

일반해면어업 오차율 감소를 위한 표본설계

2004. 6.



연구수행기관 : 수 원 대 학 교

연구책임자 : 박 진 우 (수원대학교 통계정보학과)

공동연구자 : 김 영 원 (숙명여자대학교 통계학과)

제 출 문

해양수산부 장관 귀하

본 보고서를 「일반해면어업 오차율 감소를 위한 표본 설계」 과제의 최종보고서로 제출합니다.

2004년 6월

주관연구기관명 : 수원대학교

연구책임자 : 박진우

연구원 : 김영원

머 리 말

본 보고서는 해양수산부의 수산특정연구개발과제인 「일반해면어업 오차율 감소를 위한 표본설계」의 최종보고서이다.

해양수산부의 어가대상 수산업정책수립을 위해 중요한 기초 통계를 제공하고 있는 일반해면어업 비계통생산량 조사의 경우 한정된 예산과 조사인력의 제약 하에서 표본설계가 이루어진 관계로 표본의 규모가 너무 작아 신뢰할 수 있는 통계를 생산할 수 없는 것이 현실이다. 기존 표본조사는 2002년 7월 통계청에서 2000년 어업총조사 자료를 모집단으로 1,050어를 추출한 표본어를 대상으로 하는데 2003년 1월부터 조사가 실시되고 있다. 비계통어업생산량 조사의 경우 조사대상 어종이 120여종을 상회하므로 기존 표본규모는 너무 작아 일부 주요어종에 대해서도 신뢰할 수 있는 생산량 통계를 산출할 수 없는 상황이다.

본 연구개발의 최종 목표는 기존 통계청에서 수행한 일반해면어업의 비계통생산량 통계를 위한 표본설계의 내용을 보완하는 방안을 연구하고, 이를 통해 최소한 주요 어종에 대해서는 신뢰할 만한 생산량 통계를 생산할 수 있도록 표본을 보강하는데 있다. 아울러 이런 표본 변경에 따른 추정상의 문제를 해결하기 위한 적절한 추정식을 제시하고, 동시에 과학적인 분석이 가능하도록 표본오차 산출 공식을 제공하는 데 있다.

특히 현실적으로 해양수산부에서 본 사업을 위해 확보한 소요예산을 고려하여 표본어가 확대에 있어서 가능한 기존 표본 어가수인 1,050어를 2,000어로 확대하는 것을 전제로 하였다. 추가 표본 어가를 추출을 위한 표본설계에 있어서는 2000년 어업총조사 자료와 2003년 비계통조사 표본 데이터를 바탕으로 어가의 어업형태, 주요 어획 어종 및 지역특성 등을 종합적으로 표본설계에 반영하여 대표성을 갖는 표본을 확보하는 데 중점을 두었다. 한번의 표본 추가로 끝나는 것이 아니라 추후의 변화를 감안한 사후관리의 지침까지 제시하였다.

끝으로 본 과제의 수행을 위해 필요한 모든 지원을 아끼지 않은 해양수산부 관계 공무원들께 깊은 감사를 드린다.

2004년 6월
연구책임자

< 차 례 >

제 1 장. 서 론	1
1.1 조사개요	1
1.2 연구의 개요	1
1.2.1 연구의 필요성	1
1.2.2 연구의 목적	2
1.2.3 연구의 범위	3
1.2.4 연구의 과정	4
제 2 장. 현지답사 결과	6
제 3 장. 기존 표본의 분석	9
3.1 모집단 (population)	9
3.2 추출틀 (sampling frame)	9
(1) 어업조사구 기준 모집단 어가 현황	9
(2) 시군구별 어종 분포	10
3.3 기존 표본설계 개요	13
3.4 기존 표본 통계의 문제점	14
3.4.1 표본오차	14
3.4.2 모집단과 표본의 괴리	18
3.4.3 통계 발표 양식	20
3.4.4 어업조사구 사용에 따른 조사상의 문제	21
3.4.5 표본관리 및 조사관리의 문제	22

제 4 장. 새로운 표본설계	24
4.1 새로운 표본설계를 위한 추출틀	24
(1) 추출단위로 인구주택 조사구 활용	24
(2) 인구주택 조사구 기준 조사모집단 구성	27
4.2 조사구 층화	27
(1) 시도별 사용 어법 현황	28
(2) 주요 어법 및 지역에 따른 조사구 층화	31
4.3 표본의 크기 및 목표오차	34
4.3.1 표본의 크기	34
4.3.2 표본의 배분	34
4.3.3 목표 상대표준오차	37
4.4 표본추출	39
(1) 표본 조사구의 추출	39
(2) 표본어가의 추출	40
4.5 추정	41
제 5 장. 사후관리	44
5.1 추출틀 보정	44
5.2 표본관리	45
5.3 가중값(승수) 관리	45
5.4 기타	46
제 6 장. 요약 및 제언	47

제 1 장.서 론

1.1조사개요

어업생산통계는 수산물의 업종별 및 어종별 생산량과 생산금액을 파악하여 수산물 생산, 어업경영 및 수산물 유통개선 등 수산정책 수립을 위한 기초자료를 제공하고 아울러 국가간 어업협상 등을 위한 공신력 있는 기초자료를 마련하기 위해 조사, 작성되는 통계이다.

어업생산통계 작성을 위한 조사는 크게 전국 지역수협 의 위관장과 공관장을 통해 조사되는 계통조사와 표본어가를 대상으로 조사되는 비계통조사로 구분된다. 계통조사는 수협 계통조직을 통해 전수조사가 되므로 표본오차가 존재하지 않는데 반해 비계통조사는 어가를 대상으로 하는 표본조사에 의해 이루어지는 조사이므로 표본오차가 발생하게 된다. 따라서 본 연구에서 대상으로 하는 통계는 비계통어업생산 통계가 된다.

비계통표본조사는 어가 및 업체에서 수협계통을 통하지 않고 직접 판매하는 어종의 생산량 및 판매금액을 조사하는 조사이다. 비계통조사는 1983년 처음 표본설계가 이루어져 조사된 이래 몇 차례의 표본개편이 이루어져 현재에까지 이르고 있다. 현재 사용되는 표본은 2000년 어업총조사 결과를 근거로 2002년에 이루어진 표본설계의 결과인데, 표본의 크기는 전국 1,050 어가이다. 여기서 어가란 국내에서 생계 및 판매를 목적으로 해면에서 수산 동식물의 포획, 채취 또는 양식업을 자영하는 사람(가구주 또는 가구원)이 있는 가구로 정의되었다.

비계통조사는 표본어가가 자기기입식으로 조사표를 작성하게 한 후 해양수산부 통계조사요원이 매월 정기적으로 방문하여 회수하는 방식으로 이루어진다. 조사의 질을 높이기 위해 통계조사요원은 주 2회 이상 표본어가를 방문(또는 전화)하여 조사표의 작성상태를 점검, 보완하고 있다.

1.2연구의 개요

1.2.1연구의 필요성

고도의 지식정보화 사회에서 통계는 올바른 의사결정을 위한 기초 자료를 제공하는 유용한 수단으로 확고히 자리 잡고 있다. 우리나라의 올바른 해양수산 정책을 마련하기 위해서는 보다 정확하고 신뢰할 만한 수산통계를 생산하는 것은 필수적인 일이라고 하겠다. 뿐만 아니라 국제적인 협상을 효율적으로 수행하기 위한 객관적인 기초 자료를 마련하기 위해서도 객관적인 수산업 관련 통계를 생산하는 일은 매우 중요하다.

현재 우리나라 어업생산량 관련 통계들을 살펴보면 정확성의 측면에서 믿고 쓸 수 있는 통계가 절대적으로 부족하다는 사실을 알 수 있다. 이것은 특정 조사에서 오류를 범하기 때문이 아니다. 그보다 현재의 시스템이나 여건 하에서는 아무리 잘 조사를 한다고 해도 근본적으로 해결하기 어려운 내재적인 결함을 가진 때문이라고 할 수 있다. 다시 말해 어업분야 통계생산을 위한 기반이 제대로 구축되어 있지 못하는데 그 근본원인이 있다고 하겠다.

특히 어가대상 수산업정책수립을 위해 중요한 기초 통계를 제공하고 있는 일반해면어업 비계통생산량 조사의 경우 한정된 예산과 조사인력의 제약 하에서 표본설계가 이루어진 관계로 표본의 규모가 너무 작아 신뢰할 수 있는 통계를 생산할 수 없는 것이 현실이다. 기존 표본조사는 2002년 7월 통계청에서 2000년 어업총조사 자료를 모집단으로 1,050어를 추출한 표본어를 대상으로 하는데 2003년 1월부터 조사가 실시되고 있다. 비계통어업생산량 조사의 경우 조사대상 어종이 120여종을 상회하므로 기존 표본규모는 너무 작아 일부 주요어종에 대해서도 신뢰할 수 있는 생산량 통계를 산출할 수 없는 상황이다.

이에 따라 일반해면어업 비계통생산량에 대한 일정 수준이상의 정확도를 갖는 어종별 공식통계를 산출하기 위해서는 관련된 비계통생산량 조사자료의 분석 및 전국 어가 모집단에 대한 분석을 통해 기존 표본을 대폭 보강하는 추가적인 표본설계 방안에 관한 심층적인 연구가 필요하다고 판단되며, 이를 통해 비계통어업생산량 통계의 신뢰도를 제고할 수 있는 구체적인 표본 어가 보강 작업이 수행되어야 한다.

1.2.2 연구의 목적

본 연구개발의 최종 목표는 기존 통계청에서 수행한 일반해면어업의 비계통생산량 통계를 위한 표본설계의 내용을 보완하는 방안을 연구하고, 이를 통해 최소한

주요 어종에 대해서는 신뢰할 만한 생산량 통계를 생산할 수 있도록 표본을 보강하는데 있다. 아울러 이런 표본 변경에 따른 추정상의 문제를 해결하기 위한 적절한 추정식을 제시하고, 동시에 과학적인 분석이 가능하도록 표본오차 산출 공식을 제공하는 데 있다.

특히 현실적으로 해양수산부에서 본 사업을 위해 확보한 소요예산을 고려하여 표본어가 확대에 있어서 가능한 기존 표본 어가수인 1,050어를 2,000어로 확대하는 것을 전제로 한다. 추가 표본 어가를 추출을 위한 표본설계에 있어서는 2000년 어업총조사 자료를 바탕으로 어가의 어업형태, 주요 어획 어종 및 지역특성 등을 종합적으로 표본설계에 반영하여 대표성을 갖는 표본을 확보하는 데 중점을 둘 것이다. 이를 통해 가장 효과적인 표본을 추출하여 추가함으로써 효율적인 일반해면어업 비계통생산량통계 생산이 가능하도록 하는 것을 목표로 한다.

1.2.3 연구의 범위

본 연구의 구체적인 범위를 한정하면 다음과 같다.

- 원칙적으로는 기존 표본으로 선정된 표본어가 1,050 개를 기존대로 유지한 상태에서 새로운 표본 어가를 추가적으로 추출하기로 한다. 기존 표본의 대표성을 분석하여 보다 대표성을 높일 수 있도록 하는 표본확대 방안을 마련한다.
- 표본 확대에 따른 새로운 추정식을 제시하고, 이에 따른 표본오차 산출방법을 연구한다. 아울러 추정작업이 가능한 손쉽게 이루어질 수 있도록 단계적인 가중치 계산 방법과 가중치를 이용한 추정공식을 제시한다.
- 전반적인 수산업통계생산체계의 보완에 따라 양식품종에 대한 전수조사 실시 방안을 감안하여 양식어가는 새로운 표본설계에서 제외하고, 이런 변동이 기존 표본에 미치는 영향을 분석하여 이런 변동을 추정식에 반영한다.
- 가능한 충분한 현지답사를 통해 각 지역별 특성을 파악하고, 특히 지방청 수산관리과(소) 통계담당공무원들의 의견을 수렴하여 표본설계에 반영한다. 특히 이 과정에서 지방청 담당공무원들과의 협의를 통해 무응답 및 표본교체 등의 비표본오차 발생 요인을 파악하여 이를 최소화 할 수 있는 방안을 연구한다.

1.2.4 연구의 과정

본 연구의 과정은 크게 ① 문헌조사 ② 자료수집 ③ 현지방문 ④ 기존 표본설계 분석 ⑤ 새로운 표본추출 ⑥ 추정식 연구 ⑦ 사후관리대책의 과정으로 구분할 수 있는데 각 과정별 주요 내용은 다음과 같다.

① 문헌조사

이 과정은 연구의 시작단계인데 본 연구와 관련된 기존의 참고문헌들을 조사하고 연구함으로써 전체적인 연구의 밑그림을 그리는 단계이다. 기존 조사보고서와 어업총조사 보고서를 근간으로 하고 거기에다가 유사한 여러 표본조사들의 표본설계보고서 및 최신 표본이론을 다루고 있는 문헌 및 논문들을 다수 조사하여 참고하였다.

② 자료수집

문헌연구를 통해 파악된 사항을 기초로 하여 본 연구 수행을 위해 필요한 데이터 및 기초자료를 수집하는 단계이다. 2000년 어업총조사 데이터 파일, 2003년 1월에서 11월 비계통생산량조사 데이터 파일 등이 가장 중요한 기초자료로 활용되었다.

③ 현지답사

전국에 있는 해양수산부 수산청들을 직접 방문하여 현지의 조사환경 및 조사원들의 의견을 청취하는 과정이다. 2004년 2월 2일(월)부터 2월 5일(목요일)까지 3박 4일에 걸쳐 부산, 마산, 여수, 목포, 군산 수산청을 방문하였다. 5개 지방 수산청을 방문하였지만 인근지역의 담당자들이 모두 모여 전국 각 지역의 현황을 자세히 청취하며 토론할 수 있었다. 현지답사에서 논의된 주요사항들은 다음 장에 자세히 정리되어 있다.

④ 기존의 표본설계 분석

2002년 표본설계보고서와 그에 따라 설계된 표본에서의 2003년 1월에서 11월까지의 조사자료를 기초로 하여 기존의 표본설계를 분석하는 과정이다. 기존 설계 표본 데이터의 표본오차 분석 및 모집단 추출틀인 2000년 어업총조사 데이터와의 비교 검토 등을 통해 기존 표본설계의 문제점을 파악하는 과정이다.

⑤ 새로운 표본추출

기존의 표본설계 분석 결과를 토대로 하여 새로 추가될 부분의 표본설계를 하는

과정이다. 총화, 표본의 배분, 기대오차의 계산 등이 이루어지며 최종적인 표본 추출이 이루어지는 과정이다.

⑥ 추정식 연구

새로운 표본에 의거하여 각종 통계를 생산할 수 있도록 합리적이고 효율적인 추정방법을 연구하는 과정이다. 기존의 표본을 그대로 사용하면서 950개의 표본을 새로이 추가하게 되므로 기존의 가중값을 전면적으로 새 것으로 대체하여야 한다. 또한 새로운 추정식을 제안할 경우 그에 따른 분산추정량도 제시하여야 한다. 바람직한 추정식에 요구되는 비편향성(unbiasedness), 효율성(efficiency) 등을 충족시킬 수 있는 추정법을 마련한다.

⑦ 사후관리대책

본 조사는 동일한 표본을 대상으로 매월 계속조사가 이루어지는 패널조사(panel survey)의 일종이다. 이러한 조사를 지속하다보면 모집단 상황의 변화, 표본어가의 소멸 등 다양한 상황이 생기게 되는데 각각의 경우에 대비하여 어떻게 표본을 관리해갈 것인가 하는 요령을 마련하는 과정이다. 패널조사에서 사후관리를 소홀히 할 경우 비표본오차가 커질 수 있다는 점을 감안한다면 사후관리대책의 마련에도 세심한 주의를 기울여야 한다.

제 2 장.현지답사 결과

본 연구는 새롭게 표본설계를 하는 것이 아니고 이미 한 차례 표본설계가 이루어져 조사가 진행되는 상황에서 오차를 줄이기 위해 기존의 1,050 어가의 표본에 950어가를 추가하는 연구이다. 그러므로 기존의 표본조사의 장, 단점을 보다 면밀히 파악하여 개선하는 것이 절실한 필요이다. 이를 위해서는 기존의 표본조사를 일선에서 담당하고 있는 조사원들의 다양한 의견을 듣고 각 지역별 실정을 파악하는 것은 매우 중요한 부분이라고 할 수 있다.

본 연구를 위해서 연구팀은 해양수산부 공무원과 함께 연구의 초기단계인 2004년 2월 2일에서 2월 5일까지 4박 5일 간의 지방 현지답사를 실시하였다. 답사는 부산, 마산, 여수, 목포, 군산의 5개 해양수산부 지방청을 방문하는 것으로 진행되었다. 각 지방청에서 진행되는 회의에는 해당청의 조사직 공무원 이외에도 산하 출장소에 근무하는 조사공무원들이 참석하였기 때문에 인천과 강원도 동해를 제외한 전국 모든 지역의 조사공무원을 다 만나서 의견을 청취할 수 있었다.

현지답사에서 여러 가지 논의들이 있었는데 그것들 중 본 연구에 중요한 영향을 미치게 된 논의결과들을 요약하면 다음과 같다 :

기존의 표본에 포함된 어업조사구들 중 일부는 여러 개의 인구주택조사구를 묶어서 구성한 것이므로 해당 지역의 범위가 너무 넓고 조사구 내 어가들의 특성이 서로 일치하지 않는 경향이 있다. 심한 경우에는 서너 개의 인구주택조사구가 모여서 하나의 어업조사구를 구성하는데 이 때 어업조사구 내에서 네 가구의 표본어가를 선정하다보면 네 가구가 서로 지리적으로 멀리 떨어져 있게 되어 여러 모로 낭비적인 면이 있다는 지적이 있었다.

