

연구노트

산업·직업별 고용구조 분석 및 표본설계

Analysis and Sampling Design for Occupational Employment Statistics

류제복* · 이기성** · 손창균*** · 박상현**** · 남기성*****

Jea-Bok Ryu · Gi-Sung Lee · Chang-Kyoon Son · Sang-Hyun Park · Ki-Seong

산업·직업별 고용구조조사(OES: Occupational Employment Statistics)는 국가승인통계로 산업소분류(194개)·직업세분류(392개) 수준에서의 고용구조를 파악하여 국가적 인적자원 수급정책을 위한 기본통계와 직업별 고용전망, 진로 선택, 직업훈련, 취업알선 등 노동시장 정책과 연구를 위한 기초 자료를 제공하는데 목적이 있다. 이러한 목적을 달성하기 위하여 본 연구에서는 2005년도와 2006년도에 실시된 산업·직업별 고용구조조사 자료를 특성별로 정밀 분석하고, 10% 인구주택총조사 자료를 사용하여 우리나라의 고용현황을 살펴서 이들을 바탕으로 모집단을 충분히 대표할 수 있도록 새로운 표본설계 방안을 마련하고자 한다. 표본설계방안에서는 표본배정 기준을 마련하고 아울러 가중치 부여 과정을 포함해서 추정식과 추정식의 오차계산 공식을 유도한다. 새로운 표본설계를 통해서 산업·직업별 고용구조를 보다 정확히 파악하여 국가적 인력수급 정책을 원활히 세울 수 있게 되고, 또한 고용관련 연구에도 기여를 할 것으로 기대된다.

주제어: 산업·직업별 고용구조조사, 산업별분석, 직업별분석, 표본설계

OES survey as the national official statistics aims to provide the basic data for the national labor market policy and research such as the basic statistics for human resource supply policy, the prediction of employment by occupations, the decision of occupation, the occupational training and the finding jobs et al., at the levels of industrial and occupational classifications(3-digit).

In order to achieve this objective, we analyze the OES data in 2005 and 2006

* 청주대학교 생명·유전·통계학부 통계학전공 교수.

** 교신저자(corresponding author) : 우석대학교 아동복지학과 교수 이기성.
E-mail: gisung@woosuk.ac.kr

*** 한국보건사회연구원 부연구위원.

**** 한국고용정보원 고용조사팀 팀장.

***** 한국고용정보원 고용조사팀 연구위원.

and propose the new sampling design using the long form data in Korea (10% sample data of census 2005).

In this paper, we provide the criterion of sample allocation and derive the formular for estimator and error of it including the weighting procedure. From the proposed sampling design, we would expect that it contributes to the supply policy of human resource and the research for labor market.

Key words: Occupational Employment Statistics, Industrial Analysis, Occupational Analysis, Sampling Design.

I. 서론

『산업·직업별 고용구조조사』(OES: Occupational Employment Statistics)는 국가승인통계(승인번호: 32701)로, 산업소분류(194개)·직업세분류(392개) 수준에서의 고용구조를 파악하여 국가적 인적자원 수급정책을 위한 기본통계와 직업별 고용전망, 진로선택, 직업훈련, 취업알선 등 노동시장 정책과 연구를 위한 기초자료를 제공하는 데 목적이 있다. 또한 이 조사는 부가조사인 청년패널조사의 모집단으로 사용된다. 『산업·직업별 고용구조조사』는 2001년부터 매년 조사가 실시되고 있는 가구단위 조사로, 조사내용은 종사 산업 및 직업, 임금, 성별, 학력수준, 전공, 근로형태, 고용형태, 근속년수, 경력년수, 근로시간, 이전 산업 및 직업 등이다. 고용관련 조사는 크게 사업체단위 조사와 가구단위 조사가 있다. 사업체단위의 조사로는 노동부에서 생산하고 있는 매월노동통계조사, 임금구조통계조사, 사업체근로실태조사 등이 있으며, 가구단위의 조사로는 통계청에서 생산하고 있는 인구주택총조사(10% 표본조사)와 경제활동인구조사(1963년부터 실시), 그리고 인력실태조사(2006년부터 실시) 등이 있다. 사업체단위 조사는 사업체에 취업해 있는 사람을 대상으로 조사함으로써 취업자, 특히 임금근로자의 규모는 비교적 정확히 파악할 수 있지만, 자기사업체를 운영하는 자영업자 등은 파악할 수 없는 단점이 있다. 반면 가구단위 조사는 일상적인 생활을 하는 가구를 방문하여 각 가구의 구성원을 대상으로 조사하기 때문에 사업체에 취업한 사람뿐만 아니라 주된 사업장이 없는

사람까지도 파악할 수 있다. 따라서 사업체단위 조사와 가구단위 조사는 상호 보완적인 관계에 있지만, 경제 전체의 고용상태를 파악하는 데 있어서는 가구단위 조사가 중요한 의의를 갖는다고 할 수 있다.

2005년도 『산업 · 직업별 고용구조조사』(한국고용정보원 2006) 자료에 의하면 비임금근로자가 전체의 41%를 차지하고 있으며, 1년 간 이직률이 전체 취업자의 5.5%이고, 특히 20대(10.3%)와 30대(7.3%)의 이직률이 높다. 그리고 산업 · 직업별로도 이직률에 차이가 많을 뿐만 아니라 산업 · 직업별 이동도 크므로 매년 조사가 수행되어야 산업 · 직업별 상세한 고용통계를 생산할 수 있으며 이를 기초로 정확한 국가 고용정책을 수립할 수 있을 것이다.

현행 표본설계에서 표본배분은 읍 · 면부에 비해 등부에 표본을 과다 배분하였는데, 그 근거가 명확치 않다. 이는 표본설계 당시 고용과 관련된 자료가 포함되어 있는 10% 인구주택총조사 자료의 활용이 불가능하였기 때문으로 보이며, 또한 추정과정도 불명확하다. 따라서 금번 새로운 표본설계에서는 2005년도와 2006년도에 실시된 조사자료를 특성별로 정밀 분석하고, 10% 인구주택총조사 자료를 사용하여 우리나라의 고용현황을 살펴서 이들을 바탕으로 모집단을 충분히 대표할 수 있도록 표본배정 기준을 마련한다. 아울러 가중치 부여 과정을 포함해서 추정식과 추정식의 오차계산 공식을 유도한다. 이런 과정을 거쳐 새로운 표본설계를 통해서 산업 · 직업별 고용구조를 보다 정확히 파악하여 국가적 인력수급 정책을 원활히 세울 수 있게 되고, 또한 고용관련 연구에도 기여를 할 것으로 기대된다.

