

제4장

경제활동인구조사 부가조사의 무응답 대체방법 - 근로형태별 부가조사항목 중심으로 -

최필근

제1절 서론

1. 연구배경 및 목적

통계는 각종 국가정책과 지역발전의 계획수립을 위한 중요한 자료이다. 특히 통계청에서 조사하고 있는 경제활동인구조사 부가조사는 경제활동인구조사에서 조사되지 못하는 다양하고 세부적인 내용을 파악하여 정책 수립에 필요한 기초자료를 제공하고 있다. 이러한 부가조사로는 임금근로자에 대한 세부분석 및 파견근로, 용역근로, 특수형태근로 등 다양한 근로형태에 대한 실태 및 규모를 파악하는 근로형태별 부가조사, 비임금 근로자에 대한 취업실태를 파악하는 비임금 근로자 부가조사, 청년층에 대한 경제활동상태를 정확히 파악하는 청년층 부가조사, 만55세~79세의 취업실태를 파악하는 고령층 부가조사가 있다. 이와 같이 경제활동인구조사 부가조사 자료는 매우 중요한 역할을 하므로 정확한 통계생산을 위하여 자료의 품질을 높이기 위한 노력이 계속적으로 이루어져야 할 것이다. 자료의 품질을 향상시키기 위해서는 표본의 설계에서부터 자료를 제공하기까지의 모든 단계에서 정확하게 이루어져야 한다. 하지만, 이러한 일련의 과정에서 의도하지 않은 오차들이 발생하기 마련이며 그 중 하나가 무응답으로 인해서 발생하게 된다.

통계조사에서 모든 조사대상으로부터 필요한 정보를 완벽하게 얻는 것은 쉬운 일이 아니다. 하지만 이렇게 발생된 무응답은 조사결과에 상당한 영향을 줄 수도 있어 조사의 신뢰도를 높이기 위해서는 무응답률을 낮추는 것이 매우 중요하다. 많은 노력에도 불구하고 최근의 조사환경을 고려하면 계속적으로 무응답 비율이 높아질 가능성이 있다.



1인·노인·맞벌이 가구의 증가와 같은 인구구조학적인 변화가 이루어지고 있으며, 민감한 항목에 대하여 노출하지 않으려는 의도가 강하여 항목의 특성에 따라 무응답 비율이 높아질 수 있음을 고려해야 할 것이다. 따라서 무응답에 대한 근본적인 문제해결이 쉽지 않기 때문에 불가피하게 발생하는 무응답에 대하여 사후적으로 처리하는 방법을 체계적이고 지속적으로 연구하여 경제활동인구조사 부가조사의 무응답 발생 시 정확성이 높은 대체를 할 수 있도록 대비해야 할 필요성이 있을 것이다.

무응답의 형태를 나누어 보면 단위 무응답(unit nonresponse)과 항목 무응답(item nonresponse)으로 정의할 수 있다. 단위 무응답은 조사자로부터 얻은 정보가 하나도 없는 것을 의미하고, 항목 무응답은 응답을 해야 할 항목에 대해서 응답을 하지 않거나 질문과는 무관한 응답을 함으로써 불필요한 자료가 되는 것을 의미한다. 이러한 무응답이 발생하면 적절한 방법을 통해 무응답을 처리해야 하는데 단위 무응답의 경우는 가중치 조정 방법을 사용하고, 항목 무응답의 경우는 적절한 값을 채워 넣기 위한 여러 가지 대체 방법을 이용하게 된다. 본 연구에서 사용할 경제활동인구조사 부가조사에서는 항목 무응답이 존재하지 않는다. 조사가 이루어진 자료에는 모든 항목에 응답이 있고, 조사가 되지 못한 자료는 당연히 단위 무응답의 형태로 남아 있다. 하지만, 부가조사라는 특성이 있다. 경제활동인구조사 근로형태별 부가조사는 경제활동인구조사가 이루어진 임금근로자에 대해서만 작성을 한다. 따라서 경제활동인구조사와 부가조사를 동시에 고려하면 부가조사의 항목들을 항목 무응답으로 생각할 수 있다. 그러므로 본 연구에서는 경제활동인구조사 부가조사 무응답을 항목 무응답으로 간주하고 연구를 진행하면 될 것이다.

항목 무응답 연구를 위해서는 각 항목과 연관성이 높은 항목을 찾기 위한 과정이 선행되어야 한다. 이 항목들의 정보를 이용하여 대체하고자 하는 항목의 값을 찾게 되는데 이를 일반적으로 대체군이라 한다. 대체군은 모형을 추정할 때 보조변수와 같은 역할을 하며 대체의 정확도를 결정하는 중요한 요인이라 할 수 있다. 따라서 대체군을 찾기 위한 노력도 동시에 진행되어야 할 것이다. 그 다음 과정은 자료의 특성에 가장 적합한 대체기법을 선택하는 것이다. 하지만, 본 연구에서는 대체기법을 선정하는 연구에 초점을 맞추지 않을 것이다. 경제활동인구조사 부가조사는 대부분이 범주형 항목으로 구성되어 있어 이전의 연구 결과를 고려할 때 (응용) 핫덱 대체를 사용하면 될 것으로 판단된다.

본 연구에서는 부가조사의 모든 항목을 동시에 대체를 할 것인지, 각각의 항목을 개별적으로 대체를 할 것인지에 대한 대체방식의 연구에 초점을 맞추고자 한다. 현재 경제활동인구조사 부가조사의 대체는 부가조사의 모든 항목을 동시에 대체하는 방법을 사용하고 있다. 이러한 이유는 대체가 간편하다는 측면도 있지만 더 중요한 이유는 대체 후에 발생할 수 있는 자료의 불일치(내검에서 오류 발생)를 사전에 방지하기 위함이다. 하지만 많은 항목들을 대체군으로 사용하므로 도너가 될 수 있는 자료가 적거나 없을 수

도 있으며 각 항목별로는 연관성이 없는 대체군이 사용될 수 있다는 측면에서 대체의 정확성이 떨어질 수도 있을 것이다. 물론 경제활동인구조사 부가조사처럼 각 항목별로 연관성이 많은 조사에서는 동시에 대체하는 것도 하나의 방법일 수 있다. 따라서 본 연구에서는 경제활동인구조사 부가조사 중에서 근로형태별 조사표 항목에 대하여 두 가지 방법을 비교하여 더 적절한 대체방법을 제시하고자 한다. 이 외의 부가조사의 무응답 대체도 본 연구에서 제시된 방법을 사용할 수 있을 것으로 판단된다. 그리고 본 연구를 통해서 체계적이며 정확도 높은 무응답 대체기반을 마련하여 경제활동인구조사 부가조사 자료의 품질향상에 기여할 수 있기를 기대한다.

2. 연구내용 및 방법

본 연구의 목적은 경제활동인구조사 근로형태별 부가조사 모든 항목에 대하여 무응답 대체를 위한 대체군 및 대체방법을 제시하는데 있다. 연구를 위한 자료는 경제활동인구조사와 2010년 8월 부가조사 자료를 이용할 것이다.

대체군(보조변수) 선택은 의사결정나무 방법인 CHAID 및 CART 알고리즘을 이용해 연관성분석을 실시한 후 결정할 수 있다. 두 방법은 마디를 분할하는 기준이 다를 뿐 같은 알고리즘을 가지고 있다. 이전의 연구 결과 『무응답 처리를 위한 방법론 연구(I)(통계개발원, 2009)』의 『농업총조사 항목간 연관성 분석 및 대체군(보조변수) 개발(최필근)』에서도 볼 수 있듯이 두 방법의 차이는 거의 없다고 볼 수 있다. 따라서 본 연구에서는 CHAID 알고리즘을 이용하여 각 항목의 대체군을 제시할 것이며, 이 때 자료의 특성상 대체 후의 자료의 불일치 현상이 발생되지 않게 특정 항목을 대체군에 포함시킬 것이다. 그리고 모든 항목을 동시에 대체하기 위한 대체군은 성별, 혼인상태, 종사자규모, 종사상지위, 고용계약여부, 고용시근로기간, 교육정도, 만나이 항목으로 구성할 것이다.

대체기법은 이전 연구에서 검증된 응용 핫덱 대체방법을 이용할 것이다. 그리고 부가조사 모든 항목을 동시에 대체하는 방법과 개별 항목으로 대체하는 방법의 정확도를 평가하기 위한 모의실험을 실시할 것이다. 이는 대체 전후의 구성비 변화로 판단할 수 있다. 또한 더 정확한 모의실험을 위해서 무응답자의 성향을 분석하여 이를 반영하고자 한다. 응답여부와 관련성이 있는 항목과 대체군에 포함된 항목을 비교하여 공통적으로 선택된 주요 항목을 모의실험 자료를 구축할 때 이용할 것이다.

본 연구를 위하여 제2절에서는 연관성 분석을 위한 CHAID 및 CART 알고리즘을 설명하고 각 항목에 대하여 CHAID 분석결과 및 선정된 대체군을 제시한다. 제3절에서는 본 연구에 사용하고자 하는 응용 핫덱 대체방법을 자세하게 소개한다. 제4절에서는 무응답자의 성향을 분석하여 이를 모의실험을 위한 자료 구축에 반영한다. 제5절에서는 모의

실험을 실시하여 경제활동인구조사 부가조사에 적용할 대체방법을 제시한다. 마지막으로 제6절에서는 연구의 최종적인 결론을 요약하고 본 보고서를 마무리 하고자 한다.

제2절 경제활동인구조사 부가조사의 연관성분석

이 절에서는 항목간 연관성분석에서 사용하고자 하는 의사결정나무의 두 가지 방법인 CHAID(Chi-squared Automatic Interaction Detection) 와 CART(Classification and Regression Tree) 알고리즘에 대한 이론적 배경을 간략히 소개하고, 경제활동인구조사 근로형태별 부가조사 모든 항목에 대하여 CHAID 방법으로 연관성분석을 한 결과를 제시한다. CART 방법의 결과는 CHAID 방법과 유사함을 이전의 연구에서 보였으므로 생략하기로 한다.

1. CHAID 알고리즘

의사결정나무의 분리 알고리즘 중의 하나인 CHAID는 목표변수가 범주형 자료인 경우에는 χ^2 통계량에 의한 분할, 연속형 자료인 경우에는 F검정을 이용한 분할을 수행하는 분석방법이다. 구체적인 알고리즘을 살펴보면 다음과 같다.

- **step 1** : 각 설명변수에 대하여, 목표변수와 가장 유사성(p 값으로 측정)이 큰 범주의 짝을 찾는다. p 값을 계산하는 방법은 목표변수의 자료특성에 의해 결정된다.

이 때 목표변수가 범주형인 경우는 $2 \times d$ 분할표를 통한 χ^2 검정을 사용한다. 여기서 d 는 목표변수의 범주 수이다

(예시) $2 \times d$ 분할표에서의 p 값 계산

	범주 1	범주 2	...	범주 d	합계
범주 1	f_{11}	f_{12}	...	f_{1d}	$f_{1.}$
범주 2	f_{21}	f_{22}	...	f_{2d}	$f_{2.}$
합계	$f_{.1}$	$f_{.2}$...	$f_{.d}$	$f_{..}$



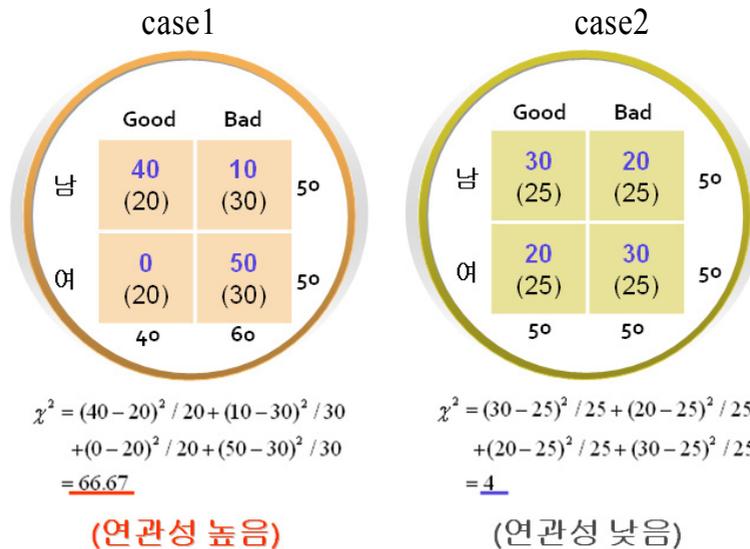
분할표에서 유사성 검정을 위한 카이제곱 통계량은

$$\chi^2 = \sum_{i,j} \frac{(f_{ij} - e_{ij})^2}{e_{ij}},$$

분포의 동일성 또는 독립성의 가설 하에서 계산된 기대도수는

$$e_{ij} = \frac{f_{i.} \times f_{.j}}{f_{..}}$$

와 같이 계산된다. 이 때 카이제곱의 값이 클수록 각 범주에 의하여 목표변수를 분리할 가능성이 커진다. 성별에 의한 선호도를 나타내고 있는 간단한 예를 살펴보면 다음과 같다.



따라서 case 1의 경우의 카이제곱 통계량이 크기 때문에 case 2의 경우보다 성별이 분리될 가능성이 커짐을 알 수 있다. 그리고 목표변수가 연속형인 경우에는 2개 이상의 그룹의 평균차이를 검정하는 분산분석표의 F검정을 사용하여 분리한다.

- **step 2** : 가장 큰 p 값을 가지는 설명변수 범주의 짝에 대하여 그 p 값과 미리 정해 놓은 α 값을 비교한다.
 - p 값이 α 값보다 클 경우에는 짝을 이루는 설명변수의 범주를 통합하고, 새로 생성된 범주에 대하여 step 1을 다시 실행한다.
 - p 값이 α 값보다 작을 경우에는 step 3으로 간다.
- **step 3** : 조정된 각 설명변수의 범주에 대하여 새로운 p 값을 계산하고, 가장 작은 p 값을 가지는 설명변수를 선택해 그 p 값과 미리 정해놓은 α 값을 비교한다.
 - p 값이 α 값보다 작거나 같을 경우에는 설명변수의 범주에 근거한 노드를 분리한다.
 - p 값이 α 값보다 클 경우에는 노드를 분리하지 않으며 이 노드는 최종노드가 된다.
- **step 4** : 더 이상 분리할 노드가 없거나 정해진 정지규칙이 만족할 때까지 위의 과정을 독립적으로 반복한다.

2. CART 알고리즘

의사결정나무의 분리 알고리즘 중의 하나인 CART는 L.Briemen (1984)에 의해 개발된 것으로 목표변수가 이산형인 경우에는 지니 지수(Gini Index), 목표변수가 연속형인 경우에는 분산의 감소량을 이용하여 이지분할을 수행한다. 기본적으로는 CHAID와 같은 알고리즘을 가지고 있으나, 마디를 분할하는데 있어서 다른 특징을 가지고 있다. CHAID에서 사용되는 카이제곱 통계량과 유사하게 지니 지수도 불순도(impurity)를 측정하는 하나의 지수이다. 임의의 한 개체가 목표변수의 i 번째 범주로부터 추출되었고, 그 개체를 목표변수의 j 번째 범주에 속한다고 오분류(misclassification)할 확률은 $P(i)P(j)$ 가 된다. 여기서 $P(i)$ 는 각 마디에서 한 개체가 목표변수의 i 번째 범주에 속할 확률이다. 이러한 오분류 확률을 모두 합하면 다음과 같다.

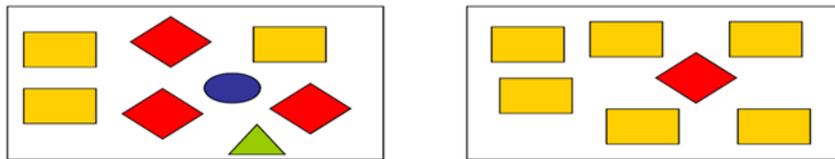


$$G = \sum_{j=1}^d \sum_{i \neq j} P(i)P(j).$$

이것은 위에서 정의한 오분류 확률의 추정치가 된다. 여기서 d 는 목표변수의 범주수를 의미한다. 일반적으로 CART는 이산형인 목표변수에 대해서는 지니 지수를 분할기준으로 사용하며 이 지수는 각 마디에서의 불순도 또는 다양도(diversity)를 재는 측도 중의 하나로써 다음과 같이 표현될 수 있다.

$$G = \sum_{j=1}^d P(j)(1 - P(j)) = 1 - \sum_{j=1}^d P(j)^2 = 1 - \sum_{j=1}^d (n_j/n)^2.$$

여기서 n 은 그 마디에 포함되어 있는 관찰치 수를 의미하며, n_j 는 목표변수의 j 번째 범주에 속하는 관찰치의 수를 의미한다. 지니 지수는 다음과 같이 구할 수 있다.



$$G = 1 - (3/8)^2 - (3/8)^2 - (1/8)^2 - (1/8)^2 = 0.69 \quad G = 1 - (6/7)^2 - (1/7)^2 = 0.24$$

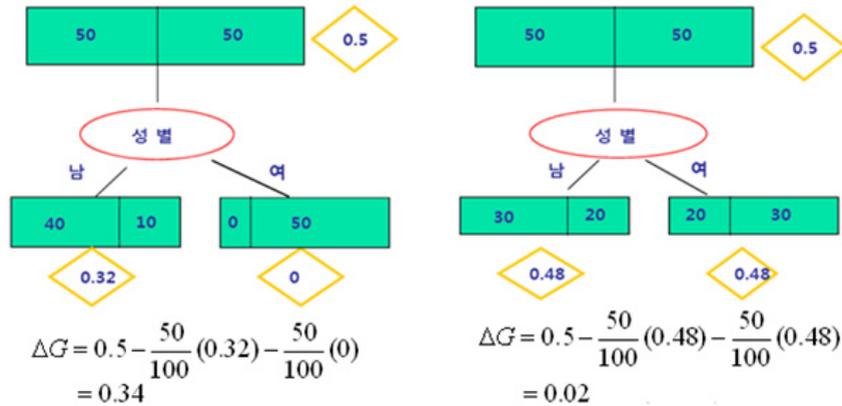
(Bad Model) (Good Model)

CART 알고리즘은 지니 지수를 가장 감소시켜주는 설명변수와 그 변수의 최적분할을 자식마디로 선택하는데, 지니 지수의 감소량은 다음과 같이 계산된다.

$$\Delta G = G - \frac{n_L}{n} G_L - \frac{n_R}{n} G_R.$$

여기서 n 은 부모마디의 관측치 수를 나타내고, n_R 과 n_L 은 각각 자식마디의 관측치 수를

의미한다. 즉, 자식마디로 분리되었을 때의 불순도가 가장 작도록 자식마디를 형성하는 것이다. 앞의 경우와 마찬가지로 성별에 의한 선호도를 나타내고 있는 간단한 예를 살펴보면 다음과 같다.

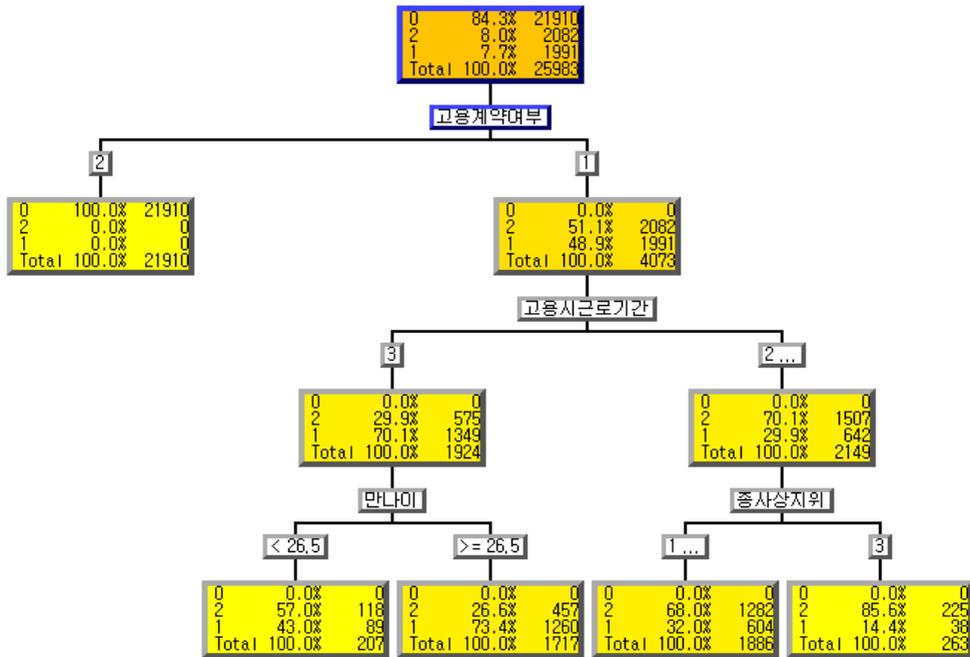


3. 부가조사 항목의 연관성분석 결과

경제활동인구조사 근로형태별 부가조사 자료에서 연관성분석을 실시할 항목은 계약기간반복갱신여부, 단기근로기간형태 등 29개 항목이다. 부가조사의 무응답은 전체 항목이 동시에 무응답이므로 대체군은 경제활동인구조사 항목에서 선택을 해야 하며, 대체 후 자료의 불일치가 발생하지 않도록 대체된 부가조사 항목을 특별 대체군으로 지정하여 사용할 것이다.

1) 계약기간반복갱신여부

계약기간반복갱신여부에 대한 연관성 분석의 결과를 [그림 4-1]과 <표 4-1>을 이용하여 설명하고자 한다. 나머지 항목들도 설명방식이 유사하기 때문에 계약기간반복갱신여부 경우에 대해서만 자세하게 설명하고, 나머지 항목들은 약식으로 설명을 할 것이다. 그리고 <표 4-1>에서 분리변수의 괄호안의 숫자는 각 가지에서의 마디 번호를 의미한다.



[그림 4-1] 계약기간반복갱신여부에 관한 연관성 모형

<표 4-1> 계약기간반복갱신여부에 관한 연관성 분석

계약기간반복갱신여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	고용계약여부	2	1
2	고용시근로기간(2)	3	1, 2, 4, 5, 6
3	만나이(3) 종사상지위(4)	26.5미만 1, 2	26.5이상 3

1 (고용계약여부)

계약기간반복갱신여부를 분리하는데 가장 연관성이 높은 항목은 고용계약여부이다. 전체 계약기간반복갱신여부의 구성비는 조사대상아님(0)이 84.3%, 예(1)가 7.7%, 아니오(2)가 8.0%이며 고용계약여부에 의하여 이를 분리하여 나간다. 고용계약여부가 정하지

않았음(2)인 경우는 좌로 분리되고, 이때의 구성비는 조사대상아님(0)이 100.0%로 모든 부분을 차지하고 있음을 알 수 있다. 즉, 고용계약여부가 정해지지 않았으면 본 항목의 조사 대상이 아님을 알 수 있다. 고용계약여부가 정했음(1)인 경우는 우로 분리되고, 이때의 구성비는 예(1)가 48.9%, 아니오(2)가 51.1%임을 볼 수 있다.

2_1

첫 번째 분리가 끝난 후 두 번째 분리가 시작된다. 이 과정은 첫 번째 분리가 된 왼쪽 부분을 다시 세부적으로 분리하는 과정이나, 이전의 과정에서 모든 분리가 끝났으므로 더 이상의 분리는 일어나지 않는다.

2_2(고용시근로기간)

이 과정은 첫 번째 분리가 된 오른쪽 부분을 다시 세부적으로 분리한다. 다음으로 연 관성이 높은 고용시근로기간 항목으로 다시 분리가 된다. 고용시근로기간이 1년(3)이면 왼쪽으로 분리되고 이때의 계약기간반복갱신여부의 구성비는 예(1)가 70.1%, 아니오(2)가 29.9%로 예(1)의 비율이 48.9%에서 70.1%로 많이 높아진 것을 볼 수 있다. 고용시근로기간이 1년(3) 이외의 경우는 오른쪽으로 분리되고 이때의 구성비는 예(1)가 29.9%, 아니오(2)가 70.1%로 아니오(2)의 비율이 50.1%에서 70.1%로 20%가 늘어난 것을 알 수 있다.

3_3(만나이)

이 과정은 두 번째 분리가 된 세 번째 부분을 다시 세부적으로 분리한다. 만나이 항목에 의해 계속적으로 분리가 된다. 고용계약여부가 정했음(1)이고 고용시근로기간이 1년(3)인 경우 만나이가 26.5(세) 이상인 사람은 계약기간반복갱신여부가 예(1)일 가능성이 조금 더 높아진다.

3_4(종사상지위)

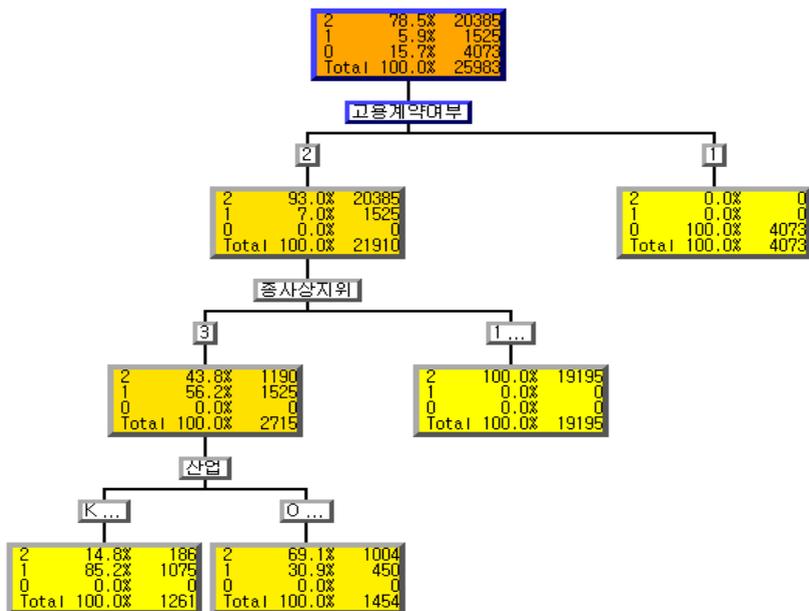
이 과정은 두 번째 분리가 된 네 번째 부분을 다시 세부적으로 분리한다. 종사상지위에 의해 다시 분리가 된다. 고용계약여부가 정했음(1)이고 고용시근로기간이 1년(3) 이외인 경우 종사상지위가 일용근로자(3)인 사람은 계약기간반복갱신여부의 아니오(2)의 구성비가 15.5%가 더 높아짐을 알 수 있다.



