The Korean Communications
in Statistics Vol. 12 No. 1, 2005
pp. 29-38

A Study on Weight Adjustment in Sampling Survey[1)]

Ran Hee Jung[2)], Sang Eun Lee[3)], and Key-Il Shin[4)]

Abstract

In sample design, determining the weights of estimates becomes usually great influence on the result. In this article, raking methods are applied to different domain and depending on the range of the domain and sample size, the results of estimates are explained and compared. For the comparison, we use the MSE, MAE, MSPE and MAPE with Actual State of Minor Enterprisers Human Resources Survey data in 2001. The simulation result shows that more elaborate method is superior to the widely used method as expected but the difference is not quite significant.

Keywords : Raking method, Neyman allocation

1. 서론

표본설계는 일반적으로 표본 추출방법과 표본규모 결정 그리고 그에 따른 추정방법을 포함한다. 이 때 표본규모와 추출방법은 표본조사 결과의 공표 범위와 허용오차에 따라 결정되게 된다. 즉 조사목적과 대상이 정해지고 조사결과의 공표범위와 허용오차도 결정되면 이 결정에 의해 표본 추출 방법과 표본규모가 결정된다. 또한 표본추출 방법에 따라 추정방법도 결정 된다. 본 논문에서는 표본 설계에서 모집단 총계 추정에 중요한 영향을 주는 승수조정에 관해 살펴보기로 한다. 표본조사에서의 총계 추정값은 모집단의 총 개수와 표본의 평균의 곱으로 나타낸다. 이는 각각의 표본에서 얻은 관측값과 승수(추출률의 역수)를 곱한 값들의 합으로 표현된다. 즉, $\hat{Y}=N\overline{Y}$ $=N/n(\sum_{i=1}^{n} y_i) = \sum_{i=1}^{n} (N/n)y_i = \sum_{i=1}^{n} wy_i$ 이다. 이때 승수($w=N/n$)는 표본 설계 시 표본으로 추출된 추출율의 역수가 되며 총계 추정의 정확성을 결정짓는 중요한 요인이 된다. 이러한 승수는 표본 설계자의 의도에 따라 다른 값을 가질 수 있다. 예를 들어 실제 표본 설계 시에는 많은 부차모집단을 고려하게 된다. 부차모집단이란 모집단이 여러 개의 층으로 분해 된 경우를 말하며, 부차 모집단 자체가 관심의 대상이 된다. 즉 전술한 데로 부차 모집단 자체가 공표의 대상이 될 수 있다. 따라서 모집단 뿐 아니라 부차모집단의 총계나 평균도 추정하게 된다. 그러나 고려하여

---

1) This research was supported by the research fund of Hankuk University of Foreign Studies, 2004
2) Graduate student, Department of Statistics, Hankuk University of Foreign Studies, Mohyun, Yongin, Kyonggi, 449-791 E-mail : nani_ju@hotmail.com
3) Associate Professor, Department of Applied Statistics, Kyonggi University, Suwon, Kyonggi, 442-720
4) Professor, Department of Statistics, Hankuk University of Foreign Studies, Mohyun, Yongin, Kyonggi, 449-791

야 할 부차 모집단의 수가 많아지는 경우 표본 설계를 할 때 이를 모두 고려할 수가 없게 된다. 따라서 표본 설계자는 중요하거나 설계 가능한 부차 모집단을 선택하게 되고 이를 기준으로 승수가 계산되며 계산된 승수가 사용된다. 물론 최종적으로 사용되는 승수는 표본 설계 시에 구해지는 것이 아니라 조사가 끝난 후 결측치 등을 고려하여 갈퀴법(Raking)을 사용하여 구하는 것이 일반적인 방법이다. 이론적으로 고려되지 않은 다른 부차 모집단의 총계나 평균을 구할 때에는 이에 맞는 다른 승수를 계산하여 사용하여야 한다. 그러나 승수 계산을 간단히 하기 위하여 또는 이미 얻어진 표본이 고려되지 않은 부차 모집단의 특성을 잘 설명할 수 없어 계산이 불가능하기 때문에 이미 계산된 승수를 그대로 적용하는 경우도 있다.