II. 현행 산업 · 직업별 고용구조 결과 분석

1. 현행 표본설계의 개요

2006년도 표본설계(대한통계협회 2006)에서는 모집단을 대표할 수 있는 가장 최근의 자료인 2005년 인구주택총조사 자료 중 10% 표본조사자료를 제외한 90% 전수조사자료를 표본추출틀로 사용하였다. 그 중 실제 조사가 매우 곤란한 섬 조사구와 일반적 특성을 갖지 못하는 기숙 및 특수사회시설 조사구를

제외한 237,680개 조사구를 표본추출틀로 사용하였다. 한편 층화과정은 7개 특·광역시와 9개 광역자치단체로 1차 층화하였다. 각 도에서는 동부와 읍·면부에 따라서 직업 및 고용형태에 차이가 보이는 점을 감안하여 동부와 읍·면부로 2차 층화하여 최종적으로 전국을 주택특성, 교육정도, 노동생산가능 인구비율 및 행정구역번호를 분류지표로 선정하여 25개 층별로 조사구를 정리 하였고, 각 층별 표본 배분현황은 <표 1>과 같다.

<표 1> 2005년 OES와 2006년 OES 조사구 표본규모 (단위: 개)

지역	2005년 OES 표본규모			2006년 OES 표본규모		
	전체	동부	읍·면부	전체	동부	읍·면부
총계	2,500	1,994	506	3,750	3,332	418
서울	547	547	-	916	916	-
부산	182	182	0	330	330	0
대구	124	124	0	231	231	0
인천	131	131	0	236	236	0
광주	68	68	-	133	133	-
대전	71	71	-	135	135	-
울산	51	51	0	99	99	0
경기	529	424	105	895	800	95
강원	81	46	35	78	48	30
충북	77	44	33	77	48	29
충남	102	33	69	100	50	50
전북	99	61	38	95	66	29
전남	108	39	69	96	50	46
경북	143	66	77	137	72	65
경남	158	88	70	161	97	64
제주	29	19	10	31	21	10

2006년도 표본설계의 특징은 다양한 산업 및 직업에 관한 표본을 얻기 위하여 2005년도 표본규모인 2,500개 조사구는 2005년 인구주택총조사 인구에 비례하여 25개 층별로 배분하고, 나머지 1,250개 조사구는 특·광역시 및 경기도부의 인구비례로 배분한 후 합산하였다는 점이다. 이로 인해 2006년도에는 2005년도에 조사되지 않은 직업들이 조사되었으며, 특히 읍·면부에 비해 동부에서는 지역별 직업세분류 직종비율이 높게 나타나는 경향이 있었다. 하지만 읍·면부의 경우에는 2005년도와 지역별 직업세분류 직종비율이 비슷한 수준을 나타나고 있었다.

2. 특성별 분석

2005년도와 2006년도 『산업·직업별 고용구조 조사』의 원자료를 이용하여 이 조사에서 가장 중요한 변수에 해당하는 취업자수에 대하여 기본적인 분석을 실시하였다.

〈표 2〉를 보면 2005년도에 비해 2006년도에는 동부의 경우에는 16개의 지역 중 14개 지역, 읍·면부의 경우에는 13개의 지역 중 6개의 지역에서 산업소분류 업종비율이 높게 나타났다. 〈표 3〉을 보더라도 2005년도에 비해 2006년도에는 동부의 경우에는 16개의 지역 중 15개 지역, 읍·면부의 경우에는 13개의 지역 중 9개의 지역에서 직업세분류 직종비율이 높게 나타났다. 이는 2005년에 비해, 2006년도에 증가된 표본 25,000가구를 모두 동부에 배정한 결과로 보인다.

〈표 4〉는 지역별 취업자수 및 가구내 취업자 비율을 2005년과 2006년을 비교할 수 있도록 작성되었다. 2005년도에 비해 2006년도 동부의 경우에는 16개의 지역 중 13개 지역, 읍·면부의 경우에는 13개의 지역 중 6개의 지역에서 취업자 비율이 높게 나타났다.

〈표 2〉 2005년 OES, 2006년 OES 지역별 취업자수, 산업소분류 표본 업종수 분포
(단위: 명, 개, %)

지역	전체				동부				읍·면부			
	2005 OES		2006 OES		2005 OES		2006 OES		2005 OES		2006 OES	
	취업자 수	가구내 취업자 비율	취업자 수	가구내 취업자 비율	산업 소분류 업종 수	산업 소분류 업종 비율	산업 소분류 업종 수	산업 소분류 업종 비율	산업 소분류 업종 수	산업 소분류 업종 비율	산업 소분류 업종 수	산업 소분류 업종 비율
총계	70,133	53.3	100,158	52.7	194	100.0	194	100.0	194	100.0	194	100.0
서울	13,971	49.1	23,524	49.4	177	91.2	182	93.8	-	-	-	-
부산	5,274	48.5	9,118	51.7	166	85.5	167	86.1	40	20.6	52	26.8
대구	3,441	48.4	6,171	51.0	146	75.2	161	83.0	74	38.1	61	31.4
인천	3,517	52.4	5,935	49.7	160	82.4	172	88.7	40	20.6	38	19.6
광주	1,929	52.6	3,382	49.7	135	69.5	156	80.4	-	-	-	-
대전	2,123	50.8	3,740	53.1	138	71.1	157	80.9	-	-	-	-
울산	1,460	48.8	2,630	53.0	118	60.8	131	67.5	73	37.6	84	43.3
경기	13,813	51.3	23,066	51.8	174	89.6	181	93.3	146	75.2	155	79.9
강원	2,355	55.1	2,111	54.4	124	63.9	122	62.9	94	48.4	93	47.9
충북	2,395	60.5	2,267	59.4	129	66.4	133	68.6	92	47.4	101	52.1
충남	3,231	63.5	3,172	63.8	117	60.3	136	70.1	130	67.0	121	62.4
전북	2,934	58.6	2,719	59.1	133	68.5	140	72.2	81	41.7	80	41.2
전남	3,404	65.3	2,817	62.8	116	59.7	118	60.8	107	55.1	91	46.9
경북	4,341	61.1	3,914	61.4	142	73.2	142	73.2	121	62.3	119	61.3
경남	4,951	58.4	4,514	59.1	149	76.8	150	77.3	130	67.0	132	68.0
제주	994	66.8	1,078	66.9	89	45.8	91	46.9	51	26.2	60	30.9

〈표 3〉 2005년 OES, 2006년 OES 지역별 취업자수, 직업세분류 표본 직종수 분포
(단위: 명, 개, %)