이상으로 계약기간반복갱신여부 항목을 분리하는 과정을 설명하였다. 요약하면, 고용계약여부, 고용시근로시간, 만나이, 종사상지위 항목이 계약기간반복갱신여부와 연관성이 있다는 것을 알 수 있다. 현재의 경제활동인구조사 자료의 모든 항목은 무응답이 존재하지 않으므로 이 항목들을 대체군으로 사용할 수 있음을 알려준다.

2) 단기근로기간형태

다음으로 단기근로기간형태 항목도 계약기간반복갱신여부 항목과 같은 방법으로 대체군을 결정하고자 한다. 단기근로기간형태의 구성비는 조사대상아님(0)이 15.7%, 예(1)가 5.9%, 아니오(2)가 78.5%이며 고용계약여부와 가장 큰 연관성을 가진다. 고용계약여부가 정했음(1)인 경우는 조사대상아님(0)이 100.0%로 모든 부분을 차지하고 있음을 알 수 있다. 즉, 고용계약여부가 정해지면 본 항목의 조사 대상이 아님을 알 수 있다. 고용계약여부가 정하지 않았음(2)인 경우의 구성비는 예(1)가 7.0%, 아니오(2)가 93.0%임을 볼 수 있다. 다음으로 연관성 있는 항목은 종사상지위인데 고용계약여부가 정하지 않았음(2)인 경우 중에서 종사상지위가 상용근로자(1) 또는 임시근로자(2)이면 단기근로기간형태의 값 모두가 아니오(2)임을 알 수 있다. 그리고 산업분류가 A, F, K, T이면 예(1)일 가능성이 상당히 높아짐을 볼 수 있다. 자세한 결과는 [그림 4-2]와 <표 4-2>를 참조하기 바란다.



[그림 4-2] 단기근로기간형태에 관한 연관성 모형

〈표 4-2〉 단기근로기간형태에 관한 연관성 분석

단기근로기간형태			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	고용계약여부	2	1
2	종사상지위(1)	3	1, 2
3	산업(1)	A, F, K, T	그 외 산업

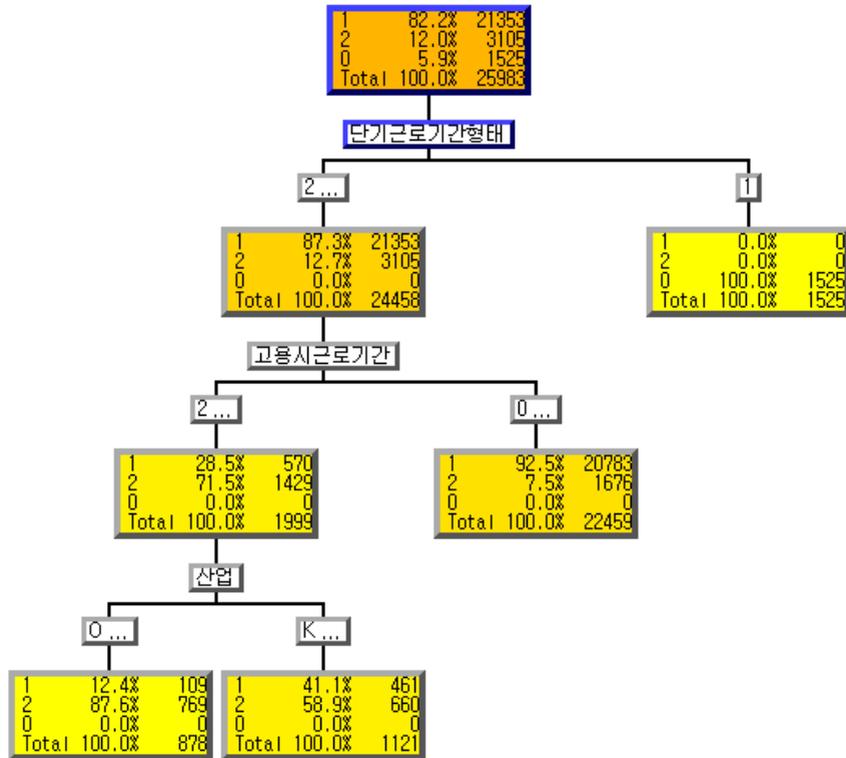
3) 계속근로가능여부

계속근로가능여부 항목은 조사대상아님(0)이 5.9%, 예(1)가 82.2%, 아니오(2)가 12.0%로 구성되어 있다. 이 항목의 경우 단기근로기간형태가 예(1)이면 조사대상아님(0)이 되어야하므로 단기근로기간형태 항목을 특별 대체군에 포함시켜 대체 후 자료의 불일치가 발생하지 않도록 한다. 따라서 단기근로기간형태가 예(1)인 경우 모두가 조사대상아님(0)인 것을 볼 수 있다. 그리고 단기근로기간형태가 예(1)이 아닌 경우에는 계속근로가능여부의 구성비는 예(1)가 87.3%, 아니오(2)가 12.7%로 나타난다. 다음은 고용시근로기간 항목과 연관성이 높는데 근로기간이 1년 또는 장기간(2년초과)일 경우에는 계속근로가능여부의 구성비 중 예(1)가 92.5%로 상당히 높아지는 것을 볼 수 있다. 또한 산업분류가 A, F, H, I, O이면 아니오(2)의 구성비가 상대적으로 높아진다. 자세한 결과는 <표 4-3>과 [그림 4-3]을 참조하기 바란다.

〈표 4-3〉 계속근로가능여부에 관한 연관성 분석

계속근로가능여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	단기근로기간형태	0, 2	1
2	고용시근로기간(1)	1, 2, 4	0, 3, 5, 6
3	산업(1)	A, F, H, I, O	그 외 산업

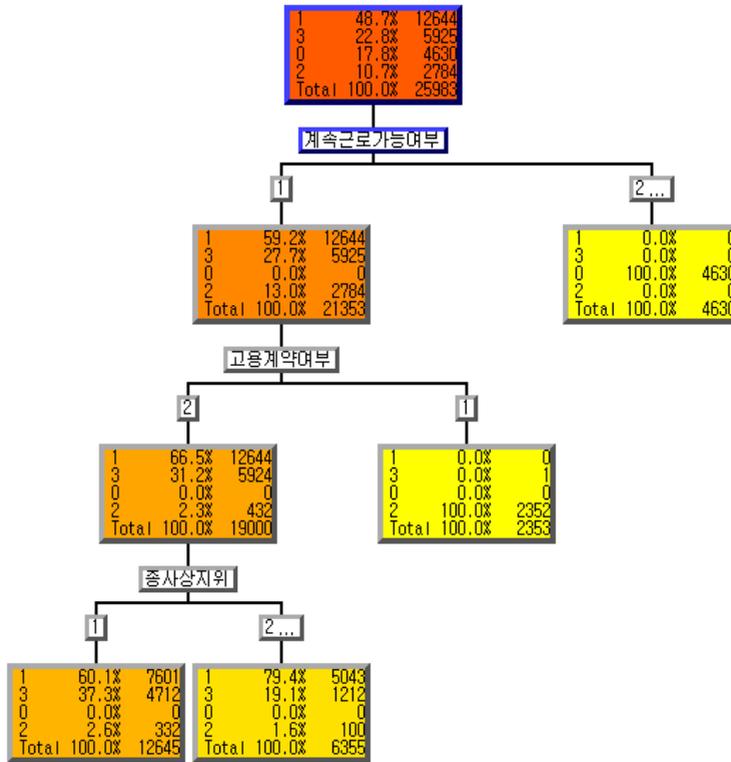




[그림 4-3] 계속근로가능여부에 관한 연관성 모형

4) 계속근로가능여부(이유)

계속근로가능여부(이유) 항목은 조사대상아님(0)이 17.8%, (1)이 48.7%, (2)가 10.7%, (3)은 22.8%로 구성되어 있다. 이 항목은 계속근로가능여부를 예(1)라고 대답한 경우에만 대답을 하는 항목으로 계속근로가능여부 항목을 특별 대체군에 포함시킨다. 따라서 계속근로가능여부가 예(1)가 아니면 모두가 조사대상아님(0)이라는 것을 알 수 있다. 다음은 고용계약여부와 연관성이 높는데 고용계약여부가 정했음(1)이면 계속근로가능여부를 예(1)라고 한 모두가 이유가 (2)라고 대답했으며, 정하지 않았음(2)이면 이유가 (1)인 경우가 66.5%로 상대적으로 높은 것을 볼 수 있다. 또한 종사상지위가 임시근로자(2) 또는 일용근로자(3)이면 이유가 (1)인 경우가 79.4%로 더 높아진다. 자세한 결과는 [그림 4-4]와 <표 4-4>를 참조하기 바란다.



[그림 4-4] 계속근로가능여부(이유)에 관한 연관성 모형

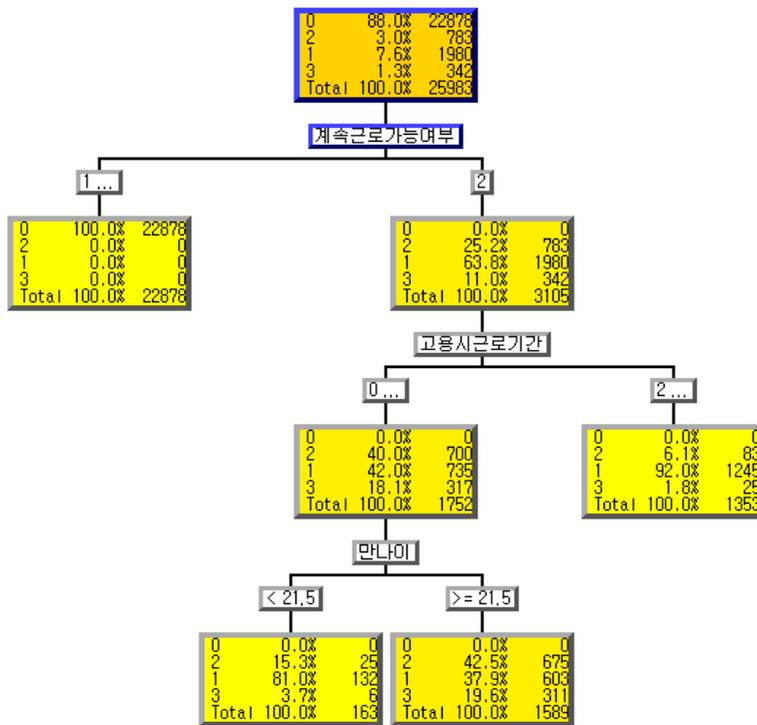
<표 4-4> 계속근로가능여부(이유)에 관한 연관성 분석

계속근로가능여부(이유)			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	계속근로가능여부	1	0, 2
2	고용계약여부(1)	2	1
3	종사상지위(1)	1	2, 3

5) 향후기대근속기간

향후기대근속기간 항목은 조사대상아님(0)이 88.0%, 1년 이하(1)이 7.6%, 1-2년(2)이 3.0%, 2년 초과(3)는 1.3%로 구성되어 있다. 이 항목은 계속근로가능여부를 아니오(2)라

고 대답한 경우에만 응답을 하는 항목으로 계속근로가능여부 항목을 특별 대체군에 포함시킨다. 따라서 계속근로가능여부가 아니오(2)가 아니면 모두가 조사대상아님(0)이라는 것을 알 수 있다. 다음은 고용시근로기간과 연관성이 높는데 고용시근로기간이 1년 미만인 경우에는 조사대상자의 92.0%가 향후기대근속기간이 1년 이하인 것으로 나타났다. 그리고 만나이가 21.5(세)미만이면 향후기대근속기간도 1년 이하일 가능성이 높은 것으로 보인다. 자세한 결과는 [그림 4-5]와 <표 4-5>를 참조하기 바란다.



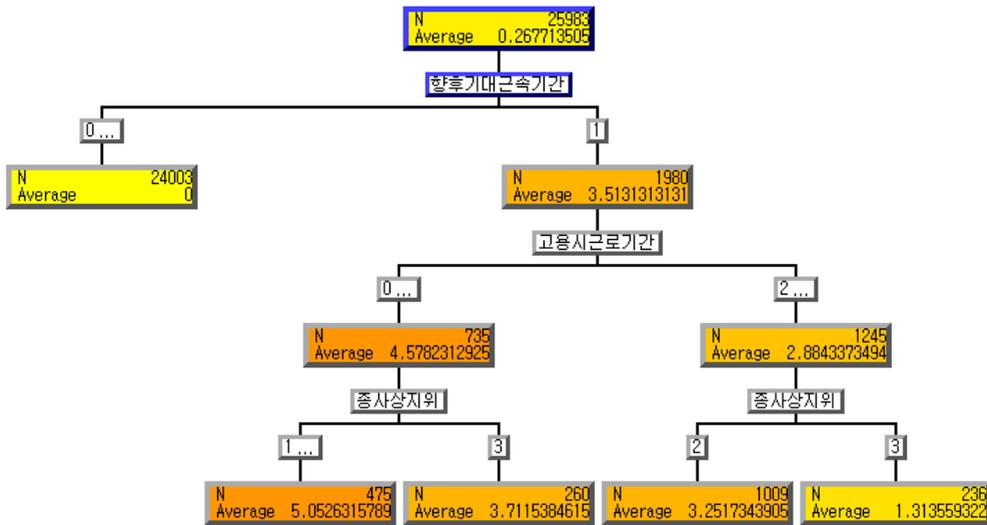
[그림 4-5] 향후기대근속기간에 관한 연관성 모형

<표 4-5> 향후기대근속기간에 관한 연관성 분석

향후기대근속기간			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	계속근로가능여부	0, 1	2
2	고용시근로기간(2)	0, 3, 4, 5, 6	1, 2
3	만나이(3)	21.5미만	21.5이상

6) 향후기대근속기간(1년이하)

향후기대근속기간(1년이하) 항목은 향후기대근속기간을 1년이하(1)라고 대답한 경우에만 응답을 하는 항목으로 향후기대근속기간 항목을 특별 대체군에 포함시킨다. 따라서 향후기대근속기간을 1년 이하(1)라고 대답한 경우의 평균은 3.51(개월)이며 그 외의 경우는 평균이 0인 것을 알 수 있다. 다음은 고용시근로기간과 연관성을 가지는데 고용시근로기간이 1년 미만이면 2.88(개월), 1년 이상이면 4.58(개월)의 향후기대근속기간을 평균적으로 갖는다. 또한 종사상지위의 경우 상용근로자(1)나 임시근로자(2)가 일용근로자(3)보다 향후기대근속기간이 평균적으로 더 길게 나타났다. 자세한 결과는 [그림 4-6]과 <표 4-6>을 참조하기 바란다.



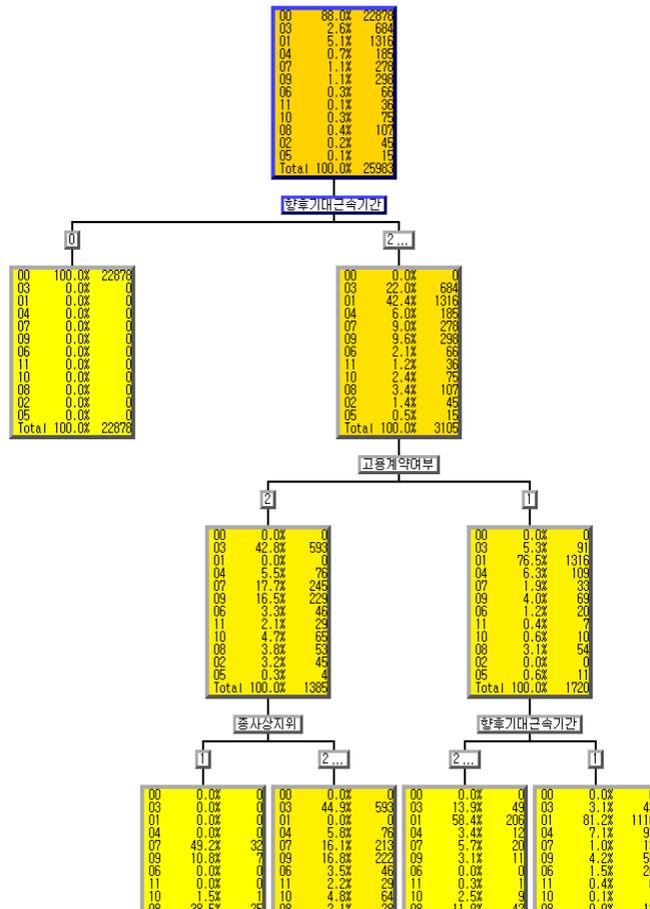
[그림 4-6] 향후기대근속기간(1년이하)에 관한 연관성 모형

<표 4-6> 향후기대근속기간(1년이하)에 관한 연관성 분석

향후기대근속기간(1년이하)			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	향후기대근속기간	0, 2, 3	1
2	고용시근로기간(2)	0, 3, 4, 5, 6	1, 2
3	종사상지위(3)	1, 2	3
	종사상지위(4)	2	3

7) 근속기간제한이유

근속기간제한이유 항목은 11개 항목으로 구성되는데 조사대상아님(0)이 88.0%, 기간만료(1)가 5.1%, 임시직(3)이 2.6%로 대부분을 차지하고 있다. 이 항목은 향후기대근속기간을 대답한 경우에만 응답을 하는 항목으로 향후기대근속기간 항목을 특별 대체군에 포함시킨다. 따라서 향후기대근속기간이 조사대상아님(0)이면 모두가 조사대상아님(0)이라는 것을 알 수 있다. 다음은 고용계약여부와 연관성이 높는데 고용계약여부가 정했음(1)이면 근속기간제한이유가 기간만료(1)인 경우가 76.5%, 정하지 않았음(2)이면 임시직(3)인 경우가 42.8%로 가장 많아진다. 그리고 종사상지위나 만나이도 근속기간제한이유와 연관성을 보이고 있다. 자세한 결과는 [그림 4-7]과 <표 4-7>을 참조하기 바란다.



[그림 4-7] 근속기간제한이유에 관한 연관성 모형

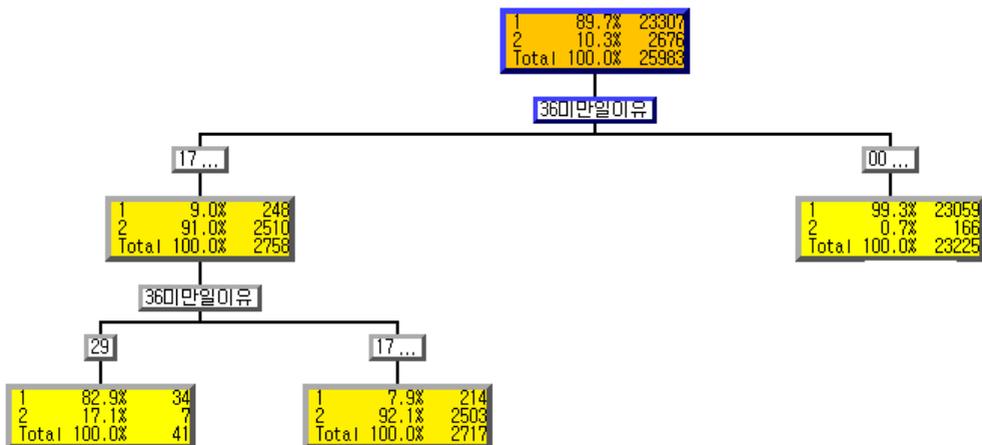
<표 4-7> 근속기간제한이유에 관한 연관성 분석

근속기간제한이유			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	향후기대근속기간	0	1, 2, 3
2	고용계약여부(2)	2	1
3	종사상지위(3) 향후기대근속기간(4)	1 2, 3	2, 3 1
4	만나이(5) 만나이(6)	51.5미만 24.5미만	51.5이상 24.5이상



8) 근로시간형태

근로시간형태의 구성비는 전일제근로(1)가 89.7%, 시간제근로(2)가 10.3%이며 36미만일이유와 가장 큰 연관성을 가진다. 36미만일이유가 11~18, 29이면 시간제근로(2)가 91.0%이면 21~23, 28, 30이면 반대로 전일제근로(1)가 99.3%로 나타났다. 이 항목은 연관성이 높은 다른 항목은 없는 것으로 보인다. 자세한 결과는 [그림 4-8]과 <표 4-8>를 참조하기 바란다.



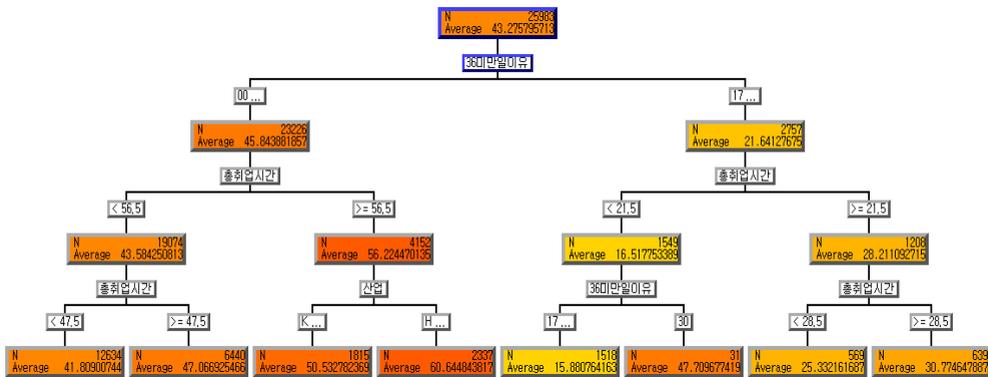
[그림 4-8] 근로시간형태에 관한 연관성 모형

<표 4-8> 근로시간형태에 관한 연관성 분석

근로시간형태			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	36미만일이유	11~18, 29	21, 22, 23, 28, 30
2	36미만일이유(1)	29	11~18

9) 평소근로시간

평소근로시간의 전체평균은 43.28(시간)이며, 평소근로시간을 분리하는데 36미만일이유가 가장 큰 역할을 한다. 이유가 21~23, 26, 28, 29이면 45.84(시간), 11~17, 30이면 21.64(시간)를 1주일간 평균적으로 일한다. 또한 총취업시간과도 큰 연관성을 보이는데 총취업시간이 많을수록 평소근로시간도 많아지는 것을 볼 수 있다. 그리고 산업분류에 의해서도 근로시간의 차이를 보인다. H(운수업)는 K(금융 및 보험업)보다 근로시간이 더 많은 것으로 나타난다. 자세한 결과는 [그림 4-9]와 <표 4-9>를 참조하기 바란다.



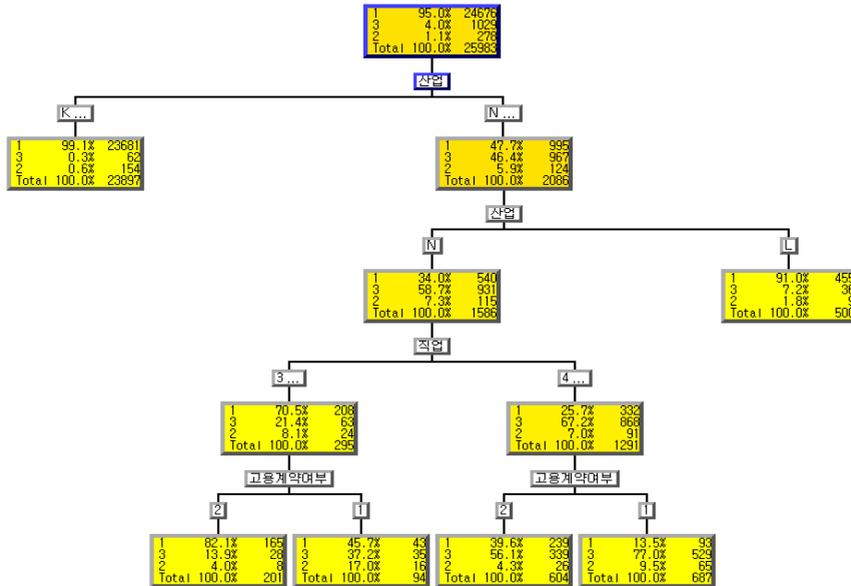
[그림 4-9] 평소근로시간에 관한 연관성 모형

<표 4-9> 평소근로시간에 관한 연관성 분석

평소근로시간			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	36미만일이유	21~23, 26, 28, 29	11~17, 30
2	총취업시간(1)	56.5미만	56.5이상
	총취업시간(2)	21.5미만	21.5이상
3	총취업시간(1)	47.5미만	47.5이상
	산업(2)	K.....	H.....
	36미만일이유(3)	11~17	30
	총취업시간(4)	28.5미만	28.5이상

10) 급여받은곳

급여받은곳 항목은 지난 주 일한 곳(1)이 95.0%, 파견업체(2)가 1.1%, 용역업체(3)가 4.0%로 구성되어 있다. 급여받은곳과 가장 큰 연관성을 가지는 항목은 산업분류이다. 산업분류가 L(부동산업 및 임대업)과 N(사업지원 서비스업)이외의 종사자는 99.1%가 지난 주 일한 곳에서 급여를 받고 있다. N(사업지원 서비스업)은 용역업체(3)에서 급여를 받는 경우가 58.7%로 상당히 높다. 또한 직업이 서비스 종사자(4), 5(판매종사자), 7(기능원), 8(장치 및 조립 종사자), 9(단순노무 종사자)이면 용역업체(3)에서 급여를 받는 경우가 58.7%에서 67.2%로 더 높아진다. 그리고 고용계약여부도 급여받은곳 항목과 연관성이 있는 것으로 나타났다. 자세한 결과는 [그림 4-10]과 <표 4-10>을 참조하기 바란다.



