

제 10 장

표본추출 (Sampling)

1. 모집단과 표본

- 모집단 (population) :
 - 연구결과가 일반화 되는 큰 집단
 - 세분화된 종류의 준거에 부합되는 대상, 사람, 사건으로 정의된 집합체
 - 신뢰도 연구에서는 무한한 측정값
- 표본 (sample) : 실제 측정이 이루어짐
 - 표본추출과정을 통해 얻어진 모집단의 하위집단 (subgroup)
 - 모집단의 결론을 위한 준거집단(reference group)

가. 모집단의 정의

- 연구결과를 일반화하기 위해 모집단의 범위를 정의
- 명확한 내적 외적 준거를 제시

나. 표본추출 오차 (Sampling bias or error)

- 표본에서 선택된 대상이 모집단의 특성을 과잉대표 하거나 최소대표 할 때 생김
- 표본의 크기(sample size)보다는 대표성을 갖도록 표본추출 절차를 거쳐야 함

다. 표적모집단과 표본틀

(Target population) (Sample frame)

- 접근가능 모집단(accessible population)=실험 모집단(experimental population)
 - 선택될 가능성이 있는 표적 모집단의 일부
 - 접근가능 모집단은 표적 모집단을 대표할 만큼 유사해야 함
 - 연구표본은 접근 가능 모집단에서 선택 됨
- 예: 현재 다운 증후군 아동: 표적 모집단
특수학교: 접근 가능모집단

2. 표본추출(Sampling)기법

- 확률적 방법과 비확률적인 방법
 - 확률 표본추출 : 모집단의 모든 기본단위가 선택 될 수 있는 동등한 기회를 의미하는 무작위 추출과정
 - 비확률적인 표본추출: 일반화에 어려움이 있음
- 표본오차(Sampling error) 또는 표본변이(Sampling variation)라고 부름
 - 표본통계량(statistics)과 모수(parameter)의 차이
- 무작위추출
 - 차이가 우연적인 것이고 사람들의 편견, 의식이나 무의식적인 개입으로 인한 것이 아니어야 한다는 점이 중요
 - 모집단의 특성을 가장 적절하게 반영하는 표본을 만듦

3. 확률 추출(Probability sampling)

- 단순무작위 추출(simple random sampling)
- 체계적인 추출(systematic sampling)
- 층화추출(stratified sampling)
- 군집추출(cluster sampling)

가. 단순무작위 추출 (simple random sampling)

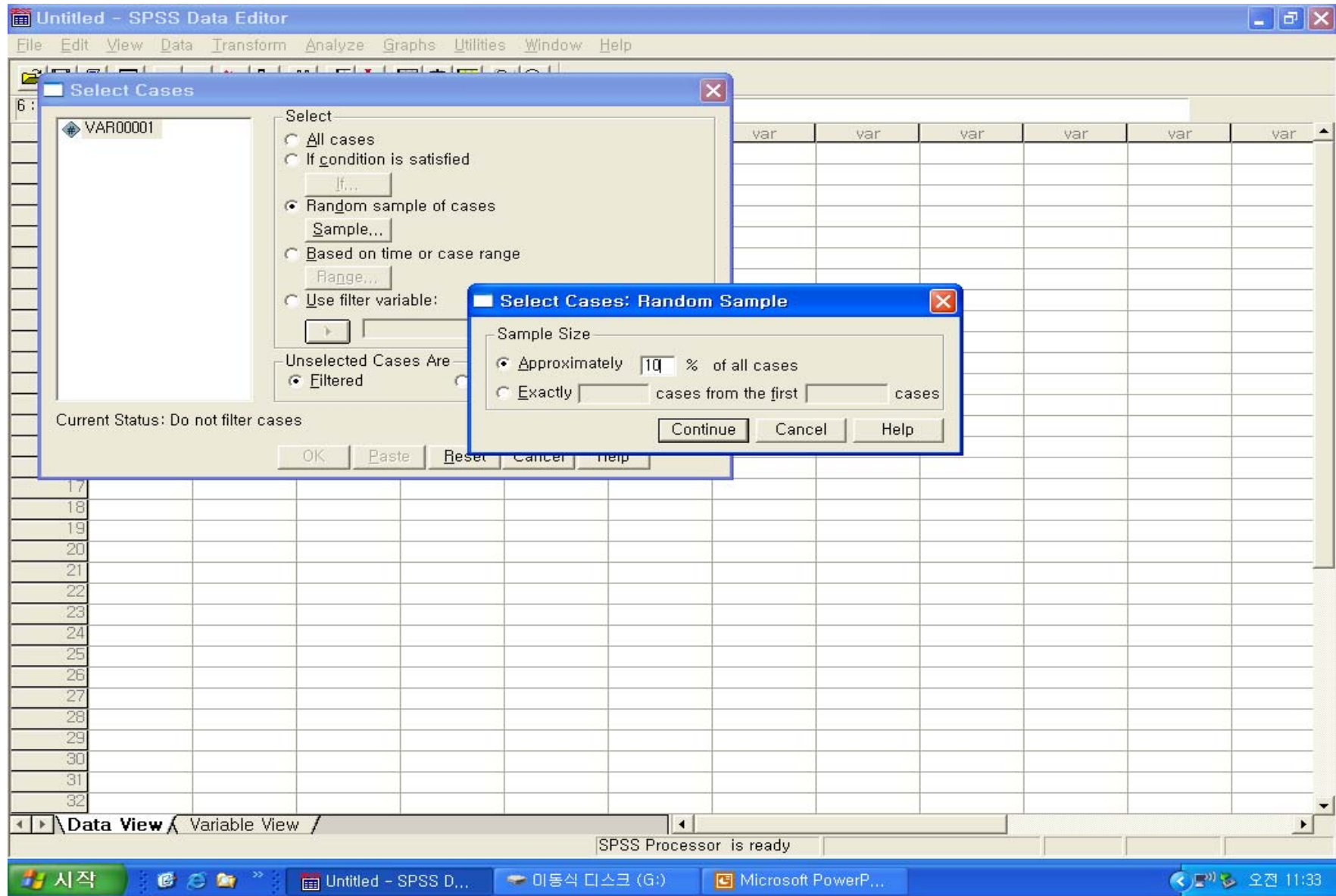
- 선택이 독립적이어서 편향(bias)되지 않는다
- 무작위 표본은 접근가능 모집단에서 추출
- 난수표사용
http://news.chosun.com/site/data/html_dir/2017/05/26/2017052600689.html
- 컴퓨터 사용 (SPSS 이용, 엑셀, 웹사이트)