지역	전체				동부				읍·면부			
	2005 OES		2006 OES		2005 OES		2006 OES		2005 OES		2006 OES	
	취업자수	가구내 취업자 비율	취업자수	가구내 취업자 비율	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류직 종수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율
총계	70,133	53.3	100,158	52.7	392	100.0	392	100.0	392	100.0	392	100.0
서울	13,971	49.1	23,524	49.4	361	92.0	374	95.4	-	-	-	-
부산	5,274	48.5	9,118	51.7	300	76.5	347	88.5	55	14.0	68	17.3
대구	3,441	48.4	6,171	51.0	282	71.9	315	80.4	97	24.7	76	19.4
인천	3,517	52.4	5,935	49.7	295	75.2	329	83.9	44	11.2	39	9.9
광주	1,929	52.6	3,382	49.7	253	64.5	293	74.7	-	-	-	-
대전	2,123	50.8	3,740	53.1	255	65.0	305	77.8	-	-	-	-
울산	1,460	48.8	2,630	53.0	197	50.2	233	59.4	92	23.4	97	24.7
경기	13,813	51.3	23,066	51.8	365	93.1	381	97.2	261	66.5	270	68.9
강원	2,355	55.1	2,111	54.4	215	54.8	197	50.3	138	35.2	142	36.2
충북	2,395	60.5	2,267	59.4	210	53.5	225	57.4	120	30.6	135	34.4
충남	3,231	63.5	3,172	63.8	193	49.2	227	57.9	190	48.4	175	44.6
전북	2,934	58.6	2,719	59.1	224	57.1	250	63.8	112	28.5	114	29.1
전남	3,404	65.3	2,817	62.8	193	49.2	211	53.8	174	44.3	153	39.0
경북	4,341	61.1	3,914	61.4	240	61.2	242	61.7	187	47.7	190	48.5
경남	4,951	58.4	4,514	59.1	251	64.0	264	67.3	191	48.7	209	53.3
제주	994	66.8	1,078	66.9	132	33.6	156	39.8	66	16.8	80	20.4

〈표 4〉 2005년 OES, 2006년 OES 지역별 취업자수, 표본가구내 취업자 비율 분포
(단위: 명, %)

지역	전체				동부				읍·면부			
	2005 OES		2006 OES		2005 OES		2006 OES		2005 OES		2006 OES	
	취업자 수	가구내 취업자 비율	취업자 수	가구내 취업자 비율	취업자 수	가구내 취업자 비율	취업자 수	가구내 취업자 비율	취업자 수	가구내 취업자 비율	취업자 수	가구내 취업자 비율
총계	70,133	53.3	100,158	52.7	52,745	50.1	86,019	51.0	17,388	66.2	14,139	66.4
서울	13,971	49.1	23,524	49.4	13,971	49.1	23,524	49.4	-	-	-	-
부산	5,274	48.5	9,118	51.7	5,151	48.3	8,980	51.7	123	57.2	138	56.3
대구	3,441	48.4	6,171	51.0	3,211	48.0	5,947	50.8	230	54.1	224	56.7
인천	3,517	52.4	5,935	49.7	3,381	51.9	5,772	49.9	136	66.7	163	43.8
광주	1,929	52.6	3,382	49.7	1,929	52.6	3,382	49.7	-	-	-	-
대전	2,123	50.8	3,740	53.1	2,123	50.8	3,740	53.1	-	-	-	-
울산	1,460	48.8	2,630	53.0	1,231	47.8	2,271	52.3	229	55.6	359	58.0
경기	13,813	51.3	23,066	51.8	10,758	50.9	20,399	51.1	3,055	52.6	2,667	57.6
강원	2355	55.1	2,111	54.4	1,250	50.3	1,151	49.8	1,105	61.9	960	61.0
충북	2,395	60.5	2,267	59.4	1,220	51.5	1,274	53.3	1,175	73.9	993	69.8
충남	3,231	63.5	3,172	63.8	899	51.6	1,473	57.2	2,332	69.7	1,699	71.0
전북	2,934	58.6	2,719	59.1	1,599	50.0	1,737	52.3	1,335	73.9	982	76.8
전남	3,404	65.3	2,817	62.8	1,008	51.3	1,345	53.7	2,396	73.8	1,472	74.3
경북	4,341	61.1	3,914	61.4	1,769	51.6	1,788	53.8	2,572	70.0	2,126	69.6
경남	4,951	58.4	4,514	59.1	2,627	50.9	2,546	53.0	2,324	70.2	1,968	69.3
제주	994	66.8	1,078	66.9	618	59.9	690	62.0	376	82.3	388	77.9

III. 새로운 산업 · 직업별 고용구조조사 표본설계

1. 새로운 표본설계의 특징

2007년도 표본설계가 갖는 특징은 다음과 같다.

- 1) 표본규모는 2006년도와 동일하게, 표본조사구는 3,750개를 유지하고 각 조사구로부터 20가구씩 총 75,000가구를 표본가구로 한다. 조사의 목적을 충실히 달성하기 위해서 합리적인 표본배정 원칙과 실사의 효율성을 높이도록 한다.
- 2) 새로운 표본설계에서의 모집단은 2005년도에 실시된 인구주택총조사에서 표본조사가 실시된 10% 자료를 조사 모집단으로 사용한다. 2005년 인구주택총조사 이후의 변화된 모집단의 특성은 사후조정방법을 사용하여 조정해 준다(한국조사연구학회 2007).
- 3) 전국단위의 산업소분류, 직업세분류별 통계 작성에 중점을 두지만, 지역별 특성별 세분류 통계생산에는 한계가 있다.
- 4) 2005년과 2006년도 OES조사자료, 2005년도 인구주택총조사 10%표본조사 자료 등을 분석해서 취업자수, 취업자 비율, 지역별, 동부와 읍·면부의 취업자 현황 등 본 조사와 관련된 사항들을 분석하여 표본배정에 사용한다. 표본배정은 비례배정과 네이만배정 등을 검토하고, 조사여건을 고려하여 최종 결정한다.
- 5) 과거에는 추정의 문제를 상세히 다루지 않았지만 본 설계에서는 가중치 부여 과정을 포함해서 추정식과 추정식의 오차계산 공식을 유도한다.
- 6) 표본추출된 조사구 목록을 작성하고 조사구 요도를 확보하며, 표본조사구 내의 가구목록을 작성해서 조사를 수행하기 전에 조사구와 가구의 변동을 반영한다. 추가로 예비표본조사구를 추출한다.

2. 모집단 분석

본 표본설계에서는 산업별, 직업별 취업자수 분포에 대한 정보가 포함된

2005년 인구주택총조사의 10% 표본조사자료를 모집단자료로 사용하였다. 따라서 2007년 표본설계의 모집단조사구는 2005년도에 실시한 인구주택총조사에서 표본조사가 실시된 10% 표본조사구 27,616개로 구성되었다. 하지만 본 표본설계에서는 기숙사 조사구, 특수시설 조사구, 그리고 관광호텔 외국인 조사구는 전수조사가 실시되었기 때문에 제외하고, 또한 섬 조사구도 조사환경을 고려하여 제외하여 총 20,385개의 조사구를 조사모집단으로 최종 설정하였다. 조사구 특성별 분포는 <표 5>와 같다.

