

## 소표본 통계단위에서의 집세 변동률 추정

박 원 란<sup>1)</sup>

요 약

도시가구의 지출 중 집세가 차지하는 비율이 높으며 그 변동에 따라 도시가구의 생활에 미치는 영향도 커서 중요한 통계자료로 인식되고 있다. 집세 계약기간이 통상 2년 단위이기 때문에 집세변동의 발생빈도가 적어서 이러한 소표본 군의 통계단위로 일반적인 집세지수를 작성하는데는 많은 어려움이 따른다. 그렇다고 해서 소표본 군의 표본을 확대하는 것도 어렵기 때문에 이러한 산술적인 표본확대가 어려운 소표본 군의 문제점을 해결하기 위해 소지역 추정법을 도입하였다. 이러한 소표본 통계단위에서의 집세 변동률 추정방법을 경기도 지역의 6개 도시에서의 집세변동을 추정하는데 적용하였으며 검토해 보았다.

주요용어 : 소지역추정법, 합성추정법, 복합추정법, 집세변동률

### 1. 서론

2002년 도시가구의 지출 중 8.2%를 차지하는 등 단일품목으로는 가장 높은 지출 비중을 차지하는 집세는 그 변동에 따라 도시가구 특히 서민들의 생활에 크게 영향을 주기 때문에 서민들은 물론이고 국가 및 지방자치단체에서도 서민생활의 안정을 위해 집세변동을 중요한 사항 중 하나로 인식하고 있다. 이처럼 집세의 정확한 측정은 중요한 사항임에도 불구하고 2년에 1번 재계약하는 관례로 인해 집세 변동율이 낮아(2002년 평균 2.8%) 하나의 통계단위에서 집세 지수를 산출하기 위해서는 통계단위별로 330가구 이상의 표본가구가 필요하나, 통계청에서 실시하는 가구표본은 다목적 표본으로 집세이외에도 경제활동인구조사, 도시가계조사 등 각종 표본조사를 실시하여야 하므로 집세조사에 필요한 표본가구를 확보하기가 어렵다.

그러나 집세의 전국 움직임과 이용자가 현재 거주하고 있는 도시의 동향과의 차이로 집세지수의 이용이 어렵고 지방자치체의 정착으로 도시별 집세지수에 대한 수요는 꾸준히 증가하고 있어 집세지수의 작성은 향후 실현되어야 할 과제임에 틀림없다. 이를 위해 표본 설계당시에는 통계 생산단위로 고려하지 않았으나, 통계조사를 실시한 후에 좀 더 세분화된 단위의 통계를 생산하는데 이용될 수 있는 통계 생산기법 중 하나인 소지역 추정법을 이용하는 방안을 검토해 보려 한다. 이에 대한 연구는 영국, 캐나다[1] 및 미국[2] 등에서 활발히 진행중이며, 우리나라의 경우 시·군·구의 실업통계를 합성추정법과 복합추정법을 이용하여 추정한 사례가 시도되고 있으며[3,4], 여타 소표본군의 통계 추정에도 적용될 필요성이 증대되고 있다.

본 연구에서는 먼저 집세조사의 특징을 살펴보고 소지역 추정법을 이용하여 경기도의 6개 도시(수원, 성남, 의정부, 안양, 부천, 고양)의 집세변동을 추정하고, 또한 직접 추정법에 의한 방법과의 비교를 통해 소표본 통계단위에서의 집세 변동률 추정의 타당성 여부를 검토하였다.

1) 통계청, 대전 서구 둔산동 정부3청사 3동 1506호

## 2. 소지역 추정법에 의한 소표본 통계단위의 집세변동을 추정

### 2.1 집세조사

부동산 임대금액의 변동에 대한 조사인 집세조사는 부동산을 임대한 임대가구를 통하여 조사하는 방법 및 부동산 공인중개사 사무실을 통해서 조사하는 방법 등 여러가지가 있다.