[그림 4-10] 급여받은곳에 관한 연관성 모형

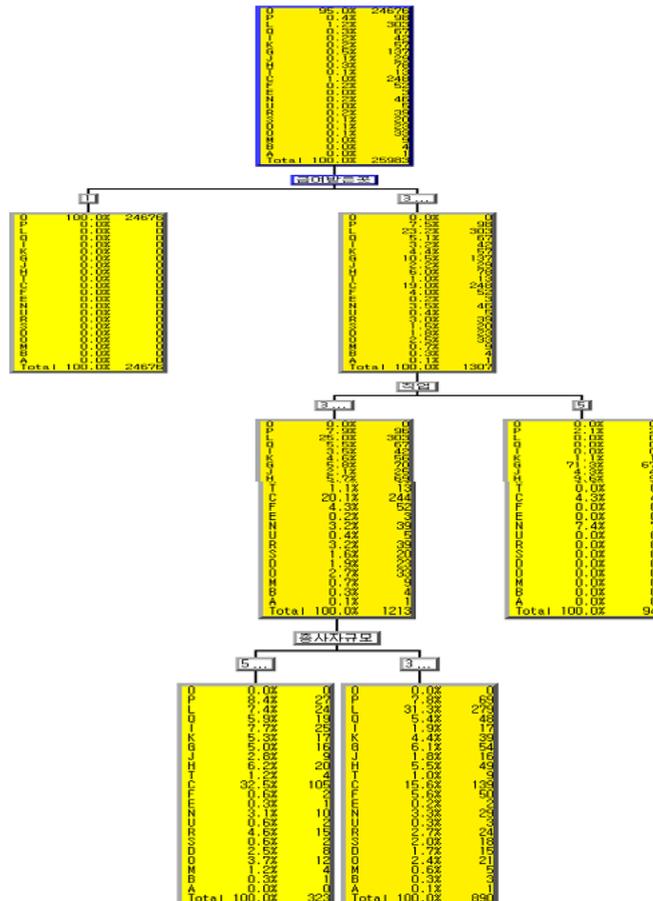
<표 4-10> 급여받은곳에 관한 연관성 분석

급여받은곳			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	산업	그 외 산업	L, N
2	산업(2)	N	L
3	직업(3)	1, 2, 3, 6	4, 5, 7, 8, 9
4	고용계약여부(5)	2	1
	고용계약여부(6)	1	1



11) 근무사업체

근무사업체 항목은 급여받은곳을 파견업체(2)나 용역업체(3)라고 대답한 경우에만 응답을 하는 항목으로 급여받은곳 항목을 특별 대체군에 포함시킨다. 따라서 급여받은곳이 지난 주 일한 곳(1)이면 모두가 조사대상아님(0)이라는 것을 알 수 있다. 급여받은곳이 파견업체(2)나 용역업체(3)이면 L(부동산 및 임대업)이나 C(제조업)에서 일했을 가능성이 높다. 여기서 직업이 판매 종사자(5)이면 근무사업체가 G(도매 및 소매업)인 경우가 71.3%나 됨을 볼 수 있다. 그리고 종사자규모도 근무사업체와 연관성을 보이고 있다. 자세한 결과는 [그림 4-11]과 <표 4-11>을 참조하기 바란다.



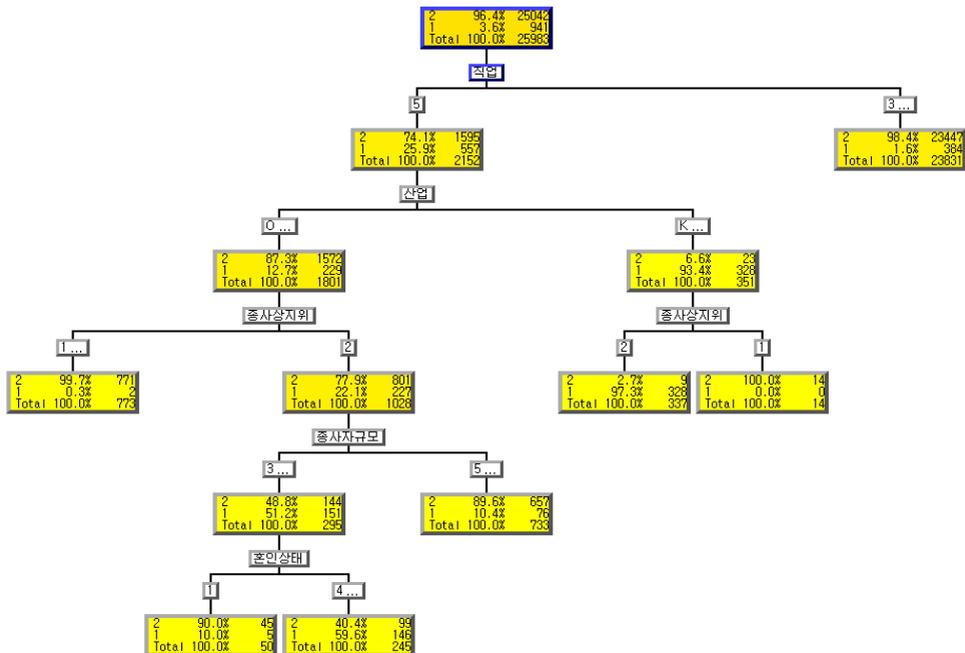
[그림 4-11] 근무사업체에 관한 연관성 모형

<표 4-11> 근무사업체에 관한 연관성 분석

근무사업체			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	급여받은곳	1	2, 3
2	직업(2)	1~4, 6~9	5
3	종사자규모(3)	4, 5	1, 2, 3, 6

12) 특수고용여부

특수고용여부 항목은 예(1)가 3.6%, 아니오(2)가 96.4%로 구성되어 있다. 이 항목은 직업과 연관성이 높는데 직업이 판매 종사자(5)인 경우 예(1)의 구성비가 25.9%로 늘어남을 볼 수 있다. 이 경우 산업이 K(금융 및 보험업), P(교육 서비스업)이면 예(1)의 구성비는 93.4%가 된다. 여기에 종사상지위가 상용근로자(1)가 되면 100%가 아니오(2)라고 응답한 것을 알 수 있다. 그리고 종사자규모와 혼인상태도 특수고용여부에 영향을 주는 항목임을 볼 수 있다. 자세한 결과는 [그림 4-12]와 <표 4-12>를 참조하기 바란다.



[그림 4-12] 특수고용여부에 관한 연관성 모형

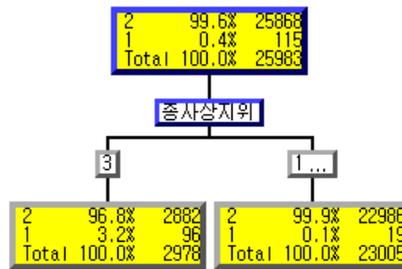


<표 4-12> 특수고용여부에 관한 연관성 분석

급여받은곳			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	직업	5	1~4, 6~9
2	산업(1)	그 외 산업	K, P
3	종사상지위(1) 종사상지위(2)	1, 3	2
		2	1
4	종사자규모(3)	3, 4, 6	1, 2, 5
5	혼인상태(5)	1	2, 3, 4

13) 가정내근로여부

가정내근로여부 항목은 가정에서(1)가 0.4%, 지정장소(2)가 99.6%로 대부분 지정장소(2)로 응답하였다. 종사상지위가 연관성이 있는데 종사상지위가 일용근로자(3)인 경우 가정에서(1) 근무할 가능성이 3.2%로 아주 조금 높아지나 큰 차이는 없는 것으로 보인다. 자세한 결과는 [그림 4-13]과 <표 4-13>을 참조하기 바란다.



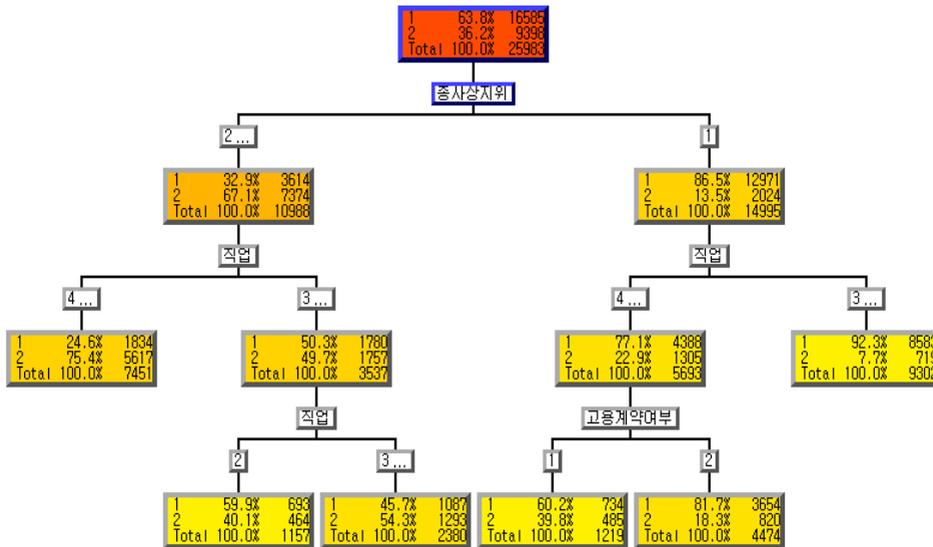
[그림 4-13] 가정내근로여부에 관한 연관성 모형

<표 4-13> 가정내근로여부에 관한 연관성 분석

가정내근로여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	3	1, 2

14) 일자리취업동기

일자리취업동기 항목은 자발적사유(1)가 63.8%, 비자발적사유(2)가 36.2%로 구성되어 있다. 이 항목과 가장 연관성이 높은 항목은 종사상지위인데 종사상지위가 상용근로자(1)인 경우 자발적사유(1)가 86.5%로 높아지며 반대로 임시근로자(2)나 일용근로자(3)면 비자발적사유(2)가 많이 늘어남을 볼 수 있다. 그리고 직업도 일자리취업동기와 높은 연관성을 보이는데 관리자(1), 전문가(2), 사무 종사자(3) 등은 자발적사유(1)가 매우 높게 나타난다. 또한 고용계약여부도 영향을 주는 항목이다. 자세한 결과는 [그림 4-14]와 <표 4-14>를 참조하기 바란다.



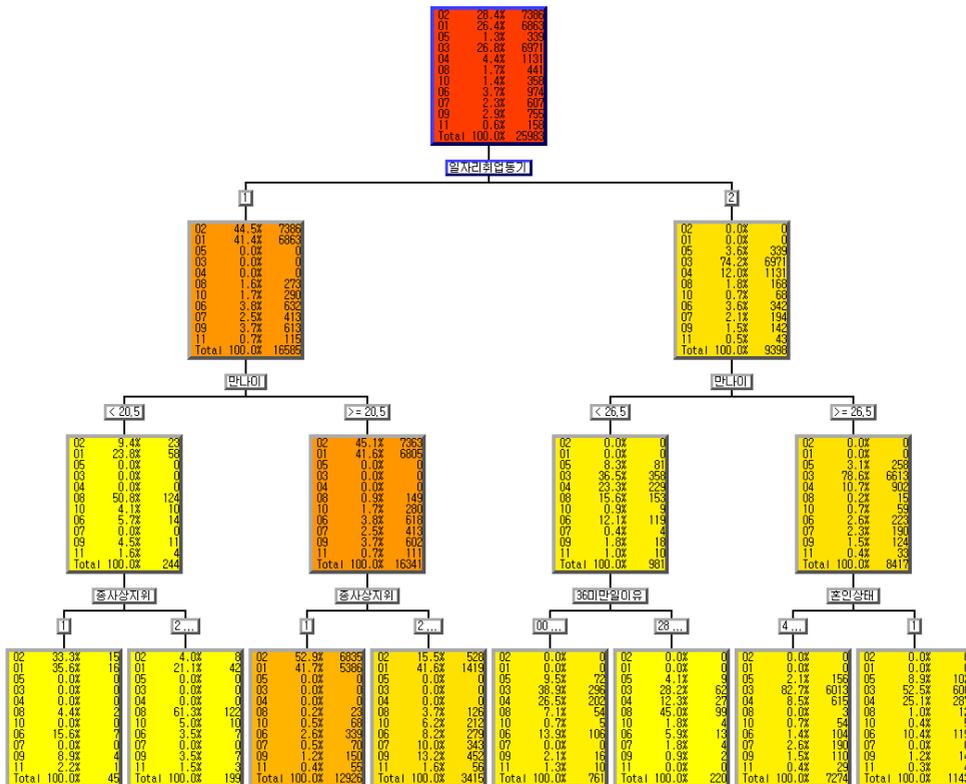
[그림 4-14] 일자리취업동기에 관한 연관성 모형

<표 4-14> 일자리취업동기에 관한 연관성 분석

일자리취업동기			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1
2	직업(1) 직업(2)	4, 6~9 4, 6~9	1, 2, 3, 5 1, 2, 3, 5
3	직업(2) 고용계약여부(3)	2 1	1, 3, 5 2

15) 주된취업사유

주된취업사유 항목은 11개 항목으로 구성되는데 근로조건만족(1)이 26.4%, 안정적인 자리(2)가 28.4%, 당장필요(3)가 26.8%로 82%정도를 차지하고 있다. 이 항목은 일자리취업동기와 연관이 되는 항목으로 일자리취업동기 항목을 특별 대체군에 포함시킨다. 일자리취업동기가 자발적사유(1)이면 당장필요(3)는 26.8%에서 0%로 되며 반대로 비자발적사유(2)이면 당장필요(3)는 74.2%로 상당히 높아짐을 볼 수 있다. 그리고 만나이도 주된취업사유와 연관성이 높는데 20.5(세)미만이면 학업 등을 병행(8)하는 이유가 많이 늘어난다. 종사상지위도 상용근로자(1)인 경우 주된취업이유가 근로조건만족(1)과 안정적인 자리(2) 비율이 높으며, 임시근로자(2)나 일용근로자(3)인 경우에는 상대적으로 타 이유가 많이 나타난다. 그리고 36미만일이유와 혼인상태도 주된취업사유와 연관성을 보이고 있다. 자세한 결과는 [그림 4-15]와 <표 4-15>를 참조하기 바란다.



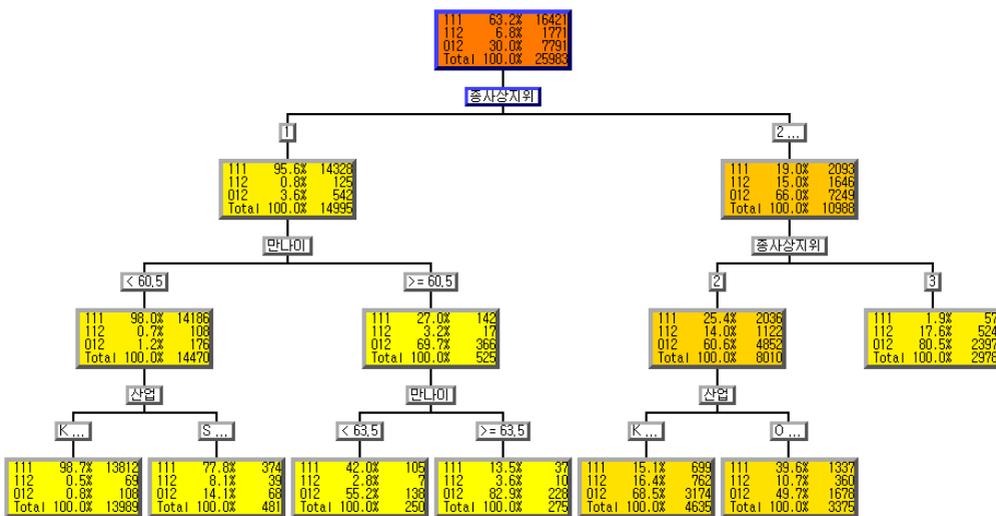
[그림 4-15] 주된취업사유에 관한 연관성 모형

<표 4-15> 주된취업사유에 관한 연관성 분석

주된취업사유			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	일자리취업동기	1	2
2	만나이(1) 만나이(2)	20.5미만 26.5미만	20.5이상 26.5이상
3	종사상지위(1) 종사상지위(2) 36미만일이유(3) 혼인상태(4)	1 1 00, 17, 23, 30 2, 3, 4	2, 3 2, 3 그 외 이유 1

16) 연금

연금가입 항목은 직장가입자(111)가 63.2%, 지역가입자(112)가 6.8%, 아니오(012)가 30.0%로 구성되어 있다. 이 항목은 종사상지위와 가장 큰 연관성을 가진다. 종사상지위가 상용근로자(1)이면 직장가입자(111)의 가능성이 95.6%가 되며, 임시근로자(2)나 일용근로자(3)인 경우에는 상대적으로 아니오(012)의 비율이 높아진다. 그리고 만나이도 연금가입과 연관성이 높는데 상용근로자(1) 중에서 60.5(세)미만이면 98.0%가 직장가입자(111)임을 알 수 있다. 또한 산업분류도 연금가입에 영향을 주는 것으로 나타났다. 자세한 결과는 [그림 4-16]과 <표 4-16>을 참조하기 바란다.



[그림 4-16] 연금에 관한 연관성 모형

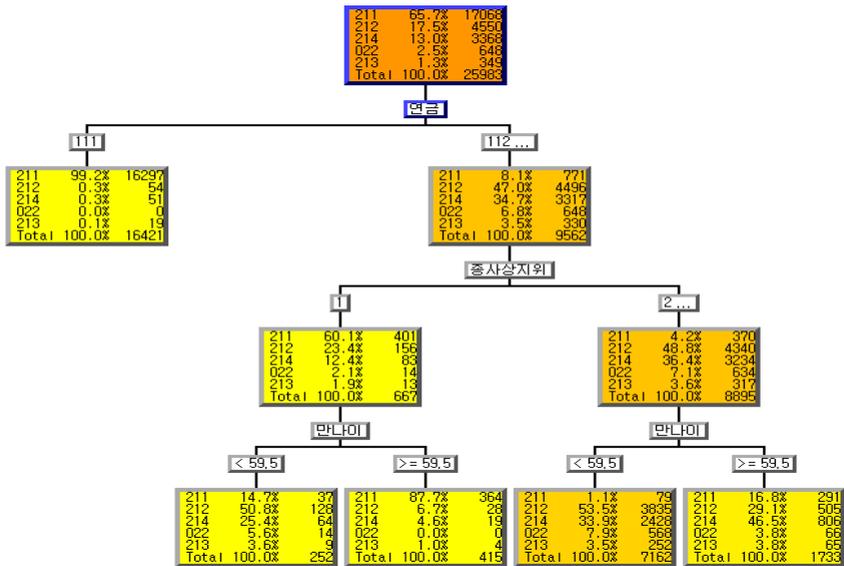


<표 4-16> 연금에 관한 연관성 분석

연금			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	1	2, 3
2	만나이(1) 종사상지위(2)	60.5미만 2	60.5이상 3
3	산업(1) 만나이(2) 산업(3)	그 외 산업 63.5미만 K.....	A, S 63.5이상 O.....

17) 건강보험

건강보험 항목은 직장가입자(211)가 65.7%, 지역가입자(212)가 17.5%, 의료수급권자(213)가 1.3%, 직장가입피부양자(214)가 13.0%, 아니오(22)가 2.5%로 구성되어 있다. 이 항목은 연금 항목과의 관련성이 크기 때문에 연금 항목을 특별 대체군에 포함시키고자 한다. 연금이 직장가입자(111)이면 건강보험도 99.2%가 직장가입자(211)인 것을 볼 수 있다. 다음으로 연관성이 높은 항목은 종사상지위인데 연금이 지역가입자(112)나 아니오(12)이면서 종사상지위가 임시근로자(2)나 일용근로자(3)이면 건강보험이 지역가입자(212) 또는 직장가입피부양자(214)일 가능성이 매우 높아진다. 그리고 만나이도 연관성이 있는 것으로 보인다. 자세한 결과는 [그림 4-17]과 <표 4-17>을 참조하기 바란다.



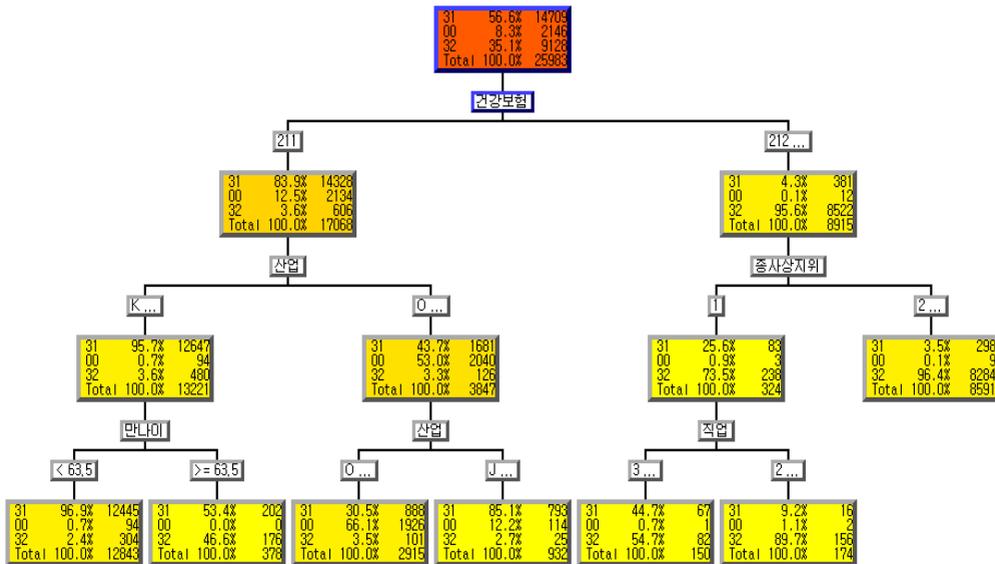
[그림 4-17] 건강보험에 관한 연관성 모형

<표 4-17> 건강보험에 관한 연관성 분석

건강보험			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	연금	111	012, 112
2	종사상지위(2)	1	2, 3
3	만나이(3) 만나이(4)	59.5미만 59.5미만	59.5이상 59.5이상

18) 고용보험

고용보험 항목은 조사대상아님(0)이 8.3%, 예(31)가 56.6%, 아니오(32)가 35.1%로 구성되어 있다. 이 항목은 건강보험 항목과의 관련성이 크기 때문에 건강보험 항목을 특별 대체군에 포함시키하고자 한다. 또한 공무원, 사립학교 교직원, 별정우체국 직원은 조사대상이 아니므로 이 직업을 인식할 수 있는 항목(직업세부코드)을 대체군에 포함시켜야 한다. 건강보험이 직장가입자(211)이면 예(1)인 경우가 83.9%이며 반대의 경우 아니오(32)인 경우가 95.6%인 것을 볼 수 있다. 건강보험이 직장가입자(211)가 아니더라도 종사상 지위가 상용근로자(1)이면 예(1)일 가능성이 더 높아진다. 그리고 산업분류도 연관성이 있는 것으로 보인다. 자세한 결과는 [그림 4-18]과 <표 4-18>을 참조하기 바란다.



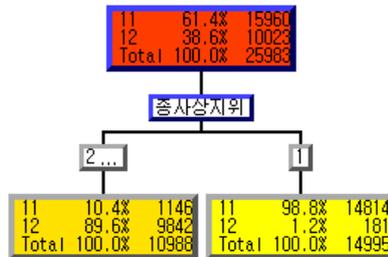
[그림 4-18] 고용보험에 관한 연관성 모형

<표 4-18> 고용보험에 관한 연관성 분석

고용보험			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	건강보험	211	022, 212, 213, 214
2	산업(1) 종사상지위(2)	k..... 1	0..... 2, 3
3	만나이(1) 산업(2) 직업(3)	63.5미만 O, P 1, 3, 4, 7~9	63.5이상 D, J, R 2, 5, 6

19) 퇴직금

퇴직금 항목의 구성비는 예(11)가 61.4%, 아니오(12)가 38.6%로 구성되어 있다. 연관성이 있는 항목은 종사상지위인데 종사상지위가 상용근로자(1)이면 98.8%가 퇴직금을 받으며, 임시근로자(2)나 일용근로자(3)인 경우에는 89.6%가 퇴직금을 받지 못하는 것으로 나타났다. 자세한 결과는 [그림 4-19]와 <표 4-19>를 참조하기 바란다.



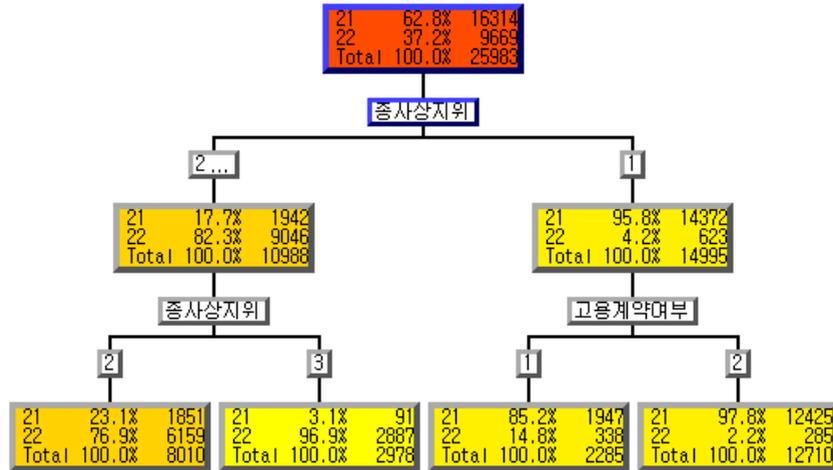
[그림 4-19] 퇴직금에 관한 연관성 모형

<표 4-19> 퇴직금에 관한 연관성 분석

퇴직금			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1

20) 상여금

상여금 항목의 구성비는 예(21)가 62.8%, 아니오(12)가 37.2%로 구성되어 있다. 이 항목은 종사상지위와 가장 큰 연관성을 가지는데 종사상지위가 상용근로자(1)이면 95.8%가 상여금을 받으며, 임시근로자(2)는 76.9%, 일용근로자(3)는 96.9%가 상여금을 받지 못하는 것으로 나타났다. 또한 상용근로자(1) 중에서 고용계약여부가 정하지 않았음(2)이면 상여금을 받을 가능성이 조금 더 높아진다. 자세한 결과는 [그림 4-20]과 <표 4-20>을 참조하기 바란다.



