

비계통어업생산통계조사 표본설계

염준근¹⁾, 이해용²⁾, 한근식³⁾, 이기성⁴⁾

요 약

본 연구에서는 2000년 어업총조사 결과를 토대로 전국의 어업별, 어종별 수산물의 비계통생산량에 관한 통계를 정도높게 산출할 수 있는 표본설계를 하고, 그에 따른 추정식을 마련함으로써 보다 편리하면서도 신뢰성이 높은 어종별 총생산량 통계를 산출하도록하였다. 본 설계에서는 가능한 한 기존의 설계에 비해 업무량이 크게 늘어나지 않는 범위내에서 효율적인 통계를 생산하는 것을 기본방향으로 하고 있으며 구체적인 방향들은 첫째, 조사 설계의 목적이 '비계통어업생산량'의 효율적인 추정인 만큼 각 어종들의 생산량 추정치의 정도를 높이는 것을 주목적으로 정하였다. 둘째, 조사 대상 어종이 114종에 이르고 어종별 비계통생산량이 상이함으로 어종별 정도를 고려하면 거의 전수조사를 해야하는 실정이므로 해양수산부에서 요청한 주요 어종과 어류 생산 어가가 많은 어류 등 41개 어로어업 어종과 11개 양식어업 어종의 생산총계가 정도높게 작성될 수 있도록 표본설계를 하였다. 셋째, 표본의 할당 및 추출은 조사원의 업무량의 범위를 크게 벗어나지 않도록 하였다.

주요용어 : 어업총조사, 어로어업, 양식어업, 어업조사구

1. 서론

비계통어업생산통계는 어가 부문에서 도매시장, 수집상, 음식점, 가공공장, 양식장, 자기매장, 기타 등 수협을 통하지 않고 출하되는 비계통출하 수산물의 생산에 관한 통계들을 작성하여 해양수산 정책수립을 위한 기초자료를 제공한다.

기존 표본은 1995년 어업총조사 결과를 기초로 설계한 것으로 그간의 우리 나라의 어업환경의 변화를 고려할 때 재설계의 필요성이 요구된다. 특히 서해안 지역은 대규모 간척사업의 여파로 많은 어가가 사라졌다. 1990년 어업총조사 결과 우리 나라의 총어가수는 121,525 가구였으며 1995년 결과에서는 104,480 가구로 약 14%정도 감소하였다. 한편 2000년 어업총조사 결과에 의하면 우리 나라의 총어가수는 81,571 가구로 급격히 감소하고 있음을 알 수 있다.

한편, 해수면 온도의 변화로 동해상의 한류성 어종의 포획량의 감소되는 등 어종에도 많은

-
- 1) 동국대학교 통계학과 교수
 - 2) 성신여자대학교 통계학과 교수
 - 3) 한신대학교 정보시스템학과 교수
 - 4) 우석대학교 전산통계학과 교수

이 논문은 통계청 학술연구지원에 의해 수행되었음

변화가 있다. 모집단의 이러한 상황변화를 표본에 시기 적절히 반영하지 않은 상태로 표본조사가 계속 시행되는 경우 비표본오차로 인해 통계의 신뢰성을 떨어뜨릴 수 있다.

본 연구의 목적은 2000년 어업총조사 결과를 토대로 전국의 어업별, 어종별 수산물의 비계통 생산량에 관한 통계를 주어진 여건 하에서 가능한 한 정도높게 산출할 수 있는 표본설계를 하고, 그에 따른 추정식을 마련함으로써 보다 편리하면서도 신뢰성이 높은 어종별 총생산량 통계를 산출하는 데에 그 목적이 있다.

2. 모집단과 층화변수

본 비계통어업생산조사를 위한 표본설계에서의 모집단은 전국 3,333개의 어업조사구에 속한 모든 어가들이다. 여기에서 비계통이란 수산물 판매중 수협을 통한 출하를 제외한 일반판매(도매시장, 수집상, 음식점, 가공공장, 양식장, 자기매장, 기타)를 통한 출하를 말한다.

본 조사는 2단추출로 이루어지게 된다. 1차추출단위는 어업조사구이고, 2차추출단위는 추출된 어업조사구내의 어가가 된다.