본 연구는 실제 표본설계에서 표본 조사결과의 공표 범위에 따라 많은 부차모집단을 고려하게 될 때 고려된 부차 모집단을 이용하여 얻어진 승수를 고려되지 않은 다른 부차 모집단의 총계 추정에 사용할 때와 새롭게 구해진 승수를 사용할 때 얻어지는 결과를 비교하였다.

본 연구에서는 부차 모집단 총계 추정에 있어 승수 적용 방법들을 실제 2003년 중소기업 인력실태조사의 표본설계를 예제로 살펴보았다. 중소기업 인력 실태조사는 제조업과 서비스업, 두 업종에서 조사가 실사되나 본 논문에서는 제조업에 관하여만 모의실험을 하였고 그 결과를 비교하였다. 2절에서는 2003 중소기업 인력실태조사와 그에 따른 표본설계를 소개하였다. 3절에서는 승수 적용에 따른 추정값의 결과를 MSE, MAE, MSPE 그리고 MAPE를 기준으로 살펴보았고 4절에 요약된 결과를 기술하였다.

2. 2003 중소기업 인력실태조사

2.1 조사 개요

중소기업 인력실태 조사는 지역 중소기업의 인력 수급 현황 및 부족인력, 교육 · 훈련, 외국인 근로자 관련 사항들을 종합적으로 조사하여 시 · 도별 자치단체의 산업인력 양성 및 중소기업 인력정책 수립 등에 필요한 기초 자료제공을 목적으로 하며, 한국표준산업분류상 제조업과 서비스업을 영위하는 종업원 수가 5인 이상 300미만인 기업체를 대상으로 하고 있다. 결과 공표단위는 전국단위의 산업 소분류, 지역단위의 산업 중분류이다.

2.2 전제 조건

조사 대상은 산업대분류상 제조업(중분류 번호 15 ~ 37) 및 서비스업(중분류 72, 74,75 및 소분류 642)에 해당하는 사업체 중, 5인이상 300인 미만 사업체이다. 사용된 샘플 프레임은 먼저 제조업의 경우 2001년 기준광공업 통계조사가 사용되었으며 서비스업의 경우 2001년 사업체기초통계조사가 사용되었다. 예상 목표오차는 전국 산업 소분류의 경우, 신뢰도 95%에 허용오차는 5% 정도 수준이고 지역별(16개시도) 산업 중분류는 신뢰도 95%에 허용오차 8% 정도수준으로 하였다. 표본사업체수는 제조업의 경우 약9000개, 서비스업은 약 1000개를 설정하였다. 16개 시도별 중분류로 1차 층화 하였고, 2차 층화는 종사자규모별로 4개층(5-19인, 20-49인, 50-99인, 100-299인)이 설정되었다. 결과 공표단위는 전국단위의 산업 소분류와 지역단위(시도별)의 산업 중분류이다.

2.3 표본 설계

종사자 규모별 특성을 파악하기 위해 일반적인 단순랜덤추출보다는 종사자 규모에 따른 층화를 하여 표본을 추출하는 층화추출이 적절하다고 판단된다. 단순임의추출은 모든 사업체에 동일한 추출확률을 부여하게 되므로 종사자규모가 큰 사업체의 정보는 상대적으로 확보하기 어려울 수 있기 때문이다. 따라서 먼저 16개의 시도로 층화한 후 각 시도별로 22개의 중분류업종을 1차층(부차모집단)으로 간주하고 각 중분류 업종에서 4개의 종사자규모 층을 2차 층으로 하는 2단 층화표본추출이 적절하다고 판단된다. 다음으로 각 종사자규모 층별 표준편차를 고려한 Neyman 배분법에 의거 업종별 표본규모 확정 및 종사자규모 층별 표본배분이 이루어 졌다.

