

**표본론 2023년 기말시험 (월, 수 9:00~10:15)**

[1] 특정 지역에 서식하는 야생 멧돼지의 총 개체수를 추정하고자 한다. 대략 500마리에서 1000마리 사이일 것으로 예상하고 100마리를 포획하여 표식을 한 다음 풀어주었다. 한 달 지난 후 표식을 한 멧돼지가 10마리 잡힐 때까지 재포획하기로 하였다. 90마리를 재포획했을 때 비로소 표식을 한 멧돼지 10마리를 채울 수 있었다고 할 때 이 지역에 서식하는 멧돼지의 총 개체수를 추정하시오. 오차의 한계도 구하시오.

[2] 전자기관(board of electronic components)에 있는 결함 수(number of defects)의 평균을 추정하고자 한다. 각 기관은 여러 개의 부품으로 구성되어 있는데 부품의 수가 많아지면 결함 수도 많아지는 경향이 있다고 하자. 총 10개의 기관 중에 4개의 기관을 부품 수에 비례하는 확률로 추출하기로 하였다. 10개 기관의 부품 수는 다음과 같았다.

10, 12, 22, 8, 16, 24, 9, 10, 8, 31

부품 수에 비례하는 확률로 추출한 결과 2, 3, 5, 10번째 기관이 추출되었고, 이 4개의 기관에서 발견한 결함 수는 각각 1, 3, 2, 5 이었다. 한 기관에 존재하는 평균 결함 수를 추정하시오. 오차의 한계도 구하시오.

[3] 12개의 원소로 이루어진 모집단이 아래와 같이 배열되어 있다고 하자. 4개에 1개씩 계통추출했을 때 (1-in-4 systematic sampling) 얻어지는 표본평균  $\bar{y}$ 를 4가지 모든 가능한 경우에 대해 구하시오. 또한  $n = 3$ 개의 원소가 단순확률추출(simple random sampling)로 얻어졌다고 가정하고 추정한 분산  $\hat{V}(\bar{y})$ 을 4가지 모든 가능한 경우에 대해 각각 구하시오. 이 예제에서  $\bar{y}$ 의 참분산  $V(\bar{y})$ 은 얼마인가? 이 예제가 계통추출 추정량의 분산에 대해 어떤 점을 말하고자 하는 것인지 설명해 보시오.

N = 12인 모집단 원소												$\bar{y}$	$\hat{V}(\bar{y})$
군집 번호	1	2	3	4	5	6	7	8	9	10	11		
1	1				5				9				
2		2				6				10			
3			3				7				11		
4				4				8				12	

[4] 크기 200인 표본을 거주지역과 소득수준으로 분류하여 보았더니 다음과 같았다. 이때 크기  $n = 200$ 인 표본은 무응답으로 인해 모집단을 대표하는 표본인지 알 수 없다고 한다.

[4-1] 거주지역 A, B의 모집단 비율이 0.6과 0.4이고 소득수준 a, b의 모집단 비율이 0.3, 0.7임을 알고 있다면, 거주지역과 소득수준으로 결정되는 네 개의 층의 크기를 갈퀴질(raking) 방법으로 추정하시오. 모집단 크기  $N$ 은 2000이다. (계산 과정에서 각 셀의 빈도를 계산할 때 소수점 첫째 자리에서 반올림해 항상 정수가 되도록 하시오.)

		소득수준		
		a	b	합
거주	A	40	70	110
지역	B	10	80	90
합		50	150	200

[4-2] 거주지역과 소득수준에 대한 참 결합분포가 아래와 같다고 한다. 이 사실로부터 갈퀴질(raking) 방법이 갖는 한계에 대해 설명하시오.

		소득수준		
		a	b	합
거주	A	0.1	0.5	0.6
지역	B	0.2	0.2	0.4
합		0.3	0.7	1.0

[5] 이번 학기 표본론 수업에서 농업자료에 대해 조별로 프로젝트를 실행하였다. 만약 관심 있는 모수가 며 재배면적의 총합 하나뿐이라면 표본추출방법을 어떻게 하고 모수 추정방법을 어떻게 하는 것이 좋을지를 최대한 구체적으로 설명하시오.