본 표본설계를 위한 추출틀은 2000년 어업총조사 결과를 이용하여 구성하였다. 2000년 어업총조사시 전국은 총 3,333개의 어업조사구로 구성되어졌는데 이는 95년 어업총조사시 1,921개 조사구에 비해 70% 가량 증가하였다. 각 시도별 조사구수는 다음과 같다.

<표 1> 지역별 어업조사구수 (2000년 어업총조사)

시도	조사구수	시도	조사구수
서울	1	충남	324
부산	120	전북	179
대구	1	전남	1,068
인천	156	경북	208
울산	37	경남	590
경기	85	제주	423
강원	141	전국	3,333

위 모집단을 우선 어업형태(어로어업과 양식어업)로 층화한 후 주요 어종(어로어업 41개, 양식어업 11개)으로 모집단을 층화하였다. 각 어종은 지역으로 층화함으로써 지역별 총계의 추정이 가능토록 하였으나 표본설계의 주 목적이 비계통 어종 총생산량의 추정에 있는 만큼 추정의 정도가 어종의 총생산량을 기준으로 설계하였으므로 지역별 어종의 총계의 추정의 정도는 높지 않을 수 있다.

위와 같이 층화하였을 때 어로어업의 제1어종별 지역별 어가수와, 양식어업의 제1어종별 지역별 어가수는 표본설계보고서의 <표 4-2>와 <표 4-3>을 참고 바란다. 이 때, 제1어종은 생산량이 가장 많은 어종이 아니라 판매금액이 제일 많은 어종을 의미한다. 생산량이 가장 많은 어종을 제1어종으로 하는 것을 고려해 볼 수 있으나, 모집단으로부터 얻을 수 있는 정보가 생산량이 아닌 판매금액 밖에 없으므로, 판매금액이 가장 많은 어종을 제1어종으로 정하였다.

3. 표본의 크기 및 할당

모집단을 구성하고 있는 총 3,333개 조사구 중에서 250개의 표본 조사구를 우선 어종별로 추출하는 데 200개의 표본 조사구는 어로어업에, 50개의 표본 조사구는 양식어업에 할당한다. 그런 다음에 어종에 할당된 조사구를 지역별로 할당한다. 지역별 할당은 비례할당과 네이만 할당을 고려하였는데 그 할당 결과가 거의 차이가 없었다. 그러나 지역별 추정도 고려해야하는 만큼 네이만할당 보다는 전남지역의 조사구를 줄이는 대신 울산, 강원 등에 1개의 조사구를 더 할당하는 비례 할당을 이용하기로 하였으며 그 결과는 다음 <표 2>와 <표 3>과 같다.

<표 2> 어로어업 어종별 지역별 표본 조사구수

지역 어종	부산	인천	울산	경기	강원	충남	전북	전남	경북	경남	제주	전국
가오리류								1				1
가자미류	2				1				2	4		9
갈치											1	1
납치류								1		3		4
농어						1		1				2
돔류								1		2	1	4
망둥어						1						1
멸치류			1					2	1			4
방어											1	1
조피볼락 (우럭)		1				3	1	1		1		7
기타볼락										1		1
삼치류	1							1				2
서대류								1				1
송어류	2	1					1	4		1		9
장어류	5		1				1	4	1	5		17
전어	1							1		1		3
쥐치류										1		1
기타어류		1					1	1		2		5
계류		1		1	1	2	1	3		1		10
대하						2						2
기타새우류						1	1	1		1		4
굴류		3				10	1	6		1		21
소라고둥											2	2
전복류								1	1			2
가무락						1	1					2

<표 2> 어로어업 어종별 지역별 표본 조사구수(계속)

지역 어종	부산	인천	울산	경기	강원	충남	전북	전남	경북	경남	제주	전국
고막류						1		2				3
동족		1				1	2					4
맛류						2	1	1				4
바지락	1			4		9	2	4		1		21
백합류							3	1				4
기타패류						1						1
낙지						2		15		1		18
문어					1			1	1	1		4
쭈꾸미						1	2	1				4
오징어류									1		2	3
기타연체동물						1		1				2
성게	2								3			5
해삼	1								1			2
기타수산동물						1		1				2
미역			1		1			2	1	1		6
툰								1				1
합계	15	8	3	5	4	40	18	60	12	28	7	200