2.4 표본규모설정

먼저 16개 시도별, 22개 중분류업종 각각을 모집단으로 Neyman 배분법에 의해 표본규모를 구하였다. 구해진 표본 규모를 각각의 시도별, 중분류업종에서 종사자수 규모 층을 모집단으로 Neyman 배분법에 의해 표본규모를 구하였다. 인력과 예산을 고려하여, 신뢰도 95%, 허용오차 8%를 적용하여 전체 8,702개 사업체를 표본규모로 최종 설정하였다.

2.5 모수 추정공식

공표를 위한 총계는 지역별 중분류와 전국 소분류이다. 표본 설계시 고려된 부차 모집단은 지역별, 중분류이므로 추정을 위한 승수는 지역별 중분류에 의해 구해진다. 본 논문에서 중요하게 다루는 부분은 표본 설계 시 고려되지 않은 부차 모집단인 전국 소분류이므로 전국 소분류 총계를 위한 추정공식을 살펴보자. 먼저 논문에서 사용되는 첨자는 다음과 같다.

i : 종사자 수, $i=1,2,3,4,$
h : 중분류, $h=1,2,\cdots,22,$
$h(s)$: 중분류 h에 속해 있는 소분류,
a : 16개 시도

이제 $x_{ah(s)i}$를 시도별, 소분류별, 종사자별 표본 총계라 하자. 그리고 w_{a*i}를 총계 추정에 사용될 승수라 하자. 그러면 지역별, 소분류별 총계 추정식은 다음과 같다.

$$\hat{X}_{ah(s)} = \sum_i w_{a*i} \cdot x_{ah(s)i} \tag{1}$$

여기서

$$w_{a*i} = \frac{N_{a*i}}{n_{a*i}} \tag{2}$$

이다. (2)에서 "*"를 한 이유는 승수를 지역별, 중분류별, 종사자별로 구할 수 있을 뿐 아니라

지역별, 소분류별, 종사자별로도 구할 수 있기 때문이다. 다음으로 전국 소분류별 총계 추정식은

$$\hat{X}_{h(s)} = \sum_a \hat{X}_{ah(s)}$$

이 된다.

3. 승수 적용방법

2절에서 표본 설계는 우선 부차 모집단, 즉 지역별 산업 중분류별로 표본 수를 결정하고 결정된 표본수를 이용하여 다시 종사자 별로 표본 수를 결정하였다. 결정된 표본 수에 따라 표본은 지역별, 중분류별, 종사자별로 추출되게 되며 따라서 승수는 지역별, 산업 중분류별 그리고 종사자별로 결정되게 된다. 물론 실제로 분석에 사용되는 승수는 칼퀴법에 의해 조정된 승수가 될 것이다. 이 때 정해진 승수는 전국 소분류별 총계를 추정할 경우에 사용될 수 있다. 그러나 전국 소분류별 총계를 구하기 위해서는 여기에 맞는 승수를 구하여 이를 적용하는 것이 타당할 것이다. 즉 지역별, 소분류별, 종사자별 승수를 구할 수 있고 여기서 얻어진 승수를 이용하여 총계를 추정하는 것이 타당하다. 이제 중분류를 기준으로 얻은 승수 적용 방법과 새롭게 소분류를 기준으로 하여 얻은 승수 적용 방법을 살펴보도록 하자.

3.1 지역별 중분류 승수 적용

부차모집단, 즉 지역별 산업 중분류에서 얻어진 승수를 조사가 끝난 시점에서 칼퀴법을 적용하여 얻어진 승수를 전국단위 산업 소분류 추정에 적용한다. 따라서 갈퀴법은 22×4 행렬, 중분류 22개 종사자 규모 4개, 이 각각 16개의 시・도 별로 적용된다. 이제 (2)식의 승수

$$w_{a*i} = \frac{N_{a*i}}{n_{a*i}}$$

는

$$w_{ahi} = \frac{N_{ahi}}{n_{ahi}} \tag{3}$$

가 되고 따라서 (1)식은 다음과 같이 된다.

$$\hat{X}_{ah(s)} = \sum_i w_{ahi} \cdot x_{ah(s)i}$$

즉 같은 종사자 층에서는 소분류에 상관없이 중분류가 같으면 같은 승수가 사용된다.

