 많이
활용됨.

문제점: 경우가 따라 모집단 전체에 대한 데이터를
구하지 못할 수 있거나 (돌고래 경우), 아니면
모집단 전체를 대상으로 하면 안되는 경우 (동선경
경우)가 존재한다.

그런 경우에는 **표본**을 선택하여 조사하는
표본조사 방식을 활용해야 함.

중조사: 표본조사와는 달리 모집단 전체를 대상으로
데이터를 조사하는 방식.

2. 표본 : 모집단에서 선택된 일부

- 예) • 지역별, 연령별, 성별을 고려하여 무작위로 선택된
성인 2,000명
- 요대양에서 작살한 비둘크 채집 조사 대상이 된
돌고래 100마리
 - 종류별로 무작위로 선택된 동선경 1,000개

3. 표본 선택방법 및 주의사항

다음의 과정을 통해 표본을 선택한다.

① 조사대상의 모집단 정의

② 모집단을 제대로 반영하는 표본 구성

↑ 어떻게? "전" 해안 환

↑ 매우 어려운 문제임.

하지만 자세한 내용은 강의에서

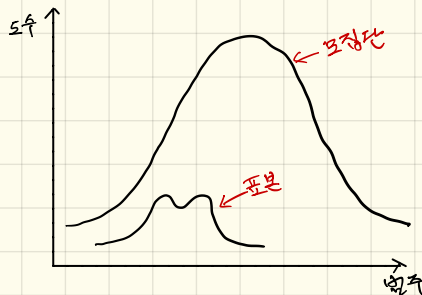
다루지 않음.

여기서는
상식적으로
판단할 것
전유!

← { 앞으로 표본이 주어졌다고
가정해서 모집단에 대한
정보를 믿는 다양한
통계 기법을 배울 것임.

4. 편향된 표본 : 모집단을 제대로 반영하지 못하는 표본

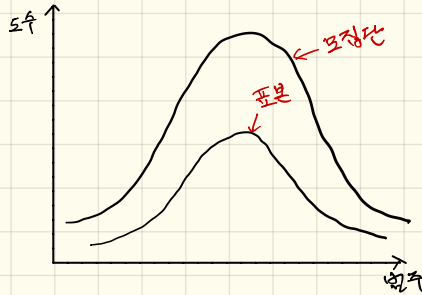
편향된 표본은 아래의 그림에서처럼 모집단과 다른 부포를
모인다.



예제: 동선점 맛의 지속시간을 알아내려 할 때
특정 맛의 점만을 표본으로 선택하면
다른 맛을 포함한 동선점에 전체에 대한
잘못된 정보를 유추할 수 있다.

5. 전체를 잘 반영하는 표본

모집단과 비슷한 분포를 갖는 표본을 선택해야 한다.



예제: 동선점의 맛 지속시간 조사하기

동선점은 네 개의 서로 다른 맛을 갖는다고 가정하자.
표본을 어떻게 구성해야 하는가?

전분당안:

- ① 모집단 설정: 개별적인 동선점 전체
- ② 맛별 동선점의 비율에 따라 무작위로 선택