조사의 편의를 감안하여 기존의 표본에는 도서지역의 어가가 누락되었다는 사실에 대해 현지의 조사원들은 심각한 문제의식을 지니고 있었다. 특히 전남 지역의 경우 반드시 추가로 설계되는 표본에 일부 어업세력이 큰 도서지역은 포함되어야 한다고 생각하고 있었다. 전남 신안군의 경우 어가인구수로 볼 때 전국에서 중요한 지역에 속하는데 도서지역으로만 구성된 군이라 표본이 전혀 없는 대표적인 경우이다. 도서지역을 표본에 포함시키려 할 때 문제가 되는 사항은 조사비용이 많이 든다는 점과 조사시간이 많이 소요된다는 점이다. 가령 정기적인 여객선이 운항되지 않는 지역의 어가가 표본에 뽑힐 경우 사실상 조사가 어려울 것으로 생각되며, 정기선이 운항되는 지역일 경우 다른 지역에 비해 여비가 과다하게 소요되므로 이에

대한 적절한 지원이 필요하다는 점이 지적되었다. 이런 점들을 설계에 반영하기 위해 각 지방청 및 과소별 관할지역내 도서현황 및 조사가능 도서현황 자료를 요청하여 구할 수 있었다.

현지의 조사원들은 표본이 추가될 경우 늘어나는 업무량에 맞는 합리적인 인원이 이루어져야 할 것을 절실하게 요구하였다. 표본의 규모가 약 두 배로 커진다고 해서 모든 지역에 골고루 두 배의 표본이 배분되는 것이 아니라 지역에 따라 많이 늘어나는 지역도 있고 적게 늘어나는 지역도 있을 것임에 대한 이해를 구했다.

새로운 표본설계가 이루어져서 각 지역별 표본 조사구가 선정되고 난 후 해당 조사구 내의 어가를 추출하는 일은 현지의 조사원이 담당하게 된다. 이 때 중앙에서는 표본어가 추출원칙을 제시해주게 된다. 현장의 조사원들은 과거의 표본추출 원칙이 너무 경직된 것이어서 현지의 상황에 따른 융통성을 발휘할 수 없음으로 인해 불편함이 있다는 점을 지적하였다. 또한 과거의 표본추출 원칙은 조사구내 어가를 2000년 총조사 당시의 어가당 어업판매금액 순으로 정렬한 후 계통표집하는 방식이었는데 조사원들은 어업판매금액이 해당어의 특성을 파악하는 것에 의문을 제기하였다. 어업판매금액은 해마다 변동이 클 수 있으므로 그보다는 어선톤수나 양식장 면적 등 다른 정보를 이용하여 어가들을 정렬하는 것이 바람직하다는 의견이었다. 아울러 어가들을 특정 변수에 의해 정렬시킨 후 계통표집할 때 기계적으로 엄격하게 하기보다는 어가의 지리적 위치 등까지 한꺼번에 고려할 수 있도록 융통성있게 추출하는 방안을 마련해줄 것을 요청하였다.

원자력 발전소의 건설, 신항만 건설 등 국가적 사업의 영향으로 어업권이 소멸되어 가거나 약화되고 있는 지역에 있는 표본어가 있는데 이런 지역의 어가들을 대상으로 조사하는 것에 조사원들이 거부감을 느끼고 있었다. 관할 지역의 다른 정상적인 어가들로 표본을 대체해야 하는 것 아닌가 하는 생각이었다. 이 점에 대해서는 먼저 해당 지역의 특수한 사정일 뿐 전국적인 상황이 아니라는 사실을 설명하였고, 아울러 그런 지역의 표본은 모집단의 변동사항을 표본에서 반영하는 부분이라고 볼 수 있으므로 표본어가 어업을 그만두지 않는 한 계속 조사를 실시하게 하였다.

조사원들은 특정 표본어가 소멸되면 해당 조사구 내에서 특성이 같은 다른 어가로 반드시 대치(substitution)해야 한다고 생각하고 있었다. 따라서 이러한 표본대치가 필요한 경우 소멸된 어가와 비슷한 특성이 해당 조사구 내에 존재하지 않는 경우가 자주 발생하게 되므로 이러한 상황을 미리 예측해서 설계하여 줄 것을 요청하였다. 이에 대해 특정 어가가 소멸되는 경우 어떤 때에는 대치해주는 것이 바람

직한 반면 어떤 경우에는 대처하지 않음으로 표본수가 줄어드는 것을 허용하는 것이 필요하다는 점을 설명하였다. 아울러 최종보고서를 작성할 때 표본의 관리요령을 따로 상세히 설명하기로 하였다.

제 3 장.기존 표본의 분석

3.1모집단 (population)

비계통어업생산통계조사는 어가를 대상으로 하는 조사이므로 이론적인 조사모집단은 현재 우리나라의 전체 어가들이다. 어가에 대해서도 명확한 정의를 내리는 것이 필요한데 본 연구에서는 통계청이 내린 정의인 “어가란 국내에서 생계 및 판매를 목적으로 해면에서 수산 동·식물의 포획, 채취 또는 양식업을 자영하는 사람(가구주 또는 가구원)이 있는 가구”를 사용하기로 한다.

3.2추출틀 (sampling frame)

(1)어업조사구 기준 모집단 어가 현황

바람직한 표본설계를 위해서는 현재의 모집단 현황을 가장 잘 반영하는 추출틀을 마련하여야 한다. 비계통어업생산통계조사를 위해 2002년도에 표본설계된 기존 표본에서는 1차 추출단위로 어업총조사에서 사용한 어업 조사구 자료를 추출틀로 사용하고 있다. 따라서 모집단 파악을 위해 우선 어업총조사 자료를 기초로 어가 현황 및 어업조사구의 특성을 정리한다.

2000년 어업총조사 자료에 의하면 총 어업조사구수는 3,333개, 총 어가수는 81,571가구인 것으로 파악된다. 이중 서울과 대구 지역에 있는 1가구씩의 어가가 있는 것으로 파악되었는데 이 두 가구는 모집단에서 제외시켰다. 우선 2000년 어업총조사에서 사용된 어업 조사구 분포와 조사구 특성, 그리고 새로이 구성한 조사구의 특성을 파악하기 위해 전국의 조사구 분포와 조사구 특성을 정리하면 <표 3-1>과 같다.

<표 3-1> 어업 조사구 기준 어가 분포 및 특성

지역	조사구수(%)	어가수	조사구 당 평균어가수	총 동력선 톤수	조사구 당 평균톤수	어가 당 평균톤수
부산	120(3.6)	3,591	29.9	15,722.4	131.0	4.38
인천	156(4.7)	3,035	19.5	12,351.1	79.2	4.07
울산	37(1.1)	1,165	31.5	3,144.6	85.0	2.70
경기	85(2.6)	1,807	21.3	1,351.1	15.9	0.75
강원	141(4.2)	4,814	34.1	16,474.8	116.8	3.42
충남	324(9.7)	9,444	29.1	11,833.9	36.5	1.25
전북	179(5.4)	4,275	23.9	11,412.7	63.8	2.67
전남	1,068(32.0)	26,936	25.2	57,534.3	53.9	2.14
경북	208(6.2)	5,778	27.8	16,435.0	79.0	2.84
경남	590(17.7)	14,009	23.7	46,802.1	79.3	3.34
제주	423(12.8)	6,715	15.9	13,382.8	31.6	1.99
전국	3,331(100.0)	81,569	24.5	206,487.3	62.0	2.53

기존 비계통어업생산통계를 위한 표본설계에서는 표본을 추출하기 위해 어업조사구를 1차 추출단위로 하는 2단 집락추출법을 사용하고 있다. 특히 기존 표본설계에서 사용되는 어업조사구는 농·어업 총조사를 동시에 고려한 것으로 하나의 조사구에 농가 또는 어가가 60-100가구(평균 80가구) 포함되도록 구성하였다. 기존 농어업총조사를 위한 조사구는 농업분야에 중점을 둔 것이기 때문에 현실적으로 어업관련 통계를 생산하기 위한 조사구로 부적합하다고 볼 수 있다(해양수산부, 2003). 특히 기존 어업 조사구의 경우 매우 광범위한 지역의 어가들이 하나의 조사구로 구성되어 있는 경우가 많기 때문에 실제 조사현장에서 조사원들이 표본을 관리하고 정확한 조사를 수행하는 데 많은 어려움이 있다. 이는 효과적인 조사 관리 및 수행을 고려해 어업생산통계조사에서 2단 집락추출법을 적용하였음에도 불구하고, 기존 어업 조사구를 1차 추출단위로 사용함으로써 표본설계에서 의도했던 목적이 충분히 달성되고 있지 못하다는 문제를 갖고 있다.

(2)시군구별 어종 분포

2000년 어업 총조사에서는 어로어업의 경우 99개 어종에 대해 어가별로 주요 어획 어종 3가지를 조사하고 있으며, 양식어업의 경우 양식어종별 양식방법 및 면적을 조사하고 있다. 따라서 본 연구에서는 이들 조사결과를 토대로 어로 및 양식 어

업을 통해 어획된 어종 현황을 정리하고자 한다. 여기서 특히 어로어업의 경우 주된 어종 3가지만을 묻고 있기 때문에 실제 각 시군구에서 어획되는 어종 중 다수의 어종은 누락되어 있다는 점에 유의할 필요가 있다. 구체적으로 어로어업에 종사하는 전체 61,400어가 중 주된 어종 3종을 모두 응답한 어가수는 38,149(62.1%)로 이는 실제적으로 3종이상의 어종을 어획하는 어가수가 상당수라는 것을 의미한다. 총조사 자료를 토대로 각 어종이 어획되는 시군구 및 어가 현황을 양식어업과 어로어업으로 구분하여 정리하면 <표 3-2>, <표3-3>과 같다.

<표 3-2> 양식어업 어종별 시군구수, 어가수 현황

어종	시군구수	어가수	어종	시군구수	어가수	어종	시군구수	어가수
굴	37	9139	넙치	33	771	기타어류	41	340
바지락	34	5897	피조개	19	674	미더덕	6	320
김	30	5628	해상종묘	14	650	대하	24	164
미역	32	3713	돔류	26	556	육상종묘	20	122
조피볼락 (우럭)	34	2360	홍합	13	535	방어	11	77
다시마	15	1803	농어	23	454	기타수산동물	11	62
고막	20	1779	기타해조류	17	419	가리비	16	60
기타패류	50	1729	우렁쟁이	13	385	기타갑각류	13	24
툰	7	1687						

<표 3-3> 어로어업 어종별 시군구수, 어가수 현황

어종	시군구수	어가수	어종	시군구수	어가수	어종	시군구수	어가수
장어류	72	8834	동족	20	1382	송어류	43	210
바지락	50	8599	대하	43	1332	다랑어류	29	171
낙지	64	8226	기타패류	63	1305	갑오징어류	24	171
기타어류	75	8048	고등어류	45	1193	청각	20	167
게류	77	7218	꽂치	49	1078	키조개	23	166
굴류	52	6957	백합류	28	1063	학공치	36	159
가자미류	63	6348	취치류	39	969	강달이류	14	156
문어	53	5708	가오리류	51	918	능성어	13	156
소라고둥	67	5241	기타수산동물	41	810	밴댕이	14	156
오징어류	50	5149	가무락	20	756	뱅어류	28	152
송어류	79	4842	망둥이	57	684	새조개	20	140
성게	37	3569	기타해조류	35	643	다시마류	16	125
조피볼락 (우럭)	66	3436	임연수어	15	612	홍어	21	110
기타새우류	63	3370	서대류	23	558	정어리	26	104
전어	60	3271	참조기	28	541	솜팽이	12	94
넙치류	63	3263	기타연체류	44	525	부세	13	90
돔류	50	3156	기타조기	34	499	꿀뚜기	13	88
쭈꾸미	51	2953	양태	17	443	전갱이류	28	87
멸치류	57	2795	도루묵	27	440	개랑조개	21	59
농어	66	2410	골뱅이	36	434	꼬시래기	11	54
기타볼락	51	2275	병어류	38	396	보리멸	24	52
갈치	50	2221	우렁쟁이	33	378	보구치	12	50
전복류	48	2108	복어류	40	375	기타가사리	12	49
미역	41	2048	대구류	21	344	상어류	11	43
고막류	30	2003	명태	23	313	도박류	4	40
해삼	50	1996	과래	21	310	가리비	18	35
청어	17	1691	양미리	24	288	눈불대	18	34
돛	27	1677	민어	25	268	매둥이류	14	29
맛류	34	1537	김	23	259	준치	12	27
오분자기	8	1495	피조개	32	249	미더덕	13	17
방어	37	1494	기타갑각류	42	227	말	5	15
우무가사리	27	1427	아귀	34	221	성대류	9	13
삼치류	44	1394	홍합	31	213	연어	2	3

3.3기존 표본설계 개요

기존 비계통어업생산통계조사를 위한 표본설계의 특성을 살펴보기 위해 2002년도에 수행된 표본설계의 주요 내용을 정리하면 다음과 같다(참조: 한국통계학회, 2002).

기존 표본설계는 1차 추출단위로 어업조사구, 2차 추출단위로 어가를 사용하는 2단 집락추출로 이루어졌다. 기존 표본은 250개의 어업조사구로 구성되어 있으며, 추출된 각 표본 어업조사구에서 4또는 5개의 어가를 추출하여 총 1,050개 어가로 표본이 구성되었다.

또한 어업조사구의 추출을 위해 전체 어업조사구는 어업형태(어로어업과 양식어업)에 따라 구분한 후, 각 조사구를 주요 어종(어로어업 99종, 양식어업 25종) 및 시도에 따라 층화하여 1차로 표본 어업 조사구를 어종 및 지역별 어가수에 비례하도록 각 층에 배분하여 층화 추출하였다. 따라서 기존 표본설계의 경우 실제 적용된 층의 개수가 너무 많아 효율적인 층화작업에 문제점이 있다고 생각된다.

전체 표본 어업조사구 250개 중 200개 조사구는 어로어업이 중요한 수입을 차지하는 어로어업 조사구에서, 나머지 50개 어업조사구는 양식어업 조사구에서 추출되었다. 기존 표본설계에서 표본의 크기는 이론적인 측면 보다는 비용과 조사원의 업무 등 현실적인 측면이 우선 고려되어 실제 각 어종별 총생산량의 추정결과의 정도(precision)는 실제 표본설계에서 반영하고 있지 못하다는 한계를 갖고 있다. 이는 실제 추정대상 어종의 수가 너무 많기 때문에 각 어종에 대해 일정 수준이상의 정도를 확보할 수 있는 표본을 추출하기 위해서는 기존 조사인력으로는 감당할 수 없는 표본 어가수를 필요로 한다는 점이 반영된 것이다.

한편 조사구를 지역 층별로 배분하는 데 있어서는 층별 어가수에 따른 비례배정을 하였으며, 특히 조사 편의를 고려하여 실제 표본추출 대상이 되는 조사모집단(survey population)은 어로어업의 경우 어가수가 8가구 이상인 880개 어업조사구로, 양식어업의 경우 역시 어가수가 8가구 이상인 336개 어업조사구로 구성하였다. 따라서 250개의 표본 어업조사구는 1,216개의 어가수가 8가구 이상인 조사구 중에서 추출된 것이다. 한편 지역층내에서 각 조사구들은 조사구내 어가들의 판매액에 따라 특성을 부여하여 나열한 후 계통추출하였다.

표본으로 추출된 어업조사구 내에서 표본 어가를 추출하기 위해 우선 조사구 내

어가들을 판매금액 및 어업수입구분(어업수입이 가장 많았던 어업 형태가 양식어업인지 어로어업인지에 따라 구분)에 따라 정렬한 어가명부를 작성하고 계통추출하여 최종 표본 어가를 선정하였다. 표본 어가수는 어로어업 특성을 갖는 조사구에서는 4가구를 양식어업 특성을 갖는 조사구에서는 5가구를 추출하였다.

3.4기존 표본 통계의 문제점

이 절에서는 기존의 표본설계에 의해 얻어진 2003년 1월 ~ 2003년 11월 조사 데이터 분석 결과를 기초로 하여 기존 표본 통계가 가지고 있는 중요한 문제점들을 지적한다. 이러한 지적 사항들을 개선하는 것이 본 연구의 중요한 과제가 될 것이다.

3.4.1 표본오차

표본오차는 통계 이용자들에게 작성되는 통계의 품질을 알려주는 가장 대표적인 척도이다. 표본조사 통계는 본질적으로 오차를 수반하게 되므로 통계의 생산자는 반드시 자신이 생산하는 통계의 표본오차를 계산하여 통계와 함께 제공해주어야 한다. 이런 점을 감안할 때 기존 표본통계의 품질을 파악하기 위해 가장 우선적으로 살펴보아야 할 것이 기존 통계의 표본오차이다.

다음의 <표 3-4>는 2003년 1월 ~ 2003년 11월 사이 각 어종별 비계통생산량 통계에 대한 상대표준오차(coefficient of variation: CV)를 계산한 표이다. 해당 기간 동안 표본어가에서 비계통으로 생산한 어종은 총 108가지 어종이었는데 각 어종별 추정생산량과 CV가 계산되어 있다. 여기서 CV값 계산불능이라고 표시한 어종들은 표본어가들 중 그 어종을 생산한 어가가 1가구에 불과하여 표본오차를 계산할 수 없는 어종들이다.