3.2 지역별 소분류 승수 적용

표본 설계시 전국단위 소분류의 허용오차는 고려되었지만 표본은 지역별 중분류에서 추출되었고 지역별 중분류에 의해 정해진 승수가 전국 소분류 총계추정에도 사용되는 것은 문제가 있다. 따라서 전국 소분류 공표의 신뢰성을 높이기 위해서는 (3)에서와 달리 지역별, 소분류별, 종사수에

따른 승수를 구하고 칼퀴법을 적용하여 최종적으로 얻어진 승수를 사용하여야 한다. 이때 갈퀴법은 69×4 행렬, 69개의 소분류와 4개의 종사자 규모를 각각 16개의 시 · 도 별로 적용한다. 즉 (2)의 승수는 다음과 같이 구해진다.

$$w_{ah(s)i} = \frac{N_{ah(s)i}}{n_{ah(s)i}} \tag{4}$$

또한 (1)식은 다음과 같이 된다.

$$\hat{X}_{ah(s)} = \sum_i w_{ah(s)i} \cdot x_{ah(s)i}$$

4. 모의 실험 및 결과

전국 지역별, 소분류 총계를 구하기 위해 사용된 승수는 3절에서 설명한 두 가지 방법으로 얻어졌다. 두 방법의 효율성을 비교하기 위하여 2001년 기준 광공업통계조사를 모집단으로 하고 이중에서 8,702개의 표본을 추출하였다. 표본 배분과 추출 방법은 2절에서 설명한 방법을 사용하였다. 지역별 중분류를 이용한 승수 계산은 상대적으로 쉽다. 식 (3)에서 N_{ahi}가 크면 일반적으로 $n_{ahi}>0$인 표본 수가 결정된다. 또한 작은 N_{ahi}에 대해서도 표본 설계시 $n_{ahi}>0$인 표본 수를 지정해 주어 승수를 쉽게 구할 수 있다.

이에 반해 소분류를 이용한 승수 계산에서는 작은 N_{ahi}에 대해서 고려할 수 없다. 즉 표본 선택은 중분류를 기준으로 먼저 얻어지기 때문에 특정 소분류는 작은 $N_{ah(s)i}$에 대해 표본이 없을 수 있다. 이 경우 승수 계산은 불가능하게 된다. 본 논문에서는 이런 경우, 즉 작은 수의 모집단을 갖고 있는 특정 소분류는 분석에서 제외하였다. 물론 N_{ahi} 또는 $N_{ah(s)i}$가 "0"인 경우는 아무런 문제가 되지 않는다. 따라서 실제 소분류는 69개 이지만 표에 나와 있는 소분류는 38개 이다. 또한 전국 소분류 모집단 수에 따라 표본의 수가 결정되기 때문에 모집단의 수가 1000이상인 소분류와 1000이하인 소분류로 나누어 각각 표를 작성하였다. 비교를 위하여 다음의 통계량을 사용하였다.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|$$

$$MSPE = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2/Y_i^2$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|/|Y_i|$$