난수표 사용

• 표

표 10-1. 난수표(Table of random numbers)

03	47	41	71	96	36	98	47	36	81	46	84	62	71	62	23	26	16	80	45	00	11	14	13	85
07	34	24	47	62	42	81	14	57	30	42	53	32	37	32	27	07	36	87	51	24	51	79	89	73
10	36	62	27	66	96	30	35	71	07	32	00	79	38	53	13	25	38	58	09	88	87	64	14	10
12	96	85	94	28	96	98	66	27	31	05	01	72	01	15	57	12	19	14	21	88	26	49	81	76
25	19	28	24	64	39	04	62	40	22	31	62	41	09	50	06	18	44	22	52	23	83	01	30	30
18	22	77	64	20	49	04	43	54	62	17	27	00	23	76	67	20	39	95	42	84	26	34	91	64
04	42	17	23	20	57	34	55	08	66	73	04	74	47	67	21	76	20	90	25	83	82	12	06	26
03	14	03	20	59	35	18	55	07	39	98	31	59	71	75	12	85	73	58	07	44	39	52	38	29
20	21	12	34	29	38	04	56	07	62	32	42	07	44	28	15	51	00	13	48	09	66	02	79	54
07	00	86	22	44	09	47	22	98	54	49	77	46	09	62	00	52	04	77	27	08	22	73	43	38
18	26	07	02	48	44	17	31	68	09	79	61	68	19	62	00	76	50	01	19	53	22	64	06	05
26	02	38	97	75	84	18	07	44	09	83	11	48	22	24	20	14	05	80	45	19	87	72	88	71
22	42	40	64	74	62	00	77	73	61	07	45	12	14	08	22	98	04	87	72	01	85	79	19	25
52	28	29	29	05	53	02	39	11	57	00	51	75	21	28	00	22	02	53	20	00	02	04	01	23
27	85	94	25	12	82	20	50	08	30	42	24	07	00	88	54	42	00	87	98	21	85	29	24	20
78	29	17	12	13	41	28	20	38	28	13	61	11	01	74	17	76	27	13	04	07	24	21	19	30
58	02	18	27	25	95	28	50	27	25	00	12	25	01	47	30	23	24	01	54	97	77	48	44	80
05	40	07	22	77	88	42	95	45	22	18	64	28	19	00	04	43	19	05	79	14	77	24	21	90
16	08	15	04	72	23	27	14	34	09	45	51	34	68	49	12	72	07	24	45	09	27	73	83	14
31	26	00	22	40	50	27	69	27	29	29	21	27	00	49	52	85	00	01	28	08	88	11	80	80
68	34	20	13	70	55	74	30	77	40	44	22	78	64	28	04	22	48	00	20	88	07	97	00	57
74	27	25	65	78	29	28	97	68	60	71	81	28	67	24	11	28	18	24	76	13	54	25	01	52
27	42	27	86	53	48	55	90	65	22	98	57	69	34	10	08	48	92	42	45	97	00	49	04	14
05	29	68	29	61	60	37	22	28	30	77	64	57	01	29	10	45	16	04	26	11	04	96	67	24
29	54	08	54	24	88	49	69	18	62	53	75	91	09	20	24	25	29	27	42	46	73	51	82	82
16	00	82	66	20	83	62	64	11	12	67	19	09	71	74	00	47	28	29	08	02	02	37	03	31
11	27	94	75	06	01	09	29	74	60	02	54	27	54	02	91	70	90	30	86	28	45	14	20	28
35	24	19	25	20	22	22	51	26	28	79	28	45	04	91	16	92	23	59	16	02	25	60	16	08
28	23	12	86	28	42	28	97	01	50	27	25	68	61	41	40	01	24	11	62	48	51	84	18	22
