9 장 확률표본추출

- 표본으로부터 얻어낸 통계량을 통하여 표본오차와 그 정도를 추정할 수 있음

제 1 절. 단순무작위표본추출법

- 크기가 N 인 모집단으로부터 모집단의 모든 표본단위가 선택될 확률이 모두 같도록 n 개의 표본단위를 선택하는 방법

1. 모집단 및 모수의 계산

(1) 측정할 모집단

- 사회학개론을 수강하는 학생들

(2) 모집단의 모수의 계산

2. 표본의 추출

(1) 표본의 추출방법

- 각 요소들은 선택될 확률이 동일함
- n 개의 일정한 수의 표본을 선정한다면 어떠한 조합이라도 선정될 확률은 동일함

표본추출과정

- 모집단 확정 및 표본프레임 작성
- 각각의 표본에 고유번호 부여
- 표본의 크기를 정하고 난수표를 이용하여 표본을 추출

(2) 표본통계량의 계산

3. 표본통계값에 의한 모집단의 모수추정

- 모집단 X 의 분포: $X \sim (\mu, \sigma^2)$
- μ 의 추정치: $\bar{X}$, σ^2 의 추정치: S^2
- 표본평균 $\bar{X}$의 분포: $\bar{X} \sim ((\mu, \sigma^2/n)$

- 점추정, 구간추정

제 2 절. 층화표본추출법

- 추정값의 표본오차를 감소시켜 표본의 대표성을 높이기 위해 사용되는 방법
- 모집단이 동질적이면 이질적인 모집단보다 표본오차가 더 적은 표본을 산출할 수 있다는 논리에 기초

1. 표본추출방법

표본추출방법

- 조사대상 모집단을 MECE 를 만족하는 소집단으로 나눈다
- 각 소집단들에서 얻을 표본의 비율을 정하여 각 소집단별 표본의 수를 정한다.

- 비례층화추출법: 소집단에서 추출할 표본의 크기를 각 소집단의 크기에 따라 비율적으로 정한다: 큰 집단에서는 많이, 적은 집단에서는 적게 추출
- 불비례층화추출법: 소집단의 크기와 분산을 모두 고려하여 표본을 추출하는 방법: 분산이 큰 소집단에

서는 많이, 분산이 적은 집단에서는 적게 추출

2. 비례층화추출법

(1) 표본의 추출

(2) 통계량추정

1) 각 소집단의 통계량

2) 전체표본의 통계량

- 소집단 평균들의 소집단의 비율에 대한 가중평균
- (소집단 분산들과 소집단의 비율의 제곱의 곱)에 대한 합

(3) 신뢰구간 추정

3. 불비례층화추출법

제 3 절. 군집표본추출법

1. 군집표본추출의 절차

- 조사대상 모집단을 MECE 를 만족하는 소집단으로 나눈다
- 분류된 소집단 중에서 무작위로 표본을 선택한다

- 1 단계 군집표본추출: 선택된 소집단의 모든 구성요소를 이용
- 2 단계 군집표본추출: 선택된 소집단에 대해 다시 표본을 추출하여 이용

2. 군집표본 추출법의 특성

- 소집단의 분포가 모집단의 분포와 유사하도록 나누는 것이 중요

- 통계적효율성: 일정한 크기의 표본에 대한 추정치의 표준오차가 적은 경우 효율성이 크다고 함

- 군집표본추출법은 층화표본추출법이나 단순무작위표본추출법보다 통계적으로 효율적이지 못한 경우가 많음

- 상대적으로 표본 추출비용이 적게 소요되므로 많은 표본을 추출하여 이용할 수 있고, 이를 통해 추정치의 표준오차를 줄일 수 있음

3. 체계적 표본추출법

(1) 표본추출법

- k 번마다 표본을 추출하는 방법

(2) 통계량추정

(3) 신뢰구간추정

- 표본선정방법이 용이함, 비용이 적게 소요됨
- 주기적 형태의 특성을 갖는 모집단에 대해서는 적합하지 않음

4. 지역표본 표본추출법

(1) 기본개념

표본추출 방법

- 서울에 있는 모든 동의 목록을 작성 (N)
- n 개의 동을 선정
- 선정된 n 개동의 모든 가구에 대해 조사를 실시

(2) 다단계지역표본추출법

표본추출 방법

- 서울에 있는 모든 동의 목록을 작성 (N)
- n 개의 동을 선정
- 선정된 n 개동에서 k 개의 가구를 선정
- 선정된 가구에 대해 조사를 실시

(3) 기타 지역표본추출법

1) 동일확률지역표본법

- 각 표본들이 뽑힐 확률이 일정한 방법

- 단순 2 단계지역표본추출법: 각 군집이 뽑힐 확률을 일정하게, 각 군집내에서 일정한 수의 표본을 추출
- 규모비례지역별 확률표본추출방법: 각 군집이 뽑힐 확률을 각 군집의 크기에 비례하게, 각 군집내에서 각 군집의 크기에 비례한 수의 표본을 추출

2) 비동일확률지역표본추출법

- 각 표본들이 뽑힐 확률이 일정하지 않은 방법

제 4 절. 표본크기의 결정

1. 신뢰구간접근법

- 오차의 양을 결정 (표본평균과 모집단평균의 차이)
- 신뢰도 수준을 결정
- 신뢰수준에 따른 Z 값을 결정
- 모집단의 표본오차를 추정
- 표본의 크기를 결정

2. 가설검증접근법

- 귀무가설과 대체가설의 수치를 결정
- Type I 오차와 Type II 오차가 일어날 확률을 결정
- Type I 오차와 Type II 오차가 일어날 확률에 따른 Z 값을 결정
- 모집단의 표본오차를 추정
- 표본의 크기를 결정

3. 그 외의 고려사항

- 연구목적
- 시간제약
- 비용제약
- 자료분석상 제약