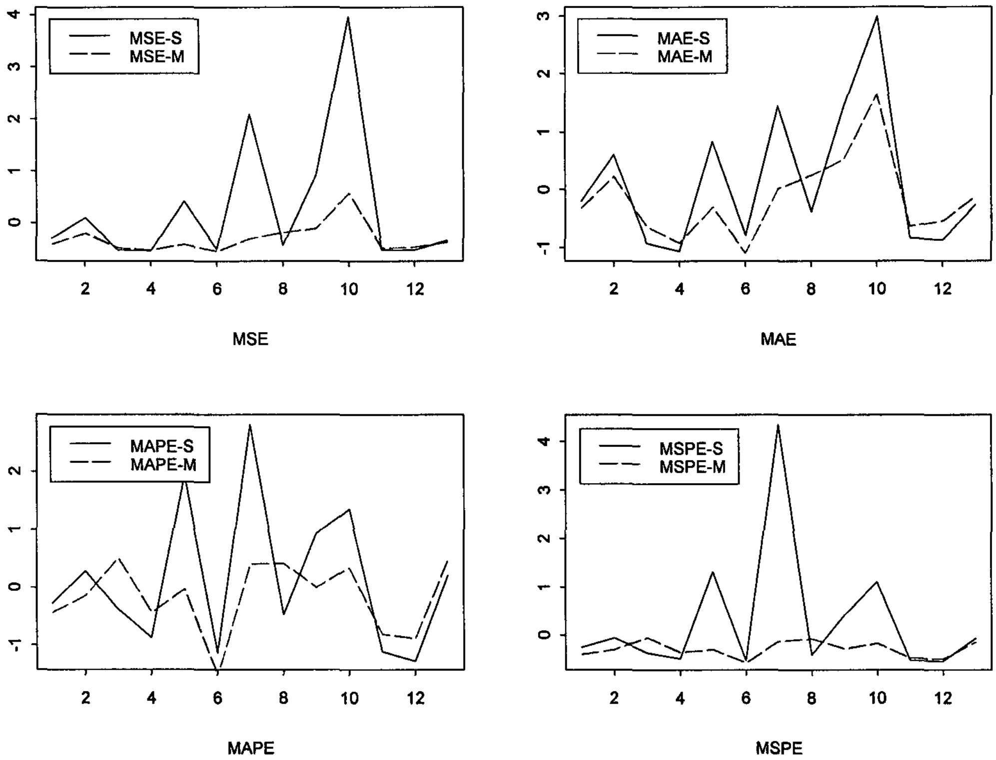
여기서 i는 i번째 반복을, Y_i는 전국 소분류 참값을, $\hat{Y}_i$는 전국 소분류 추정값을 의미한다. 그리고 반복수 $n=1,000$이 사용되었다. 각 전국 소분류별로 위에서 정의된 통계량이 구해졌으며 <표 1>과 <표 2>에 이를 정리하였다. 각 표에서 지역별 중분류 승수를 사용한 경우, MSE-M, MAE-M, MSPE-M 그리고 MAPE-M의 이름을 사용하였고 지역별 소분류 승수를 사용한 경우 MSE-S등과 같이 S를 사용하였다.

다음으로 <표 1>과 <표 2>의 결과를 쉽게 표현하기 위하여 각 통계량을 표준화 시킨 후 이를 <그림 1>과 <그림 2>에 나타내었다.

<표 1> 모집단의 수가 1000이하인 소분류 결과

소분류	모집단수	MSE-S	MSE-M	MAE-S	MAE-M	MAPE-S	MAPE-M	MSPE-S	MSPE-M
153(1)	982	8.28E6	4.71E6	1976	1782	0.136	0.123	0.039	0.022
171(2)	951	2.01E7	1.10E7	3271	2665	0.180	0.146	0.061	0.033
191(3)	256	1.09E6	2.54E6	814	1272	0.127	0.198	0.026	0.062
192(4)	596	5.34E5	1.21E6	586	817	0.088	0.123	0.012	0.027
211(5)	338	2.98E7	4.54E6	3648	1817	0.311	0.155	0.217	0.033
300(6)	560	1.58E6	4.35E5	1023	537	0.067	0.035	0.007	0.002
313(7)	566	8.17E7	7.98E6	4619	2325	0.381	0.190	0.561	0.053
315(8)	969	4.25E6	1.16E7	1681	2689	0.120	0.191	0.021	0.059
319(9)	800	4.53E7	1.38E7	4646	3148	0.232	0.158	0.114	0.035
323(10)	942	1.39E8	3.48E7	7115	4957	0.265	0.185	0.193	0.048
331(11)	885	1.33E6	2.39E6	945	1267	0.069	0.093	0.007	0.013
332(12)	985	1.27E6	3.04E6	893	1423	0.055	0.087	0.005	0.011
333(13)	544	6.94E6	5.89E6	1868	2104	0.174	0.196	0.060	0.051