<표 5> 조사구 특성별 분포

(단위: 개, %)

	합계	일반	아파트
조사구 수	20,385(100.00)	11,659(57.19)	8,726(42.81)
가구 수	1,571,617(100.00)	956,997(60.89)	614,620(39.11)
인구 수	4,436,727(100.00)	2,471,419(55.70)	1,965,308(44.30)

<표 6>은 조사구를 일반 조사구와 아파트 조사구로 구분하여 지역별 조사구수를 나타내고 있다. 동부의 조사구는 15,994개로 읍·면부 조사구 4,391개의 약 3.6배나 된다. 동부는 전체 아파트 조사구의 87.42%를 차지하고 있는 반면에 읍·면부는 상대적으로 일반 조사구의 비중이 높다. 서울, 광주, 그리고 대전에는 읍·면부가 없고 모두 동부로 구성되어 있으며, 나머지 광역시에는 약간의 읍·면부가 있다. 대부분의 지역에서 동부의 조사구가 읍·면부의 조사구보다 많다. 그러나 충남과 전남은 읍·면부의 조사구가 동부의 조사구보다 많다.

<표 7>은 지역별 만 15세 이상 인구수, 취업자수 및 취업자 비율 분포를 나타낸 것이다. 만 15세 이상 인구 중에서 동부는 76.78%를 차지하고, 읍·면부는 23.22%를 차지하고 있다. 그리고 만 15세 이상 인구 중에서 취업자수는 1,887,312명으로 취업자 비율은 52.70%로 나타났다. 또한 취업자 1,887,312명 중에서 동부는 1,370,408명으로 72.62%를 차지하고, 읍·면부는 516,904명으로 27.38%를 차지하고 있다. 취업자 비율 측면에서 살펴보면 동부의 경

우에는 49.84%이고, 읍·면부의 경우에는 62.16%로 동부보다 읍·면부의 취업자 비율이 높은 것으로 나타났다.

〈표 6〉 지역별 조사구 분포(아파트, 일반)

(단위: 개, %)

지역	전체			동부			읍·면부		
	계	일반	아파트	계	일반	아파트	계	일반	아파트
전국	20,385 (100.00)	11,659 (57.19)	8,726 (42.81)	15,994 (100.00)	8,366 (52.31)	7,628 (47.69)	4,391 (100.00)	3,293 (74.99)	1,098 (25.01)
서울	3,510	2,061	1,449	3,510	2061	1,449	-	-	-
부산	1,595	888	707	1,557	864	693	38	24	14
대구	1,157	647	510	1,083	608	475	74	39	35
인천	1,223	680	543	1,177	637	540	46	43	3
광주	674	288	386	674	288	386	-	-	-
대전	710	357	353	710	357	353	-	-	-
울산	492	254	238	404	208	196	88	46	42
경기	3,688	1,872	1,816	2,893	1,369	1,524	795	503	292
강원	792	478	314	481	231	250	311	247	64
충북	767	469	298	434	204	230	333	265	68
충남	908	569	339	328	142	186	580	427	153
전북	997	610	387	623	272	351	374	338	36
전남	873	573	300	412	179	233	461	394	67
경북	1,274	820	454	665	345	320	609	475	134
경남	1,439	854	585	832	436	396	607	418	189
제주	286	239	47	211	165	46	75	74	1

〈표 7〉 2005년 인구주택총조사 10% 자료의 지역별 만15세 이상 인구수, 취업
자수 및 취업자 비율 분포 (단위: 명, %)

지역	만15세 이상 인구 수			취업자 수			취업자 비율		
	계	동부	읍·면부	계	동부	읍·면부	전체	동부	읍·면부
전국	3,581,426 (100.00)	2,749,854 (76.78)	831,572 (23.22)	1,887,312 (100.00)	1,370,408 (72.62)	516,904 (27.38)	52.70	49.84	62.16
서울	729,292	729,292	-	365,179	365,179	-	50.07	50.07	-
부산	265,520	260,481	5,039	121,587	119,092	2,495	45.79	45.72	49.51
대구	178,836	168,717	10,119	83,325	77,389	5,936	46.59	45.87	58.66
인천	183,280	177,179	6,101	92,819	89,057	3,762	50.64	50.26	61.66
광주	101,500	101,500	-	49,085	49,085	-	48.36	48.36	-
대전	101,793	101,793	-	49,366	49,366	-	48.50	48.50	-
울산	70,708	59,198	11,510	37,164	30,659	6,505	52.56	51.79	56.52
경기	726,147	595,958	130,189	384,732	310,680	74,052	52.98	52.13	56.88
강원	124,995	64,176	60,819	65,681	30,851	34,830	52.55	48.07	57.27
충북	120,414	61,314	59,100	66,504	29,997	36,507	55.23	48.92	61.77
충남	154,153	45,048	109,105	92,803	24,609	68,194	60.20	54.63	62.50
전북	160,436	86,235	74,201	87,541	39,828	47,713	54.56	46.19	64.30
전남	158,150	55,177	102,973	94,471	26,415	68,056	59.74	47.87	66.09
경북	228,340	94,437	133,903	137,183	49,409	87,774	60.08	52.32	65.55
경남	239,917	121,332	118,585	136,776	62,464	74,312	57.01	51.48	62.67
제주	37,945	28,017	9,928	23,096	16,328	6,768	60.87	58.28	68.17

다음의 〈표 8〉은 2005년과 2006년의 OES자료, 그리고 2005년 인구주택총조사 10% 표본조사자료의 산업소분류별 분포를 비교한 것이다. 전체적으로 2005

〈표 8〉 2005년 OES, 2006년 OES와 2005년 인구주택총조사 10% 표본자료의
산업소분류 업종수 분포 (단위: 개, %)