<표 3> 양식어업 어종별 지역별 표본 조사구수

지역 어종	부산	인천	울산	경기	강원	충남	전북	전남	경북	경남	제주	전국
굴	2	1				5		8		5		21
피조개										1		1
홍합								1		1		2
바지락						6	1	3		2		12
고막								5				5
기타패류	1	1						2				4
우렁쟁이										1		1
미더덕										1		1
툰								1				1
다시마								1				1
기타해조류								1				1
합계	3	2				11	1	22		11		50

어종별로 각 지역에 대해 조사가구 표본으로 추출될 곳의 총 조사구수는 비계통어업생산통계 조사 표본설계의 <표 4-6>과 <표 4-7>에 정리하였다. 여기에서는 해당 조사구 내의 어가수가 8가구 이상인 조사구만 정리하였는데 그 이유는 추출된 조사구 내에서 4개의 어가를 2차 추출해야 하며 이 때 대체표본으로 이용할 어가를 확보하기 위해 $4 \times 2 = 8$ 개 어가 이상인 조사구만 정리하였다. 한편 양식어가에 대해서는 추출된 조사구내에서 5개의 어가를 2차 추출해야 하

며 이 때 대체표본으로 이용할 어가를 확보하기 위해 $5 \times 2 = 10$ 개 어가이상인 조사구만 정리해야 하나 10개어가 이상인 조사구만 고려할 경우 해당 어종에 대한 해당지역에 조사구가 존재하지 않을 수 있기 때문에 어로어가와 마찬가지로 8개 어가 이상을 정리하였다. 모집단을 구성하는 총 3,333개 조사구 중에서 어로어업의 경우 어가수가 8가구 이상인 조사구는 880개이고, 양식어업의 경우 어가수가 8가구 이상인 조사구는 336개로 나타났다. 어가수가 8가구 이상인 1,216개 조사구중에서 250개 조사구를 1차추출단위로 표본추출해야 하나 8개 이상의 어가로 구성된 조사구가 없는 5개 지역에 대해서는 대체표본 추출을 용이하게 하기 위해 8에 가장 가까운 어가를 보유한 조사구를 추출하였다.

4. 표본추출

본 설계에서는 어종별로 할당된 조사구를 지역별로 비례할당하여 어업조사구를 1차 추출한 후, 추출된 어업조사구내의 어가를 2차추출하는 층화 2단 집락추출법을 이용하였다.

1. 1차추출단위

각 어종별로 지역별 어가수에 따라 조사구를 비례할당하였다. 어가들을 판매금액(지난 1년간의 수산물 판매금액)에 따라 3개의 층으로 구분하였다. 특성번호 '11'층은 수산물 판매금액이 500만원 이하인 어가로 구성되고, 특성번호 '12'층은 수산물 판매금액이 500-5,000만원 이하인 어가로 구성하였다. 한편, 특성번호 '13'층은 수산물 판매금액이 5,000만원 이상인 어가들로 구성하였다.

판매금액에 따라 3개의 층으로 구분된 어가들을 각 조사구 내에서 정렬하여 500만원 이하의 수산물 판매 어가 위주로 구성된 조사구, 500-5,000만원 이하인 어가 위주로 구성된 조사구, 5,000만원 이상인 어가 위주로 구성된 조사구들로 정렬한 후 <표 2>와 <표 3>에 나타난 크기대로 계통추출하였다. 단, 조사구내에 어가수가 8가구 미만인 경우 해당 특성번호를 갖는 조사구내에서 8에 가장 가까운 어가수를 가진 조사구를 임의로 추출함으로써 대체표본 추출에 대비하였다.

조사구가 1차 추출된 후 해당 조사구가 조사가능지역인지의 여부를 파악한다. 조사가능지역 여부는 통계청에서 사전에 조사불능지역을 제시 (경기도 화성시 우정면과 서신면은 간척사업의 여파로 기존의 어업조사구가 없어질 예정) 하였으므로 이를 참고하였다. 또한, 정기적인 선편이 없는 낙도지역과 같이 선편이 없어 매달 조사가 어려운 지역과 간척사업 등으로 어가가 사라질 예정인 지역 등은 추출된 조사구의 현지 답사과정에서 조사구 교체를 요구토록 하여 해당 조사구와 인접하면서 같은 특성을 갖는 조사구로 대체시키도록 하였다.