위의 <표 3-4>에 나타난 결과를 보다 쉽게 이해할 수 있도록 CV에 크기에 따라 요약한 것이 아래의 <표 3-5>이다. 통계가 생산된 전체 108가지 어종들 중 CV가 10% 이하인 어종은 9가지에 불과하다. CV의 크기가 20% 이하인 어종 52개 정도를 제외한 나머지 55개의 어종은 표본오차가 너무 커서 통계로서의 가치가 떨어진다고 할 수 있다. 참고로 통계청에서 생산하는 어가경제통계의 주요 항목별 CV가

<표 3-4>

어종	생산량	CV	어종	생산량	CV
306	133166882	0.14	135	235647	0.15
336	36809598	0.07	230	232798	0.20
330	21163629	0.07	171	213217	0.18
354	14823265	0.13	174	190306	0.37
609	10969290	0.24	84	181371	계산불능
309	8003867	0.16	48	177249	0.09
327	7363038	0.11	509	166974	0.13
333	6314748	0.06	57	151023	0.20
339	5579021	0.12	699	137648	0.31
150	5459204	0.11	406	111685	0.22
324	5175230	0.14	621	110474	0.36
126	5093171	0.09	165	103408	0.22
69	4794499	0.19	403	103381	0.18
599	4345975	0.25	221	95616	0.18
409	3947626	0.04	9	78306	0.22
199	3596949	0.20	147	62953	0.32
156	3402605	0.19	60	57957	0.19
227	3190867	0.56	141	51927	0.16
212	3110569	0.09	499	49366	0.24
415	2900999	0.13	96	42981	0.20
399	2228347	0.12	315	42569	0.12
6	2225802	0.05	603	39908	0.45
418	1734599	0.30	75	38754	0.34
303	1666532	0.26	18	35890	0.33
203	1596205	0.11	87	34374	0.38
215	1508723	0.59	318	31663	0.48
233	1463727	0.23	627	28875	0.52
503	1349019	0.46	15	26174	0.26
299	1276240	0.20	624	25819	0.69
342	1175283	0.26	123	25713	0.20
506	1154308	0.12	42	20132	0.38
321	1133593	0.16	162	15036	0.61
99	975285	0.11	224	13751	0.15
129	972466	0.45	345	13632	0.36
209	970785	0.47	218	12057	0.13
153	961026	0.46	21	11840	0.65
618	941596	0.28	312	11797	0.10
412	866457	0.16	612	6900	0.45
102	860867	0.14	36	6838.108	0.33
3	818555	0.18	133	6570	0.33
63	702442	0.17	45	6476.91	0.21
30	604050	0.12	90	6322.197	0.32
103	547533	0.15	357	5961.689	0.34
512	497352	0.28	93	4915.05	0.43
206	474251	0.18	111	4146.851	0.27
351	459971	0.19	105	3984.498	0.22
630	386041	0.32	81	3399.412	0.73
144	360091	0.16	54	2615.833	0.32
132	329568	0.41	159	1803.2	계산불능
108	328112	0.18	66	1522.42	0.61
24	317713	0.24	33	204	계산불능
114	317229	0.18	51	115.355	계산불능

4% ~ 9% 수준인 것(한국통계학회, 2002a)과 비교한다면 비계통어업생산통계의 CV는 전체적으로 큰 편이라는 것을 쉽게 이해할 수 있을 것이다. 알 수 있다.

<표 3-5>

CV의 크기	어종수
10% 이하	9
10.1 ~ 20%	43
20.1 ~ 30%	18
30% 이상	33
계산불능	5

기존의 표본설계(한국통계학회, 2002b)에서는 어차피 1000개 내외의 표본 규모로 108개 어종의 통계를 정도높게 생산하는 것은 불가능한 일이므로 그 중 39개 주요 어종들에 대해서는 정도높은 통계를 생산할 수 있도록 하는 전략을 사용한다고 하였다. 다음의 <표 3-6>은 기존 표본설계에서 중점적으로 고려한 39개 어종의 CV를 나타내고 있는데 이 중 14개 어종의 CV가 20%를 상회하고 있다. 이 말은 현재의 표본은 표본설계에서 관심을 두고 고려한 39개 어종에 대해서도 신뢰성 높은 통계를 제공해주지 못한다는 것을 의미한다.

<표 3-6>

어종	생산량	CV	어종	생산량	CV
306	133166882	0.14	153	961026	0.46
336	36809598	0.07	618	941596	0.28
330	21163629	0.07	412	866457	0.16
309	8003867	0.16	102	860867	0.14
327	7363038	0.11	3	818555	0.18
333	6314748	0.06	63	702442	0.17
339	5579021	0.12	30	604050	0.12
150	5459204	0.11	512	497352	0.28
126	5093171	0.09	206	474251	0.18
69	4794499	0.19	630	386041	0.32
599	4345975	0.25	108	328112	0.18
409	3947626	0.04	24	317713	0.24
199	3596949	0.20	114	317229	0.18
156	3402605	0.19	78	253395	0.19
212	3110569	0.09	48	177249	0.09
415	2900999	0.13	57	151023	0.20
399	2228347	0.12	221	95616	0.18
6	2225802	0.05	9	78306	0.22
418	1734599	0.30	60	57957	0.19
203	1596205	0.11	499	49366	0.24
233	1463727	0.23	315	42569	0.12
506	1154308	0.12	54	2615.833	0.32
321	1133593	0.16	159	1803.2	계산불능
99	975285	0.11	51	115.355	계산불능
209	970785	0.47			

이처럼 현재의 표본조사를 통해 생산되는 통계의 오차가 크게 나타나는 가장 기본적인 이유는 현재 표본의 크기인 1,050개는 조사목적에 이루기에 너무 적은 수의 표본이기 때문이다. 불과 20여종의 농작물 생산량 조사를 위해 농림부에서 사용하는 표본농가수가 37,000여개인 것을 고려한다면(농업연수부, 2002) 현재의 비계통어업생산량 조사의 표본 규모는 절대적으로 부족하다는 것을 짐작할 수 있다. 위의 표본오차의 계산결과는 2002년에 이루어진 기존의 표본설계에 있어서도 예산 및 업무량에 한계가 있었기 때문에 공식통계를 산출할 정도의 정확성을 갖는 표본

설계는 이루어지지 못했다는 사실을 입증하고 있다.

3.4.2모집단과 표본의 괴리

본 조사의 근본적인 문제는 한정된 표본크기로 110여개 어종들의 생산량에 대한 통계를 생산하려는 데서 생긴다는 것은 이미 앞에서 지적한 바 있다. 이러한 문제는 또 다른 문제를 야기 시키는 데 그 중 하나가 표본이 복잡한 모집단의 특성을 제대로 반영하지 못한다는 점이다.

본 조사는 2000년 어업총조사 결과를 추출틀로 사용하는 조사이다. 2000년 어업총조사 자료를 살펴보면 각 어가에서 생산하는 어종에 관한 정보가 충분히 나타나 있지 않았다. 총조사 조사표를 보면 각 어가별로 과거 1년간 어로어업으로 어획한 주 어종 세 가지만 기록하게 했다. 그런데 2003년 표본조사 자료를 분석해본 결과 1어가 당 평균 생산하는 어종의 수는 5.6로 나타났다. 이것은 총조사에서 어가별 어종에 대한 정보가 충분치 않다는 것을 나타낸다. 어종에 관한 정보는 제한적인 반면 각 어가에서 사용하는 어법에 관한 정보는 비교적 충분히 얻을 수 있다. 2003년 표본어가를 대상으로 조사한 결과를 분석해보니 1어가 당 사용하는 어법의 수는 평균 두 가지 어법 이하인 것으로 나타났다. 각 어법별로 주로 생산되는 어종들이 일정하므로 이런 점들을 감안한다면 표본설계를 할 때 각각의 어종을 고려하는 것보다 어법을 고려하는 것이 더욱 효과적일 것으로 생각된다.

다음은 주요 어법별, 지역별 모집단 상황과 표본의 상황을 비교해보기로 한다. 우리나라에서 가장 널리 사용되는 대표적인 어법으로는 301, 281, 266, 212, 172 등의 어법을 들 수 있다. 다음의 <표 3-7>은 2000년 어업총조사 결과를 기초로 하여 작성된 주요어법별 모집단 어가수를 나타낸 표이며, <표 3-8>은 2003년 표본조사 데이터를 이용하여 작성된 주요어법별 표본 어가수를 나타내는 표이다. 두 표를 비교해보면 표본이 모집단을 잘 대표한다고 보기가 어렵다는 사실을 알 수 있다. 가령, 301 어법의 경우 모집단 어가에서는 전남이 차지하는 비중이 4.9%에 불과한 데 반해 표본에서의 전남어가의 비중은 10.4%나 되므로 전남어가가 과다하게 뽑혔다고 할 수 있다. 또한 281번 어법의 경우 모집단에서는 전남이 충남에 비해 거의 4배 가까이 많은 비율을 차지하는 데 표본에서는 오히려 충남이 전남보다 더 많은 비율을 차지하고 있음을 알 수 있다.

<표 3-7> 모집단 어법별 어가수 현황

업종		부산	인천	울산	경기	강원	충남	전북	전남	경북	경남
301	어가수	315	1269	303	830	147	3981	1866	2835	660	1074
	%	0.5	2.2	0.5	1.4	0.3	6.8	3.2	4.9	1.1	1.8
299	어가수	0	78	0	0	15	4	154	142	1	22
	%	0.0	0.1	0.0	0.0	0.0	0.0	0.3	0.2	0.0	0.0
292	어가수	2	59	12	0	38	58	29	579	209	208
	%	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.1	0.4	0.4
281	어가수	12	65	1	4	9	550	123	2194	300	373
	%	0.0	0.1	0.0	0.0	0.0	0.9	0.2	3.8	0.5	0.6
271	어가수	37	34	45	3	72	12	65	483	128	189
	%	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.8	0.2	0.3
266	어가수	644	587	169	176	1438	1204	717	5474	828	3788
	%	1.1	1.0	0.3	0.3	2.5	2.1	1.2	9.4	1.4	6.5
212	어가수	467	85	284	52	184	437	249	2618	1050	2455
	%	0.8	0.1	0.5	0.1	0.3	0.8	0.4	4.5	1.8	4.2
182	어가수	0	92	0	47	2	329	174	185	7	11
	%	0.0	0.2	0.0	0.1	0.0	0.6	0.3	0.3	0.0	0.0
172	어가수	282	385	164	196	1388	1176	527	1200	1817	2254
	%	0.5	0.7	0.3	0.3	2.4	2.0	0.9	2.1	3.1	3.9

<표 3-8> 표본 어법별 어가수 현황

업종		부산	인천	울산	경기	강원	충남	전북	전남	경북	경남
301	어가수	88	17	28	17	0	573	104	651	82	40
	%	1.4	0.3	0.5	0.3	0	9.1	1.7	10.4	1.3	0.6
299	어가수	0	23	20	2	0	15	4	60	0	124
	%	0	0.4	0.3	0.03	0	0.2	0.06	0.1	0	1.98
292	어가수	0	14	2	0	0	11	0	187	0	16
	%	0	0.22	0.03	0	0	0.18	0	2.98	0	0.25
281	어가수	0	62	0	67	11	220	203	112	0	176
	%	0	0.99	0	1.07	0.18	3.51	3.23	1.78	0	2.8
271	어가수	0	0	0	1	0	5	8	51	12	10
	%	0	0	0	0.02	0	0.08	0.13	0.81	0.19	0.16
266	어가수	172	24	26	8	2	114	37	264	17	212
	%	2.74	0.38	0.41	0.13	0.03	1.82	0.59	4.21	0.27	3.38
212	어가수	169	0	16	47	2	87	5	296	99	226
	%	2.69	0	0.25	0.75	0.03	1.39	0.08	4.72	1.58	3.6
182	어가수	34	0	0	0	0	0	0	2	4	0
	%	0.54	0	0	0	0	0	0	0.03	0.06	0
172	어가수	138	32	27	31	91	236	4	239	170	280
	%	2.2	0.51	0.43	0.49	1.45	3.76	0.06	3.81	2.71	4.46

위에서 살펴본 바와 같이 기존의 표본은 전체 모집단의 특성을 골고루 잘 나타내지 못하는 면이 있다. 그 주된 이유는 너무 작은 크기의 표본을 가지고 어로어업 99개, 양식어업 25개 전 어종을 모두 표본설계에 반영하려 한 점이다. 그 결과 과다하게 많은 층이 구성되어 오히려 표본의 대표성이 떨어지게 되었으며, 동시에 추정량의 정도를 향상 시킨다는 층화 본래의 목적도 잃어버리게 되었다. 너무 많은 어종을 중심으로 표본을 배정한 결과 표본이 모집단의 구성비율을 제대로 반영하지 못하게 되었으며 오히려 왜곡시키는 결과를 초래한 것이다.

다음으로 층화에서의 문제를 지적할 수 있다. 기존 표본설계에서는 층화변수로 어종을 사용하였는데 앞서도 언급한 바와 같이 어업총조사 정보를 참고해볼 때 어종보다는 어법을 층화변수로 선택하는 것이 더욱 효과적이었을 것이다. 또 하나 모집단과 표본 사이에 괴리가 생기게 한 원인으로 도서지역이 모두 표본에서 누락되었다는 점을 지적할 수 있다. 현지방문 결과에서도 언급한 바와 같이 전남 신안 군같은 경우는 어업세력이 상당한 규모를 가진 지역이지만 도서로만 이루어진 지역이므로 표본에서 제외되었다. 조사의 편의를 위해 도서지역을 제외하는 것을 원칙으로 하였지만 일정 규모 이상의 어업세력을 가진 도서지역에 대해서는 최소한이라도 표본에 반영해주는 것이 필요한데 이를 배제함으로써 일정 부분 표본의 대표성을 잃어버리게 되었다고 할 수 있다.

3.4.3 통계 발표 양식

현재 해양수산부에서 발표하는 어업생산통계는 업종별, 어종별, 시도별 계통 및 비계통 생산량과 생산액 통계가 월간으로 작성되고 있다. 그런데 이러한 발표 양식은 몇 가지 중요한 문제를 지니고 있다. 계통 생산량에 대한 통계는 기본적으로 전수조사에 의해 작성되므로 별 문제가 없지만 비계통 생산량 통계는 불과 1,050 개의 표본조사 자료에 의해 작성되는 것이므로 현재와 같이 발표되어서는 곤란한데 그 자세한 근거를 지적하면 다음과 같다.

표본조사에 의한 통계는 근본적으로 표본오차를 수반하게 된다. 따라서 통계를 발표할 때에는 반드시 표본오차를 함께 발표하여 통계이용자들에게 공표되는 통계의 품질을 알려주는 것이 필요하다 (Biemer and Lyberg, 2003). 통계이용자들은 일반적으로 통계를 사실인 것으로 받아들이는 경향이 있으므로 통계생산자가 미리 개별 통계의 표본오차를 살펴서 신뢰할 만하다고 판단되는 통계를 선별하여 발표해야 한다. 그러나 기존의 발표양식을 보면 표본오차에 대한 고려가 전혀 없다는

사실을 확인할 수 있다.

보다 세부적으로 기존의 발표양식을 살펴보기로 하자. 가장 먼저 어업생산량(총괄) 부분에서는 시도별 어업별 및 시도별 종류별 통계와 전국 품종별 통계를 발표하고 있는데 이 경우 비계통생산량과 계통생산량을 합한 통계이므로 상대적으로 표본오차가 크게 문제되지 않을 것으로 예상되기는 하지만 그래도 표본오차를 구하여 확인해보는 것이 필요한데 이 점이 고려되지 않았다.

다음으로 어업별, 시도별 통계가 발표되고 있는데 이 부분에서는 계통과 비계통 생산량이 각각 나누어 발표되고 있다. 이 부분의 비계통 생산량 통계의 경우 앞의 3.4.1절에서 살펴본 대로 일부 어종을 제외하고는 표본오차가 너무 커서 신뢰성이 떨어진다는 문제를 지니고 있다. 3.4.1절에서 구한 표본오차는 전국 생산량에 대한 표본오차이다. 이를 시도별로 구분한다면 각 시도별 표본오차는 전국에 비해 훨씬 표본오차가 커지게 되므로 사실상 공적으로 발표할 만한 통계가 되지 못한다. 그럼에도 불구하고 현재의 발표 양식은 이런 사실을 전혀 감안하지 않고 있는데 이는 심각한 문제로 지적될 수 있다.

참고로 통계청이나 해외의 통계보고서 등을 살펴보면 발표하는 통계에 대한 표본오차를 따로 계산하여 보고서의 부록 등에 수록하고 있다. 해양수산부에서도 그러한 모범적인 보고서를 참조하여 통계보고서의 양식을 변경하는 것이 바람직하다고 여겨진다.

3.4.4 어업조사구 사용에 따른 조사상의 문제

앞의 3.3절에서 언급한 바 있듯이 기존 어업조사구의 경우 한 조사구가 매우 광범위한 지역으로 구성된 경우가 허다하다. 그러다보니 실제 조사현장에서 조사원들이 표본을 관리하고 정확한 조사를 수행하는 데 많은 문제점을 가져오게 된다. 표본설계를 할 때에는 가능한 한 조사관리가 용이하도록 하기 위해 어업조사구를 하나의 집락으로 사용하는 이단 집락추출법을 고려하였다. 그러나 막상 집락으로 사용된 어업조사구의 크기나 범위가 서로 균일하지 못한 관계로 이런 문제가 발생하게 된 것이다.