지역	동부						읍·면부					
	2005 OES		2006 OES		2005년 인구주택총조사		2005 OES		2006 OES		2005년 인구주택총조사	
	산업 소분류	산업 소분류	산업 소분류	산업 소분류	산업 소분류	산업 소분류	산업소 분류	산업 소분류	산업 소분류	산업 소분류	산업 소분류	산업 소분류
	업종 수	업종 비율	업종 수	업종 비율	업종 수	업종 비율	업종 수	업종 비율	업종 수	업종 비율	업종 수	업종 비율
계	194	100.0	194	100.0	194	100.0	194	100.0	194	100.0	194	100.0
서울	177	91.2	182	93.8	189	97.4	-	-	-	-	-	-
부산	166	85.6	167	86.1	187	96.4	40	20.6	52	26.8	146	75.3
대구	146	75.3	161	83.0	182	93.8	74	38.1	61	31.4	158	81.4
인천	160	82.5	172	88.7	185	95.4	40	20.6	38	19.6	138	71.1
광주	135	69.6	156	80.4	179	92.3	-	-	-	-	-	-
대전	138	71.1	157	80.9	179	92.3	-	-	-	-	-	-
울산	118	60.8	131	67.5	175	90.2	73	37.6	84	43.3	163	84.0
경기	174	89.7	181	93.3	188	96.9	146	75.3	155	79.9	188	96.9
강원	124	63.9	122	62.9	175	90.2	94	48.5	93	47.9	173	89.2
충북	129	66.5	133	68.6	186	95.9	92	47.4	101	52.1	178	91.8
충남	117	60.3	136	70.1	179	92.3	130	67.0	121	62.4	183	94.3
전북	133	68.6	140	72.2	182	93.8	81	41.8	80	41.2	167	86.1
전남	116	59.8	118	60.8	174	89.7	107	55.2	91	46.9	171	88.1
경북	142	73.2	142	73.2	184	94.8	121	62.4	119	61.3	184	94.8
경남	149	76.8	150	77.3	180	92.8	130	67.0	132	68.0	181	93.3
제주	89	45.9	91	46.9	149	76.8	51	26.3	60	30.9	132	68.0

년 인구주택총조사 10% 표본조사자료가 2005년과 2006년도에 비해 산업소분류 업종비율이 높게 나타나고 있다. 이는 2005년도의 인구주택총조사의 규모가

크기 때문이다. 2005년에 비해 2006년에는 동부에서 산업소분류 업종비율이 높아진 반면에 읍·면부의 경우에는 2005년과 2006년 사이에 별 차이가 없다. 이유는 2006년도 표본설계에서 읍·면부에 비해 동부에 표본을 많이 배정하였기 때문에 동부에는 산업소분류 업종수가 비교적 많게 나타났지만 읍·면부에는 그렇지 못한 결과를 가져온 것으로 판단된다. 읍·면부에도 다양한 산업소분류 업종이 존재하는 것으로 보아 읍·면부에 비해 동부에 표본을 임의로 많이 배정하는 것은 적절치 못한 것으로 사료된다.

한편 직업세분류에 대해 2005년 OES와 2006년 OES, 그리고 2005년 인구주택총조사 10% 표본조사자료를 비교하기 위해 <표 9>를 작성하였다. OES의 경우에는 직업세분류수가 392종이고, 2005년 인구주택총조사 10% 표본조사자료의 경우에는 직업세분류수가 454종이므로 직접적인 비교는 곤란하지만 전체적인 경향을 살펴보기에는 무리가 없다고 생각된다. 이 때, 직업세분류 직종비율은 OES의 경우 직업세분류 직종수 392종이 기준이 되고, 2005년 인구주택총조사 10% 표본조사자료의 경우 직업세분류 직종수 454종이 기준이 됨을 유의해야 한다. 이는 OES의 직종별 세분류가 인구총조사 자료의 직종별 세분류에 비해 일부 직종(의회의원, 자연과학, 군인 등)을 포괄적으로 분류하기 때문인 것으로 사료된다. 2005년 인구주택총조사 10% 표본조사자료의 직업세분류 직종비율은 2006년도 OES 자료에 비해 동부의 서울, 부산, 인천, 경기를 제외하고는 전체적으로 높다.

이는 2005년도의 인구주택총조사의 규모가 크기 때문이다. 읍·면부의 경우 2005년 인구주택총조사 10% 표본조사자료의 직업세분류 직종비율이 2006년도 OES 자료보다 모든 지역에서 높게 나타나고 있다. 2006년도의 경우 동부에 비해 읍·면부의 직업세분류 직종비율이 2005년도 인구주택총조사 10% 표본조사자료에 비해 월등히 낮는데, 2006년도 표본설계에서 읍·면부에 비해 동부에 표본을 많이 배정하였기 때문으로 판단된다. 또한 <표 9>로부터 2005년과 2006년 OES 자료의 동부와 읍·면부 직업세분류수를 비교해 보면 동부가 읍·면부에 비해 상대적으로 많은 직종을 포괄하고 있지만, 도지역의 경우 읍·면부의 직종수가 동부지역에 비해 매우 낮은 수준이 아니기 때문에 동부지역에 임의로 많은 표본을 배분하는 것은 타당하지 않은 것으로 판단된다.

〈표 9〉 2005년 OES, 2006년 OES와 2005년 인구주택총조사 10% 표본자료의
직업세분류 직종수 분포 (단위: 명, 개, %)

지역	동부						읍·면부					
	2005 OES		2006 OES		2005년 인구주택총조사		2005 OES		2006 OES		2005년 인구주택총조사	
	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율	직업세 분류 직종 수	직업세 분류 직종 비율
계	392	100.0	392	100.0	454	100.0	392	100.0	392	100.0	454	100.0
서울	361	92.1	374	95.4	384	84.6	-	-	-	-	-	-
부산	300	76.5	347	88.5	380	83.7	55	14.0	68	17.3	238	52.4
대구	282	71.9	315	80.4	374	82.4	97	24.7	76	19.4	263	57.9
인천	295	75.3	329	83.9	378	83.3	44	11.2	39	9.9	221	48.7
광주	253	64.5	293	74.7	369	81.3	-	-	-	-	-	-
대전	255	65.1	305	77.8	371	81.7	-	-	-	-	-	-
울산	197	50.3	233	59.4	335	73.8	92	23.5	97	24.7	285	62.8
경기	365	93.1	381	97.2	387	85.2	261	66.6	270	68.9	379	83.5
강원	215	54.9	197	50.3	363	80.0	138	35.2	142	36.2	337	74.2
충북	210	53.6	225	57.4	361	79.5	120	30.6	135	34.4	339	74.7
충남	193	49.2	227	57.9	349	76.9	190	48.5	175	44.6	359	79.1
전북	224	57.1	250	63.8	356	78.4	112	28.6	114	29.1	321	70.7
전남	193	49.2	211	53.8	344	75.8	174	44.4	153	39.0	341	75.1
경북	240	61.2	242	61.7	369	81.3	187	47.7	190	48.5	360	79.3
경남	251	64.0	264	67.3	370	81.5	191	48.7	209	53.3	370	81.5
제주	132	33.7	156	39.8	309	68.1	66	16.8	80	20.4	227	50.0

3. 총화

2007년 표본설계에서는 1차적으로 전국단위의 산업·직업별 취업자수, 임금 등 고용관련 통계작성을 염두에 두었다. 따라서 전국의 행정구역을 총화 기준으로 하여 7대 광역시와 9개 광역자치단체로 구분하여 16개 층을 형성하였고, 각 지역을 동부와 읍·면부로 구분하여 전체 29개 층을 형성하였다.