31	56	25	84	47	96	44	23	48	13	34	68	82	52	91	00	52	41	48	25	27	55	28	89	62
26	87	40	67	14	64	05	71	95	88	11	05	65	00	68	78	82	20	27	90	57	29	00	11	81
14	00	94	45	11	25	71	68	65	90	52	27	41	14	28	22	98	12	22	68	07	22	74	81	80
98	05	01	19	00	23	95	02	75	19	07	60	62	01	25	59	22	82	43	90	49	27	28	44	20
20	46	78	23	90	97	51	40	14	02	04	02	20	21	08	20	54	16	49	26	47	26	50	11	20
44	19	58	97	79	25	05	15	03	20	01	90	10	25	02	40	78	78	89	02	02	67	74	17	23
85	26	00	20	60	22	25	85	15	19	92	09	51	59	77	59	56	29	00	22	52	11	65	20	24
67	97	19	88	21	09	98	42	98	64	61	71	62	99	15	06	51	29	39	92	28	25	77	00	51
69	71	26	69	85	54	27	60	47	54	72	22	08	11	12	44	25	02	23	16	29	56	24	29	48
28	19	01	16	02	88	27	23	80	90	42	10	28	67	42	22	17	26	81	74	14	44	67	81	84
14	03	52	88	75	67	29	28	22	41	28	28	62	00	25	13	08	27	04	20	13	29	20	20	43
17	53	73	28	71	71	41	64	20	22	12	41	94	01	25	44	25	27	28	20	02	86	74	28	63
00	28	50	21	19	23	22	23	12	08	93	02	18	29	07	62	18	20	07	24	99	22	20	20	23
41	23	52	55	00	21	04	40	68	00	10	47	48	45	88	13	41	40	89	25	07	17	14	40	17
60	20	20	81	68	31	98	23	68	28	61	20	01	76	24	20	12	48	60	19	90	10	72	24	34
14	23	28	09	94	28	29	41	28	45	27	29	01	00	00	25	27	20	12	62	62	94	00	80	80
24	20	27	74	27	98	80	23	00	04	09	77	00	19	62	74	94	00	94	04	46	67	21	66	89
85	22	04	24	41	21	01	51	94	79	23	62	45	89	28	08	21	24	46	21	23	94	13	28	47
09	79	13	77	48	21	82	97	22	21	05	01	27	28	82	72	89	44	25	60	20	89	20	54	89
88	75	80	14	22	85	25	42	48	29	22	82	22	48	02	48	07	20	21	16	04	61	67	87	87
00	16	21	20	00	39	00	01	68	00	65	65	78	28	26	94	27	20	89	21	00	89	60	26	23



나. 체계적인 표본추출(Systematic sampling)

- 전체대상자의 수를 뽑고자 하는 대상자의 순으로 나눈 수의 순서마다 대상자를 선정
- 시간이 절약됨
- 1000명 중 100명 뽑기 → 10명 당 1명씩 뽑기

다. 층화추출법(Stratified random sampling)

- 층화과정(stratification)을 통해 표본의 대표성을 향상
- 모집단의 구성원들을 동질적으로 분류하고 계층(strata)으로 나눔
- 예) 1학년 300명, 2학년 300명, 3학년 200명, 4학년 200명으로 구성되어 있을 때 무작위추출하면 학년별 특성이 반영 안됨. 모집단의 구성비율에 따라 무작위나 체계적인 비율층화방법을 이용
- 다른 추출법보다 시간이 더 소모되지만 단순 무작위 추출보다 훨씬 대표적인 표본을 제공함

라. 군집추출법(Cluster sampling)

- 모집단이 넓은 지역에 분포할 때 모집단 내에서 군집 단위를 연속적으로 추출
- 전국 시도 중에서 5개 시 혹은 도를 무작위추출 → 5개 시.도로부터 50개의 병원을 무작위 추출 → 각 병원에서 고용되어 있는 물리치료사 전부를 조사함
- 대규모 모집단을 다룰 때 편리하지만 추출오차를 감수해야 함
- 연구자들은 가구를 무작위로 추출하기 위해 다단계추출법(multistage sampling)을 적용하는 경우도 있음