이 문제의 기본 원인은 어업총조사를 할 때 어업조사구 구성에서 비롯되었다. 해양수산부(2003)는 어업조사구의 이런 문제를 해결하기 위해 새로운 어업조사구 구성 방법을 소개한 바 있다. 어쨌든 현재의 상황에서 어업조사구의 구성 자체를 변

화시키기는 어려운 형편이므로 어업조사구 대신 인구주택조사구를 집락으로 사용하면 조사 상의 문제를 완화시킬 수 있을 것이다.

3.4.5 표본관리 및 조사관리의 문제

기존의 비계통생산량 조사는 고정된 표본가구에 대해 정기적으로 조사를 반복하여 실시하는 일종의 패널조사이다. 패널조사의 경우 표본 및 조사관리 상에서 여러 가지 문제가 발생할 수 있으므로 이에 대한 적절한 대처방안을 마련하여 관리하는 것이 필요하다. 중요한 관리 항목으로는 모집단의 변화를 파악하여 추출틀에 반영하는 추출틀 관리, 표본관리 및 가중값(승수)의 조정 등을 들 수 있다. 조사 예산이나 인력 등의 한계로 인해 완벽한 추출틀 관리는 현실적으로 불가능한 측면이 있다. 하지만 표본관리나 가중값의 조정 등은 조금만 관심을 기울이면 쉽게 할 수 있는 부분이므로 마땅히 신경을 써야 할 부분이다. 그런데 기존 조사에서는 이런 부분에 대한 명확한 관리지침이 마련되어 있지 못하므로 이에 대한 보완이 필요하다.

기존의 조사관리의 기본 개념은 일반해면어업조사를 위한 표본의 크기 1,050어를 일정하게 유지하는 것이라고 여겨진다. 우리나라 어가의 경우 해가 갈수록 감소하는 경향이 두드러진다. 가령 표본어가가 이사를 간다든지 아니면 어업을 그만둔다든지 하는 경우를 예로 들 수 있다. 이 때 표본의 크기를 일정하게 유지하려면 탈락되는 표본어가를 비슷한 성격의 다른 가구로 대체(substitution)시켜야 한다. 모집단의 변화로 인해 표본에도 변화가 생기는 수가 있는데 이를 무시하고 다른 표본어가로 대체시킬 경우 표본과 모집단이 유리되는 현상이 생기게 된다. 또 소멸되는 표본어가와 같은 특성을 지니는 어가가 해당 조사구 내에 존재하지 않는 경우 어쩔 수 없이 특성이 다른 어가로 대체하게 될 텐데 이것은 또 다른 편향(bias)을 초래하게 된다. 이러한 사항들을 종합적으로 감안하여 적절한 대체 요령 및 표본관리 요령을 마련하고 조사를 보다 체계적으로 관리해야 할 필요가 있다.

마지막으로 조사의 질을 높이기 위한 방안에 대한 연구가 부족하다는 문제를 지적할 수 있다. 비계통조사의 질은 궁극적으로 응답하는 어가의 협조에 달려있다. 불성실한 응답을 기초로 작성되는 통계가 정확한 통계가 될 수는 없기 때문이다. 기존 조사는 기본적으로 응답자가 자신의 비계통생산 상황을 스스로 기입하는 자기기입식 조사방법을 따르고 있다. 이 경우 응답자가 얼마나 성실하게 응답을 하고 있는지를 파악할 뿐 아니라 응답자의 응답의 질을 높이기 위한 다각적인 노력이 절실하게 필요하다. 기존의 조사에서도 비계통조사의 질을 높이기 위해 어떤 노력을 기

울이고 있을 터인데 그러한 노력이나 연구가 객관적이고 합리적으로 검토, 관리되는 흔적을 발견할 수가 없었다. 이는 비표본오차의 문제를 충분히 인식하지 못하고 있다는 사실을 나타내므로 중요한 문제점이라고 지적할 수 있다.

제 4 장.새로운 표본설계

4.1 새로운 표본설계를 위한 추출틀

기존의 표본설계에 의해 생산되는 통계는 표본오차가 너무 크기 때문에 신뢰성 있는 비계통어업생산량 통계를 산출하기 위해서는 표본수를 확대하는 것이 절실하게 필요하다. 아울러 기존 표본에서는 도서(섬) 지역이 조사 편의를 위해 완전히 배제되어 있었는데 이로 인해 표본의 모집단 대표성에 문제가 있을 것으로 판단된다. 따라서 새로운 표본설계에서는 이런 점을 보완하는 것이 필요하다.

기존의 1차 추출단위로 사용된 어업조사구의 경우 농·어업 총조사를 동시에 고려하여 만든 것으로 하나의 조사구에 농가 또는 어가가 60-100가구(평균 80가구) 포함되도록 구성되어 현실적으로 어업관련 통계를 생산하기 위한 조사구로 부적합하다. 특히 기존 어업 조사구의 경우 우리나라 실정상 어업보다는 농업부분을 좀더 중요하게 고려하여 조사구를 구성하였기 때문에 일부 어업 조사구의 경우 매우 광범위한 지역의 어가들이 하나의 조사구로 구성되어 있는 경우가 많기 때문에 실제 조사현장에서 조사원들이 표본을 관리하고 정확한 조사를 수행하는 데 많은 문제가 있다.

현실적으로 해양수산부 (2003)에 제안한 방법과 같이 새로운 조사구를 당장 구성하여 표본설계에 반영하는 것은 불가능하기 때문에 새로운 표본설계에서는 차선책으로 1차 추출단위로 어업조사구 대신에 2000년도 인구주택 총조사를 위해 구성된 인구주택조사구를 활용하기로 한다. 기존의 어업조사구는 인구주택조사구를 재구성하여 작성된 것이므로 인구주택조사구를 추출단위로 활용하면 기존의 표본을 그대로 이용하는 동시에 새로운 표본을 추가하는 데서 오는 혼란을 줄일 수 있다는 장점이 있다.

(1)추출단위로 인구주택 조사구 활용

기존 어업생산비통계조사에서는 어업조사구를 1차 추출단위로 사용하고 있지만 현지 통계 담당자들에 의하면 어업조사구의 경우 과다하게 조사구가 큰 경우가 많아 많은 어려움을 겪고 있다. 따라서 조사 편의를 고려한 효율적인 1차 추출단위의 선택을 위해 어업조사구와 인구주택조사구 현황을 비교분석해 보고자 한다. 이를

위해 2000년 어업총조사 자료를 기준으로 조사구내 어가수를 정리하면 <표 4-1> 및 <표 4-2>와 같다.

<표 4-1> 어업조사구 기준 조사구내 어가수

조사구내 어가수	조사구수	누적조사구수(%)	어가수	누적어가수(%)
1~5	792	792(23.76)	1742	1742(2.14)
6~10	293	1085(32.55)	2321	4063(4.98)
11~20	591	1676(50.29)	9180	13243(16.23)
21~30	520	2196(65.89)	13227	26470(32.45)
31~40	424	2620(78.61)	14939	41409(50.76)
41~50	318	2938(88.15)	14337	55746(68.34)
51~60	196	3134(94.03)	10766	66512(81.54)
61~70	101	3235(97.06)	6600	73112(89.63)
71~80	51	3286(98.59)	3824	76936(94.32)
81~90	25	3311(99.34)	2164	79100(96.97)
91~100	11	3322(99.67)	1047	80147(98.25)
101~120	6	3328(99.85)	648	80795(99.05)
121~150	3	3331(99.94)	408	81203(99.55)
151~200	2	3333(100.00)	368	81571(100.00)
총계	3,333	3333(100.00)	81571	81571(100.00)

<표 4-1>에 제시된 어업 조사구 기준 조사구내 어가 현황을 살펴보면 3장에서도 언급한 것과 같이 총 어업조사구수는 3,333개(서울 및 대구 1개 어가씩 포함)이고, 조사구당 평균 어가수는 24.5어가이다. 따라서 조사구당 평균 어가수 측면에서 보면 어업조사구가 어업통계를 위한 표본설계에 적합한 추출단위로 보일 수 있지만 실제 어업조사구의 경우 조사구내 어가수가 과도하게 많은 조사구가 너무 많다는 문제점을 갖고 있다. <표 4-1>을 보면 너무 조사구가 크다고 볼 수 있는 조사구내 어가수가 40어가 이상인 조사구수가 전체에서 차지하는 비중이 21.39%이고, 특히 이들 대형 조사구에 포함된 어가는 전체어가 중 49.24%를 차지하고 있다. 실제 이런 대형 조사구를 포함한 상당수의 어업 조사구는 너무 넓은 지역을 포함하고 있다. 따라서 표본이 널리 산포되어 어업 조사구를 1차 추출단위로 사용하게 되면 현장에서 표본 조사구 및 어가를 관리하고 실제 조사를 수행하는 데 있어 상당한 어려움을 겪게 된다. 다시 말해 조사의 편의를 고려한 집락추출의 효과를 충분히 얻

을 수 없다는 문제가 발생한다. 실제 현재 표본과 관련하여 해양수산부 지방청의 통계 담당자들은 기존 표본 조사구의 크기가 너무 크기 때문에 조사에 있어서 많은 어려움을 겪고 있다는 것을 현지답사를 통해 확인할 수 있었다.

한편 <표 4-2>에 제시된 인구주택조사구 기준 조사구내 어가현황을 살펴보면 어가가 있는 인구주택조사구는 총 8,425개 이고, 조사구당 평균 어가수는 9.6어가이다. 따라서 인구주택 조사구의 경우 조사구당 평균 어가수 측면에서 보면 조사구의 크기가 너무 작은 것으로 보이지만, 소규모 조사구는 실제 표본추출 대상이 되는 추출틀 구성에 있어서 배제된다는 점을 염두에 둘 필요가 있다. 어업조사구의 경우와는 달리 인구주택조사구에서는 대형 조사구로 볼 수 있는 어가수가 40어가 이상인 조사구가 차지하는 비중이 3.6%에 지나지 않고, 이들 조사구에 포함된 어가는 전체 어가 중 8.73%에 지나지 않는다. 따라서 조사의 편의를 고려할 때 인구주택 조사구를 활용하는 것이 상당히 효과적일 것으로 판단된다. 따라서 새로운 표본설계에서는 현지 조사구 관리 및 조사 편의를 고려하여 인구주택조사구를 1차 추출단위로 사용하기로 한다. 이는 현지답사 결과 얻어진 지방청 담당자들의 의견을 표본설계에 적극 반영한 것이다.

<표 4-2>

조사구내 어가수	조사구수	누적조사구수(%)	어가수	누적어가수(%)
1~5	4857	4857(57.65)	9561	9561(11.72)
6~10	980	5837(69.28)	7636	17197(21.08)
11~20	1189	7026(83.39)	17925	35122(43.06)
21~30	722	7748(91.96)	18118	53240(65.27)
31~40	374	8122(96.40)	13057	66297(81.28)
41~50	182	8304(98.56)	8149	74446(91.27)
51~60	82	8386(99.54)	4490	78936(96.77)
61~70	30	8416(99.89)	1949	80885(99.16)
71~80	7	8423(99.98)	505	81390(99.78)
81~90	1	8424(99.99)	90	81480(99.89)
91~100	1	8425(100.00)	91	81571(100.00)
총계	8425	8425(100.00)	81571	81571(100.00)

(2)인구주택 조사구 기준 조사모집단 구성

표본설계에 있어서 모든 조사구를 실제 표본추출 대상이 되는 조사모집단(survey population)에 포함하는 것은 현실적으로 불가능하기 때문에 소규모 조사구를 제외한 조사모집단의 구성이 필요하다. 새로운 표본설계에서는 각 조사구에서 4개 어가를 표본으로 추출한다는 점을 고려하여 조사구내 어가수가 5가구 이하인 조사구는 조사모집단에서 제외하기로 한다(기존 표본설계의 경우 8가구 미만인 어업조사구를 제외하고 있음).

이에 따라 실제 조사모집단에서 제외되는 조사구수는 4,857개 조사구로 전체 조사구 중 57.65%를 차지하지만 어가수로는 9,561어가로 전체 어가 중 11.71%가 표본 추출대상 추출틀에서 제외되게 된다. 하지만 이들 조사모집단에서 누락되는 어가들은 주로 어가가 매우 적은 내륙지방 또는 어가가 밀집되어 있지 않은 지역에 거주하는 어가들로 이들 어가들이 실제 전체 비계통어업생산통계에 미치는 영향은 매우 미미할 것으로 판단된다. 따라서 어업조사구 대신 인구주택조사구를 사용함으로써 발생하는 추가적인 비포함오차(non-coverage error)는 무시할 수 있을 것으로 판단된다. 특히 인구주택조사구를 사용함으로써 발생하는 조사 업무량 감소를 통해 얻어지는 효과는 이런 약간의 비포함오차의 증가를 상쇄하고도 남는다고 볼 수 있다.

결과적으로 새로운 표본설계에서 실제 표본추출대상이 되는 조사모집단은 6어가 이상이 포함된 3,568개 인구주택조사구에 포함된 72,010어가로 구성한다.

4.2조사구 층화

효율적인 층화를 위해서는 적절한 층화변수를 선택하는 동시에 각 층의 분산을 최소화 할 수 있도록 층의 경계를 결정해야 한다. 기존 표본설계에서는 어업조사구를 층화하기 위해 전체 조사구를 각 조사구의 주요 어획 어종에 따라 어로어업의 99개 어종, 양식어업의 25개 어종으로 구분하고, 이를 다시 시도로 층화하는 방법을 적용하고 있다. 이에 따라 기존 표본설계에서는 너무 많은 수의 층을 사용하게 되기 때문에 효과적인 층화가 이루어졌다고 볼 수 없다.

기존 표본설계에서 이런 방식의 조사구 층화를 고려하게 된 이유는 어업생산량 관심대상이 되는 어로어업의 99개 어종과 양식어업의 25개 어종을 모두 표본설계

에 반영하는 것을 염두에 두었기 때문이었다. 하지만 이런 방식을 적용함으로써 과다한 층이 구성되게 되어 오히려 표본의 대표성을 확보하고 동시에 추정량의 정도를 향상시킨다는 본래 층화작업의 목적이 오히려 달성되지 못하고 있다.

2003년 표본어가를 대상으로 조사한 결과를 분석한 바에 따르면 어가당 평균 사용 어법의 수는 2보다 작은 데 반해 어가당 생산하는 평균 어종의 수는 5.6종인 것으로 나타났다. 이에 따라 새로운 표본설계에서는 어종을 직접 층화작업에 반영하는 대신 각 어가에서 사용되는 어법과 지역개념을 사용하여 층화의 효율성을 높이고자 한다. 물론 개개의 어종을 가지고 층화를 하지 않고 어법과 지역을 층화변수로 사용함에 따라 어종에 따라서는 효율이 저하되는 어종도 생길 수 있다. 그러나 대부분의 어종에 대해서는 효율을 높일 수 있을 것으로 판단된다.

(1) 시도별 사용 어법 현황

새로운 표본설계에서는 각 어가 또는 조사구에서 많이 사용하는 어법을 파악하여 이를 반영함으로써, 가능한 주요 어로어업 어종을 어획하고 있는 어가들을 균형적으로 표본어가에 포함하여 주요 어종별 비계통생산량 추정결과의 정도를 확보하고자 한다. 따라서 각 시도에서 주로 사용되는 어법종류 및 해당 어법을 사용하는 어가 및 인구주택 조사구 현황을 정리하고자 한다.

2000년 어업 총조사에서는 어로어업 어가의 경우 지난 1년간 경영한 어업별 어업의 종류를 판매금액 많은 순으로 6개까지 조사하고 있으며, 동시에 어로어업으로 어획한 주된 어종 3가지를 조사하고 있다. 어업 생산량 조사를 위해서는 각 시도별로 각 어법을 사용하는 어가를 파악하는 것이 많은 도움이 될 수 있다. 따라서 시도별로 어법종류에 따른 어가 분포현황을 정리하면 <표 4-3>과 같다. <표 4-3>에서의 조사구는 새로운 표본설계에서 추출단위로 사용하는 인구주택조사구를 기준으로 한 것이고, 각 어가는 자신들이 사용하는 어법들을 어업총조사에서 6개까지 중복 응답할 수 있기 때문에 이런 중복 응답결과를 종합하여 어가수 및 조사구수를 산출하였다는 점에 유의하기 바란다.

