

3강. 단순임의추출법(1)

◆ 담당교수 : 이기재 교수

■ 주요용어

용어	해설
복원, 비복원 추출	한번 표본으로 추출된 단위가 동일한 표본에 반복적으로 추출되는 것을 허용하는 것을 복원(復原)추출, 허용하지 않는 것을 비복원(非復原)추출이라고 한다.
단순임의추출법	크기가 N인 모집단에서 크기가 n인 표본을 추출하고자 한다면 $\binom{N}{n}$ 가지의 추출 가능한 표본들이 존재하게 된다. 이들 $\binom{N}{n}$ 가지의 표본에 대해 추출확률을 모두 동일하게 해주는 표본추출방법을 말한다.
난 수 (Random number)	확률표본으로부터 얻어지는 추정량은 표본추출변동으로 인하여 그 값이 변하지만 표본추출을 반복하면 어떤 분포를 따르게 되는데 이 분포를 표본분포라고 한다.
유한모집단수정항 (fpc)	표본조사에서의 모집단은 유한모집단이어서 표본평균의 분산은 무한모집단에서의 분산에 비해 $\frac{N-n}{N}$ 의 값을 곱해 주어야 하는데 이 값을 유한모집단 수정항(finite population correctional term : fpc)이라고 한다.

■ 실습하기

- 교재 58쪽 내용 실습
 - * n=10인 단순임의표본 추출* 표본추출방법: 분석도구 이용, rand()함수 이용
 - * 실제 얻어진 표본조사 결과를 토대로 신뢰구간 등 계산
- 교재 67쪽 연습문제 5번

■ 연습문제

- 통계학 개론에서 배운 모평균에 대한 추정량 \bar{y} 의 분산추정량은 $V(\bar{y}) = s^2/n$ 인데 반해, 단순임의추출법으로 뽑은 표본에서 구해진 표본평균 \bar{y} 의 분산추정량은 (㉠)이다. 통계학 개론에서는 무한모집단을 가정하는데 비해, 표본조사에서의 모집단

은 유한모집단이어서 (㉠)의 값을 곱해 주게 되는데 이 값을 (㉡)이라고 한다.

㉠ ㉡

㉢

- ① 표본조사 결과의 비표본오차 크기 정도
- ② 표본조사 결과의 통계적 정확도
- ③ 모집단 추출틀의 정확도
- ④ 표본조사 결과의 편향 정도

정답 : ㉠ : $\hat{V}(\bar{y}) = \frac{N-n}{N} \frac{s^2}{n}$, ㉡ : $\frac{N-n}{N}$, ㉢유한모집단 수정항(fpc)

2. 크기가 N인 모집단에서 단순임의추출법으로 크기가 n인 표본을 추출하였다. 다음 설명 중에서 잘못된 것은?

- ① 단순임의추출법은 확률추출법의 일종이다.
- ② 추출가능한 $\binom{N}{n}$ 개의 표본들에 대하여 추출확률을 모두 같게 해 주는 추출법이다.
- ③ 단순임의추출법은 조사원이 임의로 표본을 선정하기 때문에 선택편향이 발생한다.
- ④ 각 추출단위에 대한 추출확률은 모두 동일하다.

정답 : ③

해설 : 조사원의 선택편향은 조사대상의 추출을 조사원에게 맡기는 경우에 발생하기 쉽다. 단순임의추출법을 적용할 때 선택편향은 발생하지 않는다.

3. 크기가 N인 모집단에서 단순임의추출법으로 크기가 n인 표본을 추출할 경우 모평균 μ 의 추정에 대한 설명이다. 다음 설명 중에서 잘못된 것은?

- ① 표본평균 \bar{y} 는 모평균에 대한 비편향추정량이다.

② \bar{y} 의 분산 추정량은 $\hat{V}(\bar{y}) = \frac{N-n}{N} \frac{s^2}{n}$ 이다.

③ 95% 신뢰수준에서 \bar{y} 의 오차한계는 $\hat{V}(\bar{y}) = \frac{N-n}{N} \frac{s^2}{n}$ 이다.

- ④ 표본수 n을 늘리면 추정량 \bar{y} 의 분산도 커진다.

정답 : ④

해설 : 단순임의추출법에서 표본평균 \bar{y} 의 분산은 $\hat{V}(\bar{y}) = \frac{N-n}{N} \frac{s^2}{n}$ 이다.

따라서 표본수 n 을 늘리면 추정량 \bar{y} 의 분산은 작아진다.

4. 가구당 월평균 저축액을 조사하기 위해서 단순임의추출법을 적용하였다. 표본의 크기가 100인 경우에 표본평균의 표준오차(standard error)는 9(만원)이라고 한다. 표본크기를 400으로 늘리면 표준오차는 대략 얼마가 되겠는가? (단,

(단, $\frac{N-n}{N} \approx 1$ 로 가정)로 가정)

- ① 9(만원)
- ② 4.5(만원)
- ③ 3(만원)
- ④ 2.25(만원)

정답 : ②

해설 : 표본평균의 표준오차는 $\sqrt{V(\bar{y})} = \sqrt{\frac{N-n}{N} \frac{S^2}{n}} \approx \sqrt{\frac{S^2}{n}}$

5. 단순임의추출법에 대한 다음의 설명 중에서 옳은 것을 모두 고른 것은?

- I. 표본크기가 커질수록 표본에서 구한 표본표준편차는 줄어든다.
- II. 표본크기가 커질수록 표본평균은 모평균에 가까이 갈 것이다.
- III. 표본크기가 커질수록 표본평균의 표준오차는 줄어든다.

- ① I
- ② I, II
- ③ I, II, III
- ④ II, III

정답 : ④

해설 : 표본크기의 크고 작음과 표본에서 구한 표본표준편차 크기는 관련이 없다.

■ 정리하기

- 단순임의추출법이란 모든 가능한 경우의 표본들이 추출될 확률을 동일하게 해주는 표본 추출법이다. 단순임의추출법은 난수표에서 난수를 발생시켜 표

본을 추출한다.

- 모집단이 개의 기본단위로 구성되고 있고, y_i 를 i 번째 기본단위의 특성값이라고 하면 모집단의 조사변수 값은 $\{y_1, y_2, \dots, y_N\}$ 로 표시할 수 있다.

- 모수의 수학적 표현

•모평균 :
$$\mu = \frac{1}{N} \sum_{i=1}^N y_i$$

•모총계 :
$$\tau = \sum_{i=1}^N y_i = N \cdot \mu$$

•모분산 :
$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu)^2$$

- 크기가 N 인 모집단에서 반복이 없는 단순임의추출법(Simple Random Sampling Without Replacement: SRSWOR)으로 크기 n 인 표본을 추출하였다.

(i) $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 는 모평균 의 비편향추정량이며,

(ii) 추정량 \bar{y} 의 분산 추정량은 이다. 여기서

$$\hat{V}(\bar{y}) = \frac{N-n}{N} \frac{s^2}{n} \text{이다. 여기서 } s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \text{이다.}$$

■ 참고문헌

- 이계오, 박진우, 이기재, 표본조사론, 한국방송통대학교출판부, 2013. 제1장
- 통계청 홈페이지 : <http://www.nso.go.kr>