4. 비확률적인 추출(표집)

가. 편의 추출법(Convenience sampling, availability)

나. 눈덩이 추출법 또는 연대 추출법(Snowball sampling, word of mouth)

다. 할당추출법(Quota sampling, proportions, 비율)

라. 고의성 추출법(Purposeful sampling, hand-picked)

가. 편의 추출법 (Convenience sampling)

- 비확률적인 추출의 가장 일반적인 형태:
편의 추출법(Convenience sampling) 우연추출(Accidental sampling)
- 손쉽게 이용 가능한 대상만을 선택
 - 모집단과 대상자간의 특성이 다를 수 있음
 - 일반화가 어려움
- 표본추출과정에는 어느 정도의 편의성이 개입

- 연구자의 입장에서는 자신이 설정한 연구대상자의 선택기준에 맞는 사람을 찾는 것
- 연구대상자를 모집하기 위해서 대부분 신문에 광고를 내거나 광고전단을 살포
- “나이는 40세에서 77세 사이, 우측 뇌졸중 진단을 받고 보행가능한 사람을 연구대상자로 구합니다.”라고 광고
- 임상연구자가 관절염환자의 걷는 모양(gait pattern)에 대한 연구를 하려 할 때, 혹은 재가환자들로부터 연구대상자를 선택하려고 할 때 이런 표본들은 연구자의 편의에 의해 선정

단점

- 조사대상자로 참여하지 않은 사람들과 조사에 참여한 사람들의 특성이 매우 다를 수 있다는 점에서 일반화의 어려움
- 조사에 응한 지원자들은 나이, 동기, 활동수준, 건강의식 등에 있어서 응답을 하지 않은 대상자들과 특성이 전혀 다를 수 있음
- 응답자들로부터 얻어진 결과를 표적모집단에 일반화하기가 어렵다는 것

나. 눈덩이 추출법 또는 연쇄소개 추출법 (Snowball sampling)

- 특수한 특성을 가진 대상자로서 거취를 파악하기 힘들 때 유용함 (예: 매춘, 마약, 성소수자, 불법과외 등)
- 맛집, 좋은 카페는 누가 아는가?
- 졸업자 명단 확보하는 방법은?

- 대상자의 입에서 입으로 적절히 표본의 크기가 이루어질 때까지 그 수를 늘리는 방법

- 진행성 근위축증 환자 모임에 찾아가서 대상자들을 면접하고 주변에 다른 진행성근위축증 환자를 알고 있는지 계속 추천을 받아서 일정한 표본 크기가 얻어지면 표본추출을 중단
- 눈덩이추출법은 연쇄소개추출(chain referral sampling)이라고도 함

다. 할당추출법(Quota sampling, proportions, 비율)

- 모집단의 알려진 비율에 근거하여 표본을 구성
예) 대상자의 성별에 따라 응답이 달라질 수 있을 때
성비에 따라 표본을 구성
 - 모집단의 선택 대상자가 선택 가능성에 있어서 동등하지 못함
 - 표본들이 모집단의 의견을 정확히 반영하지 못할 수도 있음

- 한국의 물리치료사협회에 소속된 물리치료사들의 남녀 성비가 4 : 6라고 알려져 있고 남녀에 따라 응답이 달라질 것이라고 판단된다면 이 비율에 따라서 목표로 한 조사수가 정해진 비율로 채워질 때까지 남녀 물리치료사를 선택하여 조사

라. 고의성 추출법

(Purposeful sampling, hand-picked)

- 유의추출(purposive sampling), 판단추출(judgmental sampling)이라고도 함
- 연구대상자의 일부분은 쉽게 식별할 수 있지만 모집단 전체를 모두 확인하는 일이 거의 불가능할 경우에 사용
- 연구자가 자신의 연구능력과 사전지식을 활용하여 응답자를 선택
- 의도적으로 연구의 필요성에 맞춰 특정 질병이나 질병력을 갖고 있는 대상자만을 찾아서 연구자의 판단에 따라 연구대상자에 포함여부를 결정

- 운동프로그램의 효과를 알아보려는 연구에서 의무기록이나 면접을 통해서 자신의 연구에 적합할 만한 환자를 찾음
- 연구자는 연구기간 동안에 지시사항을 잘 따를 것 같은 환자들을 선택
- 고의성 추출법은 편의추출과 마찬가지로 일반화 제한점
- 이 표집방법은 연구계획의 초기 단계에서 설문지의 적용 가능성이나 조사도구의 타당성 등을 검토하기 위하여 많이 사용