<표 4-3>

어법 종류	총 계		부산		인천		울산		경기		강원	
	어가수	조사구	어가수	조사구	어가수	조사구	어가수	조사구	어가수	조사구	어가수	조사구
대형기선 저인망	69	61	42	39	5	5					9	5
중형기선 저인망	94	80	1	1			24	18			45	37
근해트롤	120	61	9	9							59	26
근해형망	61	42			1	1						
연안조망	451	182	3	3	2	2	2	2			7	7
근해선망	54	46	1	1	2	2	4	3			2	2
연안선망	416	225	16	12	4	2	20	10	11	4	25	16
근해안강망	366	270			34	34			5	2		
연안안강망	984	313			120	47	1	1	53	10	3	2
근해붕수망	3	3										
연안들망	385	178	42	22	1	1	26	9	11	2	2	2
근해자망	594	374	20	15	38	30	8	7	1	1	102	60
연안자망	11084	2813	460	230	515	175	198	75	230	51	1601	358
기선선인망	50	45			2	2	8	8				
연안선인망	200	96	7	7	2	2	6	5	1	1	17	9
근해통발	391	266	10	6	12	11	2	2	1	1	117	82
연안통발	9119	2527	806	370	116	55	322	78	71	26	218	112
대형정치망	301	129			3	2					87	35
중형정치망	315	143	5	5	2	2	5	3	4	3	51	20
소형정치망	1248	486	47	24	39	19	55	27	6	4	78	37
정치성구획	1262	318	9	9	73	27	14	9	2	1	39	18
이동성구획	453	104			86	17			2	1	23	18
근해연승	548	325	24	20	13	6	1	1			28	19
근해채낚기	2190	800	85	42	4	3	77	36			854	325
잠수기	417	249	2	2	20	5	2	2			150	88
연안복합 어업	19210	4307	1038	406	684	155	214	65	211	59	1741	410
마을어업	3756	391	12	2	147	75	1	1	4	4	9	5
기타어업	18623	2631	547	250	1354	142	379	108	850	60	156	71
전체	72764	17465	3186	1475	3279	822	1369	470	1463	230	5423	1764

어법 종류	충남		전북		전남		경북		경남		제주	
	어가수	조사구	어가수	조사구	어가수	조사구	어가수	조사구	어가수	조사구	어가수	조사구
대형기선 저인망					7	6			6	6		
중형기선 저인망	1	1			9	9	7	7	6	6	1	1
근해트롤			1	1	2	2	46	20	3	3		
근해형망	15	11	22	9	1	1			21	19	1	1
연안조망	143	44	97	27	149	59	5	5	31	23	12	10
근해선망	5	4	2	2	13	12	12	8	11	10	2	2
연안선망	41	15	13	10	125	50	54	28	94	66	13	12
근해안강망	55	33	89	67	181	132					2	2
연안안강망	359	90	195	49	230	96	9	6	12	10	2	2
근해봉수망			1	1	1	1			1	1		
연안들망	4	4	12	10	63	26	4	4	105	43	115	55
근해자망	22	17	44	35	109	78	130	75	53	37	67	19
연안자망	1287	260	637	185	1347	370	1987	361	2558	652	264	96
기선선인망					2	2	1	1	36	31	1	1
연안선인망	5	5	21	5	94	24	14	10	32	27	1	1
근해통발	15	13	6	6	67	23	66	44	91	74	4	4
연안통발	476	146	270	68	2840	622	1131	269	2847	763	22	18
대형정치망	3	3	3	3	12	10	163	50	29	25	1	1
중형정치망	9	7	4	4	46	17	108	42	69	29	12	11
소형정치망	13	10	73	25	515	125	141	70	221	109	60	36
정치성구획	58	12	29	8	593	101	220	67	224	65	1	1
이동성구획	6	5	154	19	156	35	2	2	24	7		
근해연승	11	9	100	60	50	37	4	2	126	79	191	92
근해채낚기	4	4	14	13	21	19	818	198	86	42	227	118
잠수기	31	12	3	3	22	19	66	44	121	74		
연안복합 어업	1277	233	781	141	6151	1114	949	255	4405	876	1759	593
마을어업	559	42	124	13	2211	162	302	41	378	38	9	8
기타어업	4076	293	1974	183	3102	475	691	162	1179	294	4315	593
전체	8475	1273	4669	947	18119	3627	6930	1771	12769	3409	7082	1677

위의 <표 4-3>에서 음영을 넣어 나타낸 어법들은 2000년 어업총조사와 2003

년 비계통어업생산량 표본조사 자료를 바탕으로 우리나라 어가 중 다수의 어가에서 사용하고 있고, 어업생산량 측면에서 중요하다고 판단되는 주요어법들을 나타낸 것으로 정리하면 다음과 같다 :

- ※ 연안복합어업, 연안자망, 연안통발, 마을어업, 정치망(대, 중, 소)
정치성구획(각망), 연안안강망, 이동성구획(기타구획), 기타어업.

(2)주요 어법 및 지역에 따른 조사구 층화

주요 어법에 따른 조사구의 특성 및 조사구의 지역적 특성을 반영하기 위해 새로운 표본설계에서는 전체 조사구를 우선 지역에 따라 다음과 같이 4개 층으로 구분한다.

지역층1: 강원, 경북, 울산

지역층2: 전남, 경남, 부산

지역층3: 전북, 충남, 경기, 인천

지역층4: 제주

또한 각 지역층에 속하는 조사구 내의 어가들이 주로 사용하는 어법에 따라 조사구를 구분하기 위해 조사모집단에 속하는 3,568개 조사구 및 72,010어를 대상으로 주요 어법에 따른 조사구 및 어가 분포 현황을 지역층 별로 정리하면 <표 4-4>와 같다. 여기서 조사구수는 조사구 내에 1개 어가라도 해당 어법을 사용하는 어가가 있는 경우 모두 계상한 것으로 실제적으로 중복 반영된 것이다.

<표 4-4> 주요어업별 조사구 및 어가 현황

어법 종류	총계		강원/경북/울산		전남/경남/부산		전북/충남/ 경기/인천		제주	
	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)
기타 어업	1718 (19.38)	17130 (29.16)	251 (13.26)	1110 (11.51)	611 (13.85)	4224 (15.10)	503 (28.18)	7946 (50.81)	353 (45.43)	3850 (70.24)
연안복 합어업	2377 (26.81)	16326 (27.79)	420 (22.19)	2435 (25.25)	1290 (29.24)	9906 (35.41)	380 (21.29)	2684 (17.16)	287 (36.94)	1301 (23.74)
연안 자망	1723 (19.43)	9628 (16.39)	511 (26.99)	3369 (34.93)	763 (17.29)	3736 (13.35)	374 (20.95)	2284 (14.60)	75 (9.65)	239 (4.36)
연안 통발	1617 (18.24)	7902 (13.45)	336 (17.75)	1518 (15.74)	1053 (23.87)	5540 (19.80)	211 (11.82)	823 (5.26)	17 (2.19)	21 (0.38)
마을 어업	302 (3.41)	3638 (6.19)	45 (2.38)	310 (3.21)	187 (4.24)	2579 (9.22)	64 (3.59)	742 (4.74)	6 (0.77)	7 (0.13)
정치망	591 (6.67)	1656 (2.82)	231 (12.20)	619 (6.42)	261 (5.92)	841 (3.01)	63 (3.53)	136 (0.87)	36 (4.63)	60 (1.09)
정치성 구획	266 (3.00)	1195 (2.03)	82 (4.33)	259 (2.69)	147 (3.33)	789 (2.82)	36 (2.02)	146 (0.93)	1 (0.13)	19 (0.02)
연안안 강망	197 (2.22)	849 (1.45)	6 (0.32)	9 (0.09)	68 (1.54)	196 (0.70)	121 (6.78)	642 (4.11)	2 (0.26)	2 (0.04)
이동성 구획	76 (0.86)	416(0.7 1)	11 (0.58)	16 (0.17)	32 (0.73)	164 (0.59)	33 (1.85)	236 (1.51)	0 (0.00)	0 (0.00)
총계	8867 (100.0)	58740 (100.0)	1893 (100.0)	9645 (100.0)	4412 (100.0)	27975 (100.0)	1785 (100.0)	15639 (100.0)	777 (100.0)	5481 (100.0)

한편 각 인구주택조사구내의 어가들은 모두 동일 어법을 사용하고 있지 않기 때문에 조사구를 어법에 따라 구분하기 위해 각 조사구내의 어가 중 가장 많은 수의 어가가 사용하는 어법을 해당 조사구의 “대표 어법”이라 하고 조사모집단을 구성하고 있는 3,568개의 조사구를 대표어법에 따라 구분하고자 한다. 각 대표어법에 속하는 지역별 인구주택조사구와 이들 조사구에 속하는 어가를 정리하면 <표 4-5>와 같다.

<표 4-5> 대표어법 기준 조사구 분류 현황

대표 어법	총계		강원/경북/울산		전남/경남/부산		전북/충남/ 경기/인천		제주	
	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)	조사구 (%)	어가수 (%)
연안복 합어법	1100 (30.83)	21476 (29.82)	170 (26.36)	2730 (27.33)	733 (39.18)	14674 (37.32)	115 (16.91)	2944 (17.26)	82 (22.04)	1128 (20.01)
연안 자망	533 (14.94)	9804 (13.61)	216 (33.49)	3430 (34.33)	221 (11.81)	4525 (11.51)	91 (13.38)	1782 (10.44)	5 (1.34)	67 (1.19)
연안 통발	1053 (29.51)	21118 (29.33)	206 (31.94)	2780 (27.83)	709 (37.89)	15210 (38.68)	126 (18.53)	2916 (17.09)	12 (3.23)	212 (3.76)
기타 어법	832 (23.32)	18131 (25.18)	39 (6.05)	710 (7.11)	86 (9.94)	4163 (10.59)	334 (49.12)	9027 (52.91)	273 (73.39)	4231 (75.04)
나머지	50 (1.40)	1481 (2.06)	14 (2.17)	340 (3.40)	22 (1.18)	749 (1.9)	14 (2.06)	392 (2.30)	0 (0.00)	0 (0.00)
계	3568 (100.0)	72010 (100.0)	645 (100.0)	9990 (100.0)	1871 (100.0)	39321 (100.0)	680 (100.0)	17061 (100.0)	372 (100.0)	5638 (100.0)

※ 여기서 “나머지” 는 대표 어법이 제시된 어법에 해당하지 않는 모든 조사구를 나타냄

<표 4-5>를 기초로 각 지역층별로 <표 4-6>에 제시된 것과 같이 각 조사구의 대표어법을 기준으로 조사구를 구분하여 표본추출에 적용한다. 여기서 대표어법에 따른 구분을 “층”이라고 부르지 않는 이유는 실제 조사구 표본추출에 있어서 이들 대표어법에 따른 구분은 층화추출에서의 “층” 개념으로 표본추출에 적용하는 것이 아니라, 각 지역층 내에서 표본 조사구를 크기비례확률(PPS) 계통추출할 때, 조사구들을 정렬하는 기준으로 사용하여 표본설계에 반영하기 때문이다. 따라서 실제 이들 대표어법구분은 “층”으로 사용되지 않지만 실제 표본추출에 있어서는 대표어법 구분은 계통추출의 정렬기준으로 활용하여 이들 대표어법으로 층을 구성하고 각 층에서 비례배분에 의해 층화추출하는 것과 거의 동일한 효과를 얻게 된다.

<표 4-6> 지역층별 대표어법에 따른 조사구 구분

지역층	구분1		구분2		구분3		구분4	
	대표어법	어가수	대표어법	어가수	대표어법	어가수	대표어법	어가수
강원/경북/울산	연안자망	3430	연안통발	2780	연안복합 어업	2730	나머지	1050
전남/경남/부산	연안통발	15210	연안복합 어업	14674	연안자망	4525	나머지	4912
전북/충남/ 경기/인천	기타어업	9027	연안복합 어업	2944	연안통발	2916	나머지	2174
제주	기타어업	4231	나머지	1407				

4.3 표본의 크기 및 목표오차

4.3.1 표본의 크기

일반적인 표본설계에서는, 표본크기를 정할 때 사전에 미리 목표예산과 목표오차를 정해놓고 거기에 합당한 표본크기를 산출하게 된다. 그러나 본 연구에서는 일반적인 경우와는 달리 연구 시작 때부터 미리 표본크기가 2,000으로 결정된 상황이다. 기존에 1,050 표본어가의 조사를 2,000 어가로 늘린다는 전제 하에서 예산이나 조사인력, 일정 등이 정해진 까닭이었다.

본 표본설계 연구는 새롭게 늘려야 할 950가구를 어떻게 배정하면 기존 표본의 약점을 보완하면서도 효율을 극대화시킬 수 있는가 하는 문제를 다루는 연구라고 할 수 있다. 다시 말하자면 본 연구의 핵심은 효과적인 층화 및 표본배분 전략을 마련하는 데 있다.

4.3.2 표본의 배분

기존의 표본 배분은 모집단 어업조사구의 주요 어종별 특성을 파악한 후 각 어종별로 골고루 배분하는 방식으로 이루어졌다. 말은 골고루 배분한다고 했지만 각 어

종별로 중요도가 다른데 이것을 객관적으로 반영하지도 못했고 또 지역별 비례배분 된 것도 아니므로 결과적으로 어정쩡한 형태의 배분이 이루어졌다. 새로운 표본설계는 기존의 표본을 그대로 사용하면서 추가적으로 950 어가를 더하는 표본설계의 성격을 띠게 된다. 따라서 가능한 한 모집단의 특성을 파악하여 기존의 표본과 모집단 특성의 괴리를 보완할 수 있도록 배분하는 것이 바람직하다.

다음의 <표 4-7>은 기존 표본 자료로부터 구한 주요 어법별 비계통생산 어가와 생산량을 나열한 표이다. 이 표에서 대부분의 어가가 몇 가지 대표적인 어법을 사용하고 있으며, 대표적인 몇 가지 어법에 의한 비계통생산량이 전체 생산량의 대부분을 차지함을 알 수 있다. 이 <표 4-7>을 이용하여 301, 281, 266, 212, 172 등의 어법이 표본설계 시에 주로 고려되어야 할 주요 어법임을 알 수 있다.

<표 4-7> 표본의 어법별 어가수
및 생산량 현황

어법	어가수	생산량
122	8	115536
151	3	16246
171	1	2231
172	250	14823724
182	26	8925255
201	10	1495474
211	29	2809217
212	294	9317567
213	7	1054451
221	10	2094323
231	1	14650
241	9	83610
251	17	1778307
266	314	11821365
271	24	633672
281	395	44172771
291	7	440067
292	31	1365025
293	10	2291327
294	12	8752265
295	6	342525
299	57	4366118
301	496	36960732

표본의 크기가 충분하지 않은 상황에서 너무 많은 변수를 고려하는 것은 바람직하지 않으므로 가장 중요한 어법이라고 생각되는 301, 266, 172 어법을 고려하여 표본을 배분하기로 한다. 앞 절의 증화에서 소개된 바대로 먼저 전국을 세 개의 주요한 지역층으로 구분한 후 각 지역층별 모집단 특성을 파악하였다. 다음의 <표 4-8>은 여러 어법별 모집단 특성에 비례하도록 표본크기를 배분한 결과이다. 어가수나 266 어법을 기준으로 비례배분할 경우 부산, 전남, 경남으로 구성되는 지역층이 가장 많은 표본이 배정되어야 한다. 반면 301 어법을 기준으로 하면 인천, 경기, 충남, 전북, 제주 지역층에 많은 표본이 배정되어야 하며, 172 어법을 기준으로 하면 상대적으로 울산, 강원, 경북의 층에 많은 표본이 배정되어야 한다. 이 말은 지역에 따라 어업의 성격이 서로 다르다는 것을 뜻한다. 따라서 표본의 배정을 어느 하나의 기준에 따라서 일률적으로 하는 것보다는 지역에 따라 서로 다른 기준을 적용시키는 것이 더욱 효율을 높일 수 있게 된다.

<표 4-8> 여러 가지 기준에 의한 표본 배분 방안

어법	층 1	층 2	층 3	합 계
	울산 강원 경북	부산 전남 경남	인천 경기 충남 전북 제주	
301	127	485	1353	2000
266	286	1162	521	2000
172	688	762	515	2000
어가수 비례	288	1092	620	2000

위의 <표 4-8>을 참조하여 층1과 층2 지역에 대해서는 상대적으로 266과 172 어법을, 층 3에 대해서는 301 어법을 고려하여 최종적인 표본의 크기를 배분한 결과가 다음의 <표 4-9>에 나와 있다. 이 표에는 비슷한 원리에 의해 결정된 지역층 내의 세부적인 시도별 표본배분 상황과 기존의 표본크기도 소개되어 있다. 기존 표본에서 상대적으로 부산, 충남, 전북, 전남 등의 지역은 표본이 과다하게 추출된 지역이므로 새로 추가될 표본이 많지 않다. 반면 강원, 제주, 경남, 경기 등의 지역은 상당히 많은 수의 표본이 추가되어야 하는 것을 알 수 있다.

<표 4-9> 최종 표본 배분 결과

	총 1			총 2			총 3					합 계
	울산	강원	경북	부산	전남	경남	인천	경기	충남	전북	제주	
최종표본수	300			1000			700					2000
소지역별	30	130	140	80	500	420	90	60	280	110	160	
기준 표본	12	16	48	75	350	167	42	20	215	77	28	1,050
추가	18	114	92	5	150	253	48	40	65	33	132	950

4.3.3 목표 상대표준오차

새로운 표본설계의 효과를 가지적으로 확인할 수 있는 방법은 새로운 설계에 의한 표본의 목표 상대표준오차(CV)를 계산하는 것이다. 문제를 아주 단순화시켜서 기존의 표본 1,050 개와 추가되는 표본 950개 모두 단순랜덤추출법(simple random sampling method)에 의해 추출했다고 가정한다면 새로운 목표상대표준오차는 다음의 식으로 계산할 수 있다 :

$$CV_{\neq w} = CV_{old} \cdot \sqrt{\frac{n_{old}}{n_{\neq w}}}$$

위의 식에서 n_{old} 는 기존의 표본크기인 1,050이고 $n_{\neq w}$ 는 추가된 이후의 표본크기 2,000이 된다. 따라서 새로운 설계의 목표 CV는 과거의 CV에 비해 약 28% 정도 줄어드는 효과가 생길 것으로 기대할 수 있다.