<표 1>을 살펴보면 소분류 333의 경우 기준에 따라 다른 결론을 주고 있으며 나머지 12개 소분류 중에서 5개가 소분류를 기준으로 승수를 구한 방법이 더 좋은 결과를 주고 있고 다른 7개 소분류에서는 중분류을 이용한 승수가 더 좋은 결과를 주고 있다. 그러나 이러한 결과는 모집단 수가 작으면 이에 따라 표본의 수도 작아지기 때문에 적은 수의 표본을 이용하여 얻은 결과를 일반화시키기는 어려울 수도 있다. 그러나 위의 결과는 표본 설계시 자료의 수가 작은 경우 정교한 설계가 필요하다는 것을 말하고 있는 것이다. 즉 표본의 수가 작은 경우 소분류를 이용한 승수를 적용할 때 각 소분류의 특징을 고려하지 않고 표본을 그대로 사용하게 되면 오히려 나쁜 결과를 줄 수도 있다는 것을 말해 주고 있다. <그림 1>을 살펴보면 중분류 승수를 사용하여 얻은 결과인 MSE-M, MAE-M등은 안정적인 결과를 주고 있는 반면 소분류 승수를 사용하여 얻은 결과인 MSE-S, MAE-S, MAPE-S 그리고 MSPE-S 등은 매우 변동이 큰 것으로 나타났다. 다음으로 모집단의 수가 1,000 이상인 소분류의 결과가 있는 <표 2>를 살펴보자. 여기서도 소분류 343은 기준에 따라 다른 결과를 주고 있다. 나머지 소분류를 살펴보면 24개 소분류에서 16개가 소분류를 이용한 승수 적용 방법이 우수한 것을 알 수 있다. 이는 모집단수가 많아 소분류의 특징을 고려하지 않고 표본을 추출하였더라도 소분류를 이용한 승수를 적용하는 것이 더 좋은 결과를 주고 있다는 것을 알 수 있다. 자료에서 모집단의 수가 큰 경우 2.2절의 표본 설계 전제 조건을 어느 정