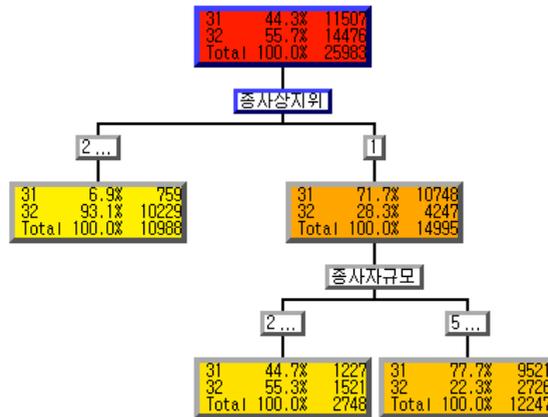
[그림 4-20] 상여금에 관한 연관성 모형

<표 4-20> 상여금에 관한 연관성 분석

상여금			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1
2	종사상지위(1)	2	3
	고용계약여부(2)	1	2

21) 시간외수당

시간외수당 항목의 구성비는 예(31)가 44.3%, 아니오(32)가 55.7%로 구성되어 있다. 이 항목은 종사상지위와 가장 큰 연관성을 가지는데 종사상지위가 상용근로자(1)이면 71.7%가 시간외수당을 받으며, 임시근로자(2)나 일용근로자(3)의 93.1%는 시간외수당을 받지 못하는 것으로 나타났다. 또한 상용근로자(1) 중에서 종사자규모가 10인 이상인 곳은 시간외수당을 받을 가능성이 조금 더 높아진다. 자세한 결과는 [그림 4-21]과 <표 4-21>을 참조하기 바란다.



[그림 4-21] 시간외수당에 관한 연관성 모형

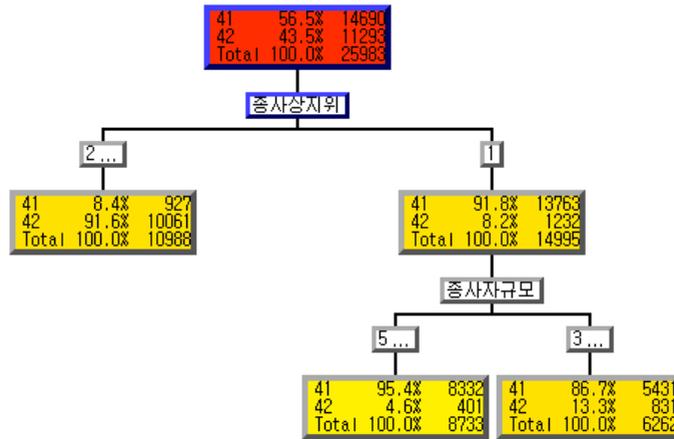
<표 4-21> 시간외수당에 관한 연관성 분석

시간외수당			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1
2	종사자규모(2)	1, 2	3, 4, 5, 6

22) 유급휴가

유급휴가 항목의 구성비는 예(41)가 56.5%, 아니오(42)가 43.5%로 구성되어 있다. 이

항목은 종사상지위와 가장 큰 연관성을 가지는데 종사상지위가 상용근로자(1)이면 91.8%가 유급휴가를 받으며, 임시근로자(2)나 일용근로자(3)의 91.6%는 유급휴가를 받지 못하는 것으로 나타났다. 또한 상용근로자(1) 중에서 종사자규모가 크면 유급휴가를 받을 가능성이 조금 더 높아진다. 자세한 결과는 [그림 4-22]와 <표 4-22>를 참조하기 바란다.



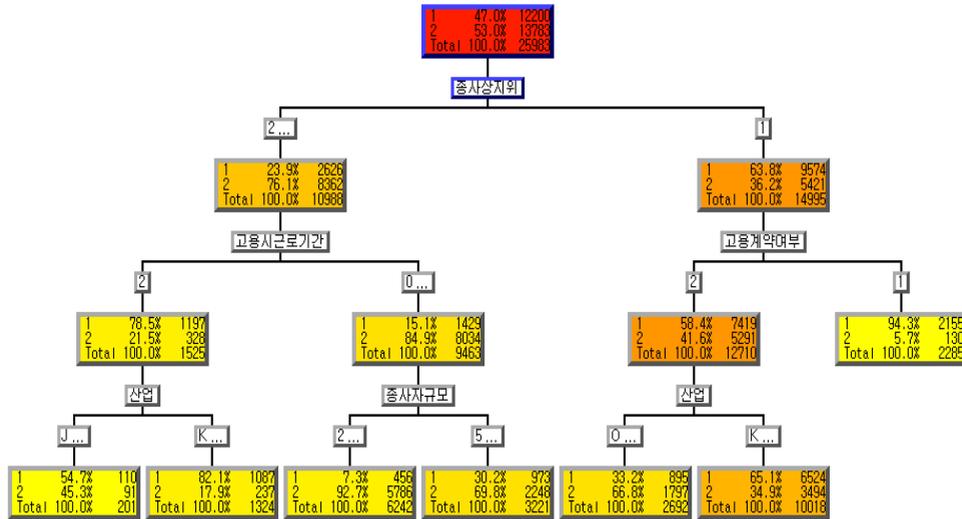
[그림 4-22] 유급휴가에 관한 연관성 모형

<표 4-22> 유급휴가에 관한 연관성 분석

유급휴가			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1
2	종사자규모(2)	4, 5, 6	1, 2, 3

23) 근로계약서작성여부

근로계약서작성여부 항목은 예(1)가 47.0%, 아니오(2)가 53.0%로 구성되어 있다. 이 항목은 종사상지위와 가장 큰 연관성을 가지는데 종사상지위가 상용근로자(1)이면 63.8%가 근로계약서를 작성하였고, 임시근로자(2)나 일용근로자(3)의 76.16%는 근로계약서를 작성하지 못한 것으로 나타났다. 그리고 상용근로자(1) 중에서 고용계약여부가 정했음(1)이면 근로계약서 작성 비율이 94.5%로 상당히 높아진다. 또한 산업분류나 종사자규모도 근로계약서작성여부와 연관성이 나타난다. 자세한 결과는 [그림 4-23]과 <표 4-23>을 참조하기 바란다.



[그림 4-23] 근로계약서작성여부에 관한 연관성 모형

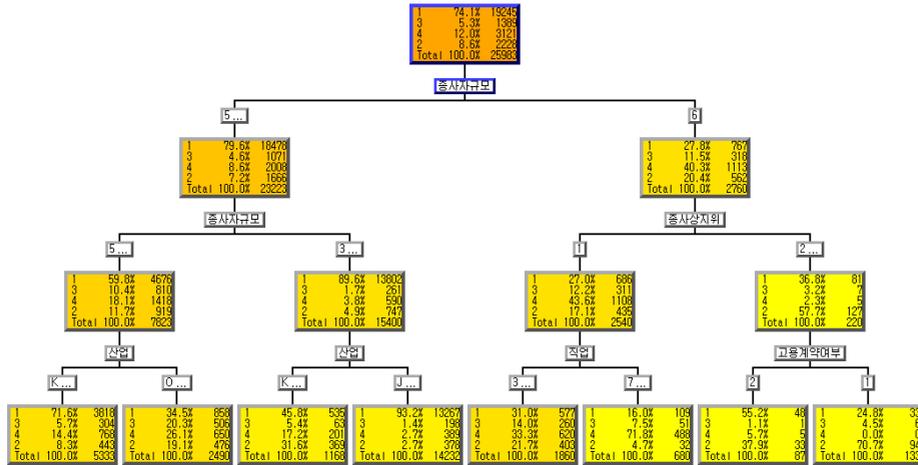
<표 4-23> 근로계약서작성여부에 관한 연관성 분석

근로계약서작성여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1
2	고용시근로기간(1) 고용계약여부(2)	2 2	0, 1~6 1
3	산업(1) 종사자규모(2) 산업(3)	J..... 1, 2 O, P, S, U	K..... 3~6 그 외 산업

24) 노동조합가입여부

노동조합가입여부 항목은 4개 항목으로 구성되는데 노동조합이 없음(1)이 74.1%, 가입대상아님(2)이 8.6%, 가입안함(3)이 5.3%, 가입함(4)이 12.0%를 차지하고 있다. 이 항목과 가장 연관성이 높은 항목은 종사자규모인데 종사자규모가 300이상이면 가입비율이 40.3%로 매우 높아짐을 볼 수 있다. 여기에 종사상지위가 상용근로자(1)이면 가입함(1)의 비율이 더 높아지며, 임시근로자(2)나 일용근로자(3)인 경우에는 가입대상아님(2)의

비율이 57.7%로 높아진다. 또한 고용계약여부를 정한 경우 가입대상아님(2)의 비율이 70.7%로 더 높아진다. 그리고 산업분류도 노동조합가입여부와 연관성을 보이고 있다. 자세한 결과는 [그림 4-24]와 <표 4-24>를 참조하기 바란다.



[그림 4-24] 노동조합가입여부에 관한 연관성 모형

<표 4-24> 노동조합가입여부에 관한 연관성 분석

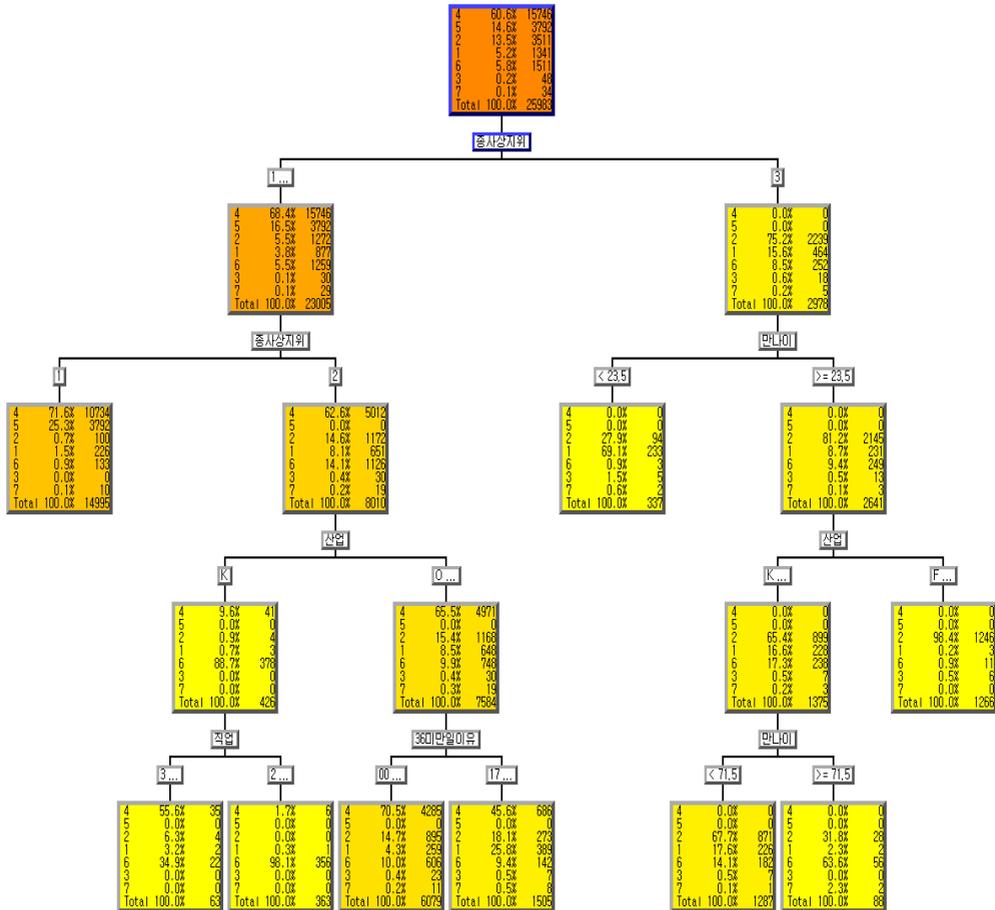
노동조합가입여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	중사자규모	1~5	6
2	중사자규모(1) 중사상지위(2)	4, 5 1	1, 2, 3 2, 3
3	산업(1) 산업(2) 직업(3) 고용계약여부(4)	그 외 산업 D, K, O 1~6, 9 2	D, H, M, O, P 그 외 산업 7, 8 1

25) 임금형태

임금형태 항목은 7개 항목으로 구성되는데 시급제(1)는 5.2%, 일급제(2)는 13.5%, 주급제(3)는 0.2%, 월급제(4)는 60.6%, 연봉제(5)는 14.6%, 실적급제(6)는 5.8%, 기타(7)는



0.1%를 차지하고 있다. 이 항목과 가장 연관성이 높은 항목은 종사상지위인데 종사상 지위가 일용근로자(3)이면 임금형태가 일급제(2)의 비율이 75.2%로 높아진다. 반면에 상용근로자(1)는 월급제(4)의 비율이 더 높아진다. 만나이도 임금형태와 연관성이 있는데 23.5(세)의 경우 시급제(1)의 비율이 상대적으로 높아진다. 산업분류에 따라서도 차이가 생기는데 K(금융 및 보험업)의 임시근로자의 88.7%는 실적급제(6)의 임금형태를 보인다. 그리고 직업에서의 연관성도 나타난다. 자세한 결과는 [그림 4-25]와 <표 4-25>를 참조하기 바란다.



[그림 4-25] 임금형태에 관한 연관성 모형

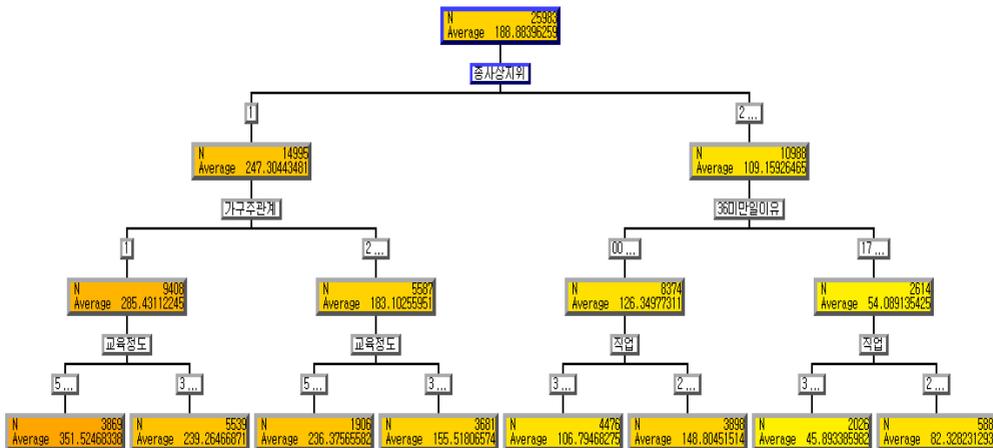
<표 4-25> 임금형태에 관한 연관성 분석

임금형태			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	1, 2	3
2	종사상지위(1) 만나이(2)	1 23.5미만	2 23.5이상
3	산업(2) 산업(4)	K 그 외 산업	그 외 산업 A, F
4	직업(3) 36미만일이유(4) 만나이(7)	1, 3, 4 그 외 이유 71.5미만	2, 5, 9 11, 13~17, 22 71.5이상



26) 최근3개월평균임금

최근3개월임금의 전체평균은 188.88(만원)이며, 종사상지위가 가장 큰 연관성을 가진다. 종사상지위가 상용근로자(1)이면 247.30(만원), 임시근로자(2)나 일용근로자(3)이면 109.16(만원)의 임금을 평균적으로 받는다. 상용근로자(1)이면서 가구주(1) 경우에는 평균 임금이 285.43(만원)으로 올라간다. 또한 교육정도가 대졸이상인 경우가 고졸이하인 경우보다 훨씬 많은 임금을 받는 것을 볼 수 있다. 그리고 직업, 36미만일이유, 만나이드 임금이 많은 영향을 주는 것을 알 수 있다. 자세한 결과는 [그림 4-26]과 <표 4-26>을 참조하기 바란다.



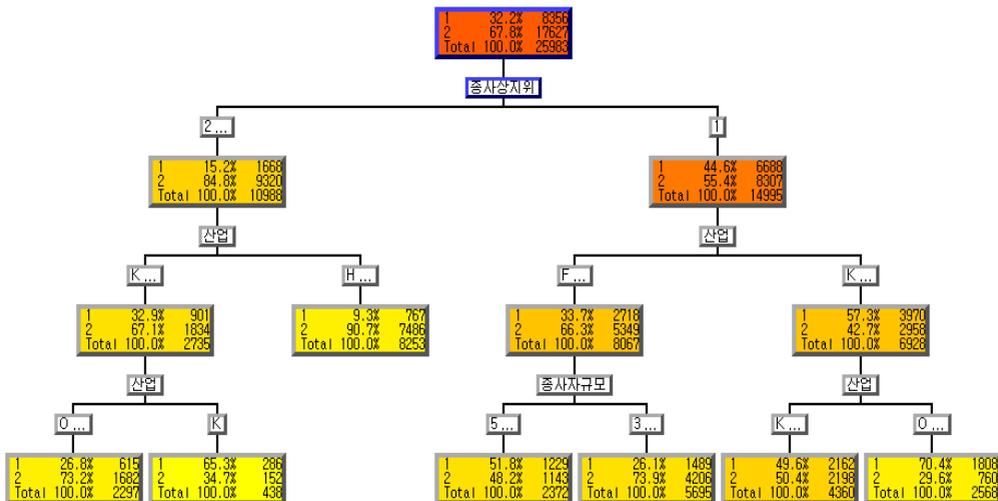
[그림 4-26] 최근3개월평균임금에 관한 연관성 모형

<표 4-26> 최근3개월평균임금에 관한 연관성 분석

최근3개월임금			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	1	2, 3
2	가구주관계(1) 36미만일이유(2)	1 그 외 이유	2, 3, 4, 6, 8, 9 11~18, 21
3	교육정도(1) 교육정도(2) 직업(3) 직업(4)	5, 6 5, 6 3, 4, 6, 9 3, 4, 5, 6, 9	1~4 1~4 1, 2, 5, 7, 8 1, 2, 7, 8
4	종사자규모(2) 만나이(4)	1~5 37.5미만	6 37.5이상

27) 교육훈련경험여부

교육훈련경험여부 항목은 예(1)가 32.2%, 아이오(2)가 67.8%로 구성되어 있다. 교육훈련경험여부와 가장 큰 연관성을 가지는 항목은 종사상지위이다. 종사상지위가 임금근로자(1)이면 예(1)일 가능성은 44.6%로 올라가며, 임시근로자(2)나 일용근로자(3)이면 아니오(2)일 가능성이 84.8%로 나타났다. 그리고 산업분류도 교육훈련경험여부와 높은 연관성을 보이고 있다. 또한 종사자규모가 100인 이상인 경우 교육훈련경험이 있을 가능성이 높아짐을 볼 수 있다. 자세한 결과는 [그림 4-27]과 <표 4-27>을 참조하기 바란다.



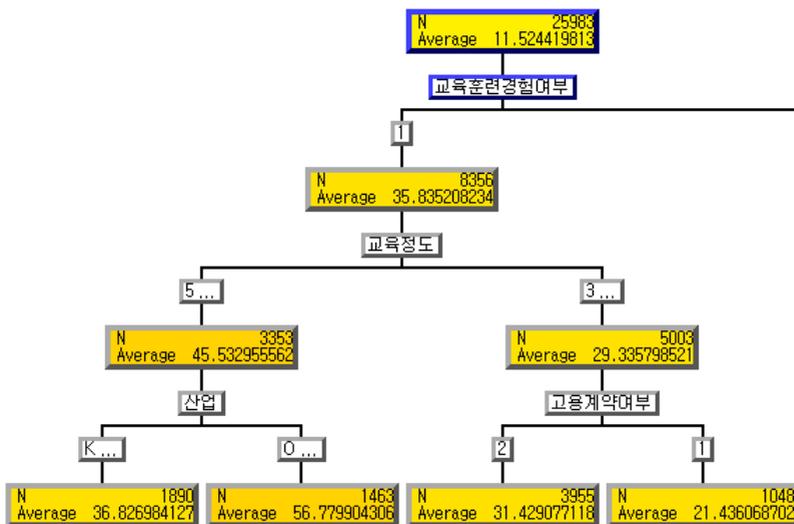
[그림 4-27] 교육훈련경험여부에 관한 연관성 모형

<표 4-27> 교육훈련경험여부에 관한 연관성 분석

교육훈련경험여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	종사상지위	2, 3	1
2	산업(1) 산업(2)	K..... F.....	H..... K.....
3	산업(1) 종사자규모(3) 산업(4)	J, L, O, P, Q 5, 6 K.....	K 1~4 O, P

28) 총교육시간

총교육시간 항목은 교육훈련경험여부를 예(1)라고 대답한 경우에만 응답을 하는 항목으로 교육훈련경험여부 항목을 특별 대체군에 포함시킨다. 따라서 교육훈련경험여부를 예(1)라고 대답한 경우의 평균은 35.84(시간)이며, 아니오(2)인 경우는 평균이 0인 것을 알 수 있다. 다음은 교육정도와 연관성을 가지는데 교육정도가 대학교 이상이면 45.53(시간), 고졸 이하이면 29.34(시간)의 교육시간을 평균적으로 갖는다. 또한 산업분류나 고용계약여부도 총교육시간과 연관성이 있는 것으로 나타났다. 자세한 결과는 [그림 4-28]과 <표 4-28>을 참조하기 바란다.



[그림 4-28] 총교육시간에 관한 연관성 모형

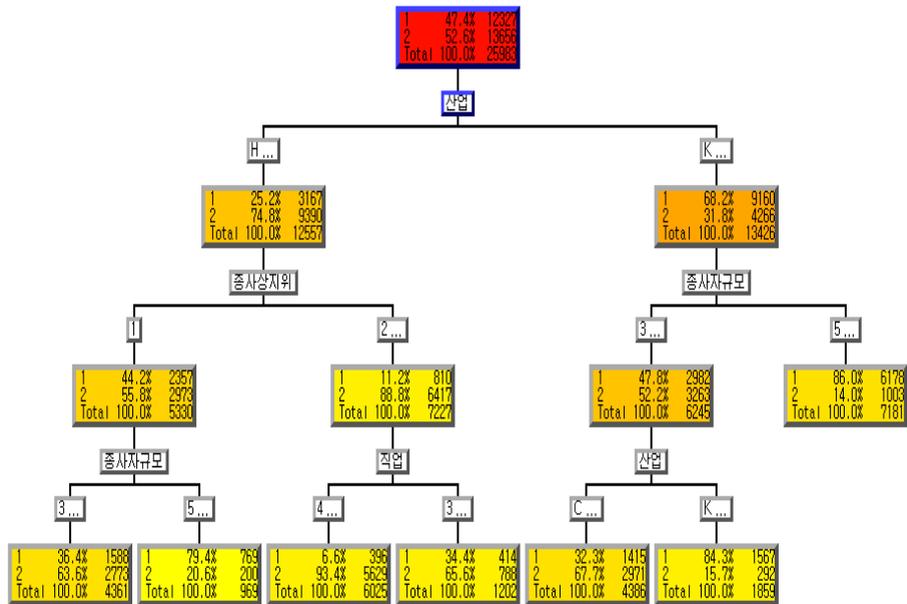


<표 4-28> 총교육시간에 관한 연관성 분석

총교육시간			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	교육훈련경험여부	1	2
2	교육정도(1)	5, 6	1~4
3	산업(1) 고용계약여부(2)	그 외 산업 2	A, F, I, O, P 1

29) 주5일근로여부

주5일근로여부 항목은 실시함(1)이 47.4%, 실시안함(2)이 52.6%로 구성되어 있다. 주5일근로여부와 가장 큰 연관성을 가지는 항목은 산업분류이다. 그리고 산업분류에 따라서 종사상지위와 종사자규모도 큰 영향을 주는데 종사상지위가 임시근로자(2)나 일용근로자(3)이면 실시안함(2)일 가능성이 높으며, 종사자규모가 100인 이상인 경우에는 실시함(1)일 가능성이 높은 것으로 나타났다. 또한 직업과 총취업시간도 조금의 영향을 주는 것을 알 수 있다. 자세한 결과는 [그림 4-29]와 <표 4-29>를 참조하기 바란다.



[그림 4-29] 주5일근로여부에 관한 연관성 모형

〈표 4-29〉 주5일근로여부에 관한 연관성 분석

주5일근로여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	산업	H.....	K.....
2	종사상지위(1) 종사자규모(2)	1 1, 2, 3	2, 3 4, 5, 6
3	종사자규모(1) 직업(2) 산업(3)	1~4 1, 4-9 B, C, E, N, Q	5, 6 2, 3 D, J, K, M, O
4	총취업시간(5)	41.5미만	41.5이상

이상으로 연관성분석을 통하여 각 항목에 대한 대체군을 제시하였다. 각 항목들은 종사상지위, 산업, 종사자규모, 직업, 만나이, 고용계약여부 등 주요 항목들에 대하여 공통적인 대체군을 갖는 것으로 나타났다. 이는 동시 대체를 위한 대체군과 매우 유사하게 선택되었음을 알 수 있다.

그리고 이 결과를 일반화하기 위해서 타 연도 자료를 이용한 결과와 비교하고자 한다. 본 자료는 2010년 8월 부가조사로 2010년 3월 자료와 2011년 3월 자료를 이용하여 대체군을 선정함으로써 연도와 계절적인 상황이 대체군에 영향을 주는지를 검토하고자 한다. 대체군 선정의 결과가 <표 4-30>에 주어졌다.