본 연구에서 이용되는 집세조사자료는 임대가구를 통해 실시된 것이며, 임대가구를 통하여 조사하는 방법은 조사대상가구 전체의 임대금액의 변동을 조사하여 소비자가 부담하고 있는 임대금액의 변동을 파악할 수 있으나, 발생빈도가 적어 표본규모가 커야하는 단점이 있다. 집세를 결정하는 요인으로 여기에서는 주택의 유형(단독, 공동주택 2가지로 분류) 및 점유형태의 2가지 요인으로 압축하여 살펴보았다.

### 2.2 소지역 추정법

소지역 추정법은 통계생산 단위의 표본수가 적거나, 표본이 없는 경우에 통계생산단위와 통계적 특성이 유사하다고 판단되는 주변의 조사정보를 통합하여 세분화된 통계 생산단위의 통계를 생산하는 방법으로서 미국이나 캐나다 등에서는 센서스 중간 년도에 해당되는 해의 주 또는 county의 인구추정 시 사용되며 또한 우리나라에서는 주로 소지역의 실업자 추정에 대한 연구에 많이 적용되었다[5,6,7,8].

본 논문에서는 해당 소지역에서 조사된 자료에 의한 직접추정법과 유사한 소지역들의 정보를 추가시킨 합성추정법, 그리고 이들을 가중 평균한 복합추정법에 의하여 집세변동을 산출하였으며, 9개도 중 도내 및 도시간의 특징이 비슷한 경기도가 도시별 편향이 적다고 판단되어 경기도에 대하여 집세 변동률을 추정해 보았다.

먼저 직접추정량은 보통 해당 소지역에서 조사된 자료만을 이용하여 추정하며 불편추정량이나 소지역의 표본의 크기가 작을 경우에는 분산이 커지기 때문에 신뢰성에 문제가 있을 수 있다.

$$\widehat{R}_i = \sum_{s=1}^2 w_s {}_sR_i = \sum_{s=1}^2 w_s \left( \frac{\sum_{h=1}^{n_i} {}_sY_{ih}}{\sum_{h=1}^{n_i} {}_sX_{ih}} - 1 \right) \quad (1)$$

여기에서  $s$ 는 점유형태(전세, 월세)를 나타내는 첨자

$\widehat{R}_i$ 는 소지역  $i$ 의 집세변동률

$w_s$ 는 점유형태별(전세, 월세)가중치

$n_i$ 는  $i$ 지역의 집세대상 가구수

${}_sX_{ih}$ 는 소지역  $i$ 의 임대가구의 전월 집세금액

${}_sY_{ih}$ 는 소지역  $i$ 의 임대가구의 금월 집세금액

표 1은 경기도의 6개 도시의 최근 2002년 및 2003년 1월, 2월에 대한 집세변동률에 대한 조사된 자료에 기초한 직접추정법으로 산출한 집세지수이다. 집세변동의 발생주기가 2년으로 해당 조사기간 안에 변동되는 표본의 월별 반영효과가 매우 적다. 그림 1은 이러한 직접추정에 의한 월별 집세지수의 변동추이를 나타낸 것으로서 안양시의 증가율이 매우 크며, 반면에 고양

시의 증가율이 둔한 편이다. 경기도의 6개 도시의 표본 내 임대가구의 집세 변동률은 표본의 크기가 작기 때문에 집세변동이 발생하지 않은 달이 많으며 집세변동 발생시 임대금액의 변동이 큰 경우는 지수가 급격히 움직이기 때문에 매월의 움직임이 잘 반영된다고 보기는 어렵다. 이러한 소표본에 의한 분산의 증가로 인해 직접추정법에 의한 통계자료의 신뢰성을 증대시키기 위하여 합성추정법을 도입하게 된다.