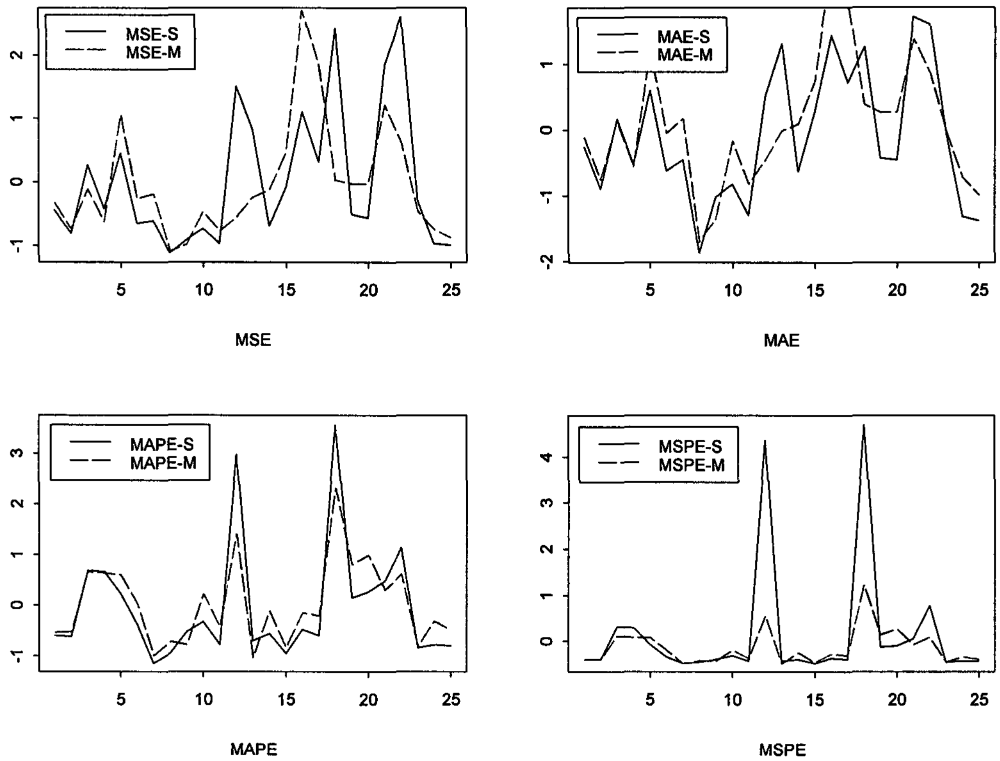


<그림 1> 모집단의 수가 1000이하인 경우

도 만족하고 있다는 것을 말해주고 있다. 물론 전제 조건은 CV를 기준으로 하고 있으나 <표 2>의 MAPE-S와 MAPE-M을 보면 자료의 수가 3,000개 이상일 경우 약 5% 내외인 것을 알 수 있다. 이를 고려한다면 어느 정도 전제 조건을 만족한다고 할 수 있다. 그러나 자료의 수가 1000-3000개인 경우 MAPE-S와 MAPE-M가 높게 나와 전제 조건을 만족한다고 하기는 어려울 것이다. 물론 <표 1>에서도 모집단의 수가 1,000 이하이므로 전제조건을 만족한다고 보기는 어렵다.

<표 2> 모집단의 수가 1000이상인 소분류 결과

소분류	모집단수	MSE-S	MSE-M	MAE-S	MAE-M	MAPE-S	MAPE-M	MSPE-S	MSPE-M
151(1)	3296	1.66E7	1.92E7	3284	3513	0.050	0.053	0.004	0.004
154(2)	2601	7.86E6	9.61E6	2269	2499	0.049	0.054	0.004	0.004
172(3)	1869	3.11E7	2.43E7	3953	3911	0.115	0.114	0.028	0.021
173(4)	1794	1.70E7	1.23E7	2858	2810	0.114	0.113	0.027	0.020
174(5)	1904	3.74E7	5.15E7	4640	5602	0.092	0.111	0.015	0.020
179(6)	3101	1.17E7	2.07E7	2703	3619	0.061	0.082	0.006	0.011
181(7)	9207	1.26E7	2.24E7	2964	3955	0.022	0.030	0.001	0.001
193(8)	1509	8.00E5	1.59E6	733	1011	0.032	0.044	0.002	0.003
212(9)	2499	5.55E6	3.96E6	2075	1550	0.054	0.041	0.004	0.003
221(10)	1910	9.74E6	1.60E7	2383	3425	0.064	0.092	0.007	0.011
222(11)	3699	4.27E6	8.85E6	1642	2390	0.041	0.059	0.003	0.005
251(12)	1035	6.19E7	1.38E7	4509	2965	0.231	0.152	0.162	0.036
252(13)	6931	4.59E7	2.11E7	5778	3663	0.045	0.028	0.003	0.001
281(14)	3529	1.06E7	2.39E7	2683	3836	0.052	0.075	0.004	0.009
289(15)	9095	2.47E7	3.74E7	4136	4873	0.032	0.038	0.001	0.002
291(16)	5616	5.25E7	9.03E7	5983	7833	0.056	0.073	0.005	0.008
293(17)	6379	3.42E7	6.98E7	4824	6806	0.050	0.070	0.004	0.007
295(18)	1041	8.34E7	2.75E7	5704	4310	0.261	0.198	0.173	0.058
311(19)	1410	1.51E7	2.61E7	3013	4118	0.088	0.121	0.013	0.022
312(20)	1810	1.37E7	2.61E7	2970	4121	0.094	0.131	0.014	0.026
321(21)	2048	6.99E7	5.51E7	6429	5900	0.105	0.096	0.019	0.015
322(22)	1466	8.75E7	4.16E7	6247	5077	0.138	0.112	0.043	0.020
343(23)	3178	2.01E7	1.62E7	3636	3672	0.038	0.039	0.002	0.002
361(24)	3190	4.31E6	9.49E6	1622	2555	0.041	0.065	0.003	0.006
369(25)	2684	3.73E6	6.43E6	1512	2122	0.040	0.056	0.003	0.004