성별, 연령 등과 같은 특성별 추정 시에는 조사 후 사후층화방법을 적용해서 추정치를 보정해 주는 것이 바람직하다.

4. 표본크기 결정 및 배분

본 조사의 목적은 산업소분류(194개)·직업세분류(392개) 수준에서의 고용구조를 파악하여 국가적 인적자원 수급정책을 위한 기본통계와 직업별 고용전망, 진로선택, 직업훈련, 취업알선 등 노동시장 정책과 연구를 위한 기초자료를 제공하는 데 있다.

이러한 목적을 달성하기 위해 표본크기의 결정은 한국고용정보원의 실사여건을 고려해서 3,750조사구를 표본조사구로 하였고, 표본가구는 조사구당 20가구로 정하였다. 이에 따라 표본크기는 75,000가구이다.

한편 3,750개의 표본조사구를 배정하는 방법으로 비례배정과 네이만배정을 고려하였다. 표본배정 공식은 다음의 식(1)과 식(2)를 사용하였다(김영원의 3인 2006; Cochran 1977).

① 비례배정

$$n_h = n \frac{N_h}{N} \quad (1)$$

② 네이만배정

$$n_h = n \frac{N_h \sqrt{P_h Q_h}}{\sum N_h \sqrt{P_h Q_h}} \quad (2)$$

비례배정은 층별 조사구수를 기준으로 하였고, 네이만배정의 경우는 층별 조사구수와 취업자 비율의 변동을 함께 고려하였다. 동부의 경우에는 서울과 경기도 그리고 읍·면부의 경우에는 경기, 경남, 경북이 많이 배정되었으나 이 중에서 특히 경기 지역에 표본조사구가 많이 배정되었다. 비례배정과 네이만배정에 의한 표본조사구수는 <표 10>과 같다.

2007년 표본설계에서는 취업자 비율에 대한 정보가 포함되어 있는 2005년 인구주택총조사 10% 표본조사자료를 이용한 네이만배정을 적용하여 표본조사

<표 10> 2006년 OES 조사자료, 2005년 인구주택총조사 10% 표본조사자료를 이용한 2007년 조사구 표본규모 (단위: 개)

구분	전체				동부				읍·면부			
지역	2006년 표본수	비례 배정	네이만 배정		2006년 표본수	비례 배정	네이만 배정		2006년 표본수	비례 배정	네이만 배정	
			2006년 OES 자료	2005년 인구 주택 총조사 자료			2006년 OES 자료	2005년 인구 주택 총조사 자료			2006년 OES 자료	2005년 인구 주택 총조사 자료
계	3,750	3,750	3,750	3,750	3,332	2,942	2,989	2,958	418	808	761	792
서울	916	646	657	650	916	646	657	650	-	-	-	-
부산	330	293	298	294	330	286	291	287	0	7	7	7
대구	231	213	216	213	231	199	202	200	0	14	14	13
인천	236	225	229	226	236	217	220	218	0	8	9	8
광주	133	124	126	125	133	124	126	125	-	-	-	-
대전	135	131	133	131	135	131	133	131	-	-	-	-
울산	99	90	92	91	99	74	76	75	0	16	16	16
경기	895	678	688	681	800	532	541	535	95	146	147	146
강원	78	146	147	146	48	88	90	89	30	57	57	57
충북	77	141	138	141	48	80	81	80	29	61	57	61
충남	100	167	159	166	50	60	61	61	50	107	98	105
전북	95	183	176	182	66	115	117	115	29	69	59	67
전남	96	161	152	157	50	76	77	76	46	85	75	81
경북	137	234	229	231	72	122	124	123	65	112	105	108
경남	161	265	260	264	97	153	155	154	64	112	105	110
제주	31	53	50	52	21	39	38	39	10	14	12	13

구를 배정하는 것이 적절하다고 판단된다. 한편 <표 10>을 보면 네이만배정 결과와 비례배정 결과가 거의 비슷한 값으로 나타나고 있는데, 이는 각 지역별 취업자비율이 비슷한 관계로 표본배정에 큰 영향을 미치지 않았기 때문으로 분석해 볼 수 있다.

5. 표본추출법

1) 표본조사구 추출

각 층으로부터의 표본조사구 추출은 계통추출법을 사용한다. 계통추출을 위해서 조사구를 행정구역으로 정렬하고 아파트 조사구와 일반 조사구 순으로 정돈한 후에 각각 가구수에 비례하는 층화확률비례계통추출법을 사용한다. 이러한 추출방법은 표본조사구를 지역별로 고르게 뽑을 수 있고 또한 행정구역 내에서 아파트 조사구와 일반 조사구가 균형있게 추출되어서 표본의 대표성이 커진다. 모집단조사구는 2005년도 인구주택총조사 10%표본조사자료를 사용하는 관계로 표본으로 추출된 일부의 조사구에 변동이 있을 수 있다. 표본으로 추출된 조사구를 조사에 사용할 수 없는 경우에는 예비표본조사구를 사용한다.

2) 표본가구 추출

표본조사를 위한 조사대상 가구는 표본조사구로부터 추출한다. 조사모집단 자료에 의하면, 조사구는 평균 57.63개의 가구로 구성되어 있다. 『산업·직업별 고용구조조사』에서는 표본조사구에 있는 가구들로부터 20가구를 표본가구로 추출한다. 표본가구를 추출하기 위해서는 통계청으로부터 제공받은 표본조사구에 대한 가구목록을 사용한다. 통계청에서 제공한 가구목록을 사용하기 전에 조사됨은 표본조사구를 사전에 방문해서 가구목록과 실제 거주하고 있는 가구들을 확인한다. 이때 목록에 누락되었거나 추가되는 가구를 파악해서 가구목록을 새로이 작성하여 실제 조사에 사용한다. 표본가구 추출은 비복원 단순확률추출법을 사용한다. 표본조사구 내의 가구목록을 작성한 후 난수를 발생하여, 발생된 난수에 해당되는 가구를 표본가구로 한다. 예를 들어 표본조사구 내의 가구가 60가구일 경우, EXCEL이나 SAS 등을 이용해서 1~60

사이의 비복원 난수 20개를 발생하여 번호에 해당되는 가구를 표본가구로 한다. 한편 표본가구 중에서 표본으로 사용할 수 없는 가구가 있을 경우에는 교체표본을 사용하는데 이를 위해서 예비표본가구를 선정한다. 예비표본가구는 주어진 표본조사구에서 표본가구로 선정된 가구를 제외하고 나머지 가구들을 대상으로 필요한 수의 예비표본을 표본가구의 선정과 동일한 비복원 단순확률 추출법을 사용하여 추출한다.