<표 1> 직접추정법에 의해 산출된 지수 (2002.1=100.0)

|     | 2002년 |       |       |       |       |       |       |       |       |       |       |       | 2003년 |       |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     | 1월    | 2월    | 3월    | 4월    | 5월    | 6월    | 7월    | 8월    | 9월    | 10월   | 11월   | 12월   | 1월    | 2월    |
| 수원  | 100.0 | 100.1 | 100.7 | 101.6 | 101.8 | 101.9 | 102.0 | 102.6 | 102.7 | 103.0 | 103.4 | 103.6 | 104.6 | 105.1 |
| 성남  | 100.0 | 101.6 | 102.7 | 103.1 | 103.3 | 104.8 | 104.9 | 105.2 | 105.6 | 105.8 | 105.9 | 106.3 | 106.3 | 106.3 |
| 의정부 | 100.0 | 101.1 | 102.9 | 102.9 | 103.0 | 103.6 | 104.0 | 105.2 | 106.3 | 106.3 | 106.3 | 106.4 | 106.4 | 106.4 |
| 안양  | 100.0 | 100.9 | 101.7 | 102.6 | 103.6 | 103.6 | 107.7 | 107.8 | 109.8 | 111.1 | 111.2 | 111.3 | 111.4 | 112.1 |
| 부천  | 100.0 | 100.0 | 99.9  | 102.8 | 103.2 | 103.5 | 103.5 | 103.5 | 103.5 | 103.6 | 103.6 | 103.6 | 103.8 | 103.9 |
| 고양  | 100.0 | 100.5 | 101.0 | 101.4 | 101.6 | 101.6 | 101.6 | 101.6 | 101.6 | 101.6 | 101.9 | 101.9 | 101.9 | 102.0 |

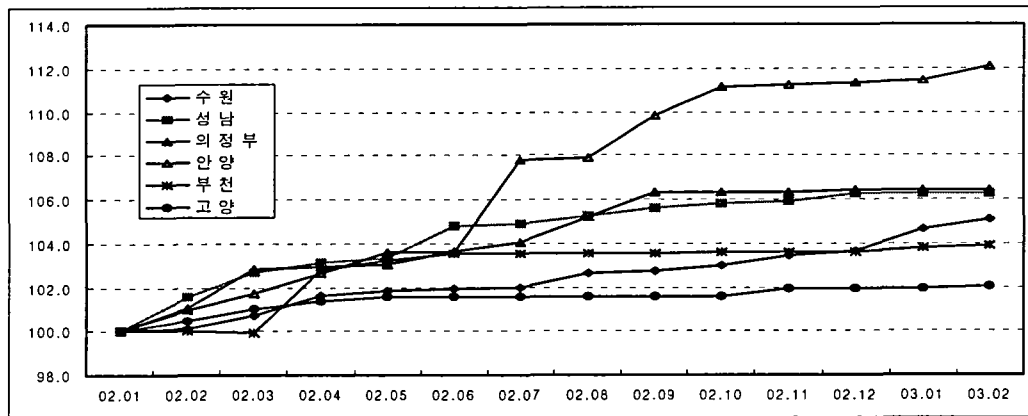


그림 1. 직접추정에 의한 월별 집세지수 변동추이

합성추정법은 추정하고자 하는 소지역과 유사한 소지역들의 정보를 이용하여 추정값의 정도를 높이고자 하는 추정방식으로, 분산은 직접추정량의 분산에 비해 작으나, 전체한 가정이 성립하지 않을 경우에는 심각한 편향이 발생할 수 있다.

$$\bar{R}_i = \sum_j w_{ij} R_{.j} = \sum_j w_{ij} \left[ \frac{\sum_{h=1}^{n_j} Y_{.jh}}{\sum_{h=1}^{n_j} X_{.jh}} - 1 \right] \quad (2)$$

여기에서,  $R_{.j}$ 는  $j$ 유형(단독전세, 단독월세, 공동주택전세, 공동주택월세)별 집세변동률

$X_{.jh}$ 는  $j$ 유형 임대가구의 전월 집세금액

$Y_{.jh}$ 는  $j$ 유형 임대가구의 금월 집세금액

$w_{ij}$ 는  $i$  번째 소지역의  $j$ 유형별 가중치

(2000년 인구주택총조사 주택유형별 구성비와 가구표본에서 나온 주택 유형별 평균 임차료로부터 산출)