그러나 본 연구에서 사용한 표본추출의 방법은 단순랜덤추출법보다 훨씬 복잡한 형태의 다변량 복합설계이다. 따라서 정확한 목표 CV를 계산하는 것은 쉽지 않은 일이다. 뿐만 아니라 본 연구에서 사용한 표본설계 방법은 모든 어종에 대해 동일한 효과가 나타나도록 고려한 것이 아니라 가장 대표적인 어법의 어종들을 보다 잘 반영할 수 있도록 하였기 때문에 어종에 따라서는 기존 CV에 비해 약 28% 정도 CV 감소효과가 나타날 수도 있는 반면 또 경우에 따라서는 그보다 작은 양의 CV 감소효과가 나타날 수도 있다. 다만 우리가 목표하는 바는 생산량이 많고 상대적으로 중요성이 큰 어종에 있어서는 새로운 설계에 의한 CV 감소 효과가 보다 크게

나타나도록 하는 것이었다.

다음의 <표 4-10>은 몇 가지 대표적인 주요 어종들의 기존 CV와 새로운 목표 CV를 계산하여 나타낸 표이다. 이 표에서 CV 감소율이 28% 이상인 어종은 새로운 표본설계에서 상대적으로 보다 중요하게 고려된 어종이라고 볼 수 있다. 오징어, 쭈꾸미 등의 어종은 상대적으로 표본오차의 크기가 많이 줄어든 어종이라고 할 수 있다.

<표 4-10> 주요 어종들의 목표 표준상대오차(CV)

어종	과거 CV	새 목표CV	CV감소율 (%)
가오리	0.18	0.137	24
가자미	0.05	0.035	30
멸치	0.19	0.129	33
조피볼락(우럭)	0.11	0.083	25
삼치류	0.18	0.157	13
송어	0.09	0.073	19
장어	0.11	0.096	13
전어	0.19	0.133	30
게	0.09	0.065	28
기타새우	0.23	0.159	31
굴류	0.14	0.094	33
소라고둥	0.16	0.112	30
가무락	0.16	0.112	30
고막	0.11	0.077	30
바지락	0.07	0.048	32
백합	0.12	0.084	30
낙지	0.04	0.028	30
문어	0.16	0.112	30
쭈꾸미	0.13	0.094	28
오징어	0.3	0.102	66
성게	0.12	0.068	44

과거 CV의 크기가 10% 이내인 어종수가 9개에 불과하였는데 본 연구의 결과 확대된 표본으로 조사를 할 경우 대략 40 내지 50 개의 어종으로 늘어날 것으로 기대된다. 특히 현지답사 과정에서 각 지역별로 주요 어종이라고 파악된 어종들의 CV를 낮추는데 주력하였기 때문에 실질적인 효용성이 많이 증가된 통계를 생산할 수 있을 것으로 기대된다.

표본의 크기를 2,000 개로 늘이는 것으로 인해 표본오차의 감소 효과가 뚜렷함에도 불구하고 이것만으로는 우리나라의 어가에서 생산되는 110 여종 모두에 대해 신뢰할 만한 비계통생산 통계를 작성할 수 없다는 한계를 지닌다. 아직도 40 여개 어종들에 대해서는 표본오차를 줄이기 위해 표본크기를 더욱 늘릴 필요가 있다. 이론적으로 볼 때, 아주 제한된 지역에서 소량으로 생산되어 전수조사 이외의 방법으

로는 통계를 생산하기 어려운 몇몇 어종을 제외한 나머지 대부분 어종들의 상대표본오차의 크기를 20% 이내가 되게 하려면 표본의 크기를 3,150 개, 15% 이내가 되게 하려면 5,500 개로 늘이는 것이 필요하다. 표본의 크기를 무작정 늘이는 것이 현실적으로 어려운 점을 감안할 때, 굳이 비계통생산량 조사를 하지 않아도 될 만한 일부 어종을 제외하고 지역적 사정을 잘 고려한다면 본 조사의 목적을 이루기 위해 충분한 표본의 크기는 3,500 개에서 4,000 개 사이가 바람직한 것으로 판단된다.

4.4 표본추출

새로운 표본설계에서는 앞에서 설명한 바와 같이 먼저 조사구를 지역층에 따라 층화하고, 1차 추출단위가 인구주택조사구이고 2차 추출단위가 어가인 층화 2단집락 추출법을 적용한다.

새로운 표본설계에서는 각 지역층내에서 인구주택조사구를 해당 조사구의 어가수를 MOS(measure of size)로 하는 크기비례확률 계통추출한 후, 표본으로 추출된 각 조사구에서 4어가씩 표본으로 추출하는 것을 원칙으로 하고 있다. 따라서 각 시도 지역층내에서는 표본어가의 추출확률이 동일하게 되는 동일확률추출법(epsem; equal probability selection method)이 적용된 것이다. 한편 일부 어가수가 많은 조사구의 경우 현지 조사업무량을 고려하여 4어가 대신 8어를 표본 어가로 선정하고 있는데, 이 경우에는 해당 조사구의 표본어가들의 경우 선택확률이 다른 조사구에 비해 2배로 증가하기 때문에 이를 가중값에 반영해 주고 있다.

(1) 표본 조사구의 추출

표본 조사구 추출방법을 구체적으로 설명하면 우선 각 지역층내의 조사구들을 앞에서 정의한 어법구분에 따라 분류 정렬하고, 다시 각 어법구분내의 조사구들을 지역코드순으로 정렬한 후 조사구의 어가수를 MOS로 하는 크기비례 계통추출법을 적용하여 표본 조사구를 추출하였다.

하지만 기존 표본설계는 실제적으로는 2002년도에 수행된 어업생산통계조사 표본을 확대 개편하는 의미를 갖고 있기 때문에 이미 기존 표본으로 선정되어 있는

어가를 표본에 그대로 유지하는 것이 현실적으로 바람직하다. 이는 통계의 연속성 측면에서 뿐만 아니라 현재 해양수산부 지방청의 통계 조사 인력의 업무량을 고려할 때 불가피한 선택이었다.

따라서 기존 표본 어가를 표본에 그대로 유지하기 위해 기존 표본어가 소속되어 있는 인구주택조사구를 현지 조사를 통해 확인하고 이들 인구주택조사구는 신규 표본에 포함하는 것을 원칙으로 하였다. 실제 현황 파악결과 기존 표본설계에서는 인구주택조사구 보다 큰 개념의 어업조사구를 추출단위로 사용하고 있기 때문에 기존 표본이 소속된 많은 인구주택조사구의 경우 실제 표본 어가의 수가 4개 미만인 경우가 많이 발생한다. 이런 경우 해당 조사구내에서 추가로 표본 어가를 추출하여 각 인구주택조사구에서 4개 어가를 선정하는 기준을 만족하도록 조정하였다.

결과적으로 기존 표본 조사구를 유지하면서 표본 조사구를 추가하여 최종적으로 500개의 조사구를 표본으로 선정하기 위해 본 표본설계에서는 우선 신규 조사구 선정방법에 따라 모든 500개의 인구주택 조사구를 일단 추출한 후에 이들 중 기존 조사구와 행정구역상 가까운 위치에 있는 조사구는 기존 조사구로 대체 하는 방식을 통해 기존 표본 조사구를 최종 표본에 유지하도록 했다.

아울러 기존 표본의 경우 조사상의 어려움 때문에 도서지역에 해당하는 조사구는 모두 표본에서 제외하였다. 이는 도서지역에서 많은 어업활동이 이루어지고 있다는 점을 고려할 때, 기존 어업생산량조사 표본의 경우 도서지역이 제외됨으로써 상당 부분 추정결과에 편향이 발생할 소지가 있다고 생각된다. 이에 따라 해양수산부 담당자들과의 협의를 통해 어업세력이 큰 일부 도서지역에 대해서는 유의적인 방법으로 이들 도서지역의 조사구를 표본에 추가하였다. 여기서도 마찬가지로 정상적인 방법으로 추출된 500개 조사구 중 표본추출 대상 도서지역과 지역적으로 인접한 조사구를 대체하는 방식으로 도서지역 조사구를 최종 표본 조사구로 선정한다.

아울러 현재 해양수산부 지방청 조사인력의 업무량을 고려하여 일부 어가수가 많은 표본 조사구에 대해서는 8개 표본어를 최종적으로 추출하는 방식을 취하고 있다. 여기서 일정규모 이상인 조사구의 경우 8개 어가를 추출한 것은 전적으로 현지 조사원의 업무량을 고려한 배려이다.

결과적으로 최종적으로 선정된 표본 조사구를 유형에 따라 정리해 보면, 우선 4개 어가를 추출하는 표본조사구 수는 318개(이 중 기존 표본 조사구는 192개)이고, 8개 어가를 추출하는 표본조사구 수는 92개(이 중 기존 표본 조사구는 63개)이며, 총 409개 조사구 중 기존 표본 조사구 수는 255개이고, 추가된 표본 조사구 수는

134개이다. 한편 어가수를 기준으로 정리해 보면, 기존 표본 조사구에 속하는 어가수는 1,272어가($4 \times 192 + 8 \times 63 = 1,272$)이고, 신규로 추가된 표본 조사구에 속하는 어가수는 728어가($4 \times 126 + 8 \times 28 = 728$)이다. 최종 표본 조사구는 첨부된 “[부록1] 표본조사구 목록”을 참고하기 바란다.

(2) 표본어가의 추출

표본조사구에서 표본 어가의 추출은 조사구당 4개 어가를 추출하는 것을 원칙으로 하고, 일부 대규모 조사구의 경우 이들 조사구를 가상적으로 2개의 조사구로 간주하여 8개 조사구를 추출하도록 하였다. 여기서 1개 인구주택조사구에서 8어를 선택하도록 처리한 것은 현실적으로 표본 확대에 큰 장애가 되고 있는 조사인력의 부족을 염두에 둔 것이다.

한편 조사구내에서 표본 어가를 추출은 우선 각 표본 인구주택조사구에서 정렬기준에 따라 조사구내 어가를 순서대로 정렬한 명부를 작성하고, 정렬된 어가 명부에서 계통추출에 의해 표본 어가를 선정하는 방식을 취하고 있다. 이와 관련된 구체적인 절차는 “[부록2] 표본어가 추출방법”을 참고하기 바란다.

4.5 추 정

본 표본조사는 복합조사(complex survey)이며 최종 추출단위인 어가들의 추출확률이 지역층별로 다르므로 각 어가별 조사자료에 대해 서로 다른 가중값을 사용하게 된다. 추정식을 소개하기 전에 먼저 사용하게 될 기호들을 정의하면 다음과 같다 :

· $h = 1, 2, \dots, H$; 지역층의 수

· $i = 1, 2, \dots, n_h$; h 번째 층의 i 번째 표본조사구

· $j = 1, \dots, k$ ($k = 4$ 또는 8) ; 표본조사구내 j 번째 표본가구

· $n = \sum_{h=1}^H kn_h$; 총 표본가구 수

· w_{hij} ; h 번째 층의 i 번째 조사구의 j 번째 표본가구에 대한 가중값

· y_{hij} : h 번째 층의 i 번째 조사구의 j 번째 표본가구에 대한 조사값

본 조사에서는 가능한 단순한 형태의 추정 방식을 사용하는 것을 원칙으로 하며 따라서 보조변수를 이용하는 비(ratio) 또는 회귀(regression) 추정 등은 고려하지 않기로 한다. 그 이유는, 본 연구는 전적으로 새로운 표본설계를 하는 것이 아니며 기존의 표본에 새로운 일부 표본을 추가하는 연구인 까닭에 기존의 통계와의 시계열을 유지할 수 있도록 하기 위해서 가능한 한 기존에 사용하던 추정식의 양식을 그대로 사용하려 한 때문이다. 통계를 생산하는 관리자 입장에서 추정식을 보다 간편하게 이해하고 활용할 수 있도록 하기 위해 각 조사단위별 조사값과 조사단위의 가중값(weight, 일명 승수(乘數))만 가지면 편리하게 계산할 수 있는 식으로 모집단 평균에 대한 추정량과 추정량의 분산추정량을 표현했는데 그 식은 아래의 식(4.1)과 식(4.2)에 나타나있다 :

$$\widehat{Y} = \frac{\left(\sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^k w_{hij} y_{hij} \right)}{w_{...}} \quad (4.1)$$

$$\widehat{V}(\widehat{Y}) = \sum_{h=1}^H \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (e_{hi.} - \bar{e}_{h..})^2 \quad (4.2)$$

여기서 $w_{...} = \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^k w_{hij}$ 로서 표본 조사구내 모든 조사단위들의 가중값의 합을 나타내고, $e_{hi.} = \left(\sum_{j=1}^k w_{hij} (y_{hij} - \widehat{Y}) \right) / w_{...}$ 이고 $\bar{e}_{h..} = \left(\sum_{i=1}^{n_h} e_{hi.} \right) / n_h$ 이다.

식(4.2)로부터 모집단 평균에 대한 추정량의 표준오차를 구하면 식(4.3)과 같이 된다.

$$SE(\widehat{Y}) = \sqrt{\widehat{V}(\widehat{Y})} \quad (4.3)$$

한편, 모집단 총계에 대한 추정량과 추정량의 분산추정량 및 표준오차는 식(4.4)부터 식(4.6)으로 나타낼 수 있다:

$$\widehat{Y} = \sum_{h=1}^H \sum_{i=1}^{n_h} \sum_{j=1}^k w_{hij} y_{hij} \quad (4.4)$$

$$\widehat{V}(\widehat{Y}) = \sum_{h=1}^H \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_{h..})^2 \quad (4.5)$$

$$SE(\widehat{Y}) = \sqrt{\widehat{V}(\widehat{Y})} \quad (4.6)$$

여기서 $y_{hi} = \sum_{j=1}^k y_{hij}$ 이고 $\bar{y}_{h..} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$ 이다.

참고로 추정량의 식을 위의 식(4.1)에서 식(4.6)과 같이 표현한 또 하나의 큰 이유는 모든 계산을 통계 계산용 소프트웨어인 SAS에서 제공하는 PROC SURVEY 모듈을 이용하여 간편하게 할 수 있도록 하기 위해서이다. 과거에는 추정식을 직접 프로그래밍하여 계산하였으므로 통계의 관리자와 계산을 하는 프로그래머가 이원화되어 통계 생산에 어려움이 많았다. 그러나 이제는 통계 관리자가 직접 쉽게 각종 통계를 직접 계산하여 생산할 수 있으므로 통계의 작성 및 관리에 큰 도움이 될 것으로 기대할 수 있다.

제 5 장.사후관리

5.1추출틀 보정

어업비계통생산량조사는 동일한 표본을 대상으로 매월 정기적으로 조사하는 패널 조사이다. 본 연구의 결과 얻어진 표본은 직전의 어업총조사 시점인 2000년 12월 1일의 어가를 추출틀로 하고 있으므로 2000년 당시의 모집단을 반영하고 있는 셈이다. 그렇기 때문에 그 이후의 어가 현황의 변화를 나타내지는 못하고 있다. 따라서 모집단이 계속 변할 경우 표본은 모집단의 상황을 제대로 반영하지 못하게 되어 통계량에 편향을 가져오게 될 것이다. 이런 편향이 생기지 않게 하려면 모집단의 변화상황을 파악하여 표본에 반영해야 한다. 이를 위해 추출틀을 관리하고 보정하는 것이 필요하다.

우리나라의 어가 현황은 해가 경과함에 따라 점차 줄어드는 추세이다. 다음의 <표 5-1>은 1995년과 2000년 어업총조사 결과 나타난 우리나라 어가수와 어가인구수 현황을 나타내고 있다. 5년 사이에 어가수로는 22,909 호 (21.9%), 어가인구수 기준으로는 95,861 명(27.6%)가 감소하였다. 이렇게 어가수나 어업인구수가 감소하는 요인은 크게 두 가지로 나눌 수 있는데 하나는 자연감소에 의한 것이고 다른 하나는 대규모 간척사업 등의 이유로 인해 특정 지역에서 집중적으로 감소가 일어나는 경우이다. 어업총조사를 수시로 할 수는 없으므로 자연감소에 의한 변화까지를 다 반영하기는 현실적으로 어렵다. 그러나 특정한 사회적 이유로 어느 지역에 집중적인 변화가 생기는 경우에는 해당지역에 대해서만 제한적으로 추출틀을 보정하여 관리하는 것이 바람직하다.

<표 5-1> 어가수와 어가인구수의 변화

연도	어가(호)	어가인구(명)
1995	104,480	347,210
2000	81,571	251,349

5.2 표본관리

표본으로 선정된 어가가 어업을 포기하는 경우나 다른 지역으로 이사를 가는 상황이 생길 수 있는데 이에 대해 적절한 관리요령을 마련하여 체계적으로 관리해가는 것이 필요하다. 지금까지는 표본어가에 변동이 생겨 더 이상 조사를 하기가 어려운 경우 일률적으로 대처하는 식으로 관리해왔다. 그러나 이렇게 할 경우 모집단 어가의 변동으로 인해 표본에도 자연스럽게 변화가 생기는 부분을 인위적으로 무시하는 결과가 초래되어 표본의 대표성이 왜곡될 가능성이 있다.