IV. 추정

1. 가중치 산정

일반적으로 추정단계에서 가중치를 이용하면 모집단에 대한 특성치인 모수에 대한 비편향추정량(unbiased estimator)을 얻을 수 있다. 통계분석 과정에서 가중치를 무시하고 계산한 추정치는 심각한 편향이 발생할 수 있다. 표본의 크기가 큰 대규모 조사에서 문제가 되는 것은 추정량의 편향이기 때문에 추정 과정에서 반드시 가중치를 이용해야 한다(한국조사연구학회 2004). 일반적으로 복합표본조사의 가중치는 설계 가중치, 무응답에 대한 조정, 사후증화를 통한 조정 등의 세 가지 요인을 통합하여 산정된다.

본 조사에서는 전국의 행정구역을 층화변수로 하여 7대 광역시와 9개 광역자치단체로 구분하여 16개 층을 형성하였고, 각 층을 동부와 읍·면부로 구분하여 전체 29개 층을 형성하였다. 『산업·직업별 고용구조조사』의 가중치는 각 지역 내에서 동부와 읍·면부로 모집단 취업자수와 표본의 조사 취업자수의 복원배율로 계산된다. 따라서 2007년 표본설계에서는 전국 또는 특성별(지역별, 동부, 읍·면부별)로 취업자수 총계에 대한 추정치 계산이 가능토록 하였고, 추정량과 추정량의 표준오차공식을 제시하였다. 『산업·직업별 고용구조조사』에서 각종 통계치는 사후증화 조정을 통해서 얻어진 가중치를 이용하여 계산된 것으로 추정량의 분산 계산이 복잡하다. 최근에는 복합표본조사 데이터를 전문적으로 분석할 수 있는 다양한 통계소프트웨어들이 개발되어 널리 활용되고 있다. 널리 사용되고 있는 대표적인 복합표본조사 분석용 통계소프

트웨어는 SAS 8.0판 이상, SUDAAN, WestVarPC, Stata 등이다. 『산업·직업별 고용구조조사』자료에 대한 고급분석을 위해서는 복합표본조사 분석용 통계소프트웨어를 활용하는 것이 필요하다.

본 설계에서 사용된 조사모집단은 2005년 인구주택총조사의 10% 표본조사 자료를 사용한 관계로 전국 또는 지역별 추정에 있어서는 <표 11>의 가중치를 사용해야 한다. 가중치는 가구수 기준과 인구수 기준의 두 경우를 사용할 수 있는 데 본 설계에서는 취업자수가 주요 관심사이므로 가구기준 가중치보다는 인구기준 가중치의 사용이 바람직하다고 생각된다. 그리고 특성별(산업별, 직업별 등) 전국 추정치 계산을 위해서는 통계청의 자료(<http://www.nso.go.kr>)를 이용하여 <표 11>과 같은 방법으로 계산한 가중치를 사용한다.

<표 11> 동부, 읍·면부 가중치(인구기준)

(단위: 명)

지역	전체조사자료		조사모집단 (10%표본자료)		가중치(w_h)	
	동부	읍·면부	동부	읍·면부	동부	읍·면부
전국	38,514,753	8,764,198	3,444,065	992,662	11.18	8.83
서울	9,820,171	-	880,181	-	11.16	-
부산	3,450,551	73,031	313,568	6,255	11.00	11.68
대구	2,313,792	150,755	208,746	13,174	11.08	11.44
인천	2,461,218	70,062	224,094	7,013	10.98	9.99
광주	1,417,716	-	129,923	-	10.91	-
대전	1,442,856	-	128,990	-	11.19	-
울산	883,812	165,365	76,933	14,933	11.49	11.07
경기	8,608,876	1,806,523	768,340	164,644	11.20	10.97
강원	894,695	569,864	80,430	72,407	11.12	7.87
충북	906,993	553,460	78,393	70,088	11.57	7.90
충남	703,946	1,185,549	59,204	129,977	11.89	9.12
전북	1,239,698	544,315	110,923	85,531	11.18	6.36
전남	791,595	1,028,224	72,236	120,278	10.96	8.55
경북	1,358,743	1,248,898	120,086	153,981	11.31	8.11
경남	1,830,441	1,225,915	155,439	141,916	11.78	8.64
제주	389,650	142,237	36,579	12,465	10.65	11.41

실제 조사를 마치고 얻은 자료들에는 무응답(예를 들면 부재, 응답거절 등)으로 인해 결측치가 생긴다. 또한 조사 결과의 특성치별(예를 들면, 성별, 학력별 등) 비율이 모집단의 비율을 따르지 않는 경우 통계청의 보조정보(추계인구 등) 자료를 이용하여 사후중화 방법에 의해 이들을 가중치 조정에 반영해 주어야 추정의 정확성을 높일 수 있다.

2. 특성별 추정

1) 전국 취업자수 총계 추정

복합표본조사 데이터를 분석할 때 가중치를 무시하고 분석하면 모수 추정에 심각한 편향이 발생할 수 있고, 추정량의 분산이 과소평가되어 문제가 된다. 따라서 『산업 · 직업별 고용구조조사』에서 모집단의 특성치에 대한 추정은 가중치를 이용한 통계치를 이용해야 하고, 만약 단순총계를 사용하면 추정치에 편향이 발생할 수 있다.

『산업 · 직업별 고용구조조사』에서 모집단의 특성치에 대한 추정은 가중치를 이용하는데, 이 때 사용될 기호들은 다음과 같다.

$h=1,2,\dots,L$: 층을 나타내는 첨자

$i=1,2,\dots,n_h$: h 층 내의 표본조사구를 나타내는 첨자

$j=1,2,\dots,m_{hi}$: h 층의 i 번째 표본조사구 내의 표본가구를 나타내는 첨자

L : 층의 개수

n_h : h 층 내의 표본조사구 개수

m_{hi} : h 층의 i 번째 표본조사구 내의 표본가구 개수

$n = \sum_{h=1}^L \sum_{i=1}^{n_h} m_{hi}$: 표본크기

w_{hij} : h 층의 i 번째 표본조사구 내의 j 번째 표본가구의 가중치

y_{hij} : h 층의 i 번째 표본조사구 내의 j 번째 표본가구의 관찰값

f_h : h 층의 추출률

『산업·직업별 고용구조조사』에서 전국 취업자수 총계($\hat{Y}$)는 다음과 같다.

$$\hat{Y} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$$

그리고 전국 취업자수 총계의 분산추정량은 다음과 같다.

$$\widehat{Var}(\hat{Y}) = \sum_{h=1}^L \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi.} - \bar{y}_{h..})^2$$

여기서, $y_{hi.} = \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$, $\bar{y}_{h..} = \left(\sum_{i=1}^{n_h} y_{hi.} \right) / n_h$ 이다.