소표본 통계단위에서의 집세 변동률 추정

표 2는 2000년 인구주택총조사 자료결과에 의한 점유형태별, 주택유형별 집세 구성비를 나타낸 것으로 전세 및 월세의 점유형태와 단독주택 및 공동주택의 주택유형별로 분류하였다. 표 3은 이러한 주택유형별 가중치를 고려하여 소지역 조사자료를 추정된 합성추정법에 의해 산출된 집세지수이다. 그림 2는 합성추정에 의한 월별 집세지수 변동추이로서 타 도시에 비해 상대적으로 고양시가 증가율이 크게 나타나며, 이는 직접추정법에 의한 고양시의 타도시에 상대적으로 둔화된 증가율과는 다른 양상을 보인다.

<표 2> 점유형태별, 주택유형별 구성비 (%)

|     | 전 세   |       | 월 세   |       |
|-----|-------|-------|-------|-------|
|     | 단 독   | 공동주택  | 단 독   | 공동주택  |
| 수 원 | 54.27 | 45.73 | 85.61 | 14.39 |
| 성 남 | 64.63 | 35.37 | 72.44 | 27.56 |
| 의정부 | 51.62 | 48.38 | 80.74 | 19.26 |
| 안 양 | 48.22 | 51.78 | 80.71 | 19.29 |
| 부 천 | 60.88 | 39.12 | 83.39 | 16.61 |
| 고 양 | 32.06 | 67.94 | 62.74 | 37.26 |

※ 자료 : 2000년 인구주택총조사 결과

<표 3> 합성추정법에 의해 산출된 지수 (%)

|     | 2002년 |       |       |       |       |       |       |       |       |       |       |       | 2003년 |       |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     | 1월    | 2월    | 3월    | 4월    | 5월    | 6월    | 7월    | 8월    | 9월    | 10월   | 11월   | 12월   | 1월    | 2월    |
| 수원  | 100.0 | 100.6 | 101.4 | 102.2 | 102.6 | 103.0 | 103.6 | 103.9 | 104.3 | 104.6 | 104.8 | 105.0 | 105.1 | 105.3 |
| 성남  | 100.0 | 100.6 | 101.4 | 102.2 | 102.6 | 103.0 | 103.5 | 103.9 | 104.2 | 104.5 | 104.8 | 105.0 | 105.1 | 105.3 |
| 의정부 | 100.0 | 100.6 | 101.5 | 102.2 | 102.6 | 103.1 | 103.7 | 104.0 | 104.4 | 104.6 | 104.9 | 105.1 | 105.2 | 105.4 |
| 안양  | 100.0 | 100.7 | 101.5 | 102.2 | 102.6 | 103.1 | 103.7 | 104.1 | 104.4 | 104.7 | 104.9 | 105.1 | 105.2 | 105.4 |
| 부천  | 100.0 | 100.6 | 101.4 | 102.2 | 102.5 | 103.0 | 103.5 | 103.9 | 104.2 | 104.5 | 104.8 | 105.0 | 105.1 | 105.3 |
| 고양  | 100.0 | 100.8 | 101.6 | 102.4 | 102.8 | 103.3 | 104.0 | 104.4 | 104.7 | 105.0 | 105.2 | 105.4 | 105.6 | 105.7 |