표본관리는 두 가지 방향으로 이루어지는 것이 필요한데 하나는 표본어가를 다른 어가로 대체(substitution)하는 방향이고, 다른 하나는 아예 표본어가를 제거하는 방향이다. 모집단의 특별한 변동에 의해서라기보다 해당 표본어가가 이사를 한다든지 하는 특수한 사정이 생겨 더 이상 응답대상이 되기 어려운 경우에는 종전과 같이 가능한 한 해당 어가와 특성이 비슷한 어가를 추출하여 대체시키면 된다. 그러나 이사가 아니라 해당 어가가 더 이상 어업에 종사하지 않게 되는 경우, 다시 말해 가구는 그대로 있는데 조사대상 어가로 합당하지 않게 그 속성이 변하는 경우에는 대체시키지 말고 표본에서 제거하면 된다. 이런 식으로 표본을 관리해가면 자연스럽게 시간이 경과함에 따라 표본어가의 수는 감소하게 될 것이다.

5.3 가중값(승수) 관리

새로운 표본설계에 의한 가중값은 2000년 12월 1일 기준의 어업총조사 자료를 기초로 하여 작성된 것이다. 표본조사가 매월 지속되는 과정에서 표본어가의 소멸, 모집단의 변동, 무응답 등의 상황이 생길 때 그 때마다 적절히 가중값을 변경해주어야 한다. 가장 좋은 것은 본 통계의 작성 담당자가 표본이론을 이해하고 새로운 상황이 발생할 때마다 수시로 거기에 맞도록 가중값을 변경해가도록 하는 것일 텐데 현실적인 여건상 그러기는 어려울 것으로 생각된다. 하지만 최소한 1년에 한번 정도는 가중값을 보정하여 불필요한 편향을 제거시키는 작업이 필요하다.

가중값을 보정하기 위해서는 먼저 매년 발표되는 어업기본통계조사를 이용하여 2000년 어업총조사 이후의 모집단 변동상황을 파악해야 한다. 또 다른 측면으로는 표본설계 시점 이후의 표본어가의 변동사항을 파악해야 한다. 이 결과를 종합하여 계수조정법(calibration method)을 활용하면 효과적인 가중값 조정 방안을 마련할

수 있을 것이다.

5.4기타

위에서 언급한 관리 방안 이외에도 조사의 신뢰도에 영향을 미치는 부분들이 많은데 무엇보다 중요한 것은 응답의 정확성을 높이기 위해 노력하는 일일 것이다. 이 부분은 통계적인 분야라기 보다는 조사방법론적 분야로 볼 수 있으므로 본 연구의 범위에 포함되는 부분은 아니라고 생각되지만 그래도 매우 중요한 부분이므로 간단하게나마 짚어본다.

비계통생산량조사는 기본적으로 응답자들이 매번 생산, 판매한 비계통 어업생산량을 직접 조사표에 기재해나가야 하는 조사이므로 응답자들의 적극적인 협조가 필수적인 조사이다. 그러므로 조사 관리자의 입장에서는 응답자들이 빠뜨림없이 성실하게 기재할 수 있도록 동기를 부여하고 도움을 주는 방안들을 개발해가야 한다. 현재까지의 사용하고 있는 조사방법의 문제점을 파악하고 그에 대한 개선안을 찾는 노력은 항시적으로 이루어져야 한다.

제 6 장.요약 및 제언

본 연구는 일반해면어업 어업생산통계의 오차율을 감소시키기 위해 기존의 비계통 표본가구 외에 추가적인 표본을 설계하는 것을 목적으로 하여 이루어졌다. 이를 위해 먼저 기존의 표본설계가 가지고 있는 장, 단점을 면밀히 검토한 후 단점을 효과적으로 보완할 수 있는 방안을 마련하였다. 본 연구의 중요한 결과를 요약하면 다음과 같다.

첫째, 기존의 표본어가인 1,050 가구에 추가적으로 950 가구의 표본어를 추출하여 전체 표본의 크기가 2,000 가구가 되게 하였다.

둘째, 기존의 표본설계와는 달리 1차 추출단위 (primary sampling unit:psu)를 어업조사구가 아닌 인구주택조사구로 정하였다. 그리하여 총 500 개의 psu를 추출한 후, 각 psu에서 4가구의 어가를 추출하는 층화 2단 집락추출법을 사용하였다. 아울러 추가된 표본설계를 반영하여 전체 표본어의 가중값(weight)을 새롭게 변경하였다.

셋째, 층화변수로는 지역과 어법을 고려하였다. 과거에는 어종별로 조사구를 층화한 후 어종별로 적절하게 표본을 배정하였는데 현재 규모의 표본수로는 이런 배정이 비효율적이라는 것을 알아내었다. 그래서 어종보다는 가장 많은 어가들이 이용하는 주요 어법들을 층화변수로 고려하였다. 어법은 지역과 밀접한 관련을 가지는 것으로 나타났기 때문에 지역 특성에 따라 층화변수로 사용하는 어법은 서로 다르다.

넷째, 표본배분에 있어서는 모든 어종을 동일하게 고려하는 것보다 상대적으로 생산량이 많고 중요한 어종이라고 여겨지는 어종 통계량의 효율을 높일 수 있도록 하는 배분법을 사용하였다.

다섯째, 표본의 확대에 의한 변경 부분을 추정식에 반영하였다. 추정식은 따로 프로그래밍을 하지 않아도 SAS라는 통계 분석용 소프트웨어를 활용하여 손쉽게 계산할 수 있도록 하기 위해 가중값과 조사값만 있으면 계산이 가능하도록 하는 추정식을 사용하였다.

여섯째, 본 연구의 결과로 어종별 생산량 통계의 기대 표준오차가 과거의 오차에 비해 약 30% 내외로 감소할 것으로 계산되었다. 과거에는 비교적 신뢰도가 높은 어종 통계의 수가 불과 10개 내외였는데 비해 새로운 표본설계의 결과로 40 내지 50 어종으로 늘어났다.

일곱째, 조사의 지속적인 관리를 위한 사후관리 지침을 제시하였다. 비계통표본조

사는 매월 정기적으로 조사가 이루어지는 일종의 패널조사이므로 한번의 설계로 끝나서는 안되며 수시로 추출틀이나 표본을 관리해가야 하는데 이를 위해 필요한 지침을 소개하였다.

다음으로는 향후 어업생산량통계 조사가 계속적으로 높은 품질의 조사로 유지, 발전하도록 하기 위해 중요하다고 생각되는 몇 가지 사항을 제안하고자 한다.

첫째, 본 표본확대 연구 이후에는 통계분석의 단계에서 전문성을 제고하여 올바르고 수준높은 통계를 생산하는 일에 각별한 관심을 기울이는 것이 필요하다. 기존의 어업생산통계 보고자료를 보면 한결같이 표본오차를 무시하고 보고서가 작성되었다. 다시 말해 통계적으로 문제가 있는 통계들을 과다하게 생산해온 것이다. 본 연구의 결과 어업생산통계가 일정 부분 개선될 것으로 기대되는데 앞으로는 통계적으로 면밀한 검토를 거쳐 전문가 입장에서 볼 때 결함이 없는 통계를 생산하는데 노력을 기울여야 할 것이다. 이를 위해서는 통계전문인력을 확보하는 것이 절실한 필요이다.

둘째, 110여 가지의 모든 어종에 대해 신뢰성있는 통계를 생산하는 것은 현실적으로 불가능한 일이다. 예산의 측면에서도 그렇지만 경제성 측면에서도 굳이 그럴 필요가 없다는 생각이 든다. 따라서 향후 비계통어업생산량이 총 어업생산량에서 차지하는 비중이 상대적으로 높고, 아울러 어가경제에 상대적으로 큰 영향을 주는 어종들을 선별하여 그 어종에 대해서만 통계를 작성하는 방향으로 나가는 것이 바람직하다고 여겨진다. 즉, 통계생산어종을 추려서 그 어종 통계의 정확도가 높아지도록 조사를 이끌어가는 것이 바람직할 것이다.

셋째, 본 조사는 매월 정기적으로 조사가 반복되는 계속조사이므로 조사에서 가장 중요한 점은 응답자들의 협력을 높이도록 노력하는 것이다. 응답의 부실로 인한 비표본오차를 최소화시키기 위해 다양한 방법을 개발하고 연구하는 노력을 기울여야 한다.

넷째, 어업총조사에서 어업생산량통계를 위해 필요한 정보들이 충분히 얻어질 수 있도록 해야 한다. 가령 2000년 어업총조사에서 만든 어업조사구는 여러 어업 관련 통계를 생산하기에 비효과적인 형태로 만들어졌다. 어업총조사 결과가 표본설계를 위한 추출틀로 사용되는데 이것이 비효율적인 상황에서는 표본설계를 아무리 잘 한다고 해도 근본적인 한계를 지니게 되는 것이다. 따라서 어업총조사의 기획단계에 어업생산량 통계를 위해 필요한 사항들을 충분히 감안하여 반영하도록 해야 한다.

참 고 문 헌

- Biemer, P. P. and Lyberg, L. E. (2003), <Introduction to Survey Quality>, Wiley.
- Cochran, W. G. (1977), <Sampling Techniques>, Wiley.
- Sarndal, C. E., Swensson, B. and Wretman, J (1992), <Model Assisted Survey Sampling>, Springer.
- 박진우(1997), “수산물 비계통생산량조사를 위한 표본설계연구”, 「한국통계학회 논문집」, 한국통계학회.
- 박진우, 정동명, 조영일 (2002), “어업기본통계의 Calibration 추정에 대한 연구”, 「통계분석연구」, 통계청.
- 박홍래(2000), 「통계조사론」, 영지문화사.
- 염준근, 이해용, 한근식, 이기성 (2002). “비계통어업생산통계조사 표본설계”, 「통계분석연구」, 통계청.
- 최성애, 장학봉 (2003). 「어가경제통계 개선에 관한 연구」, 한국해양수산개발원.
- 통계청 (2001), 「2000 어업총조사 분석보고서」.
- 통계청(2000), 「어가경제통계」.
- 한국통계학회 (2002), 「비계통어업생산조사 표본설계 학술연구용역 최종보고서」.
- 한국통계학회(2002), 「농가경제조사 표본설계 최종보고서」.
- 한국통계학회(2002), 「어가경제조사 표본설계」.
- 해양수산부 (2003), 「어업생산통계개선에 관한 연구용역 최종보고서」.

< 부 록 >

[부록 1] 표본 조사구 목록

[부록 2] 표본어가 추출방법

[부록 1] 표본 조사구 목록

부산광역시 (표본 조사구수:19, 표본 어가수:80)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
2104066034	4	추가	897.750		2104066035
2109054093	4	추가	897.750		2109054094
2110056069	4	추가	897.750		2110056070
2110062030	4	추가	897.750		2110062031
2110062067	4	추가	897.750		2110062069
2110062078	4	추가	897.750		2110062079
2110062226	4	추가	897.750		2110062225
2112054028	4	추가	897.750		2112054029
2112056021	4	추가	897.750		2112056033
2112057001	4	추가	897.750		2112057003
2112057006	8	추가	448.875		2112057007
2112057015	4	추가	897.750		2112057016
2114060045	4	추가	897.750		2114060050
2131011174	4	추가	897.750		2131011175
2131011189	4	추가	897.750		2131011190
2131011196	4	추가	897.750		2131011195
2131012058	4	추가	897.750		2131012059
2131031004	4	추가	897.750		2131031026
2131031034	4	추가	897.750		2131031037

인천광역시 (표본 조사구수:18, 표본 어가수:88)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
2301063001	4	추가	758.750		2301063008
2301063004	4	추가	758.750		2301063005
2301063014	8	현행	379.375		
2301063015	4	현행	758.750		
2301063016	8	현행	379.375		
2331033024	4	현행	758.750		
2331033029	4	현행	758.750		
2331033032	4	현행	758.750		
2331034018	4	추가	758.750		2331034014
2331041015	4	추가(도서)	758.750	삼산면 강화도	
2332031002	4	현행	758.750		
2332031003	4	현행	758.750		
2332035001	4	현행	758.750		
2332035003	4	현행	758.750		
2332035004	4	현행	758.750		
2332035005	8	현행	379.375		
2332036xxx	4	추가(도서)	758.750	자월면 대이작도	
2332037014	8	추가	379.375		2332037016

울산광역시 (표본 조사구수:6, 표본 어가수:32)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
2604054028	4	현행	291.250		
2604054029	4	현행	291.250		
2604054030	8	현행	145.625		
2631031013	8	추가	145.625		2631031040
2631031041	4	현행	291.250		
2631031042	4	현행	291.250		

경기도 (표본 조사구수:12, 표본 어가수:60)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3109072007	4	추가	451.750		3109072006
3109072019	8	추가	225.875		3109072023
3109072024	4	추가	451.750		3109072026
3115061322	4	현행	451.750		
3133037004	8	추가	225.875		3133037003
3133037036	4	현행	451.750		
3133040001	8	추가	225.875		3133040015
3133040074	4	현행	451.750		
3133040076	4	현행	451.750		
3133040077	4	현행	451.750		
3133040091	4	추가	451.750		3133040089
3133040092	4	추가	451.750		3133040095

강원도 (표본 조사구수:28, 표본 어가수:128)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3203011058	4	추가	1203.500		3203011056
3203011060	4	추가	1203.500		3203011065
3203011082	4	추가	1203.500		3203011083
3203034032	4	현행	1203.500		
3203034035	4	현행	1203.500		
3203035001	4	추가	1203.500		3203035002
3203037003	4	추가	1203.500		3203037006
3204059010	4	추가	1203.500		3204059014
3204059033	4	현행	1203.500		
3204059034	4	현행	1203.500		
3206051002	4	현행	1203.500		
3206056070	4	추가	1203.500		3206056069
3206059006	4	추가	1203.500		3206059008
3206060005	4	추가	1203.500		3206060010
3207012004	4	추가	1203.500		3207012003
3207012013	4	추가	1203.500		3207012011
3207012014	4	추가	1203.500		3207012022
3207031031	4	추가	1203.500		3207031032
3207031044	4	추가	1203.500		3207031043
3240012047	8	추가	601.750		3240012045
3240012049	4	추가	1203.500		3240012050
3240012053	4	추가	1203.500		3240012055
3240031005	8	추가	601.750		3240031007
3240031017	4	추가	1203.500		3240031015
3240032003	4	추가	1203.500		3240032006
3240033020	8	추가	601.750		3240033018
3241032008	8	현행	601.750		
3241034013	4	추가	1203.500		3241034017

충청남도 (표본 조사구수:61, 표본 어가수:280)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3403011006	4	현행	2361.000		
3403011007	4	현행	2361.000		
3403011008	4	현행	2361.000		
3403011009	4	현행	2361.000		
3403011010	4	현행	2361.000		
3403032007	8	현행	1180.500		
3403032008	4	현행	2361.000		
3403032009	4	현행	2361.000		
3403033xxx	4	추가(도서)	2361.000	오천면 삼시도	
3403033xxx	4	추가(도서)	2361.000	오천면 녹도	
3403033023	4	현행	2361.000		
3403033026	4	현행	2361.000		
3403033031	4	현행	2361.000		
3403033032	4	현행	2361.000		
3403034004	4	현행	2361.000		
3403034009	4	현행	2361.000		
3403037013	4	현행	2361.000		
3403037014	4	현행	2361.000		
3403056008	4	현행	2361.000		
3405011006	4	현행	2361.000		
3405011008	4	현행	2361.000		
3405011046	4	현행	2361.000		
3405032037	8	현행	1180.500		
3405033001	4	현행	2361.000		
3405033003	4	현행	2361.000		
3405034012	4	현행	2361.000		
3405034014	4	현행	2361.000		
3405034015	4	현행	2361.000		
3405034016	4	현행	2361.000		
3434031002	8	현행	1180.500		
3434031004	4	현행	2361.000		
3434031012	4	현행	2361.000		
3434031013	4	현행	2361.000		
3434031020	4	현행	2361.000		
3434040015	8	현행	1180.500		
3434040025	4	현행	2361.000		
3434041011	4	추가	2361.000		3434041010
3434041022	4	추가	2361.000		3434041023
3436037001	8	현행	1180.500		
3436037012	4	추가	2361.000		3436037011
3436037022	8	현행	1180.500		

(충청남도 계속)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3438012004	4	추가	2361.000		3438012005
3438012010	4	추가	2361.000		3438012008
3438031003	4	현행	2361.000		
3438031006	4	현행	2361.000		
3438031008	8	현행	1180.500		
3438032003	8	추가	1180.500		3438032007
3438033010	4	현행	2361.000		
3438033021	4	현행	2361.000		
3438033025	4	현행	2361.000		
3438033033	4	현행	2361.000		
3438034011	8	현행	1180.500		
3438034018	4	현행	2361.000		
3438034026	4	현행	2361.000		
3438036002	4	현행	2361.000		
3438036014	4	현행	2361.000		
3439032027	4	현행	2361.000		
3439032036	4	현행	2361.000		
3439032037	4	현행	2361.000		
3439038064	4	추가	2361.000		3439038065
3439039008	4	현행	2361.000		