세 자료를 이용하여 대체군을 선정한 결과는 매우 유사하게 나타났다. 대체군 중 높은 연관성을 가진 항목일수록 대체군의 변화는 없는 것으로 보인다. 단지 연관성이 낮은 항목들에서 기존의 대체군에서 제외되던지 또는 새로운 대체군으로 편입되는 현상이 몇 항목에서 나타났다. 몇 가지 예를 들면 계속근로가능여부(이유) 항목의 경우에는 기존의 대체군에서 산업 항목이 추가되었으며, 향후기대근속기간(1년이하) 항목의 경우에는 종사상지위 항목이 제외된 것을 볼 수 있다. 그리고 평소근로시간 항목은 종사자규모 항목이, 건강보험 항목은 가구주와의관계 항목이, 시간외수당 항목은 산업 항목이 기존의 대체군에서 추가되었다. 하지만, 추가되거나 제외된 항목들은 대체군으로써의 역할이 크지 않는 연관성이 낮은 항목이므로 대체의 정확도에는 큰 차이가 없을 것으로 판단된다. 그러므로 본 연구에서 제시되는 대체군을 타 연도 자료에 계속 적용하더라도 큰 차이가 없을 것으로 생각된다.



〈표 4-30〉 각 항목에 대한 대체군 선정 결과

항목	연도	대체군
계약기간반복갱신여부	2010(3월)	고용계약여부, 고용시근로기간, 만나이
	2010(8월)	고용계약여부, 고용시근로기간, 만나이, 종사상지위
	2011(3월)	고용계약여부, 고용시근로기간, 만나이, 종사상지위
단기근로기간형태	2010(3월)	고용계약여부, 종사상지위, 산업
	2010(8월)	고용계약여부, 종사상지위, 산업
	2011(3월)	고용계약여부, 종사상지위, 산업
계속근로가능여부	2010(3월)	단기근로기간형태, 고용시근로기간, 산업
	2010(8월)	단기근로기간형태, 고용시근로기간, 산업
	2011(3월)	단기근로기간형태, 고용시근로기간, 산업
계속근로가능여부 (이유)	2010(3월)	계속근로가능여부, 고용계약여부, 종사상지위, 산업
	2010(8월)	계속근로가능여부, 고용계약여부, 종사상지위
	2011(3월)	계속근로가능여부, 고용계약여부, 종사상지위, 산업
향후기대근속기간	2010(3월)	계속근로가능여부, 고용시근로기간
	2010(8월)	계속근로가능여부, 고용시근로기간, 만나이
	2011(3월)	속근로가능여부, 고용시근로기간, 만나이
향후기대근속기간 (1년이하)	2010(3월)	향후기대근속기간, 고용시근로기간
	2010(8월)	향후기대근속기간, 고용시근로기간, 종사상지위
	2011(3월)	향후기대근속기간, 고용시근로기간
근속기간제한이유	2010(3월)	향후기대근속기간, 고용계약여부, 종사상지위, 만나이
	2010(8월)	향후기대근속기간, 고용계약여부, 종사상지위, 만나이
	2011(3월)	향후기대근속기간, 고용계약여부, 종사상지위, 만나이
근로시간형태	2010(3월)	36미만일이유
	2010(8월)	36미만일이유
	2011(3월)	36미만일이유, 총취업시간(주업시간)
평소근로시간	2010(3월)	36미만일이유, 총취업시간(주업시간), 산업, 종사자규모
	2010(8월)	36미만일이유, 총취업시간, 산업
	2011(3월)	36미만일이유, 총취업시간(주업시간), 산업, 종사자규모
급여받은곳	2010(3월)	산업, 직업, 고용계약여부, 고용시근로기간
	2010(8월)	산업, 직업, 고용계약여부
	2011(3월)	산업, 직업, 고용계약여부
근무사업체	2010(3월)	급여받은곳, 직업, 만나이
	2010(8월)	급여받은곳, 직업, 종사자규모
	2011(3월)	급여받은곳, 직업, 만나이, 종사자규모

〈표 4-30〉 각 항목에 대한 대체군 선정 결과(계속)

항목	연도	대체군
특수고용여부	2010(3월)	직업, 산업, 종사상지위, 종사자규모, 혼인상태, 고용계약여부
	2010(8월)	직업, 산업, 종사상지위, 종사자규모, 혼인상태
	2011(3월)	직업, 산업, 종사상지위, 종사자규모, 혼인상태
가정내근로여부	2010(3월)	종사상지위, 산업
	2010(8월)	종사상지위
	2011(3월)	종사상지위, 산업
일자리취업동기	2010(3월)	종사상지위, 직업, 고용계약여부
	2010(8월)	종사상지위, 직업, 고용계약여부
	2011(3월)	종사상지위, 직업, 고용계약여부, 만나이
주된취업사유	2010(3월)	일자리취업동기, 만나이, 종사상지위, 혼인상태, 활동상태
	2010(8월)	일자리취업동기, 만나이, 종사상지위, 36미만일이유, 혼인상태
	2011(3월)	일자리취업동기, 만나이, 종사상지위, 36미만일이유, 혼인상태
연금	2010(3월)	종사상지위, 만나이, 산업, 고용계약여부
	2010(8월)	종사상지위, 만나이, 산업
	2011(3월)	종사상지위, 만나이, 산업, 직업
건강보험	2010(3월)	연금, 종사상지위, 만나이, 가구주와의관계
	2010(8월)	연금, 종사상지위, 만나이
	2011(3월)	연금, 종사상지위, 만나이, 가구주와의관계
고용보험	2010(3월)	건강보험, 산업, 종사상지위, 만나이, 산업
	2010(8월)	건강보험, 산업, 종사상지위, 만나이, 직업
	2011(3월)	건강보험, 산업, 종사상지위, 만나이, 고용시근로기간
퇴직금	2010(3월)	종사상지위
	2010(8월)	종사상지위
	2011(3월)	종사상지위
상여금	2010(3월)	종사상지위, 고용계약여부
	2010(8월)	종사상지위, 고용계약여부
	2011(3월)	종사상지위, 고용계약여부
시간외수당	2010(3월)	종사상지위, 종사자규모, 산업
	2010(8월)	종사상지위, 종사자규모
	2011(3월)	종사상지위, 종사자규모, 산업
유급휴가	2010(3월)	종사상지위, 종사자규모, 고용시근로기간
	2010(8월)	종사상지위, 종사자규모
	2011(3월)	종사상지위, 종사자규모



〈표 4-30〉 각 항목에 대한 대체군 선정 결과(계속)

항목	연도	대체군
근로계약서작성여부	2010(3월)	종사상지위, 고용시근로기간, 고용계약여부, 산업, 종사자규모
	2010(8월)	종사상지위, 고용시근로기간, 고용계약여부, 산업, 종사자규모
	2011(3월)	종사상지위, 고용시근로기간, 고용계약여부, 산업, 종사자규모
노동조합가입여부	2010(3월)	종사자규모, 종사상지위, 산업, 직업, 고용계약여부
	2010(8월)	종사자규모, 종사상지위, 산업, 직업, 고용계약여부
	2011(3월)	종사자규모, 종사상지위, 산업, 직업, 고용계약여부
임금형태	2010(3월)	종사상지위, 만나이, 산업, 직업, 36미만일이유, 종사자규모
	2010(8월)	종사상지위, 만나이, 산업, 직업, 36미만일이유
	2011(3월)	종사상지위, 만나이, 산업, 직업, 36미만일이유
최근3개월평균임금	2010(3월)	종사상지위, 가구주관계, 성별, 교육정도, 직업, 종사자규모, 만나이
	2010(8월)	종사상지위, 가구주관계, 36미만일이유, 교육정도, 직업, 종사자규모, 만나이
	2011(3월)	종사상지위, 가구주관계, 성별, 36미만일이유, 교육정도, 직업, 종사자규모, 만나이
교육훈련경험여부	2010(3월)	종사상지위, 산업, 종사자규모, 직업, 고용계약여부
	2010(8월)	종사상지위, 산업, 종사자규모
	2011(3월)	종사상지위, 산업, 종사자규모, 직업, 고용시근로기간
총교육시간	2010(3월)	교육훈련경험여부, 산업, 고용계약여부, 직업
	2010(8월)	교육훈련경험여부, 교육정도, 산업, 고용계약여부
	2011(3월)	교육훈련경험여부, 교육정도, 산업, 고용계약여부
주5일근로여부	2010(3월)	산업, 종사상지위, 종사자규모, 직업, 총취업시간(주업시간)
	2010(8월)	산업, 종사상지위, 종사자규모, 직업, 총취업시간
	2011(3월)	산업, 종사상지위, 종사자규모, 직업, 총취업시간(주업시간)

제3절 무응답 대체방법

무응답 대체방법은 무응답 항목의 대체값으로 한 개의 값을 부여하는 단일 대체방법(single imputation)과 여러 개의 값을 대체하는 다중 대체방법(multiple imputation,

Rubin(1987))으로 구분되며, 단일 대체방법은 무응답 항목에 유일하게 결정된 대체값을 대입하는 결정적 대체방법(deterministic imputation)과 대체값을 확률적으로 결정하여 대입하는 확률적 대체방법(stochastic imputation)으로 구분된다. 이러한 방법들에 대한 자세한 설명은 『무응답 처리를 위한 방법론 연구(Ⅰ)(통계개발원, 2009)』를 참조하기 바란다.

이 절에서는 본 연구에서 사용될 방법인 응용 핫덱 대체방법에 대하여 설명하고자 한다. 이전에 언급하였듯이 경제활동인구조사 부가조사는 대부분의 항목이 범주형 자료이므로 핫덱 방법으로 대체를 실시하면 될 것으로 판단된다. 또한 여러 핫덱 방법 중에서 응용 핫덱 대체방법의 우수성을 이전의 연구에서 검증하였으므로 이 방법을 이용하여 대체를 하고자 한다. 그리고 응용 핫덱 방법과 다른 방법들을 비교하여 설명하기 위해 본 절에서는 몇 가지 방법을 추가하여 간략하게 예를 들어 설명할 것이다.

<표 4-31> 설명을 위한 예시 자료

사업체	대체항목(연속형)	대체군1(범주형)	대체군2(연속형)
1	25	A	94
2	29	B	206
3	24	B	66
4	39	A	295
5	37	B	232
6	28	A	172
7	40	B	370
8	27	A	81
9	40	B	292
10	34	B	206
11	30	A	93
12	39	B	310
13	38	B	236
14	28	A	111
15	missing	B	301
16	38	B	303
17	40	B	315
18	26	A	166
19	24	A	77
20	34	B	222

본 예제에서는 대체해야 하는 연속형 항목에 대하여 범주형과 연속형 항목의 대체군을 가지고 있는 경우이다. 경제활동인구조사 부가조사는 3개의 항목만이 연속형 자료이지만 범주형 항목으로 설명하지 않는 이유는 다양한 방법들을 예제에서 소개하기 위함이다. 만약, 범주형 항목으로 설명한다면 회귀, 비 평균 대체방법들로 설명할 수 없으므로 다양한 방법으로 설명 가능한 연속형 항목의 대체과정을 보여주고자 한다.

1. 응용 핫덱 대체방법(applied hot-deck imputation)

랜덤 핫덱 방법을 응용한 것으로 연속형 항목에도 자유롭게 적용할 수 있게 한 방법이다. 대체군의 적용시 범주형 항목은 항목값의 일치여부를 판단하여 점수를 부여하고, 연속형 항목은 범주화하지 않고 신뢰구간의 개념을 이용하여 그 구간 안에 들어가면 비슷한 개체로 판단하여 점수를 부여한다. 모든 대체군의 항목들에 대하여 비교한 후 가장 점수가 높은 개체들을 선택하여 그 중에서 하나의 개체를 임의로 대체하는 방법이다.

예제의 경우 대체군1이 B인 사업체에 1점을 부여하고, 대체군2의 값이 286~316사이에 있는 사업체에 1점을 부여하게 된다. 이는 평균 대체방법에서 설명한 내용으로 대체군2의 경우 연속형 항목이므로 301이 여러 개 존재하지 않는다면 301과 유사($\pm 5\%$)한 값들을 하나의 집단으로 묶어 같은 대체군으로 간주하였다. 마지막으로 점수가 가장 높은 사업체를 도너로 선택하게 되는데 9, 12, 16, 17번째 사업체가 2점으로 가장 높은 점수를 획득하였다. 이 때 최고 점수가 2개 이상이면 이들 중 랜덤하게 하나의 사업체를 선택하여 대체하게 된다. 따라서 15번째 사업체는 40, 39, 38, 40 중 하나의 값을 가지게 되며 40은 2개이므로 선택될 가능성이 더 높다는 것을 알 수 있다.

2. 평균 대체방법(mean imputation)

무응답 항목에 목표변수의 전체 평균을 대입하거나 또는 대체군 내의 평균을 대입하는 방법이다. 이 방법은 간단하여 이용되기 쉬운 장점이 있으며, 항목이 양적 변수이고 구하고자 하는 통계량이 평균일 때 유용하다. 그러나 대체 후의 값들은 평균값의 빈도수가 지나치게 많아져 응답값들의 분포가 왜곡되고, 중위수나 백분위수와 같은 평균이 아닌 통계량을 구할 때에는 효율이 저하되는 단점이 있다.

예제의 경우 15번째 사업체는 대체군1은 B, 대체군2는 301의 값을 가진다. 따라서 이 두 값을 가지고 있는 사업체를 찾아 대체하고자 하는 값의 평균을 구하면 될 것이다. 여

기서 한 가지 고려할 사항은 대체군2의 경우 연속형 항목이므로 301이 여러 개 존재하지 않는다면 301과 유사($\pm 5\%$)한 값들을 하나의 집단으로 묶어 생각하면 될 것이다. 이것은 이후에 설명할 응용 핫데크 방법에서 사용한 것으로 효율성은 입증되었다.

이러한 절차를 통하여 최종적으로 선택되는 도너는 9, 12, 16, 17번째 사업체이며, 이들의 평균값 $(40+39+38+40)/4=39.25(39)$ 이 최종 대체값이 됨을 알 수 있다.

3. 회귀 대체방법(regression imputation)

회귀 대체는 대체하고자 하는 항목에 대하여 가능한 보조변수를 이용하여 회귀모형에 적합시키는 방법으로 목표변수와 보조변수의 관계가 절편이 있는 직선 관계이고 목표변수의 분산이 동일할 때 유용한 방법이다. Kalton(1982)은 i 번째 결측값의 대체값을 다음과 같이 구하였다.

$$\hat{y} = \hat{\beta}_0 + \sum_{k=1}^K \hat{\beta}_k x_{ki} + \epsilon_i$$

여기서 $\hat{\beta}$ 는 일반적 최소제곱법에 의해 추정된 계수이다. 이 방법은 미국의 인구조사(CPS: current population survey)에서 발생하는 결측값을 대체하기 위해서 이용하였으며, 그 결과 대체된 값과 실제값의 평균절대편차를 비교할 때 다른 대체 방법에 비해 매우 적절함을 보였다.

예제의 경우 대체군1이 B인 11개의 자료를 이용하여 회귀모형에 적합시킨다. 실제 자료에서는 11개보다 훨씬 많은 자료를 이용하므로 자료의 수에 대한 문제는 고려하지 않는다. 이러한 과정에 의해 추정된 모형은 $\hat{y} = 21.214 + 0.058x$ 가 된다. 따라서 x 의 값에 301을 넣어 얻은 최종 대체값은 38.672(39)가 됨을 알 수 있다.

4. 비 대체방법(ratio imputation)

비 대체방법은 이용 가능한 보조변수가 있을 때 목표변수와 보조변수의 관계를 이용하여 무응답을 대체하는 방법으로 목표변수와 보조변수가 원점을 지나는 직선관계이며, 분산이 보조변수에 비례하는 경우 효과적인 것으로 나타났다. 유한 모집단에서 예측이론에 의하면 비 대체방법이 매우 우수한 것으로 알려져 있으나 항목들이 양적변수일 때에만 사용가능하다. 만일 목표변수와 보조변수가 비례관계이면 다음과 같다.



$$\hat{y} = \left(\frac{\bar{y}_R}{\bar{x}_R} \right) x_k$$

여기서 R은 응답이 있는 자료이다. 회귀 대체방법과 마찬가지로 대체군1이 B인 11개의 자료를 이용하여 대체값을 추정한다. 이 때 \bar{y}_R 은 35.73이며, \bar{x}_R 은 250.73이다. 따라서 15번째 사업체의 x 값이 301이므로 추정값은 $y = (35.73/250.73)*301 = 42.89(43)$ 이 됨을 알 수 있다.

5. 최빈수 대체방법(mode imputation)

최빈수 대체방법은 평균 및 응용 핫덱 대체방법과 유사하게 도너를 선택한다. 하지만 최종 도너를 선택하는 과정에서 최빈수를 찾아 무응답 항목에 대체하는 것이다. 그러나 이 방법도 평균 대체방법처럼 무응답 비율이 높아질수록 응답값들의 분포가 왜곡될 가능성이 커지는 단점을 가지고 있다.

예제의 경우 9, 12, 16, 17번째 사업체가 도너로 선택이 되며, 이 때 대체 가능한 값은 40, 39, 38, 40이므로 최빈수 40이 실제 대체가 됨을 알 수 있다.

제4절 모의실험 자료 구축

이 절에서는 더 정확한 모의실험을 위해서는 실제 자료와 유사한 형태의 무응답 자료를 만드는 것이 중요한데 이를 위해서 주어진 실제 자료로부터 무응답자들의 성향을 파악하여 모의실험 자료의 구축에 반영할 것이다. 하지만 주어진 자료의 무응답 비율이 대략 2.5%정도이므로 무응답자의 성향 파악을 위한 연구에 한계가 있음을 밝혀 둔다.

1. 응답자와 무응답자의 차이분석

응답자와 무응답자의 성향의 차이를 분석하기 위하여 경제활동인구조사의 항목을 이용한다. 각 항목에의 구성비 차이 및 카이제곱 검정을 통해 두 집단 간에 차이가 존재하는지의 여부를 판단할 것이다. 총 자료의수는 26,664명으로 응답자는 25,983(97.5%)명이며 무응답자는 681(2.5%)명이다. 먼저 몇 가지 항목에 대한 구성비의 차이를 살펴보자.

〈표 4-32〉 혼인상태 항목에 대한 응답자와 무응답자의 구성비 차이

세부항목	구성비(응답자)	구성비(무응답자)	차이(응답자-무응답자)
1(미혼)	23.5%	38.0%	-14.5%
2(유배우)	68.2%	55.5%	12.7%
3(사별)	4.3%	2.2%	2.1%
4(이혼)	4.0%	4.3%	-0.3%

혼인상태 항목에 대한 응답자와 무응답자의 구성비 차이가 <표 4-32>에 정리되어 있다. 응답자 중에서 미혼의 비율은 23.5%이고 무응답자 중에서 미혼의 비율은 38.0%로 무응답자위 구성비가 14.5% 높게 나타났다. 반면에 유배우의 경우는 반대의 결과를 보이고 있다. 응답자 중에서 유배우의 비율은 68.2%이고 무응답자 중에서 유배우의 비율은 55.5%로 응답자의 구성비가 12.7% 더 높게 나타났다. 하지만 사별과 이혼에 대해서는 구성비도 작고 큰 차이를 보이지 않는다. 그러나 전체적으로 판단할 때 부가조사의 응답률은 혼인상태와 연관성이 있으며 무응답자의 미혼의 비율은 응답자에서보다 더 높아지는 것을 볼 수 있다.

종사상지위 항목에 대한 응답자와 무응답자의 구성비 차이가 <표 4-33>에 정리되어 있다. 응답자 중에서 상용근로자의 비율은 57.7%이고 무응답자 중에서 상용근로자의 비율은 51.7%로 응답자의 구성비가 6.0% 높게 나타났다. 반면에 임시근로자의 경우는 반대의 결과를 보이고 있다. 응답자 중에서 임시근로자의 비율은 30.8%이고 무응답자 중에서 임시근로자의 비율은 38.0%로 무응답자의 구성비가 7.2% 더 높게 나타났다. 하지만 일용근로자에 대해서는 구성비도 작고 큰 차이를 보이지 않는다. 그러나 전체적으로 판단할 때 부가조사의 응답률은 종사상지위와 연관성이 있으며 무응답자의 임시근로자의 비율은 응답자에서보다 더 높아지는 것을 볼 수 있다.

〈표 4-33〉 종사상지위 항목에 대한 응답자와 무응답자의 구성비 차이

세부항목	구성비(응답자)	구성비(무응답자)	차이(응답자-무응답자)
1(상용근로자)	57.7%	51.7%	6.0%
2(임시근로자)	30.8%	38.0%	-7.2%
3(일용근로자)	11.5%	10.3%	1.2%

마지막으로 성별의 경우를 살펴보자. 성별 항목에 대한 응답자와 무응답자의 구성비 차이가 <표 4-34>에 정리되어 있다. 응답자 중에서 남자의 비율은 55.1%이고 무응답자 중에서 남자의 비율은 57.7%로 무응답자의 구성비가 2.6% 높게 나타났다. 반면에 여자의 경우는 반대의 결과를 보이고 있다. 하지만 구성비의 차이는 크지 않음을 알 수 있다. 따라서 부가조사의 응답률은 성별과는 연관성이 크지 않은 것으로 판단된다.



〈표 4-34〉 성별 항목에 대한 응답자와 무응답자의 구성비 차이

세부항목	구성비(응답자)	구성비(무응답자)	차이(응답자-무응답자)
1(남자)	55.1%	57.7%	-2.6%
2(여자)	44.9%	42.3%	2.6%

혼인상태, 종사상지위, 성별 항목에 대하여 응답자와 무응답자의 구성비 차이를 통하여 응답률에 영향을 주는지를 살펴보았다. 이러한 차이를 모든 항목에 대하여 살펴보기 위해 카이제곱 검정을 실시하였다. 이전에 언급하였듯이 무응답자의 비율이 너무 낮아 카이제곱 검정 시 자료수가 적은 셀이 많이 발생하므로 피셔의 정확도 검정(Fisher's Exact Probability Test)을 사용하여 분석하였다. 결과가 <표 4-35>에 정리되어 있다.

부가조사의 응답여부와 경제활동인구조사 항목과의 연관성은 매우 높은 것으로 나타났다. 성별, 활동상태, 취업여부, 일시휴직여부, 가구구분과 같은 항목을 제외한 항목들은 응답에 영향을 주는 것으로 보인다. 특히 가구주와의관계, 혼인상태, 산업, 만나이 등은 더 높은 연관성을 보이고 있다. 무응답자의 비율이 낮아 본 결과의 정확도가 조금 떨어질 수도 있지만 연관성의 유무에는 큰 영향을 주지 않을 것으로 판단된다.

〈표 4-35〉 카이제곱 검정 결과

항목	카이제곱	유의확률
시군	6.601	0.010
가구주와의관계	100.679	0.000
성별	1.761	0.185
혼인상태	81.416	0.000
활동상태	1.669	0.402
취업여부	4.912	0.086
일시휴직여부	2.986	0.811
부업여부	6.922	0.031
총취업시간	19.030	0.001
취업구분	20.800	0.000
36미만일이유	31.528	0.025
추가취업	26.217	0.000
산업	52.238	0.000
직업	18.720	0.016
종사자규모	13.673	0.018
종사상지위	16.100	0.000
고용계약여부	16.380	0.000
고용시군로기간	17.480	0.004
가구구분	0.457	0.499
교육정도	24.933	0.000
만나이	50.109	0.000

2. 모의실험에 사용할 항목선정

정확한 모의실험을 위해 자료를 구축할 때 무응답자의 성향을 반영할 것이다. 따라서 응답여부에 영향을 주는 항목들을 살펴보았다. 이러한 항목들 중에서 가장 영향을 줄 수 있는 항목을 선정할 것이다. 이 때 모의실험 자료를 구축하는 과정이 복잡하지 않게 최소한(1~2개)의 항목만을 선택하고자 한다. 그리고 대체군으로 선정된 항목에서만 선정하여 모의실험의 효과가 더 커지게 할 것이다.

먼저 응답여부에 큰 영향을 주면서 대체군으로 사용되는 항목들은 혼인상태, 총취업시간, 산업, 종사상지위, 고용계약여부, 교육정도, 만나이 항목이다. 이 7개 항목 중에서 모의실험에 사용할 항목을 선택하고자 한다. 선택을 위한 방법은 응답여부를 종속변수로 두고 로지스틱 회귀분석과 CHAID 알고리즘을 이용하여 분석한다.

〈표 4-36〉 로지스틱 회귀분석 우도비 검정 결과

항목	카이제곱	유의확률
절편	0.000	.
총취업시간	0.596	0.440
만나이	0.704	0.402
혼인상태	30.73	0.000
산업	44.73	0.000
종사상지위	16.47	0.000
고용계약여부	12.27	0.000
교육정도	9.64	0.086

응답여부와 응답여부에 영향을 주는 항목들의 로지스틱 회귀분석 우도비 검정 결과가 <표 4-36>에 정리되어 있다. 본 검정은 최종모형과 축소모형(최종모형에서 효과 하나를 생략하여 만든 모형) 사이의 차이를 이용하여 항목의 유의성을 검정하는 것이다. 검정 결과 유의확률이 0.000인 혼인상태, 산업, 종사상지위, 고용계약여부 4개의 항목이 본 모형에서 매우 유의한 변수인 것으로 판단된다.

다음으로 CHAID 알고리즘을 이용하여 응답여부에 대한 연관성 분석을 실시하였다. 응답여부와 가장 연관성이 있는 항목은 혼인상태로 나타났다. 미혼(1)과 미혼아님(2,3,4)의 차이가 나타났다. 그리고 고용계약여부에서도 정했음(1)과 정하지 않았음(2)은 응답여부에 영향을 주는 것으로 보인다. 또한 종사상지위와 교육정도도 응답여부와 연관성이 높게 나타남을 볼 수 있다. 자세한 결과는 <표 4-37>을 참조하기 바란다.



<표 4-37> 응답여부에 관한 연관성 분석

응답여부			
깊이(연관성)	분리변수	분리지점(좌)	분리지점(우)
1	혼인상태	2, 3, 4	1
2	고용계약여부(1)	2	1
3	종사상지위(1) 교육정도(2)	1, 3 1, 3, 5, 6	2 2, 4

두 가지 방법으로 응답여부에 많은 영향을 주는 항목들을 살펴본 결과 혼인상태, 종사상지위, 고용계약여부 항목이 공통적으로 제시되었다. 하지만 본 연구에서는 모의실험 자료구축의 과정이 복잡하지 않게 혼인상태와 종사상지위 2개 항목을 이용하기로 한다. 따라서 <표 4-38>의 무응답자의 구성비를 바탕으로 모의실험 자료를 구축할 것이다.