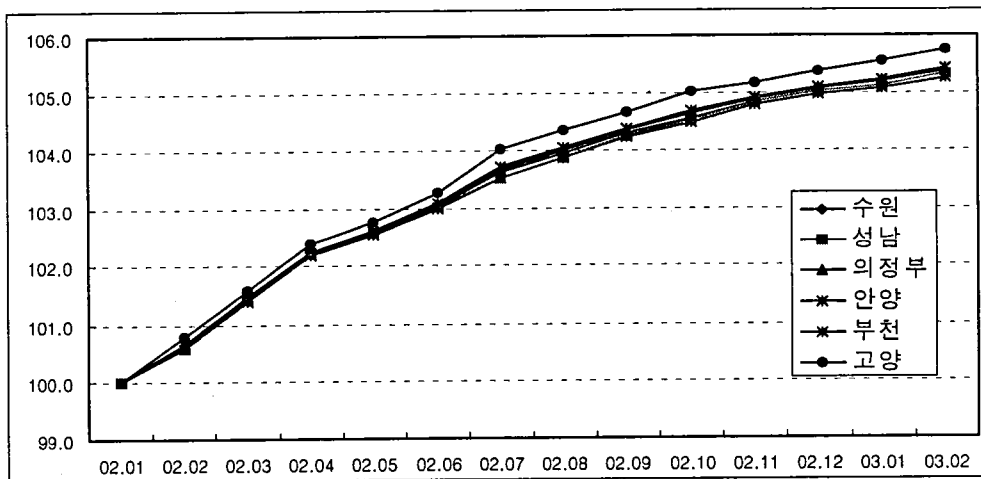


그림 2. 합성추정에 의한 월별 집세지수 변동추이

소지역에 배정된 표본수가 적기 때문에 표본조사만을 이용한 직접추정량의 불안정에서 오는 낮은 신뢰성과 합성추정량의 편향을 보완하기 위해서 직접추정량과 합성추정량의 가중평균을 사용하는 복합 추정량은 가중치에 따라 도시의 특성과 무변동을 어느 정도 수정할 수 있다.

$$\widehat{R}_i^c = w_i \widehat{R}_i + (1 - w_i) \widehat{R}_i^s \quad (3)$$

여기에서  $\widehat{R}_i$ 는 표본조사에서 직접 계산한 추정량이며,  $\widehat{R}_i^s$ 는 합성추정량을 나타낸다.  $w_i$ 는 가중값으로 0과 1사이의 값이다.

여기서, 평균제곱오차  $MSE(\widehat{R}_i^c)$ 를 최소화하는 가중치  $w_i$ 는 아래와 같다.

$$w_{i(opt)} = \frac{MSE(\widehat{R}_i^s)}{MSE(\widehat{R}_i^s) + V(\widehat{R}_i)} \quad (4)$$

여기서 또한,  $\widehat{R}_i^s$ 의 편향이 없다고 가정하고  $MSE(\widehat{R}_i^s)$  대신 분산  $V(\widehat{R}_i^s)$ 을 비추정법에 의하여 분산값을 계산해본 결과 평균 가중치  $w_i$ 가 0.562로 계산되었다.

그림 3은 평균가중치 0.56을 갖는 복합추정에 의한 월별 집세지수 변동추이를 나타낸 것으로 14개월의 기간동안에 평균적으로 완만히 증가하는 추세를 나타낸다. 특히 안양시의 증가율은 직접추정법에 의한 경우보다 감소되어지고 고양시의 경우는 약간 증가되어지는 양상을 보인다.