전라북도 (표본 조사구수:24, 표본 어가수:112)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3502039006	8	추가	534.375		3502039005
3502039013	4	현행	1068.750		
3502039014	8	추가	534.375		3502039019
3502040005	4	현행	1068.750		
3502051005	4	현행	1068.750		
3502072011	4	현행	1068.750		
3502073090	4	현행	1068.750		
3506031003	4	현행	1068.750		
3506039002	4	현행	1068.750		
3506039009	4	현행	1068.750		
3537035003	4	현행	1068.750		
3537035021	4	현행	1068.750		
3537036001	4	현행	1068.750		
3537039017	4	현행	1068.750		
3537039020	4	현행	1068.750		
3538034010	4	현행	1068.750		
3538034020	4	현행	1068.750		
3538034022	4	현행	1068.750		
3538034024	4	현행	1068.750		
3538036031	4	현행	1068.750		
3538040001	8	현행	534.375		
3538040013	8	현행	534.375		
3538042003	4	추가(도서)	1068.750	부안군 위도	
3538042010	4	추가(도서)	1068.750	부안군 위도	

전라남도 (표본 조사구수:90, 표본 어가수:500)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3601066018	4	현행	6734.000		
3602011024	4	현행	6734.000		
3602011033	4	현행	6734.000		
3602011035	4	현행	6734.000		
3602011053	4	현행	6734.000		
3602011054	4	현행	6734.000		
3602011056	4	현행	6734.000		
3602032030	8	추가	3367.000		3602032028
3602033041	4	현행	6734.000		
3602033046	4	현행	6734.000		
3602033048	4	현행	6734.000		
3602033049	8	현행	3367.000		
3602035xxx	4	추가(도서)	6734.000	화정면 백야도	
3602035xxx	4	추가(도서)	6734.000	화정면 여자도	
3602057001	4	현행	6734.000		
3602064001	4	현행	6734.000		
3602066112	4	현행	6734.000		
3602066115	4	현행	6734.000		
3602066116	4	현행	6734.000		
3602069035	4	현행	6734.000		
3602069037	4	현행	6734.000		
3603035042	8	현행	3367.000		
3603035043	8	현행	3367.000		
3603035044	8	현행	3367.000		
3603037109	4	현행	6734.000		
3635033xxx	4	추가(도서)	6734.000	금산면 거금도	
3635033xxx	4	추가(도서)	6734.000	금산면 거금도	
3635038024	8	현행	3367.000		
3635039001	8	현행	3367.000		
3635039002	8	현행	3367.000		
3635041014	8	현행	3367.000		
3635043023	4	현행	6734.000		
3635043024	4	현행	6734.000		
3635044001	8	현행	3367.000		
3635044002	8	현행	3367.000		
3635044003	8	현행	3367.000		
3635044007	8	현행	3367.000		
3635044026	8	현행	3367.000		
3636012044	8	현행	3367.000		
3636039016	4	현행	6734.000		
3638012033	4	현행	6734.000		
3638012034	8	현행	3367.000		

(전라남도 계속)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3638012038	8	현행	3367.000		
3638013026	8	현행	3367.000		
3638013031	4	현행	6734.000		
3638031021	4	현행	6734.000		
3638032018	8	현행	3367.000		
3638032025	8	현행	3367.000		
3638037022	8	현행	3367.000		
3639032002	8	현행	3367.000		
3639032013	4	현행	6734.000		
3639033001	8	현행	3367.000		
3640032031	8	추가	3367.000		3240032030
3640035004	4	현행	6734.000		
3640035016	8	현행	3367.000		
3640036019	4	현행	6734.000		
3640036020	8	현행	3367.000		
3640040011	4	현행	6734.000		
3642031001	4	현행	6734.000		
3642033032	4	현행	6734.000		
3642033034	4	현행	6734.000		
3642034031	4	현행	6734.000		
3642034032	8	현행	3367.000		
3642034039	4	현행	6734.000		
3642035013	4	현행	6734.000		
3642035017	4	현행	6734.000		
3643011002	8	현행	3367.000		
3644012043	4	현행	6734.000		
3644012045	8	현행	3367.000		
3644036xxx	4	추가(도서)	6734.000	염산면 향화도	
3644036001	8	현행	3367.000		
3644036003	8	현행	3367.000		
3646011030	4	현행	6734.000		
3646012001	4	현행	6734.000		
3646013001	8	현행	3367.000		
3646013006	4	현행	6734.000		
3646013023	4	현행	6734.000		
3646032001	4	현행	6734.000		
3646033004	4	현행	6734.000		
3646033007	4	현행	6734.000		
3646033022	4	현행	6734.000		
3646033035	8	현행	3367.000		
3646034003	4	현행	6734.000		
3646038013	4	추가	6734.000		3646038014
3647011003	4	현행	6734.000		

(전라남도 계속)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3647033034	4	현행	6734.000		
3647033035	8	현행	3367.000		
3647036xxx	4	추가(도서)	6734.000	조도면 조도	
3648036xxx	8	추가(도서)	3367.000	흑산면 흑산도	
3648043xxx	4	추가(도서)	6734.000	압해면 압해도	

경상북도 (표본 조사구수:32, 표본 어가수:140)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3701111019	4	추가	1444.500		3701111016
3701111032	4	추가	1444.500		3701111031
3701111076	4	현행	1444.500		
3701111077	4	현행	1444.500		
3701132079	8	현행	722.250		
3701133034	8	추가	722.250		3701133041
3701133038	4	현행	1444.500		
3701134005	4	현행	1444.500		
3701134006	4	현행	1444.500		
3701134007	4	현행	1444.500		
3701134011	4	현행	1444.500		
3701134017	4	현행	1444.500		
3701232034	4	추가	1444.500		3201232035
3701264001	4	추가	1444.500		3701264002
3702011007	4	현행	1444.500		
3702011008	4	현행	1444.500		
3702011043	4	추가	1444.500		3702011044
3702032025	4	추가	1444.500		3702032024
3735011070	4	추가	1444.500		3735011071
3735031039	4	추가	1444.500		3735031043
3735032025	4	추가	1444.500		3235032021
3735035017	4	추가	1444.500		3735035015
3735036052	4	현행	1444.500		
3735036053	4	현행	1444.500		
3742011022	4	현행	1444.500		
3742011023	4	현행	1444.500		
3742035025	4	추가	1444.500		3742035028
3742037046	4	추가	1444.500		3742037045
3742038024	4	추가	1444.500		3742038009
3743011008	4	추가(도서)	1444.500	울릉읍 울릉도	
3743011019	8	추가(도서)	722.250	울릉읍 울릉도	
3743031003	4	추가(도서)	1444.500	울릉읍 울릉도	3743031002

경상남도 (표본 조사구수:84, 표본 어가수:420)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3802131xxx	4	추가(도서)	3502.250	구산면 실리도	
3802131008	8	현행	1751.125		
3802131012	8	현행	1751.125		
3802131013	4	현행	3502.250		
3802131021	4	추가	3502.250		3802131020
3802131028	4	현행	3502.250		
3802131032	8	현행	1751.125		
3802134032	4	추가	3502.250		3802134034
3802151014	8	현행	1751.125		
3804062054	4	현행	3502.250		
3804062055	8	현행	1751.125		
3804064021	4	현행	3502.250		
3804064024	4	현행	3502.250		
3804066006	4	추가	3502.250		3804066029
3804066036	4	추가	3502.250		3804066037
3805011020	4	추가	3502.250		3805011021
3805011033	4	추가	3502.250		3805011027
3805011034	4	현행	3502.250		
3805031005	4	현행	3502.250		
3805031006	4	현행	3502.250		
3805031013	4	현행	3502.250		
3805031014	4	현행	3502.250		
3805031019	4	현행	3502.250		
3805031023	4	현행	3502.250		
3805032022	4	현행	3502.250		
3805033013	4	현행	3502.250		
3805033014	4	현행	3502.250		
3805033015	4	현행	3502.250		
3805034003	8	추가(도서)	1751.125	육지면 육지도	
3805034016	4	추가(도서)	3502.250	육지면 육지도	3805034019
3805035005	8	추가(도서)	1751.125	한산면 한산도	
3805035024	4	추가(도서)	3502.250	한산면 한산도	
3805036003	4	추가(도서)	3502.250	사랑면 사랑도	
3805036011	4	추가(도서)	3502.250	사랑면 사랑도	
3805057033	8	현행	1751.125		
3806032016	4	추가	3502.250		3806032015
3806033001	4	현행	3502.250		
3806033008	4	현행	3502.250		
3806033014	4	현행	3502.250		
3806033023	4	추가	3502.250		3806033022
3806035028	4	현행	3502.250		
3806037010	4	추가	3502.250		3806037003

(경상남도 계속)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3806037030	8	추가	1751.125		3806037031
3806059001	8	현행	1751.125		
3806059002	8	현행	1751.125		
3809031011	4	추가	3502.250		3809031028
3809033001	4	추가	3502.250		3809033006
3809033007	4	추가	3502.250		3809033008
3809034xxx	4	추가(도서)	3502.250	거제면 산달도	
3809034043	4	현행	3502.250		
3809034044	4	현행	3502.250		
3809035017	4	추가	3502.250		3809035019
3809036xxx	4	추가(도서)	3502.250	사동면 가조도	
3809036023	8	현행	1751.125		
3809036038	4	현행	3502.250		
3809036040	4	현행	3502.250		
3809038006	4	현행	3502.250		
3809038007	8	현행	1751.125		
3809039008	8	현행	1751.125		
3809039009	8	현행	1751.125		
3809039010	4	현행	3502.250		
3809039028	4	추가	3502.250		3809039027
3834011135	4	현행	3502.250		
3834011136	4	현행	3502.250		
3834032014	8	추가	1751.125		3834032016
3834042019	8	현행	1751.125		
3834042025	4	추가	3502.250		3834042026
3834043028	4	현행	3502.250		
3835011009	4	현행	3502.250		
3835011010	4	현행	3502.250		
3835032015	4	추가	3502.250		3835032016
3835033029	4	추가	3502.250		3835033030
3835034011	4	추가	3502.250		3835034013
3835035035	4	추가	3502.250		3835035032
3835036018	4	현행	3502.250		
3835036019	4	현행	3502.250		
3835037004	8	추가	1751.125		3835037003
3835038013	4	현행	3502.250		
3835039005	4	추가	3502.250		3835039006
3835039034	4	추가	3502.250		3835039036
3836036017	8	추가	1751.125		3836036013
3836036028	8	현행	1751.125		
3836037014	4	추가	3502.250		3836037013
3836038037	8	현행	1751.125		

제주도 (표본 조사구수:35, 표본 어가수:160)

표본조사구번호	표본어가수	유형	가중값	비고	예비표본조사구
3901058011	4	현행	1678.750		
3901068001	4	현행	1678.750		
3901068002	4	현행	1678.750		
3902051007	4	현행	1678.750		
3902051009	4	현행	1678.750		
3902055024	4	추가	1678.750		3902055025
3902061020	4	추가	1678.750		3902061021
3931011011	8	추가	839.375		3931011013
3931011085	4	추가	1678.750		3931011084
3931011105	4	추가	1678.750		3931011098
3931012029	4	추가	1678.750		3931012040
3931013001	4	추가	1678.750		3931013006
3931013014	4	추가	1678.750		3931013013
3931013021	8	추가	839.375		3931013024
3931013033	4	추가	1678.750		3931013032
3931013061	4	추가	1678.750		3931013060
3931013068	4	추가	1678.750		3931013066
3931014043	4	추가	1678.750		3931014044
3931014088	4	추가	1678.750		3931014087
3931031007	4	추가	1678.750		3931031006
3931031023	4	추가	1678.750		3931031021
3931033004	8	추가(도서)	839.375	우도면 우도	
3931033007	8	추가(도서)	839.375	우도면 우도	
3932011049	4	추가	1678.750		3932011048
3932011089	8	추가	839.375		3932011088
3932013002	4	현행	1678.750		
3932013003	4	현행	1678.750		
3932013026	4	추가	1678.750		3932013025
3932013060	4	현행	1678.750		
3932013061	4	현행	1678.750		
3932013063	4	현행	1678.750		
3932013064	4	현행	1678.750		
3932013075	4	추가	1678.750		3932013072
3932031008	4	현행	1678.750		
3932032015	4	추가	1678.750		3932032014

[부록 2] 표본어가 추출방법

표본으로 선정된 표본 인구주택조사구에서 다음과 같은 방법으로 표본 어가를 선정한다. 구체적인 적용방법은 <예시>를 참고하기 바란다.

(1) 특성에 따라 정렬된 어가 명부의 작성

: 각 표본 인구주택조사구에서 다음 정렬기준에 따라 조사구내 어가를 순서대로 정렬한 명부를 작성한다.

<어가명부 정렬 기준>

어가 특성 구분	특성별 정렬 기준	비고
어선 사용 어가	어선보유톤수 큰 어가부터 순서에 따라 정렬	
비어선 어가	지난 1개월 동안 총 작업 일수에 많은 어가부터 순서에 따라 정렬	맨손어업 등
양식 어가	양식면적이 큰 어가부터 순서에 따라 정렬	

* 여기서 어가 특성 구분은 어가의 수입이 가장 많은 어업형태에 따라 구분

(2) 정렬된 어가 명부에서 계통추출에 의한 어가 선정

: <어가명부 정렬 기준>에 따라 작성된 조사구 어가 명부에서 각 조사구에 할당된 표본 어가 4가구(또는 8가구)를 계통추출한다. 단, 계통추출로 선정된 어가의 조사가 현실적으로 불가능한 경우 명부상 가장 인접해 있는 어가로 대체한다.

(3) 현행 표본조사구에서 추가 표본어의 선정

: 일단 (2)의 방법에 따라 계통추출로 4가구(또는 8가구)를 추출하고, 이들 추출된 표본어를 목록에서 가장 가까운 위치에 있는 현행 표본어로 대체하는 방식으로 최종 표본어를 선정한다. 따라서 현재 2개 표본어가 있는 조사구의 경우 새로 선정된 4개 어가 중 2개 어가는 현행 표본어로 대체하게 되고 나머지 2개 어가는 새로운 어가를 표본으로 사용하게 된다.

<표본어가 추출 예시>

이해를 돕기 위해 23개 어가로 구성된 조사구에서 4개 어가를 표본으로 선정하는 사례를 제시하면 다음과 같다.

[단계 1] 어가명부의 작성

23개 어가 중 10어가는 “어선사용어가”, 6어가는 “비어선어가”, 나머지 7어가는 양식어가라고 하면, 어가명부 정렬기준에 따라 아래 <예시 표>와 같이 23개 어가가 정렬된다.

[단계 2] 계통추출방법

23어가 중 4어를 추출하는 계통추출을 위한 추출간격은 $23/4 = 5.75$ 로 계산되지만, 편의상 소수점 아래는 무조건 버리고 추출간격을 5로 한다(따라서 실제적으로 23어가 중 마지막에 정렬된 3개 어가는 표본추출대상에서 제외한 것임). 계통추출을 위해 1-5사이에서 하나의 난수를 난수표를 이용해 선정한다. 예를 들어 난수로 3이 선정되었다면 4개의 표본어가는 <예시 표>와 같이 일련번호가 “3”, “8”, “13”, “18”인 어가를 표본어가로 선정한다. 신규 표본 조사구의 경우 이들 4개 어가가 최종 표본어가가 된다. 단, 이들 표본어가 중 현실적으로 조사가 불가능한 어가가 발생하는 경우 해당어의 위 또는 아래 어가로 표본어를 대체한다.

[단계 3] 현행 표본 조사구에서 현행 표본어의 선정

만약 이 조사구가 현재 표본으로 사용되고 있는 조사구로 현재 2개 어가가 표본어가로 선정되어 있다면(<예시 표>에서 일련번호가 “11”과 “14”인 어가) [단계 2]에서 선정된 4어가 중 2어를 현행 표본어가로 대체하여 사용한다. 예를 들어 아래 <예시 표>에서 현행 표본어가인 “14”는 가장 가까운 위치에 있는 신규 계통추출어가 “13” 대신, 어가 “11”은 계통추출어가 “8” 대신 표본어가로 사용한다. 따라서 최종 표본어가는 현행 표본어가인 “11”, “14” 과 추가 표본어가인 “3”, “18” 등 4개 어가로 구성된다.

참고1: 8개 조사구가 배정된 조사구의 경우 위와 동일한 요령으로 표본 어가를 선정한다. 예를 들어 45개 어가에서 8개 어가를 계통추출하는 경우 추출간격으로는 5를 사용한다. ($45/8=5.625 \Rightarrow$ 소수점 아래 무시하고 추출간격으로 “5” 사용)

참고2: 현행 표본 조사구 중 5개 표본어가가 선정되어 있는 데, 새로운 표본에서 4개 어가만 배정된 경우, 제시된 계통추출방법에 의한 표본과 가능한 유사한 형식의 표본이 선정될 수 있도록 현행 표본 어가 중 1개 어가를 표본에서 제외시킨다.

<예시 표> 조사구에서 표본어가 추출 예시

	어가 일련번호	신규 계통추출 표본어가	현행 표본 조사구인 경우	
			현행 표본 어가	최종 표본어가
어선사용어가 (어선톤수가 큰 어가부 터 순서에 따라 정렬)	1			
	2			
	3	표본어가-1		표본어가-1
	4			
	5			
	6			
	7			
	8	표본어가-2		
	9			
	10			
비어선어가 (지난 1개월 동안 작업 일수에 많은 어가부터 순서에 따라 정렬)	11		현행표본-1	표본어가-2
	12			
	13	표본어가-3		
	14		현행표본-2	표본어가-3
	15			
	16			
양식어가 (양식면적이 큰 어가부 터 순서에 따라 정렬)	17			
	18	표본어가-4		표본어가-4
	19			
	20			
	21			
	22			
	23			