<표 4-38> 응답여부*혼인상태*종사상지위의 구성비

응답여부			혼인상태			
			미혼	유배우	사별	이혼
응답	종사상지위	상용근로자	13.5%	42.0%	0.9%	1.4%
		임시근로자	7.6%	19.1%	2.3%	1.8%
		일용근로자	2.4%	7.1%	1.1%	0.8%
무응답	종사상지위	상용근로자	19.5%	30.4%	0.6%	1.2%
		임시근로자	15.0%	18.9%	1.5%	2.6%
		일용근로자	3.5%	6.2%	0.1%	0.4%

응답자의 경우 미혼인 상용근로자의 구성비는 13.5%, 유배우인 상용근로자는 42.0%, 미혼인 임시근로자는 7.6%이지만 무응답자의 경우에는 19.5%, 30.4%, 15.0%로 나타났다. 각 셀에 따라서 무응답자의 구성비는 응답자와 6~12%정도의 차이를 보이고 있다. 따라서 모의실험 자료는 이러한 결과를 반영하여 실제 자료와 더 유사한 자료로 구축하여 모의실험 결과의 정확성을 높일 것이다.

제5절 모의실험

이 절에서는 제시된 대체군을 이용하여 항목별로 모의실험을 실시하여 대체의 결과를 검토하고자 한다. 모의실험에 사용할 자료는 2010년 8월 기준 경제활동인구조사 근로형태별 부가조사 자료이다. 데이터 파일은 부가조사 자료 중 응답한 자료만을 이용했으며, 모의실험을 위해서 <표 4-38>의 무응답자의 구성비를 바탕으로 5%, 10%, 20%의 결측치를 발생시켰다. 그리고 각 항목별로 대체의 정확 정도를 판단하기 위하여 대체전 후의 구성비 변화를 살펴볼 것이다. 대체방법은 응용 핫덱 대체방법을 이용하며, 부가조사의 모든 항목을 동시에 대체하는 방법과 각 항목별로 대체하는 방법을 비교할 것이다. 모의실험의 결과를 살펴보면 모든 항목들에서 유사한 결과를 보이고 있으므로 몇 가지 항목에 대해서만 설명을 하고 나머지 항목들은 결과만을 제시할 것이다.

1. 계약기간반복갱신여부

<표 4-39> 계약기간 반복갱신여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	21,910 (84.32%)	21,910 (0)	21,910 (0)	14,111,853 (85.27%)	14,111,853 (0)	14,111,853 (0)
1	1,991 (7.66%)	1,995 (0.02)	1,993 (0.01)	1,198,811 (7.24%)	1,200,057 (0.01)	1,200,297 (0.01)
2	2,082 (8.01%)	2,078 (0.01)	2,080 (0.01)	1,239,485 (7.49%)	1,238,239 (0.01)	1,237,998 (0.01)

(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	21,910 (84.32%)	21,910 (0)	21,910 (0)	14,111,853 (85.27%)	14,111,853 (0)	14,111,853 (0)
1	1,991 (7.66%)	2,001 (0.04)	1,999 (0.03)	1,198,811 (7.24%)	1,204,760 (0.04)	1,204,826 (0.04)
2	2,082 (8.01%)	2,072 (0.05)	2,074 (0.04)	1,239,485 (7.49%)	1,233,536 (0.04)	1,233,469 (0.04)



(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	21,910 (84.32%)	21,910 (0)	21,910 (0)	14,111,853 (85.27%)	14,111,853 (0)	14,111,853 (0)
1	1,991 (7.66%)	1,975 (0.06)	1,979 (0.04)	1,198,811 (7.24%)	1,189,984 (0.05)	1,191,239 (0.04)
2	2,082 (8.01%)	2,098 (0.06)	2,094 (0.05)	1,239,485 (7.49%)	1,248,311 (0.05)	1,247,056 (0.05)

부가조사 항목을 동시에 대체하기 위한 대체군은 성별, 혼인상태, 종사자규모, 종사상지위, 고용계약여부, 고용시근로기간, 교육정도, 만나이, 산업, 직업 항목을 사용하였으며, 계약기간반복갱신여부 항목을 대체하기 위한 대체군은 고용계약여부, 고용시근로기간, 만나이, 종사상지위 항목을 사용하였다. 계약기간반복갱신여부 항목에 대한 모의실험 결과가 <표 4-39>에 정리되어 있다.

먼저 동시대체와 개별대체의 결과를 비교해보면 표본수를 고려한 결과에서 개별대체의 정확도가 조금 더 높은 것으로 보이지만 추정수를 고려한 결과에서는 차이가 거의 없는 것으로 판단된다. 이는 사용된 대체군에서의 정보가 거의 유사하기 때문인 것으로 판단된다. 그리고 두 경우 모두 대체전후의 구성비 차이는 매우 작다는 것을 알 수 있다. 무응답비율이 5%인 경우의 구성비차이는 최대 0.01%정도가 발생하며 무응답비율이 높아질수록 차이는 커지지만 20%의 무응답이 발생하더라도 계약기간반복갱신여부 항목의 구성비는 최대 0.05% 이내로 추정할 수 있는 것으로 나타났다. 이 경우 추정 오차율 또한 최대 1%를 초과하지 않는 것으로 나타났다. 실제 적용을 할 때 본 모의실험의 결과보다 좋은 경우도 있고 나쁜 경우도 있겠지만 평균적으로 이 정도의 차이가 발생할 가능성이 높은 것으로 이해하면 될 것이다. 그리고 표본에서의 구성비 변화비율과 추정에서의 구성비 변화비율도 차이가 거의 발생하지 않는 것으로 보인다. 실제 정확도는 추정에서의 결과를 이용하면 될 것이다.

2. 연금

연금 항목을 대체하기 위한 대체군은 종사상지위, 만나이, 산업 항목을 사용하였다. 연금 항목에 대한 모의실험 결과가 <표 4-40>에 정리되어 있다. 계약기간반복갱신여부 항목의 결과와 유사하게 동시대체와 개별대체의 차이는 보이지 않는다. 추정수를 고려한 결과를 살펴보면 무응답비율이 5%인 경우 두 방법 모두 구성비 차이가 0.02~0.03%

정도로 매우 안정적인 대체가 되고 있다. 무응답비율이 20%인 경우에는 구성비 차이가 최대 0.15%까지 커지지만 이 정도의 차이는 매우 작은 변화라고 볼 수 있다. 실제 추정 오차율을 보더라도 0.9% 정도이므로 이전에 연구된 다른 조사의 결과에 비해서 매우 좋은 것을 알 수 있다. 일반적으로 무응답비율이 20% 정도일 경우 추정 오차율은 1~2%는 발생하는 경우가 많다. 이러한 결과는 부가조사라는 특성 때문인 것으로 생각된다. 즉, 본조사의 내용을 안다면 부가조사의 내용을 정확성이 높게 추정 가능한 것으로 여겨진다. 이는 대부분의 다른 항목에서도 나타나는 현상이다. 그리고 이전의 경우와 마찬가지로 표본에서의 구성비 변화비율과 추정에서의 구성비 변화비율도 거의 차이가 나지 않는 것으로 보이며 실제 정확도는 추정에서의 결과를 이용하면 될 것이다.

〈표 4-40〉 연금의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
111	16,421 (63.20%)	16,407 (0.05)	16,409 (0.05)	10,774,390 (65.10%)	10,770,067 (0.02)	10,769,755 (0.03)
112	1,771 (6.82%)	1,777 (0.02)	1,777 (0.02)	1,100,795 (6.65%)	1,103,120 (0.02)	1,104,843 (0.02)
012	7,791 (29.98%)	7,799 (0.04)	7,797 (0.03)	4,674,963 (28.25%)	4,676,962 (0.01)	4,675,551 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
111	16,421 (63.20%)	16,436 (0.06)	16,434 (0.05)	10,774,390 (65.10%)	10,779,735 (0.03)	10,777,743 (0.02)
112	1,771 (6.82%)	1,763 (0.03)	1,766 (0.02)	1,100,795 (6.65%)	1,104,066 (0.02)	1,103,300 (0.02)
012	7,791 (29.98%)	7,784 (0.02)	7,783 (0.03)	4,674,963 (28.25%)	4,666,348 (0.05)	4,669,106 (0.04)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
111	16,421 (63.20%)	16,448 (0.10)	16,440 (0.07)	10,774,390 (65.10%)	10,788,704 (0.09)	10,787,103 (0.08)
112	1,771 (6.82%)	1,784 (0.05)	1,786 (0.05)	1,100,795 (6.65%)	1,110,910 (0.06)	1,107,048 (0.04)
012	7,791 (29.98%)	7,751 (0.15)	7,757 (0.13)	4,674,963 (28.25%)	4,650,534 (0.15)	4,655,999 (0.12)



3. 근로계약서작성여부

근로계약서작성여부 항목을 대체하기 위한 대체군은 종사상지위, 고용시근로기간, 고용계약여부, 산업, 종사자규모 항목을 사용하였다. 근로계약서작성여부 항목에 대한 모의실험 결과가 <표 4-41>에 정리되어 있다. 이전의 결과와 유사하게 동시대체와 개별대체의 차이는 거의 보이지 않는다. 개별대체의 결과에서 구성비 차이가 0.01%정도 작은 것으로 보이지만 두 방법에 대해서 유의적으로 차이가 있다고 말할 수는 없을 것으로 보인다. 추정수를 고려한 결과를 살펴보면 무응답비율이 5%인 경우 두 방법 모두 구성비 차이가 0.01% 정도로 매우 안정적인 대체가 되고 있다. 그리고 무응답비율이 20%인 경우에는 구성비 차이가 최대 0.05% 정도로 추정을 할 수 있으므로 무응답비율이 높지만 매우 정확성이 높은 대체가 되는 것으로 나타났다. 실제 추정 오차율을 보더라도 0.1% 정도밖에 되지 않음을 알 수 있다. 이 항목의 경우에는 추정에서의 구성비 변화비율이 표본에서의 구성비 변화비율보다 조금 작게 나타났지만 유의적으로 작다고는 할 수 없을 것이다.

<표 4-41> 근로계약서작성여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	12,200 (46.95%)	12,204 (0.02)	12,204 (0.02)	7,973,496 (48.18%)	7,975,008 (0.01)	7,975,934 (0.01)
2	13,783 (53.05%)	13,779 (0.02)	13,779 (0.02)	8,576,653 (51.82%)	8,575,141 (0.01)	8,574,215 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	12,200 (46.95%)	12,186 (0.05)	12,208 (0.03)	7,973,496 (48.18%)	7,968,305 (0.03)	7,970,307 (0.02)
2	13,783 (53.05%)	13,797 (0.05)	13,775 (0.03)	8,576,653 (51.82%)	8,581,844 (0.03)	8,579,842 (0.02)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	12,200 (46.95%)	12,226 (0.10)	12,179 (0.08)	7,973,496 (48.18%)	7,981,841 (0.05)	7,966,936 (0.04)
2	13,783 (53.05%)	13,757 (0.10)	13,804 (0.08)	8,576,653 (51.82%)	8,568,308 (0.05)	8,583,213 (0.04)

4. 교육훈련경험여부

교육훈련경험여부 항목을 대체하기 위한 대체군은 종사상지위, 산업, 종사자규모 항목을 사용하였다. 교육훈련경험여부 항목에 대한 모의실험 결과가 <표 4-42>에 정리되어 있다. 이전의 결과와 유사하게 동시대체와 개별대체의 차이는 거의 보이지 않는다. 추정수를 고려한 결과를 살펴보면 무응답비율이 5%인 경우 구성비 차이가 0.05~0.06% 정도로 매우 안정적인 대체가 되고 있다. 그리고 무응답비율이 20%인 경우에는 구성비 차이가 최대 0.13~0.14%까지 커지지만 이 정도의 차이는 작은 변화라고 볼 수 있다. 실제 추정 오차율을 보더라도 0.4% 정도이므로 대체의 정확도는 높은 것으로 생각된다. 이 항목의 경우에는 추정에서의 구성비 변화비율이 표본에서의 구성비 변화비율보다 조금 크게 나타났지만 앞의 결과와 마찬가지로 유의적으로 크다고는 할 수 없을 것이다.

<표 4-42> 교육훈련경험여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	8,356 (32.16%)	8,366 (0.04)	8,364 (0.03)	5,175,905 (31.27%)	5,185,187 (0.06)	5,183,807 (0.05)
2	17,627 (67.84%)	17,617 (0.04)	17,619 (0.03)	11,374,243 (68.73%)	11,364,962 (0.06)	11,366,342 (0.05)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	8,356 (32.16%)	8,342 (0.05)	8,347 (0.04)	5,175,905 (31.27%)	5,188,442 (0.08)	5,189,022 (0.08)
2	17,627 (67.84%)	17,641 (0.05)	17,636 (0.04)	11,374,243 (68.73%)	11,361,707 (0.08)	11,361,127 (0.08)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	8,356 (32.16%)	8,382 (0.10)	8,387 (0.12)	5,175,905 (31.27%)	5,196,753 (0.13)	5,198,970 (0.14)
2	17,627 (67.84%)	17,601 (0.10)	17,596 (0.12)	11,374,243 (68.73%)	11,353,396 (0.13)	11,351,179 (0.14)



5. 연속형 항목

본 부가조사 항목 중에서 연속형 항목은 평소근로시간, 최근3개월평균임금, 총교육시간 3개 항목이다. 평소근로시간 항목의 대체군은 36미만일이유, 총취업시간, 산업이며, 최근3개월평균임금 항목의 대체군은 종사상지위, 가구주와의관계, 36미만일이유, 교육정도, 직업, 종사자규모, 만나이이며, 총교육시간 항목의 대체군은 교육훈련참석여부, 교육정도, 산업, 고용계약여부 항목을 사용하였다. 연속형 항목의 경우 절대평균차이의 오차비율을 정확성 여부를 결정하는 측도로 사용하였다.

연속형 항목들에 대한 모의실험 결과가 <표 4-43>, <표 4-44>, <표 4-45>에 정리되어 있다. 3개의 항목 모두 동시대체와 개별대체의 차이는 거의 보이지 않는다. 가중평균의 결과를 살펴보면 무응답비율이 높아짐에 따라 절대평균차이의 오차비율도 커지는 것을 볼 수 있으나, 무응답비율이 10%인 경우에도 오차비율이 0.5%가 넘지 않으며 무응답비율이 20%까지 높아져도 오차비율은 1%를 넘지 않는 것으로 나타났다. 이 결과는 범주형 항목들과 마찬가지로 다른 조사에서의 결과보다 정확도가 더 높은 것을 알 수 있다. 이러한 결과도 본조사와 부가조사라는 특성의 영향인 것으로 판단된다.

<표 4-43> 평소근로시간의 모의실험 결과

무응답 비율	표본 평균	동시대체 절대평균차이 (오차비율)	개별대체 절대평균차이 (오차비율)	가중 평균	동시대체 절대평균차이 (오차비율)	개별대체 절대평균차이 (오차비율)
5%	43.28시간	0.04 (0.09)	0.05 (0.12)	43.38시간	0.05 (0.12)	0.05 (0.12)
10%		0.10 (0.23)	0.11 (0.25)		0.11 (0.25)	0.10 (0.23)
20%		0.22 (0.51)	0.20 (0.46)		0.21 (0.48)	0.20 (0.46)

〈표 4-44〉 최근3개월평균임금의 모의실험 결과

무응답 비율	표본 평균	동시대체 절대평균차이 (오차비율)	개별대체 절대평균차이 (오차비율)	가중 평균	동시대체 절대평균차이 (오차비율)	개별대체 절대평균차이 (오차비율)
5%	188.80만원	0.24 (0.13)	0.23 (0.12)	195.27만원	0.29 (0.15)	0.30 (0.15)
10%		0.51 (0.27)	0.51 (0.27)		0.57 (0.29)	0.58 (0.30)
20%		1.11 (0.59)	1.09 (0.58)		1.15 (0.59)	1.14 (0.58)

〈표 4-45〉 총교육시간의 모의실험 결과

무응답 비율	표본 평균	동시대체 절대평균차이 (오차비율)	개별대체 절대평균차이 (오차비율)	가중 평균	동시대체 절대평균차이 (오차비율)	개별대체 절대평균차이 (오차비율)
5%	35.84시간	0.05 (0.14)	0.05 (0.14)	36.51시간	0.07 (0.19)	0.08 (0.22)
10%		0.12 (0.33)	0.13 (0.36)		0.13 (0.36)	0.13 (0.36)
20%		0.25 (0.70)	0.26 (0.73)		0.28 (0.77)	0.28 (0.77)

6. 그 외의 항목들

앞에서 설명하지 못한 항목들의 결과를 살펴보면 대부분의 항목에서 비슷한 결과를 볼 수 있다. 동시대체와 개별대체의 차이는 거의 보이지 않으며, 무응답비율이 20%인 경우에도 구성비 차이가 크지 않다는 것을 알 수 있다. 하지만, 구성비가 매우 작은 세부 항목의 경우에는 구성비 차이는 없어 보이지만 구성비 변화비율의 오차는 매우 커질 수 있음을 명심해야 할 것이다. 예를 들면, 총 10,000개 중 세부항목의 도수가 30인 경우 대체결과 25가 되었다고 하자. 이 때 실제 구성비 변화비율은 0.05%가 줄어들었지만, 구성비 변화비율의 오차는 16.7%가 발생하게 된다. 따라서 이러한 경우에는 그 결과에 대하여 주의 깊게 사용해야 할 것으로 판단된다. 모든 항목의 모의실험 결과가 <표 4-46>에서 <표 4-67>까지 주어져 있으니 참조하기 바란다.



<표 4-46> 단기근로기간형태의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	4,073 (15.68%)	4,073 (0)	4,073 (0)	2,438,295 (14.73%)	2,438,295 (0)	2,438,295 (0)
1	1,525 (5.87%)	1,522 (0.01)	1,530 (0.02)	847,947 (5.12%)	845,432 (0.01)	850,534 (0.02)
2	20,385 (78.46%)	20,388 (0.01)	20,380 (0.02)	13,263,906 (80.14%)	13,266,423 (0.02)	13,261,319 (0.01)

(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	4,073 (15.68%)	4,073 (0)	4,073 (0)	2,438,295 (14.73%)	2,438,295 (0)	2,438,295 (0)
1	1,525 (5.87%)	1,531 (0.02)	1,532 (0.03)	847,947 (5.12%)	852,944 (0.03)	848,749 (0.01)
2	20,385 (78.46%)	20,379 (0.03)	20,378 (0.03)	13,263,906 (80.14%)	13,258,910 (0.03)	13,263,104 (0.01)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	4,073 (15.68%)	4,073 (0)	4,073 (0)	2,438,295 (14.73%)	2,438,295 (0)	2,438,295 (0)
1	1,525 (5.87%)	1,532 (0.03)	1,530 (0.02)	847,947 (5.12%)	857,771 (0.06)	855,658 (0.05)
2	20,385 (78.46%)	20,378 (0.03)	20,380 (0.02)	13,263,906 (80.14%)	13,254,083 (0.06)	13,256,196 (0.04)

〈표 4-47〉 계속근로가능여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	1,525 (5.87%)	1,522 (0.01)	1,530 (0.02)	847,947 (5.12%)	845,432 (0.01)	850,534 (0.02)
1	21,353 (82.18%)	21,348 (0.02)	21,348 (0.02)	13,809,743 (83.44%)	13,803,963 (0.03)	13,807,011 (0.01)
2	3,105 (11.95%)	3,113 (0.03)	3,105 (0)	1,892,459 (11.43%)	1,900,754 (0.05)	1,892,604 (0.01)

(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	1,525 (5.87%)	1,531 (0.02)	1,532 (0.03)	847,947 (5.12%)	852,944 (0.03)	848,749 (0.01)
1	21,353 (82.18%)	21,362 (0.04)	21,360 (0.03)	13,809,743 (83.44%)	13,818,618 (0.06)	13,820,660 (0.07)
2	3,105 (11.95%)	3,090 (0.06)	3,091 (0.05)	1,892,459 (11.43%)	1,878,587 (0.08)	1,880,740 (0.07)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	1,525 (5.87%)	1,532 (0.03)	1,530 (0.02)	847,947 (5.12%)	857,771 (0.06)	855,658 (0.05)
1	21,353 (82.18%)	21,323 (0.11)	21,371 (0.07)	13,809,743 (83.44%)	13,798,102 (0.07)	13,815,614 (0.04)
2	3,105 (11.95%)	3,128 (0.09)	3,082 (0.09)	1,892,459 (11.43%)	1,894,276 (0.02)	1,878,877 (0.08)



〈표 4-48〉 계속근로가능여부(이유)의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	4,630 (17.82%)	4,635 (0.02)	4,635 (0.02)	2,740,406 (16.56%)	2,746,186 (0.03)	2,743,138 (0.01)
1	12,644 (48.66%)	12,652 (0.03)	12,647 (0.01)	8,328,692 (50.32%)	8,326,807 (0.01)	8,316,301 (0.07)
2	2,784 (10.71%)	2,768 (0.06)	2,774 (0.03)	1,723,456 (10.41%)	1,710,340 (0.08)	1,718,938 (0.02)
3	5,929 (22.82%)	5,928 (0.01)	5,927 (0.01)	3,757,595 (22.70%)	3,766,817 (0.06)	3,771,771 (0.08)

(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	4,630 (17.82%)	4,621 (0.04)	4,623 (0.03)	2,740,406 (16.56%)	2,731,531 (0.06)	2,729,489 (0.07)
1	12,644 (48.66%)	12,667 (0.09)	12,664 (0.08)	8,328,692 (50.32%)	8,330,550 (0.02)	8,330,872 (0.02)
2	2,784 (10.71%)	2,792 (0.04)	2,792 (0.04)	1,723,456 (10.41%)	1,729,141 (0.04)	1,730,285 (0.04)
3	5,929 (22.82%)	5,903 (0.10)	5,904 (0.10)	3,757,595 (22.70%)	3,758,928 (0.01)	3,759,503 (0.02)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	4,630 (17.82%)	4,660 (0.11)	4,612 (0.07)	2,740,406 (16.56%)	2,752,047 (0.07)	2,734,534 (0.04)
1	12,644 (48.66%)	12,604 (0.15)	12,683 (0.15)	8,328,692 (50.32%)	8,313,784 (0.09)	8,334,764 (0.04)
2	2,784 (10.71%)	2,771 (0.05)	2,783 (0.01)	1,723,456 (10.41%)	1,715,422 (0.05)	1,728,989 (0.04)
3	5,929 (22.82%)	5,948 (0.07)	5,905 (0.09)	3,757,595 (22.70%)	3,768,896 (0.07)	3,751,862 (0.03)

〈표 4-49〉 향후기대근속기간의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	22,878 (88.05%)	22,870 (0.03)	22,878 (0)	14,657,690 (88.57%)	14,649,395 (0.05)	14,657,545 (0.01)
1	1,980 (7.62%)	1,985 (0.02)	1,970 (0.04)	1,159,615 (7.01%)	1,158,757 (0.01)	1,151,455 (0.05)
2	783 (3.01%)	788 (0.02)	790 (0.03)	514,678 (3.11%)	522,691 (0.05)	523,980 (0.06)
3	342 (1.32%)	340 (0.01)	345 (0.01)	218,165 (1.32%)	219,306 (0.01)	217,168 (0.01)

(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	22,878 (88.05%)	22,893 (0.06)	22,892 (0.06)	14,657,690 (88.57%)	14,671,562 (0.08)	14,669,408 (0.07)
1	1,980 (7.62%)	1,962 (0.07)	1,960 (0.08)	1,159,615 (7.01%)	1,148,306 (0.07)	1,144,130 (0.10)
2	783 (3.01%)	789 (0.03)	782 (0.01)	514,678 (3.11%)	516,762 (0.01)	517,107 (0.01)
3	342 (1.32%)	339 (0.02)	349 (0.02)	218,165 (1.32%)	213,519 (0.03)	219,503 (0.01)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	22,878 (88.05%)	22,855 (0.09)	22,901 (0.09)	14,657,690 (88.57%)	14,655,873 (0.02)	14,671,272 (0.07)
1	1,980 (7.62%)	1,969 (0.08)	1,954 (0.10)	1,159,615 (7.01%)	1,146,336 (0.08)	1,144,810 (0.09)
2	783 (3.01%)	810 (0.10)	784 (0.01)	514,678 (3.11%)	525,143 (0.06)	520,191 (0.03)
3	342 (1.32%)	349 (0.02)	344 (0.01)	218,165 (1.32%)	222,797 (0.03)	213,875 (0.03)



〈표 4-50〉 향후기대근속기간(1년이하)의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	24,003 (92.38%)	23,998 (0.02)	24,013 (0.04)	15,390,533 (92.99%)	15,391,391 (0.01)	15,398,693 (0.05)
1	753 (2.90%)	750 (0.01)	751 (0.01)	418,394 (2.53%)	417,047 (0.01)	417,098 (0.01)
2	207 (0.80%)	207 (0)	205 (0.01)	123,846 (0.75%)	124,582 (0.01)	122,213 (0.01)
3	222 (0.85%)	223 (0.01)	220 (0.01)	119,175 (0.72%)	117,908 (0.01)	117,364 (0.01)
4	232 (0.89%)	236 (0.02)	229 (0.01)	131,642 (0.80%)	132,182 (0.01)	129,840 (0.02)
5	100 (0.38%)	98 (0.01)	99 (0.01)	62,413 (0.38%)	61,566 (0.01)	61,948 (0.01)
6	195 (0.75%)	197 (0.01)	193 (0.01)	130,423 (0.79%)	131,060 (0.01)	129,351 (0.01)
7	47 (0.18%)	46 (0.01)	46 (0.01)	28,692 (0.17%)	27,383 (0.01)	27,926 (0.01)
8	43 (0.17%)	45 (0.01)	44 (0.01)	23,756 (0.14%)	24,915 (0.01)	24,109 (0.01)
9	24 (0.09%)	24 (0)	25 (0.01)	17,401 (0.11%)	17,401 (0)	17,846 (0.01)
10	106 (0.41%)	108 (0.01)	107 (0.01)	72,283 (0.44%)	73,435 (0.01)	72,409 (0.01)
11	36 (0.14%)	35 (0.01)	35 (0.01)	23,141 (0.14%)	22,325 (0.01)	22,467 (0.01)
12	15 (0.06%)	16 (0.01)	16 (0.01)	8,450 (0.05%)	8,953 (0.01)	8,886 (0.01)