또한, 전국 취업자수 총계의 상대표준오차는 다음의 식을 통해서 계산한다.

$$\widehat{CV}(\hat{Y}) = \frac{\sqrt{\widehat{Var}(\hat{Y})}}{\hat{Y}} \times 100$$

2) 각 특성별 취업자수 총계 추정

『산업·직업별 고용구조조사』에서 각 특성별(지역별, 동부, 읍·면부별) 취업자수 총계($\hat{Y}_h$)는 다음과 같다.

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$$

그리고 각 특성별 취업자수 총계의 분산추정량은 다음과 같다.

$$\widehat{Var}(\hat{Y}_h) = \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (y_{hi.} - \bar{y}_{h..})^2$$

또한, 각 특성별 취업자수 총계의 상대표준오차는 다음의 식을 통해서 계산한다.

$$\widehat{CV}(\hat{Y}_h) = \frac{\sqrt{\widehat{Var}(\hat{Y}_h)}}{\hat{Y}_h} \times 100$$

V. 제언

우리나라의 고용구조를 산업소분류(194개) · 직업세분류(392개) 수준에서 파악하여 국가 고용정책을 수립하는 데 활용되고 있는 『산업 · 직업별 고용구조조사』의 발전과 자료의 효율적인 활용을 위해서 다음과 같은 몇 가지 사항에 대한 연구검토를 제안한다.

1) 새로운 표본설계 방안 검토

매년 실시되는 『산업 · 직업별 고용구조조사』는 표본조사구만 3,750개(표본가구: 75,000)로 조사규모가 국내 최대이다. 따라서 매년 새로운 표본조사구를 사용하는 것은 조사 수행에 많은 어려움이 있으므로 현행 표본조사구를 일부만 교체하는 계속(순환)표본추출(successive sampling 또는 rotation sampling) 방법의 적용도 검토할 필요가 있다. 이 방법은 전체적인 고용상태를 파악하면서 동시에 고용의 변화를 살피는 데 유용하다.

2) 조사자료에 대한 다양한 통계분석

조사자료의 처리와 분석에 있어서 통계전문가들을 활용하여 정확하고 다양한 통계 생산이 가능토록 한다. 조사자료에 대해, 초기 가중치는 물론 조사과정에서 발생하는 무응답에 따른 가중치, 그리고 표본이 모집단을 충분히 대표해 주지 못하는 상황에서 가중치 조정을 통해 보다 정확한 추정과 추정의 신뢰성을 측정해 주기 위한 추정오차도 계산되어야 한다. 또한 우리나라의 고용구조는 여러 원인에 의해 변하므로 조사결과에 대한 다양한 통계적 방법을 사용하여 인과관계를 규명하는 것이 필요하다.

3) 조사자료 발표의 한계성

과거 조사자료를 분석한 결과에 의하면 산업소분류와 직업세분류에 표본이 하나도 없거나 1 또는 2인 경우가 있다. 특히 직업세분류의 경우 2006년 표본조사 자료에서 읍 · 면부에 취업자가 한 명도 없는 직종이 57개, 1명의 취업자만 있는 직종이 51개, 취업자가 2명만 있는 직종이 25개로 파악되었다. 따라서

모든 산업소분류와 직업세분류 수준에서 통계를 공표하는 데 주의를 기울여야 한다. CV가 높은 경우는 공표 시에 표시를 해서 이용자들이 통계 사용에 혼선이 없도록 하는 것이 바람직하다. 예를 들어 CV가 20%~30% 수준의 경우는 ‘*’를 표시하고, CV가 30%를 초과하는 경우는 ‘**’를 표시하는 것도 하나의 방법이다(Groves et al. 2004).

4) 유사통계와의 비교연구

『산업·직업별 고용구조조사』는 국내 최대의 조사임에도 불구하고 통계청의 『경제활동인구조사』 결과를 모집단으로 간주하여 사후보정하는 것은 바람직하지 않다. 다만, 2006년부터 통계청에서 실시하고 있는 인력실태조사(표본가구: 60,000)와의 비교연구는 필요할 것으로 보인다.

5) 철저한 실사 관리

본 조사는 표본조사구에서 20가구를 표본가구로 선정해서 조사가 실시된다. 그러나 2006년도 조사결과를 분석해 보면 조사거절, 부재, 부적절 가구 등의 이유로 표본조사구 내에 있는 거의 모든 가구를 대상으로 조사가 이루어졌다. 이는 표본가구가 표본으로 선정되지 못하고 조사 협조가 용이한 가구들을 대상으로 조사가 이루어져서 조사결과의 편향이 발생할 가능성이 높게 된다(한국통계학회 조사통계연구회 2000). 그러므로 사전에 표본가구에 대해 조사를 홍보하고, 면접원들에 대한 철저한 교육, 그리고 조사방법의 개선을 통해 초기 표본의 조사 성공률을 높이는 노력이 필요하다(Lessler & Kalsbeek 1992).

본 조사목적이 산업별, 직업별 고용구조를 파악하기 위한 것이므로 취업 상태에 있지 않은 가구원들에 대한 기본적인 사항에 대한 조사도 필요하다. 또한 응답률과 조사 실패의 다양한 원인들을 꼼꼼히 살펴서 향후 조사에 반영할 필요가 있다(Bierner & Lyberg 2003).

참고문헌

- 김영원 · 류제복 · 박진우 · 홍기학 공역. 2006. << 표본조사의 이해와 활용 >>. 톰슨코퍼레이션코리아.
- 대한통계협회. 2006. << 산업 · 직업별 고용구조조사 표본설계 >>.
- 박흥래. 2000. << 통계조사론 >>. 영지문화사.
- 한국고용정보원. 2006. << 2006 산업 · 직업별 고용구조조사 - 기초분석보고서 5차년도 자료분석 >>.
- 한국조사연구학회. 2004. << 제7차 전국 장내 기생충감염 실태조사 표본설계 최종보고서 >>.
- 한국조사연구학회. 2007. << 농촌통계조사 표본설계 최종보고서 >>.
- 한국통계학회 조사통계연구회. 2000. << 무응답 오차 >>. 자유아카데미.
- Biemer, P. P. and Lyberg, L. E. 2003. *Introduction to Survey Quality*. John Wiley and Sons.
- Cochran. W. G. 1977. *Sampling Techniques*. John Wiley & Sons, Inc.
- Groves, R. M., Dillman, D. A. Eltinge, J. L., and Little, R. J. A. 2002. *Survey Nonresponse*. John Wiley and Sons.
- Groves, R. M., Fowler, Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. 2004. *Survey Methodology*. John Wiley and Sons.
- Lessler, J. T. and Kalsbeek, W. D. 1992. *Nonsampling Error in Surveys*. John Wiley & Sons, Inc.
- SAS/STAT User's Guide Version 8. SAS Publishing.