2. 조사구내의 어가 추출

다음으로 표본으로 추출된 어업조사구로부터 2차 추출단위인 어가를 추출하는 방법을 설명하겠다.

조사구내에서 판매금액별로 구분된 각 어가는 다시 어업수입구분(지난 1년간 가구전체의 어업수입이 가장 많았던 어업 형태는 양식어업과 어로어업 중 어느 것입니까?)에 따라 '21'층과 '22'층으로 구분하였다.

여기에서 층'21'은 어로어업 수입이 양식어업수입보다 큰 어가를 의미하며 층'22'는 양식어업

으로 인한 수입이 어로어업보다 큰 어가를 의미한다. 어가명부의 작성은 다음과 같은 기준을 이용하였다.

판매금액	$\left\{ \begin{array}{l} 500\text{만원 이하} \\ 500-5,000\text{만원 이하} \\ 5,000\text{만원 이상} \end{array} \right.$	$\left\{ \begin{array}{l} : 11 \\ : 12 \\ : 13 \end{array} \right.$	어업수입구분	$\left\{ \begin{array}{l} \text{어로어업} \geq \text{양식어업} : 21 \\ \text{어로어업} < \text{양식어업} : 22 \end{array} \right.$		
					어업수입구분	$\left\{ \begin{array}{l} \text{어로어업} \geq \text{양식어업} : 21 \\ \text{어로어업} < \text{양식어업} : 22 \end{array} \right.$
						$\left\{ \begin{array}{l} \text{어로어업} \geq \text{양식어업} : 21 \\ \text{어로어업} < \text{양식어업} : 22 \end{array} \right.$

조사구내의 어가들은 어업수입에 의한 층으로 정렬되고 정렬된 층내에서 어가특성이 같은 경우 제2어종과 제3어종의 크기에 비례하여 어가를 추출하였다. 어가수는 어로어업의 경우 4가구를 양식어업의 경우 5가구를 계통추출하여 최종 표본을 완성하였다.

5. 추정

본 조사의 결과 어종별 비계통생산량에 대한 추정이 이루어지게 된다. 어종별 추정은 층화집락추출에 따라 추정하게 된다.

각 어종별 비계통생산량은 어로어업과 양식어업으로 나누어 다음의 공식으로 추정한다.

$$\text{어로어업} : \hat{\tau}_h = \sum_{i=1}^{11} \frac{N_{hi}}{n_{hi}} \sum_{j=1}^{n_{hi}} \frac{M_{hij}}{4} \sum_{k=1}^4 y_{hijk}, \quad \text{양식어업} : \hat{\tau}_h = \sum_{i=1}^{11} \frac{N_{hi}}{n_{hi}} \sum_{j=1}^{n_{hi}} \frac{M_{hij}}{5} \sum_{k=1}^5 y_{hijk}$$

여기서 사용된 기호들은 각각

$\hat{\tau}_h = h$ 어종의 비계통생산량의 추정량

$h =$ 어종을 나타내는 첨자 ($h = 1, 2, \dots, 114$)

$i =$ 지역을 나타내는 첨자 ($i = 1, 2, \dots, 11$)

$j =$ 표본조사구를 나타내는 첨자 ($j = 1, 2, \dots, N_{hi}$)

$k =$ 조사구 내 표본어가를 나타내는 첨자 ($k = 1, 2, 3, 4$ 또는 5)

$N_{hi} = h$ 어종의 i 지역 내에 속한 모집단 조사구수

$n_{hi} = h$ 어종의 i 지역 내에 속한 표본 조사구수

$M_{hij} = h$ 어종의 i 지역 내 j 번째 표본 조사구 내의 총어가 수

$y_{hik} =$ 각 표본어가에서 조사된 특정어류의 생산량 데이터를 의미한다.

참고문헌

1. 대한통계협회(1997), 수산물 비계통생산고조사 표본설계.
2. 박홍래(2000), 통계조사론, 영지문화사.
3. 통계청(2001), 어업생산통계.
4. 통계청(2001), 2000 어업총조사 분석보고서.
5. Lessler, J. T. and Kalsbeek, W. D.(1992), Nonsampling Error in Surveys, John Wiley and Sons.
6. Little, R. J. A. and Rubin, D. B.(1987), Statistical Analysis with Missing Data, John Wiley and Sons