

IV. 추측통계학(Inferential Statistics)

<생각해보자>

Q. 인간은
결정이나
예측을
얼마나 잘 할
수 있을까?

만약 결정능력이 매우 뛰어나다 학지라도 실제로는 그렇지 않을 수 있다.

단순히 개인적인 추측보다 더 믿을만한 ▷ 통계적 추론 방법을 제공하는 것이 수리 통계학자들의 몫이다. 통계적 추론은 모집단을 결정하는 수치적 기술측도인 모수(parameters)에 대한 예측이나 의사결정이다.

어떤 모집단에 관한 속성 A에 관심

Q1. 성인(30세 이상)의 일일 평균 나트륨 섭취량이 궁금하다. Q2. 성인(30세 이상)의 일일 평균 나트륨 섭취량이 3300mg이 넘는지 궁금하다.

Û

모집단을 대표하는 표본을 추출 & 표본의 속성 분석

30세 이상 성인 100명을 선발하여 하루 나트륨 섭취량 관측 : 평균 3400mg, 표준편차 1100mg

//

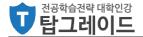
모수의 값을 추정하거나 예측하는 것

주장 : 하루 나트륨 섭취량은 3400mg이다. 모수의 값이 얼마가 될 것이라는 예상에 근거하여 모수의 값에 대한 의사결정을 내리는 것

> 주장 : 3300mg을 넘을 것이다' (그렇다/그렇지 않다)

추정 (Estimation) 가설검정

(Hypothesis Testing)



[1] 추정(Estimation)

- ① 모집단에서 추출한 표본에 기초하여 모평균, 모분산, 모비율과 같은 모집단의 확률분포의 모수가 어떤 값을 갖는가를 추측하는 과정
 - ① 추정량(estimator) : 모집단 특성을 추정하기 위하여 이용되는 통계량, 즉, 확률변수
 - © 추정치(estimate) : 추정량을 통하여 계산된 구체적인 수치
 - (예) 중학교 1학년의 평균몸무게에서

추정량: 표본평균 $\mu = \overline{X}$

추정치: 실제로 500명의 학생을 뽑아 관측한 몸무게의 평균값 45.5kg

- ② 점추정(point estimation)
 - ① 표본자료에 근거하여, 모수의 값으로 예상되는 하나의 값을 추정하는 것
 - © 점추정값(point estimate)
 - (예) 제품의 불량률 추정에서 불량률이 1%일 것이다.
- ③ 구간추정(interval estimation)
 - ① 표본자료에 근거하여, 모수가 속할 것으로 예상되는 구간을 추정하는 것
 - © 구간추정값 또는 신뢰구간 (confidence interval)
 - (예) 제품의 불량률 추정에서 불량률이 구간(0.5%, 1.5%)에 있을 것이다.

[2] 가설검정

- ① 모집단의 분포의 특성, 혹은 모수의 참값에 대하여 어떤 가설을 수립하고 모집단에서 추출한 표본에 기초하여 가설의 채택이나 기각을 결정하는 통계적 방법
 - ⊙ 실제 분포에 의한 통계적 검정
 - ① 집단간의 평균을 비교하고 연관성을 비교하는 검정
- ② Z- 검정, t- 검정, F- 검정은 종속변수가 양적변수일 때 집단간 평균비교를 위한 통계방법이다.
- ③ Z-검정과 t-검정은 두 집단 이하의 비교에 쓰이는 통계방법이다.
- ④ F-검정은 비교집단이 세 집단 이상일 때 집단간 평균비교를 위하여 쓰이는 통계방법으로 분산분석(ANOVA)이라 하다.
- ⑤ $\chi^2 \text{검정은 독립변수와 종속변수가 질적변수로 집단 간의 차 혹은 두 변수간의 연관성을 검정하기 위하여 사용되다.$
- ⑥ 독립변수와 종속변수가 원래 양적변수이나 어떤 기준에 의하여 그 양적변수를 범주화(categorize)하여 구분하였을 때 χ^2 검정을 사용할 수 있다.