<그림 2> 모집단의 수가 1000이상인 경우

특히 모집단의 수가 작아 MAPE-S가 큰 경우 MAPE-S와 MAPE-M를 비교하면 MAPE-M이 더 작은 것을 알 수 있다. 중분류를 이용한 승수 적용 방법이 우수한 경우를 살펴보면 많은 경우 자료의 수가 상대적으로 작은 경우임을 확인 할 수 있다.

모집단 수의 크기에 따른 결과를 살펴보자. 전술한 데로 24개 소분류중 16개가 소분류를 이용한 승수 적용 방법이 우수하였다. 이제 모집단의 수가 3,000이상인 소분류의 결과를 살펴보자. 두 승수 적용 방법의 비교를 쉽게 하기 위하여 각 기준에서 얻어진 결과의 비를 이용하였다. 예를 들어 $MSE-R=\frac{MSE-M}{MSE-S}$로 구해지며 결과는 <표 3>에 나와있다.

<표 3> 각 통계량의 비

소분류	MSE-R	MAE-R	소분류	MSE-R	MAE-R
151	1.158	1.070	289	1.516	1.178
179	1.762	1.339	291	1.719	1.309
181	1.777	1.334	293	2.041	1.411
222	2.071	1.456	343	0.807	1.010
252	0.460	0.634	361	2.204	1.576
281	2.257	1.430			

<표 3>에서 전술한 데로 소분류 343은 기준에 따라 다른 결과를 주고 있다. 이를 제외하면 10개 소분류 중에서 9개의 소분류가 1 보다 큰 결과를 나타내고 있다. 이는 모집단의 수가 큰 경우 소분류를 기준으로 얻은 승수를 사용해야 한다는 결론을 말해 주고 있다.

4. 결론

전국 소분류 총계 추정을 할 때 사용되는 승수는 지역별 중분류 총계를 구할 때 사용되었던 승수를 그대로 적용할 수 있다. 이때 지역별 중분류를 이용한 승수는 지역별 중분류의 특징을 이용하여 구한 것으로 이를 전국 소분류 총계 추정에 바로 적용하는 것은 문제가 될 수 있다. 그러나 전국 소분류의 경우 소분류의 수가 69개로 많고 따라서 정해진 표본수를 가지고 소분류의 특징을 살릴 수 있는 표본 할당은 거의 불가능하다. 3절에서 얻은 결과를 살펴보면 전국 소분류 중에서 모집단의 수가 많은 경우 소분류를 이용한 승수를 적용하면 좋은 결과를 얻을 수 있는 것으로 판단된다. 그러나 모집단의 수가 작은 경우 추출된 표본이 소분류의 특징을 잘 설명하지 못하게 되어 오히려 나쁜 결과를 주고 있다. 결론 적으로 모집단의 수가 작을 때 소분류를 이용한 승수를 적용할 경우에는 많은 주의가 필요하게 된다.

참고문헌

[1] 박홍래 (1999). 「통계조사론」, 영지문화사.
[2] 박재수 (1989). 「표본조사법」, 박영사.
[3] 신민웅, 이상은 (2001). 「표본조사를 위한 표본설계」, 교우사.
[4] 산업연구원 (2003). 「중소기업인력실태 조사」, 중소기업청.

[2004년 8월 접수, 2004년 11월 채택]