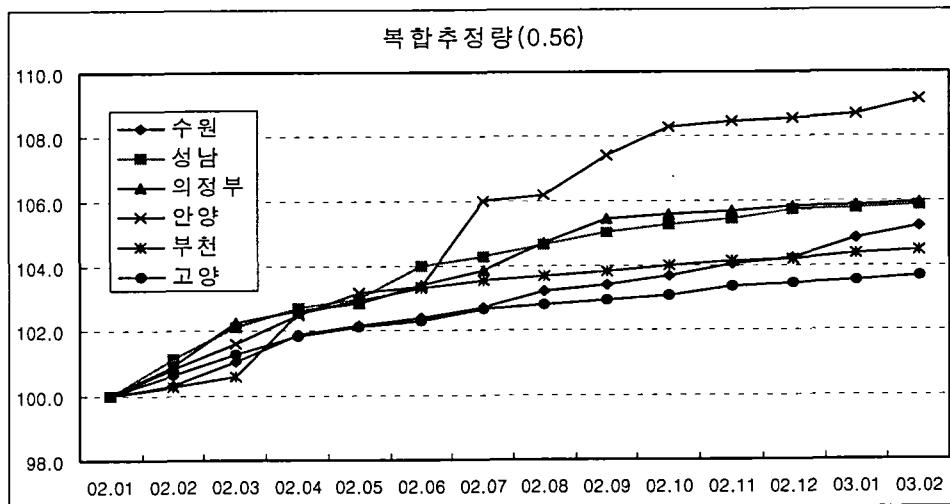


그림 3. 복합추정에 의한 월별 집세지수 변동추이

또한  $w_i$ 를 바꾸어 가면서 시뮬레이션 한 결과 각도시의 개별 움직임이 나타나기 위해서는 0.5이상의 가중치가 적당하였고, 이 중 집세 무변동이 계속되는 현상을 제거하기 위해서는 가중치로가 0.6이하가 적당함을 확인하였다.

## 소표본 통계단위에서의 집세 변동률 추정

검토 결과 소표본 통계단위에서는 직접추정법에 의한 집세변동률은 매월의 움직임을 반영하기 어려우며, 합성추정량은 매월의 움직임을 잘 반영하고 있는 듯하나 도시별 움직임을 반영하지 못하고 경기도 평균을 반영함으로써 6개 도시가 거의 같은 움직임을 보인다. 매월의 집세변동이 도시별로 특성을 가지면서 매월 집세변동을 어느 정도 반영할 수 있는 복합추정량이 집세지수 산출 시 가장 적당하다고 판단되며 복합추정량을 활용함으로써 오히려 표본규모를 확대하기 어려운 현실을 고려해 볼 때 소표본 도시의 집세변동을 추정의 한계와 어느 정도의 신뢰수준을 갖춘 통계를 생산하기 위해서는 이러한 소표본추정법의 도입이 필요함을 확인하였다.

### 3. 결론

본 연구에서는 소표본 통계단위에서의 집세 변동률 추정을 경기도 지역의 6개 도시에서의 집세변동을 추정하는데 적용하였다. 발생빈도가 2년주기를 갖는 집세를 도시별로 특성을 가지면서 매월 집세변동에 어느 정도 반영하기 위해서는 복합추정법의 적용이 집세지수 산출 시 가장 적당함을 확인하였다. 특히 복합추정량 산출을 위해 사용되는 가중치의 산출방법으로 전월 대비 가구별 집세 변동율을 분산으로 상관시키는 시도를 하였다. 표본특성이 비슷한 지역에 대해서는 광역시 및 도의 경계에 관계없이 좀 더 세분화된 표본군으로 묶어 통계생산이 가능하다고 사료된다. 결론적으로, 발생주기가 비교적 긴 실업자지수나 집세지수 등 표본규모의 확대가 어려운 소규모 표본군에 대한 한계를 극복하기 위해서는 소지역추정법에 의한 통계생산이 유용함을 확인하였으며, 이를 다양한 분야에서 확대 적용할 수 있도록 가중치의 산출방법에 대한 체계적인 이론의 정립이 선행되어야 실용화가 확대될 수 있겠다.

### 참고문헌

- [1] Drew, J. D., Sinho, M. P., and Choudhry, G.H.(1982), "Evaluation of small area techniques for the Canadian Labour Force Survey." Survey Methodology, 8,17-47
- [2] Tiller,R.B. and Brown,S.(1993), 'Bureau of Labor Statistics' State and Local Area Estimates of Employment and Unemployment", Indirect Estimators in Fedral Programs, 5-1~19
- [3] 이계오, 류계복, 이귀형 (1999), "소지역추정법을 이용한 충북 시·군·구의 실업자 추정", 소지역통계워크샵, 37-56.
- [4] 이계오 (2000), "소지역 추정법", 공군사관학교 전산통계학과
- [5] 이계오, 김규영 (2002), "시군구 실업자 추정에서 분산 추정", 2002년 춘계학술발표회 논문집, 7-12.
- [6] 이계오 (2002) "소지역 추정법에 의한 시군구 실업통계개발", 통계청 연구보고서
- [7] 통계청 (1997), "지역통계작성의 현황과 발전방향"
- [8] 통계청 (2001), "소지역통계 추정법", 통계기획국 조사관리과