(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	24,003 (92.38%)	24,021 (0.07)	24,023 (0.08)	15,390,533 (92.99%)	15,401,843 (0.07)	15,406,019 (0.09)
1	753 (2.90%)	746 (0.03)	747 (0.03)	418,394 (2.53%)	415,138 (0.02)	415,336 (0.02)
2	207 (0.80%)	211 (0.01)	205 (0.01)	123,846 (0.75%)	125,945 (0.01)	122,440 (0.01)
3	222 (0.85%)	217 (0.01)	218 (0.01)	119,175 (0.72%)	116,146 (0.01)	117,737 (0.01)
4	232 (0.89%)	225 (0.02)	228 (0.01)	131,642 (0.80%)	128,310 (0.01)	128,855 (0.02)
5	100 (0.38%)	95 (0.01)	98 (0.01)	62,413 (0.38%)	58,556 (0.03)	59,733 (0.02)
6	195 (0.75%)	198 (0.01)	196 (0.01)	130,423 (0.79%)	131,512 (0.01)	129,728 (0.01)
7	47 (0.18%)	45 (0.01)	46 (0.01)	28,692 (0.17%)	28,132 (0.01)	28,353 (0.01)
8	43 (0.17%)	44 (0.01)	45 (0.01)	23,756 (0.14%)	24,337 (0.01)	24,568 (0.01)
9	24 (0.09%)	24 (0)	25 (0.01)	17,401 (0.11%)	16,373 (0.01)	17,914 (0.01)
10	106 (0.41%)	106 (0)	101 (0.02)	72,283 (0.44%)	73,267 (0.01)	68,826 (0.02)
11	36 (0.14%)	36 (0)	36 (0)	23,141 (0.14%)	22,140 (0.01)	22,190 (0.01)
12	15 (0.06%)	15 (0)	15 (0)	8,450 (0.05%)	8,450 (0)	8,450 (0)



(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	24,003 (92.38%)	24,014 (0.04)	24,029 (0.10)	15,390,533 (92.99%)	15,403,813 (0.08)	15,405,338 (0.09)
1	753 (2.90%)	746 (0.03)	748 (0.02)	418,394 (2.53%)	412,710 (0.04)	412,913 (0.04)
2	207 (0.80%)	210 (0.01)	202 (0.02)	123,846 (0.75%)	126,272 (0.01)	120,894 (0.02)
3	222 (0.85%)	225 (0.02)	217 (0.01)	119,175 (0.72%)	123,321 (0.03)	116,079 (0.02)
4	232 (0.89%)	224 (0.03)	225 (0.02)	131,642 (0.80%)	126,906 (0.03)	129,293 (0.02)
5	100 (0.38%)	104 (0.02)	104 (0.02)	62,413 (0.38%)	62,550 (0.01)	64,552 (0.01)
6	195 (0.75%)	189 (0.02)	190 (0.02)	130,423 (0.79%)	125,399 (0.03)	126,929 (0.02)
7	47 (0.18%)	50 (0.01)	49 (0.01)	28,692 (0.17%)	30,583 (0.01)	30,824 (0.02)
8	43 (0.17%)	42 (0.01)	45 (0.01)	23,756 (0.14%)	20,899 (0.01)	26,550 (0.02)
9	24 (0.09%)	24 (0)	22 (0.01)	17,401 (0.11%)	16,855 (0.01)	16,353 (0.01)
10	106 (0.41%)	101 (0.02)	103 (0.01)	72,283 (0.44%)	67,879 (0.03)	70,817 (0.01)
11	36 (0.14%)	40 (0.01)	35 (0.01)	23,141 (0.14%)	25,241 (0.01)	22,245 (0.01)
12	15 (0.06%)	14 (0.01)	14 (0.01)	8,450 (0.05%)	7,719 (0.01)	7,362 (0.01)

〈표 4-51〉 근속기간제한이유의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	22,878 (88.05%)	22,870 (0.03)	22,878 (0)	14,657,690 (88.57%)	14,649,395 (0.05)	14,657,545 (0.01)
1	1,316 (5.06%)	1,330 (0.06)	1,322 (0.03)	751,909 (4.54%)	763,830 (0.08)	755,800 (0.02)
2	45 (0.17%)	45 (0)	44 (0.01)	27,294 (0.16%)	26,925 (0.01)	26,698 (0.01)
3	684 (2.63%)	691 (0.03)	679 (0.02)	433,697 (2.62%)	436,760 (0.02)	429,873 (0.02)
4	185 (0.71%)	183 (0.01)	183 (0.01)	105,605 (0.64%)	105,005 (0.01)	105,070 (0.01)
5	15 (0.06%)	15 (0)	16 (0.01)	7,014 (0.04%)	6,527 (0.01)	6,779 (0.01)
6	66 (0.25%)	66 (0)	67 (0.01)	32,435 (0.20%)	33,641 (0.01)	34,031 (0.01)
7	278 (1.07%)	276 (0.01)	284 (0.02)	203,385 (1.23%)	202,935 (0.01)	208,957 (0.03)
8	107 (0.41%)	109 (0.01)	106 (0.01)	69,329 (0.42%)	70,061 (0.01)	69,317 (0.01)
9	298 (1.15%)	289 (0.04)	296 (0.01)	195,573 (1.18%)	189,730 (0.03)	191,196 (0.02)
10	75 (0.29%)	72 (0.01)	73 (0.01)	45,586 (0.28%)	44,132 (0.01)	44,926 (0.01)
11	36 (0.14%)	37 (0.01)	35 (0.01)	20,632 (0.12%)	21,208 (0.01)	19,957 (0.01)



(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	22,878 (88.05%)	22,893 (0.06)	22,892 (0.05)	14,657,690 (88.57%)	14,671,562 (0.08)	14,669,409 (0.07)
1	1,316 (5.06%)	1,301 (0.05)	1,307 (0.03)	751,909 (4.54%)	744,704 (0.04)	745,075 (0.04)
2	45 (0.17%)	41 (0.01)	40 (0.02)	27,294 (0.16%)	23,550 (0.02)	23,080 (0.02)
3	684 (2.63%)	689 (0.02)	682 (0.01)	433,697 (2.62%)	435,346 (0.01)	432,442 (0.01)
4	185 (0.71%)	186 (0.01)	183 (0.01)	105,605 (0.64%)	105,448 (0.01)	104,260 (0.01)
5	15 (0.06%)	16 (0.01)	16 (0.01)	7,014 (0.04%)	8,127 (0.01)	7,550 (0.01)
6	66 (0.25%)	68 (0.01)	67 (0.01)	32,435 (0.20%)	34,297 (0.01)	32,884 (0.01)
7	278 (1.07%)	279 (0.01)	285 (0.03)	203,385 (1.23%)	204,723 (0.01)	207,173 (0.02)
8	107 (0.41%)	104 (0.01)	106 (0.01)	69,329 (0.42%)	66,458 (0.02)	69,201 (0.01)
9	298 (1.15%)	296 (0.01)	296 (0.01)	195,573 (1.18%)	191,817 (0.02)	194,450 (0.01)
10	75 (0.29%)	74 (0.01)	73 (0.01)	45,586 (0.28%)	43,603 (0.02)	44,009 (0.01)
11	36 (0.14%)	36 (0)	36 (0)	20,632 (0.12%)	20,514 (0.01)	20,615 (0.01)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	22,878 (88.05%)	22,855 (0.09)	22,901 (0.09)	14,657,690 (88.57%)	14,655,873 (0.02)	14,671,272 (0.07)
1	1,316 (5.06%)	1,324 (0.04)	1,307 (0.03)	751,909 (4.54%)	755,118 (0.02)	748,396 (0.02)
2	45 (0.17%)	57 (0.05)	47 (0.01)	27,294 (0.16%)	33,974 (0.05)	28,241 (0.01)
3	684 (2.63%)	673 (0.04)	671 (0.05)	433,697 (2.62%)	426,029 (0.05)	425,171 (0.05)
4	185 (0.71%)	184 (0.01)	179 (0.02)	105,605 (0.64%)	102,676 (0.02)	102,777 (0.02)
5	15 (0.06%)	13 (0.01)	16 (0.01)	7,014 (0.04%)	6,181 (0.01)	7,583 (0.01)
6	66 (0.25%)	72 (0.03)	66 (0)	32,435 (0.20%)	35,518 (0.01)	32,613 (0.01)
7	278 (1.07%)	289 (0.04)	286 (0.03)	203,385 (1.23%)	208,331 (0.03)	210,645 (0.04)
8	107 (0.41%)	111 (0.02)	110 (0.01)	69,329 (0.42%)	72,394 (0.02)	72,005 (0.02)
9	298 (1.15%)	297 (0.01)	295 (0.01)	195,573 (1.18%)	190,162 (0.03)	189,153 (0.04)
10	75 (0.29%)	71 (0.02)	68 (0.03)	45,586 (0.28%)	42,664 (0.02)	41,448 (0.03)
11	36 (0.14%)	37 (0.01)	37 (0.01)	20,632 (0.12%)	21,229 (0.01)	20,844 (0.01)



<표 4-52> 근로시간형태의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	23,307 (89.70%)	23,307 (0)	23,309 (0.01)	14,970,369 (90.45%)	14,971,224 (0.01)	14,969,222 (0.01)
2	2,676 (10.30%)	2,676 (0)	2,674 (0.01)	1,579,780 (9.55%)	1,578,925 (0.01)	1,580,927 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	23,307 (89.70%)	23,306 (0.01)	23,311 (0.02)	14,970,369 (90.45%)	14,966,965 (0.02)	14,975,549 (0.04)
2	2,676 (10.30%)	2,677 (0.01)	2,672 (0.02)	1,579,780 (9.55%)	1,583,184 (0.02)	1,574,600 (0.04)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	23,307 (89.70%)	23,315 (0.03)	23,313 (0.02)	14,970,369 (90.45%)	14,978,886 (0.06)	14,976,553 (0.04)
2	2,676 (10.30%)	2,668 (0.03)	2,670 (0.02)	1,579,780 (9.55%)	1,571,263 (0.06)	1,573,596 (0.04)

<표 4-53> 급여받은곳

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	24,676 (94.97%)	24,680 (0.02)	24,681 (0.02)	15,760,680 (95.23%)	15,759,141 (0.01)	15,762,693 (0.01)
2	278 (1.07%)	284 (0.02)	274 (0.01)	200,354 (1.21%)	204,279 (0.02)	199,435 (0.01)
3	1,029 (3.96%)	1,019 (0.04)	1,028 (0.01)	589,115 (3.56%)	586,729 (0.02)	588,020 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	24,676 (94.97%)	24,676 (0)	24,679 (0.01)	15,760,680 (95.23%)	15,762,548 (0.01)	15,762,757 (0.01)
2	278 (1.07%)	285 (0.03)	285 (0.03)	200,354 (1.21%)	202,544 (0.01)	202,871 (0.01)
3	1,029 (3.96%)	1,022 (0.03)	1,019 (0.04)	589,115 (3.56%)	585,056 (0.02)	584,521 (0.03)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	24,676 (94.97%)	24,693 (0.07)	24,687 (0.04)	15,760,680 (95.23%)	15,770,615 (0.06)	15,766,801 (0.04)
2	278 (1.07%)	274 (0.02)	271 (0.03)	200,354 (1.21%)	193,945 (0.04)	194,074 (0.04)
3	1,029 (3.96%)	1,016 (0.05)	1,025 (0.02)	589,115 (3.56%)	585,590 (0.02)	589,275 (0.01)

〈표 4-54〉 근무사업체의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	24,676 (94.97%)	24,680 (0.02)	24,681 (0.02)	15,760,680 (95.23%)	15,759,141 (0.01)	15,762,693 (0.01)
A	1 (0.00%)	1 (0)	1 (0)	130 (0.00%)	130 (0)	130 (0)
B	4 (0.02%)	4 (0)	4 (0)	1,461 (0.01%)	1,461 (0)	1,461 (0)
C	248 (0.95%)	246 (0.01)	246 (0.01)	135,114 (0.82%)	134,156 (0.01)	133,948 (0.01)
D	23 (0.09%)	20 (0.01)	22 (0.01)	12,597 (0.08%)	11,048 (0.01)	13,068 (0.01)
E	3 (0.01%)	3 (0)	3 (0)	1,760 (0.01%)	1,760 (0)	1,760 (0)
F	52 (0.20%)	54 (0.01)	51 (0.01)	33,183 (0.20%)	33,856 (0.01)	32,859 (0.01)
G	137 (0.53%)	140 (0.01)	137 (0)	83,993 (0.51%)	86,446 (0.01)	83,668 (0.01)
H	78 (0.30%)	75 (0.01)	76 (0.01)	43,869 (0.27%)	42,292 (0.01)	42,276 (0.01)
I	42 (0.16%)	45 (0.01)	44 (0.01)	20,369 (0.12%)	22,207 (0.01)	21,051 (0.01)
J	29 (0.11%)	32 (0.01)	29 (0)	20,167 (0.12%)	22,542 (0.02)	20,094 (0.01)
K	57 (0.22%)	61 (0.02)	56 (0.01)	43,035 (0.26%)	46,709 (0.02)	42,411 (0.01)
L	303 (1.17%)	298 (0.02)	306 (0.01)	201,371 (1.22%)	196,942 (0.04)	203,095 (0.01)
M	9 (0.03%)	8 (0.01)	9 (0)	4,384 (0.03%)	4,112 (0.01)	4,026 (0.01)
N	46 (0.18%)	44 (0.01)	47 (0.01)	28,527 (0.17%)	27,652 (0.01)	28,547 (0.01)
O	33 (0.13%)	33 (0)	32 (0.01)	20,754 (0.13%)	20,815 (0.01)	20,876 (0.01)
P	98 (0.38%)	98 (0)	98 (0)	55,483 (0.33%)	55,172 (0.01)	55,424 (0.01)
Q	67 (0.26%)	69 (0.01)	69 (0.01)	40,982 (0.25%)	43,147 (0.01)	42,475 (0.01)
R	39 (0.15%)	39 (0)	38 (0.01)	20,318 (0.12%)	20,946 (0.01)	20,046 (0.01)
S	20 (0.08%)	18 (0.01)	18 (0.01)	10,569 (0.06%)	9,317 (0.01)	9,317 (0.01)
T	13 (0.05%)	11 (0.01)	12 (0.01)	7,498 (0.05%)	6,619 (0.01)	7,245 (0.01)
U	5 (0.02%)	4 (0.01)	4 (0.01)	3,904 (0.02%)	3,677 (0.01)	3,677 (0.01)



(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	24,676 (94.97%)	24,676 (0)	24,679 (0.01)	15,760,680 (95.23%)	15,762,548 (0.01)	15,762,757 (0.01)
A	1 (0.00%)	1 (0)	2 (0.01)	130 (0.00%)	130 (0)	721 (0.01)
B	4 (0.02%)	4 (0)	4 (0)	1,461 (0.01%)	1,461 (0)	1,461 (0)
C	248 (0.95%)	235 (0.05)	243 (0.01)	135,114 (0.82%)	128,093 (0.05)	135,615 (0.01)
D	23 (0.09%)	23 (0)	22 (0.01)	12,597 (0.08%)	12,075 (0.01)	11,423 (0.01)
E	3 (0.01%)	4 (0.01)	3 (0)	1,760 (0.01%)	3,176 (0.01)	1,760 (0)
F	52 (0.20%)	55 (0.01)	56 (0.02)	33,183 (0.20%)	34,971 (0.01)	35,175 (0.01)
G	137 (0.53%)	135 (0.01)	132 (0.02)	83,993 (0.51%)	82,549 (0.01)	80,670 (0.02)
H	78 (0.30%)	81 (0.01)	80 (0.01)	43,869 (0.27%)	46,182 (0.01)	45,432 (0.01)
I	42 (0.16%)	44 (0.01)	44 (0.01)	20,369 (0.12%)	21,886 (0.01)	21,146 (0.01)
J	29 (0.11%)	31 (0.01)	28 (0.01)	20,167 (0.12%)	21,681 (0.01)	19,627 (0.01)
K	57 (0.22%)	57 (0)	56 (0.01)	43,035 (0.26%)	42,400 (0.01)	42,268 (0.01)
L	303 (1.17%)	307 (0.01)	305 (0.01)	201,371 (1.22%)	201,180 (0.01)	202,403 (0.01)
M	9 (0.03%)	8 (0.01)	8 (0.01)	4,384 (0.03%)	4,125 (0.01)	4,125 (0.01)
N	46 (0.18%)	49 (0.01)	47 (0.01)	28,527 (0.17%)	30,134 (0.01)	29,312 (0.01)
O	33 (0.13%)	33 (0)	31 (0.01)	20,754 (0.13%)	20,295 (0.01)	19,359 (0.01)
P	98 (0.38%)	100 (0.01)	98 (0)	55,483 (0.33%)	56,493 (0.01)	54,784 (0.01)
Q	67 (0.26%)	63 (0.02)	67 (0)	40,982 (0.25%)	38,796 (0.02)	39,713 (0.01)
R	39 (0.15%)	38 (0.01)	40 (0.01)	20,318 (0.12%)	19,145 (0.01)	19,664 (0.01)
S	20 (0.08%)	19 (0.01)	19 (0.01)	10,569 (0.06%)	10,218 (0.01)	10,218 (0.01)
T	13 (0.05%)	15 (0.01)	14 (0.01)	7,498 (0.05%)	8,705 (0.01)	8,611 (0.01)
U	5 (0.02%)	5 (0)	5 (0)	3,904 (0.02%)	3,904 (0)	3,904 (0)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
0	24,676 (94.97%)	24,693 (0.07)	24,687 (0.04)	15,760,680 (95.23%)	15,770,615 (0.06)	15,766,801 (0.04)
A	1 (0.00%)	2 (0.01)	1 (0)	130 (0.00%)	441 (0.01)	130 (0)
B	4 (0.02%)	6 (0.01)	5 (0.01)	1,461 (0.01%)	2,316 (0.01)	1,743 (0.01)
C	248 (0.95%)	230 (0.06)	237 (0.04)	135,114 (0.82%)	130,935 (0.03)	130,656 (0.03)
D	23 (0.09%)	21 (0.01)	21 (0.01)	12,597 (0.08%)	10,100 (0.02)	11,199 (0.01)
E	3 (0.01%)	3 (0)	3 (0)	1,760 (0.01%)	1,760 (0)	1,760 (0)
F	52 (0.20%)	54 (0.01)	48 (0.02)	33,183 (0.20%)	34,312 (0.01)	30,443 (0.02)
G	137 (0.53%)	140 (0.01)	142 (0.02)	83,993 (0.51%)	86,020 (0.01)	86,368 (0.01)
H	78 (0.30%)	79 (0.01)	81 (0.01)	43,869 (0.27%)	44,986 (0.01)	45,831 (0.01)
I	42 (0.16%)	41 (0.01)	40 (0.01)	20,369 (0.12%)	18,813 (0.01)	18,750 (0.01)
J	29 (0.11%)	27 (0.01)	27 (0.01)	20,167 (0.12%)	18,977 (0.01)	18,874 (0.01)
K	57 (0.22%)	62 (0.02)	56 (0.01)	43,035 (0.26%)	47,688 (0.03)	42,752 (0.01)
L	303 (1.17%)	299 (0.02)	310 (0.02)	201,371 (1.22%)	197,009 (0.03)	205,006 (0.02)
M	9 (0.03%)	11 (0.01)	9 (0)	4,384 (0.03%)	5,230 (0.01)	4,384 (0)
N	46 (0.18%)	47 (0.01)	48 (0.01)	28,527 (0.17%)	29,263 (0.01)	29,470 (0.01)
O	33 (0.13%)	33 (0)	34 (0.01)	20,754 (0.13%)	19,976 (0.01)	20,656 (0.01)
P	98 (0.38%)	94 (0.02)	96 (0.01)	55,483 (0.33%)	51,880 (0.02)	54,961 (0.01)
Q	67 (0.26%)	66 (0.01)	67 (0)	40,982 (0.25%)	39,248 (0.01)	39,855 (0.01)
R	39 (0.15%)	35 (0.02)	34 (0.02)	20,318 (0.12%)	17,850 (0.01)	17,794 (0.01)
S	20 (0.08%)	21 (0.01)	20 (0)	10,569 (0.06%)	10,954 (0.01)	11,633 (0.01)
T	13 (0.05%)	15 (0.01)	13 (0)	7,498 (0.05%)	8,703 (0.01)	8,012 (0.01)
U	5 (0.02%)	4 (0.01)	4 (0.01)	3,904 (0.02%)	3,072 (0.01)	3,072 (0.01)



〈표 4-55〉 특수고용여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	941 (3.62%)	950 (0.04)	946 (0.02)	558,758 (3.38%)	563,793 (0.03)	563,847 (0.03)
2	25,042 (96.38%)	25,033 (0.04)	25,037 (0.02)	15,991,391 (96.62%)	15,986,356 (0.03)	15,986,302 (0.03)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	941 (3.62%)	930 (0.04)	928 (0.05)	558,758 (3.38%)	552,789 (0.04)	551,099 (0.05)
2	25,042 (96.38%)	25,053 (0.04)	25,055 (0.05)	15,991,391 (96.62%)	15,997,360 (0.04)	15,999,050 (0.05)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	941 (3.62%)	924 (0.06)	926 (0.06)	558,758 (3.38%)	550,972 (0.05)	552,995 (0.04)
2	25,042 (96.38%)	25,059 (0.06)	25,057 (0.06)	15,991,391 (96.62%)	15,999,177 (0.05)	15,997,154 (0.04)

〈표 4-56〉 가정내근로여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	115 (0.44)	115 (0)	116 (0.01)	68,192 (0.41%)	69,648 (0.01)	69,087 (0.01)
2	25,868 (99.56%)	25,868 (0)	25,867 (0.01)	16,481,957 (99.59%)	16,480,501 (0.01)	16,481,062 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	115 (0.44)	110 (0.02)	112 (0.01)	68,192 (0.41%)	64,410 (0.02)	65,545 (0.01)
2	25,868 (99.56%)	25,873 (0.02)	25,871 (0.01)	16,481,957 (99.59%)	16,483,739 (0.02)	16,484,604 (0.01)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	115 (0.44)	106 (0.03)	108 (0.02)	68,192 (0.41%)	62,070 (0.03)	63,856 (0.02)
2	25,868 (99.56%)	25,877 (0.03)	25,875 (0.02)	16,481,957 (99.59%)	16,488,079 (0.03)	16,486,293 (0.02)

〈표 4-57〉 일자리취업동기의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	16,585 (63.83%)	16,578 (0.03)	16,590 (0.02)	10,830,417 (65.44%)	10,825,479 (0.03)	10,833,701 (0.02)
2	9,398 (36.17%)	9,405 (0.03)	9,393 (0.02)	5,719,732 (34.56%)	5,724,670 (0.03)	5,716,448 (0.02)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	16,585 (63.83%)	16,570 (0.06)	16,594 (0.03)	10,830,417 (65.44%)	10,819,780 (0.06)	10,839,358 (0.05)
2	9,398 (36.17%)	9,413 (0.06)	9,389 (0.03)	5,719,732 (34.56%)	5,730,368 (0.06)	5,710,791 (0.05)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	16,585 (63.83%)	16,607 (0.08)	16,612 (0.10)	10,830,417 (65.44%)	10,849,724 (0.12)	10,850,476 (0.12)
2	9,398 (36.17%)	9,376 (0.08)	9,371 (0.10)	5,719,732 (34.56%)	5,700,425 (0.12)	5,699,673 (0.12)

〈표 4-58〉 주된취업사유의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	6,863 (26.41%)	6,857 (0.02)	6,860 (0.01)	4,521,455 (27.32%)	4,527,281 (0.03)	4,527,451 (0.03)
2	7,386 (28.43%)	7,397 (0.04)	7,380 (0.03)	4,805,828 (29.04%)	4,808,636 (0.01)	4,795,385 (0.06)
3	6,971 (26.83%)	6,970 (0.01)	6,965 (0.02)	4,108,873 (24.83%)	4,113,509 (0.02)	4,109,394 (0.01)
4	1,131 (4.35%)	1,123 (0.03)	1,137 (0.02)	737,297 (4.45%)	728,677 (0.05)	741,522 (0.03)
5	339 (1.30%)	338 (0.01)	336 (0.01)	238,105 (1.44%)	240,510 (0.01)	236,164 (0.01)
6	974 (3.75%)	977 (0.01)	970 (0.02)	731,443 (4.42%)	728,776 (0.02)	727,307 (0.03)
7	607 (2.34%)	610 (0.01)	611 (0.01)	347,923 (2.10%)	348,375 (0.01)	350,192 (0.02)
8	441 (1.70%)	440 (0.01)	442 (0.01)	308,751 (1.87%)	307,832 (0.01)	307,599 (0.01)
9	755 (2.91%)	763 (0.04)	760 (0.01)	458,858 (2.77%)	461,794 (0.02)	462,743 (0.02)
10	358 (1.38%)	353 (0.02)	360 (0.01)	204,687 (1.24%)	198,216 (0.04)	202,763 (0.01)
11	158 (0.61%)	155 (0.01)	162 (0.01)	86,928 (0.53%)	86,543 (0.01)	89,628 (0.01)



(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	6,863 (26.41%)	6,850 (0.05)	6,855 (0.03)	4,521,455 (27.32%)	4,514,716 (0.04)	4,520,367 (0.01)
2	7,386 (28.43%)	7,379 (0.03)	7,399 (0.05)	4,805,828 (29.04%)	4,801,590 (0.03)	4,808,562 (0.01)
3	6,971 (26.83%)	6,983 (0.05)	6,960 (0.04)	4,108,873 (24.83%)	4,112,659 (0.02)	4,101,615 (0.05)
4	1,131 (4.35%)	1,143 (0.05)	1,144 (0.05)	737,297 (4.45%)	749,030 (0.08)	747,233 (0.06)
5	339 (1.30%)	339 (0)	341 (0.01)	238,105 (1.44%)	240,710 (0.01)	239,632 (0.01)
6	974 (3.75%)	968 (0.02)	965 (0.04)	731,443 (4.42%)	727,752 (0.02)	724,644 (0.04)
7	607 (2.34%)	603 (0.02)	600 (0.03)	347,923 (2.10%)	346,808 (0.01)	345,702 (0.01)
8	441 (1.70%)	445 (0.01)	443 (0.01)	308,751 (1.87%)	308,849 (0.01)	311,127 (0.01)
9	755 (2.91%)	758 (0.01)	759 (0.01)	458,858 (2.77%)	456,441 (0.01)	458,542 (0.01)
10	358 (1.38%)	359 (0.01)	360 (0.01)	204,687 (1.24%)	205,571 (0.01)	207,414 (0.01)
11	158 (0.61%)	156 (0.01)	157 (0.01)	86,928 (0.53%)	86,024 (0.01)	85,312 (0.01)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	6,863 (26.41%)	6,881 (0.07)	6,883 (0.08)	4,521,455 (27.32%)	4,541,097 (0.12)	4,540,071 (0.11)
2	7,386 (28.43%)	7,391 (0.02)	7,371 (0.06)	4,805,828 (29.04%)	4,808,994 (0.04)	4,805,792 (0.01)
3	6,971 (26.83%)	6,946 (0.10)	6,955 (0.06)	4,108,873 (24.83%)	4,085,592 (0.14)	4,089,076 (0.12)
4	1,131 (4.35%)	1,134 (0.01)	1,124 (0.02)	737,297 (4.45%)	745,306 (0.05)	728,684 (0.05)
5	339 (1.30%)	342 (0.02)	346 (0.03)	238,105 (1.44%)	235,379 (0.02)	242,580 (0.03)
6	974 (3.75%)	986 (0.04)	983 (0.03)	731,443 (4.42%)	738,543 (0.04)	738,789 (0.04)
7	607 (2.34%)	589 (0.07)	594 (0.05)	347,923 (2.10%)	337,235 (0.06)	338,117 (0.06)
8	441 (1.70%)	445 (0.01)	447 (0.02)	308,751 (1.87%)	312,409 (0.02)	315,842 (0.04)
9	755 (2.91%)	760 (0.01)	764 (0.03)	458,858 (2.77%)	465,215 (0.04)	465,418 (0.04)
10	358 (1.38%)	354 (0.02)	361 (0.01)	204,687 (1.24%)	196,052 (0.06)	201,006 (0.03)
11	158 (0.61%)	155 (0.01)	155 (0.01)	86,928 (0.53%)	84,326 (0.02)	84,138 (0.02)

〈표 4-59〉 건강보험의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
211	17,068 (65.69%)	17,056 (0.05)	17,058 (0.04)	11,121,761 (67.20%)	11,118,070 (0.02)	11,119,439 (0.01)
212	4,550 (17.51%)	4,552 (0.01)	4,559 (0.04)	2,844,068 (17.18%)	2,845,850 (0.02)	2,846,681 (0.02)
213	349 (1.34%)	348 (0.01)	345 (0.01)	179,499 (1.08%)	179,671 (0.01)	177,030 (0.01)
214	3,368 (12.96%)	3,378 (0.04)	3,375 (0.03)	1,966,728 (11.88%)	1,972,312 (0.04)	1,970,637 (0.03)
022	648 (2.49%)	649 (0.01)	646 (0.01)	438,092 (2.65%)	434,246 (0.03)	436,362 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
211	17,068 (65.69%)	17,057 (0.04)	17,064 (0.02)	11,121,761 (67.20%)	11,122,097 (0.01)	11,120,155 (0.01)
212	4,550 (17.51%)	4,560 (0.04)	4,543 (0.03)	2,844,068 (17.18%)	2,842,189 (0.01)	2,837,107 (0.04)
213	349 (1.34%)	354 (0.02)	357 (0.03)	179,499 (1.08%)	182,305 (0.02)	182,812 (0.02)
214	3,368 (12.96%)	3,374 (0.03)	3,363 (0.02)	1,966,728 (11.88%)	1,973,346 (0.04)	1,965,031 (0.01)
022	648 (2.49%)	638 (0.03)	656 (0.03)	438,092 (2.65%)	430,213 (0.05)	445,045 (0.04)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
211	17,068 (65.69%)	17,098 (0.11)	17,093 (0.10)	11,121,761 (67.20%)	11,136,011 (0.09)	11,129,061 (0.04)
212	4,550 (17.51%)	4,570 (0.08)	4,577 (0.11)	2,844,068 (17.18%)	2,852,372 (0.05)	2,855,797 (0.08)
213	349 (1.34%)	333 (0.06)	340 (0.03)	179,499 (1.08%)	173,531 (0.03)	179,497 (0.01)
214	3,368 (12.96%)	3,330 (0.14)	3,339 (0.11)	1,966,728 (11.88%)	1,954,769 (0.07)	1,958,400 (0.05)
022	648 (2.49%)	652 (0.02)	634 (0.05)	438,092 (2.65%)	433,465 (0.03)	427,394 (0.07)



〈표 4-60〉 고용보험의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
31	14,709 (56.61%)	14,704 (0.02)	14,703 (0.02)	9,696,488 (58.59%)	9,697,349 (0.01)	9,697,389 (0.01)
32	9,128 (35.13%)	9,135 (0.03)	9,130 (0.01)	5,588,664 (33.77%)	5,589,756 (0.01)	5,586,408 (0.01)
00	2,146 (8.26%)	2,144 (0.01)	2,150 (0.01)	1,264,997 (7.64%)	1,263,045 (0.01)	1,266,352 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
31	14,709 (56.61%)	14,718 (0.03)	14,718 (0.03)	9,696,488 (58.59%)	9,700,379 (0.02)	9,699,653 (0.02)
32	9,128 (35.13%)	9,132 (0.02)	9,129 (0.01)	5,588,664 (33.77%)	5,586,738 (0.01)	5,587,733 (0.01)
00	2,146 (8.26%)	2,133 (0.05)	2,136 (0.04)	1,264,997 (7.64%)	1,263,032 (0.01)	1,262,764 (0.01)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
31	14,709 (56.61%)	14,732 (0.09)	14,726 (0.07)	9,696,488 (58.59%)	9,709,485 (0.08)	9,708,655 (0.07)
32	9,128 (35.13%)	9,118 (0.04)	9,120 (0.03)	5,588,664 (33.77%)	5,585,019 (0.02)	5,585,387 (0.02)
00	2,146 (8.26%)	2,133 (0.05)	2,137 (0.04)	1,264,997 (7.64%)	1,255,645 (0.05)	1,256,107 (0.05)

〈표 4-61〉 퇴직금의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
11	15,960 (61.42%)	15,966 (0.03)	15,955 (0.01)	10,469,567 (63.26%)	10,473,618 (0.02)	10,467,722 (0.01)
12	10,023 (38.58%)	10,017 (0.03)	10,028 (0.01)	6,080,582 (36.74%)	6,076,531 (0.02)	6,082,427 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
11	15,960 (61.42%)	15,950 (0.03)	15,950 (0.03)	10,469,567 (63.26%)	10,464,351 (0.03)	10,463,986 (0.03)
12	10,023 (38.58%)	10,033 (0.03)	10,033 (0.03)	6,080,582 (36.74%)	6,085,798 (0.03)	6,086,163 (0.03)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
11	15,960 (61.42%)	15,946 (0.05)	15,941 (0.07)	10,469,567 (63.26%)	10,458,868 (0.06)	10,460,728 (0.05)
12	10,023 (38.58%)	10,037 (0.05)	10,042 (0.07)	6,080,582 (36.74%)	6,091,281 (0.06)	6,089,421 (0.05)

〈표 4-62〉 상여금의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
21	16,314 (62.79%)	16,312 (0.01)	16,314 (0)	10,731,564 (64.84%)	10,728,532 (0.02)	10,733,822 (0.02)
22	9,669 (37.21%)	9,671 (0.01)	9,669 (0)	5,818,585 (35.16%)	5,821,617 (0.02)	5,816,327 (0.02)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
21	16,314 (62.79%)	16,318 (0.01)	16,320 (0.02)	10,731,564 (64.84%)	10,736,887 (0.03)	10,737,051 (0.04)
22	9,669 (37.21%)	9,665 (0.01)	9,663 (0.02)	5,818,585 (35.16%)	5,813,262 (0.03)	5,813,098 (0.04)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
21	16,314 (62.79%)	16,329 (0.05)	16,299 (0.06)	10,731,564 (64.84%)	10,740,879 (0.06)	10,720,414 (0.06)
22	9,669 (37.21%)	9,654 (0.05)	9,684 (0.06)	5,818,585 (35.16%)	5,809,270 (0.06)	5,829,735 (0.06)



<표 4-63> 시간외수당의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
31	11,507 (44.29%)	11,511 (0.01)	11,510 (0.01)	7,375,570 (44.56%)	7,381,149 (0.04)	7,379,611 (0.03)
32	14,476 (55.71%)	14,472 (0.01)	14,473 (0.01)	9,174,579 (55.44%)	9,169,000 (0.04)	9,170,538 (0.03)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
31	11,507 (44.29%)	11,493 (0.06)	11,520 (0.05)	7,375,570 (44.56%)	7,367,609 (0.04)	7,381,823 (0.04)
32	14,476 (55.71%)	14,490 (0.06)	14,463 (0.05)	9,174,579 (55.44%)	9,182,540 (0.04)	9,168,326 (0.04)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
31	11,507 (44.29%)	11,526 (0.07)	11,530 (0.09)	7,375,570 (44.56%)	7,390,665 (0.10)	7,391,851 (0.10)
32	14,476 (55.71%)	14,457 (0.07)	14,453 (0.09)	9,174,579 (55.44%)	9,159,484 (0.10)	9,158,298 (0.10)

<표 4-64> 유급휴가의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
41	14,690 (56.54%)	14,687 (0.01)	14,686 (0.02)	9,747,362 (58.90%)	9,744,705 (0.02)	9,745,772 (0.01)
42	11,293 (43.46%)	11,296 (0.01)	11,297 (0.02)	6,802,787 (41.10%)	6,805,444 (0.02)	6,804,377 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
41	14,690 (56.54%)	14,683 (0.03)	14,683 (0.03)	9,747,362 (58.90%)	9,741,309 (0.04)	9,741,470 (0.04)
42	11,293 (43.46%)	11,300 (0.03)	11,300 (0.03)	6,802,787 (41.10%)	6,808,839 (0.04)	6,808,679 (0.04)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
41	14,690 (56.54%)	14,703 (0.05)	14,681 (0.04)	9,747,362 (58.90%)	9,758,384 (0.06)	9,737,857 (0.06)
42	11,293 (43.46%)	11,280 (0.05)	11,302 (0.04)	6,802,787 (41.10%)	6,791,765 (0.06)	6,812,292 (0.06)

〈표 4-65〉 노동조합가입여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	19,245 (74.07%)	19,239 (0.03)	19,241 (0.02)	12,437,818 (75.15%)	12,427,911 (0.06)	12,435,039 (0.01)
2	2,228 (8.57%)	2,230 (0.01)	2,232 (0.02)	1,315,270 (7.95%)	1,319,387 (0.02)	1,320,619 (0.03)
3	1,389 (5.35%)	1,388 (0.01)	1,386 (0.02)	905,401 (5.47%)	907,424 (0.01)	901,145 (0.02)
4	3,121 (12.01%)	3,126 (0.02)	3,124 (0.01)	1,891,660 (11.43%)	1,895,426 (0.02)	1,893,346 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	19,245 (74.07%)	19,260 (0.06)	19,257 (0.04)	12,437,818 (75.15%)	12,427,706 (0.06)	12,436,058 (0.01)
2	2,228 (8.57%)	2,218 (0.03)	2,219 (0.03)	1,315,270 (7.95%)	1,320,340 (0.03)	1,309,772 (0.04)
3	1,389 (5.35%)	1,378 (0.05)	1,392 (0.01)	905,401 (5.47%)	902,060 (0.02)	907,754 (0.01)
4	3,121 (12.01%)	3,127 (0.02)	3,115 (0.02)	1,891,660 (11.43%)	1,900,043 (0.05)	1,896,565 (0.03)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	19,245 (74.07%)	19,275 (0.11)	19,270 (0.09)	12,437,818 (75.15%)	12,456,226 (0.11)	12,450,741 (0.08)
2	2,228 (8.57%)	2,206 (0.08)	2,203 (0.09)	1,315,270 (7.95%)	1,304,702 (0.07)	1,297,651 (0.11)
3	1,389 (5.35%)	1,365 (0.10)	1,370 (0.08)	905,401 (5.47%)	887,197 (0.11)	895,029 (0.06)
4	3,121 (12.01%)	3,137 (0.06)	3,140 (0.07)	1,891,660 (11.43%)	1,902,024 (0.06)	1,906,728 (0.12)



〈표 4-66〉 임금형태의 모의실험 결과

5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	1,341 (5.16%)	1,346 (0.02)	1,345 (0.02)	824,483 (4.98%)	828,097 (0.02)	828,929 (0.03)
2	3,511 (13.51%)	3,504 (0.02)	3,505 (0.02)	1,958,361 (11.83%)	1,951,599 (0.04)	1,957,996 (0.01)
3	48 (0.18%)	49 (0.01)	46 (0.01)	35,440 (0.21%)	35,798 (0.01)	34,466 (0.01)
4	15,746 (60.60%)	15,741 (0.02)	15,749 (0.01)	9,827,467 (59.38%)	9,838,064 (0.06)	9,830,901 (0.02)
5	3,792 (14.59%)	3,797 (0.02)	3,793 (0.01)	2,954,194 (17.85%)	2,950,223 (0.02)	2,951,515 (0.02)
6	1,511 (5.82%)	1,512 (0.01)	1,512 (0.01)	931,954 (5.63%)	928,248 (0.02)	928,653 (0.02)
7	34 (0.13%)	34 (0)	33 (0.01)	18,250 (0.11%)	18,120 (0.01)	17,689 (0.01)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	1,341 (5.16%)	1,335 (0.02)	1,333 (0.03)	824,483 (4.98%)	822,562 (0.01)	821,439 (0.02)
2	3,511 (13.51%)	3,507 (0.01)	3,520 (0.04)	1,958,361 (11.83%)	1,961,497 (0.02)	1,963,062 (0.03)
3	48 (0.18%)	49 (0.01)	47 (0.01)	35,440 (0.21%)	34,134 (0.01)	33,263 (0.01)
4	15,746 (60.60%)	15,755 (0.04)	15,740 (0.02)	9,827,467 (59.38%)	9,840,206 (0.08)	9,832,677 (0.03)
5	3,792 (14.59%)	3,801 (0.04)	3,804 (0.05)	2,954,194 (17.85%)	2,949,922 (0.03)	2,956,330 (0.01)
6	1,511 (5.82%)	1,504 (0.03)	1,507 (0.02)	931,954 (5.63%)	924,406 (0.04)	925,877 (0.04)
7	34 (0.13%)	32 (0.01)	32 (0.01)	18,250 (0.11%)	17,422 (0.01)	17,502 (0.01)

(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	1,341 (5.16%)	1,353 (0.05)	1,355 (0.05)	824,483 (4.98%)	822,348 (0.01)	833,902 (0.06)
2	3,511 (13.51%)	3,491 (0.07)	3,489 (0.08)	1,958,361 (11.83%)	1,950,307 (0.05)	1,948,395 (0.06)
3	48 (0.18%)	43 (0.01)	46 (0.01)	35,440 (0.21%)	32,824 (0.01)	33,757 (0.01)
4	15,746 (60.60%)	15,766 (0.08)	15,731 (0.06)	9,827,467 (59.38%)	9,851,389 (0.14)	9,811,325 (0.10)
5	3,792 (14.59%)	3,779 (0.05)	3,798 (0.03)	2,954,194 (17.85%)	2,943,424 (0.07)	2,964,203 (0.06)
6	1,511 (5.82%)	1,514 (0.01)	1,527 (0.06)	931,954 (5.63%)	930,326 (0.01)	937,114 (0.03)
7	34 (0.13%)	37 (0.01)	37 (0.01)	18,250 (0.11%)	19,531 (0.01)	21,451 (0.02)

〈표 4-67〉 주5일근로여부의 모의실험 결과

(5%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	12,327 (47.44%)	12,322 (0.02)	12,328 (0.01)	8,102,018 (48.95%)	8,104,995 (0.02)	8,104,413 (0.02)
2	13,656 (52.56%)	13,661 (0.02)	13,655 (0.01)	8,448,131 (51.05%)	8,445,515 (0.02)	8,445,736 (0.02)
(10%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	12,327 (47.44%)	12,315 (0.04)	12,318 (0.03)	8,102,018 (48.95%)	8,095,171 (0.04)	8,096,327 (0.03)
2	13,656 (52.56%)	13,668 (0.04)	13,665 (0.03)	8,448,131 (51.05%)	8,454,748 (0.04)	8,453,822 (0.03)
(20%)	표본수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)	추정수 (구성비)	동시대체 (구성비차이)	개별대체 (구성비차이)
1	12,327 (47.44%)	12,353 (0.10)	12,346 (0.08)	8,102,018 (48.95%)	8,120,909 (0.12)	8,117,901 (0.10)
2	13,656 (52.56%)	13,630 (0.10)	13,637 (0.08)	8,448,131 (51.05%)	8,429,624 (0.12)	8,432,247 (0.10)



제6절 결론

본 연구에서는 2010년 8월 기준 경제활동인구조사 근로형태별 부가조사 항목들에 대하여 연관성 분석을 실시하고 각 항목의 대체군을 제시하였으며, 응용 핫텍 대체방법으로 모의실험을 실시하여 부가조사의 모든 항목을 동시에 대체하는 방법과 각 항목별로 대체하는 방법을 비교하였다. 그리고 정확한 모의실험을 실시하기 위하여 응답자와 무응답자의 차이분석을 하였으며, 이를 통하여 모의실험 자료를 구축할 때 무응답자의 성향을 반영할 항목들을 선정하였다. 이러한 작업을 바탕으로 확인한 연구의 결과를 몇 가지로 정리하면 다음과 같다.

- 동시대체와 개별대체의 차이는 거의 존재하지 않는 것으로 나타났다. 이러한 결과는 사용된 대체군에서의 정보가 거의 유사하기 때문인 것으로 판단된다. 각 개별 항목의 대체군 대부분이 동시대체에서 사용한 대체군에 포함되기 때문에 대체에 필요한 더 유용한 정보를 얻을 수가 없기 때문이다. 항상 동시대체와 개별대체의 차이가 없는 것은 아니지만 본 부가조사에서는 조사의 특성상 두 방법의 차이는 존재하지 않는 것으로 결론을 내릴 수가 있다. 그리고 대체의 정확도도 다른 조사에 비해 매우 높은 것으로 나타났다. 무응답비율이 높아짐에 따라서 대체의 정확도가 낮아지는 것은 다르지 않으나, 무응답비율이 5%인 경우에는 구성비 차이가 거의 발생하지 않는 것을 볼 수 있다. 향후 무응답비율의 증가를 대비하여 실시한 20%의 경우의 실험에서도 구성비의 변화는 0.1% 내외였고, 구성비 변화비율의 오차도 대부분의 항목에서 1%이내인 것으로 나타났다. 이는 본조사와 부가조사의 관계에서 본 조사의 내용을 잘 활용하면 부가조사의 내용을 정확성 높게 추정 가능하다는 것을 보여 준다.
- 개별대체를 위해 제시된 각 항목별 대체군 사용의 일반화를 위해서 2010년 3월 자료와 2011년 3월 자료를 이용하여 대체군을 선정함으로써 연도와 계절적인 상황이 대체군에 영향을 주는지 검토하였다. 검토결과 세 자료 모두에서 대체군 항목들이 거의 일치하였다. 따라서 대체군 사용의 일반성은 확인이 되었다. 그러나 앞선 결과에서 개별대체가 동시대체보다 좋은 결과를 보여주지 않는 것으로 판단되어 각 항목별 대체군의 사용은 필요가 없을 것 같다. 하지만 대체군 선정과정에서의 연구결과는 타 조사의 연구에 많은 도움을 줄 수 있을 것이다.
- 본 연구의 모의실험에서는 응용 핫텍 대체방법이 사용되었다. 이 방법은 핫텍, 계층적 핫텍, 순차적 핫텍 등과 마찬가지로 범주형 자료의 대체에 매우 적절한 방법으로 알려져 있다. 단지, 대체군 항목들이 연속형 항목이 많을 경우(예: 농업총조사)에는

다른 핫택류의 대체방법보다 정확성이 높게 나타났다. 하지만 대체 과정에서의 시간은 가장 많이 사용된다. 그리고 본 자료처럼 대부분의 항목이 범주형인 경우에는 다른 핫택류의 방법들과 유사한 결과를 보이는 것을 이전 연구에서 확인하였다. 범주형 항목이 대부분인 인구주택총조사의 대체에 계층적 핫택 대체방법을 사용하는 것도 이러한 이유에서다. 따라서 대체방법은 기존에 사용되고 있는 계층적 핫택 대체방법을 이용하여도 본 연구의 결과와 유사할 것이며 대체 시간은 더 적게 필요할 것으로 판단된다.

- 본 부가조사 자료의 경우 응답되지 않은 자료는 전체 항목이 무응답으로 단위 무응답으로 간주할 수 있다. 하지만 부가조사의 특성상 본조사라는 좋은 자료를 가지고 있기 때문에 이 두 자료를 합친다면 항목 무응답 형태가 된다. 따라서 항목 무응답 대체방법으로 대체가 가능해진다. 여기서 한 가지 고려해야 할 사항은 항목 무응답 비율이다. 각 개인으로 본다면 항목 무응답 비율이 50% 정도로 볼 수 있다. 만약 하나의 독립적인 조사에서 항목 무응답 비율이 50%인 자료가 있다면 이 자료를 사용하지 않을 가능성이 높을 것이다. 그러나 본 자료는 본조사와 부가조사의 관계가 있어 본조사의 정보로 부가조사를 매우 정확하게 추정할 수 있는 것으로 나타났다. 따라서 항목 무응답 비율이 높다고 할지라도 항목 무응답으로 간주한 대체방법은 문제되지 않을 것으로 판단된다.
- 표본에서의 구성비 변화비율과 추정에서의 구성비 변화비율의 차이는 거의 발생하지 않는 것으로 보인다. 대체결과를 이용할 때 대체된 표본의 값 자체를 사용하지 않고 가중치를 이용한 추정값을 사용하므로 본 내용은 큰 의미가 없을 것이다. 하지만 이 결과로 볼 때 적절한 방법에 의해서 대체가 진행된다면 표본에서의 대체결과와 가중치를 고려한 대체결과의 정확도가 유사하게 된다는 것을 알 수 있다.
- 본 연구에서는 근로형태별 부가조사 자료로 연구를 진행하였으나 타 부가조사 자료도 이와 유사한 형태를 가지며 대체결과 또한 유사할 것으로 판단되므로 적절한 대체군을 선택하여 동시대체를 실시하면 될 것으로 판단된다.
- 구성비가 매우 작은 세부 항목의 경우에는 구성비 차이는 없어 보이지만 구성비 변화비율의 오차는 매우 커질 수 있음을 명심해야 할 것이다. 따라서 이러한 항목에 대한 대체결과를 사용할 경우에는 항상 주의를 기울여야 할 것이다.

최근 들어 통계조사의 환경이 계속적으로 악화되고 있어 각종 조사의 무응답 비율이 높아지고 있다. 이와 같은 현실에서 무응답 대체기법 연구는 매우 중요하며, 본 연구가 경제활동인구조사 부가조사의 무응답 대체 과정에 많은 부분 도움이 되기를 기대한다.



참고문헌

- 김규성(2000), "무응답 대체 방법과 대체 효과", 「조사연구」, 제1권 2호, pp.1-14.
- 김영원·이주원(2003), "CART를 활용한 결측값 대체방법: 인구주택총조사 혼인상태 항목을 중심으로", 「조사연구」, 제4권 2호, pp.1-21.
- 김영원·조선경(1996), "표본조사에서 항목 무응답 대체 방법", 「한국통계학회논문집」, 제3권 3호, pp.145-159.
- 송순관(2005), 「2005 인구주택총조사 무응답 처리방법 연구 및 읍면동 통계작성 가능성 검토」, 통계청 인구조사과.
- 이현정·최필근(2009), "인구주택총조사 무응답 처리기법", 연구보고서, 통계개발원.
- 최필근(2008), "농업총조사 항목간 연관성 분석 및 대체군(보조변수) 개발", 연구보고서, 통계개발원.
- 최필근(2008), "농업총조사 무응답 대체기법 연구(I)", 연구보고서, 통계개발원.
- 최필근(2010), "출생전후기 사망통계의 무응답 대체기법", 연구보고서, 통계개발원.
- 최필근(2011), "경제총조사 항목 무응답 대체방법 연구", 연구보고서, 통계개발원.
- 통계교육원(2005), 「무응답처리 실무론」.
- _____ (2009), 「무응답 자료처리 및 분석」.
- Afifi, A. A. and R. M. Elashoff(1966), "Missing Observations in Multivariate Statistics I: Review of the Literature", *J. Am. Statist. Assoc.*, Vol. 61, pp.595-604.
- Berry, M. J. A. and G. S. Linoff(1997), *Data Mining Techniques*, John Wiley & Sons, New York.
- Kalton, G. and D. Kasprzyk(1986), "The Treatment of Missing Survey Data", *Survey Methodology*, Vol. 12, pp.1-16.
- Kass, G.(1980), "An Exploratory Technique for Investigating Large Quantities of Categorical Data", *Applied Statistics*, Vol. 29, No. 2, pp.119-127
- Quinlan, J. R.(1986), "Induction of Decision Tree", *Machine Learning*, 1, pp.81-106.
- Rubin, D. B. and J. A. Little(1986), *Statistical Analysis with Missing Data*, John Wiley & Sons, New York.
- Sande, I. G.(1979), "A Personal View of Hot Deck Imputation Procedures", *Survey Methodology*, Vol. 5, pp.238-258.