The Korean Journal of Applied Statistics (2018) DOI: http://dx.doi.org/10.5351/KJAS.2018.31.3.397
31(3), 397–408

A bias adjusted ratio-type estimator

Jung-Taek Oh[a] · Key-Il Shin[a,1]

[a]Department of Statistics, Hankuk University of Foreign Studies

(Received May 28, 2018; Revised June 7, 2018; Accepted June 8, 2018)

Abstract

Various methods for accurate parameter estimation have been developed in a sample survey and it is also common to use a ratio estimator or the regression estimator using auxiliary information. The ratio-type estimator has been used in many recent studies and is known to improve the accuracy of estimation by adjusting the ratio estimator. However, various studies are under way to solve it since the ratio-type estimator is biased. In this study, we propose a generalized ratio-type estimator with a new parameter added to the ratio-type estimator to remove the bias. We suggested a method to apply this result to the parameter estimation under the error assumption of heteroscedasticity. Through simulation, we confirmed that the suggested generalized ratio-type estimator gives good results compared to conventional ratio-type estimators.

Keywords: bias, unbiased estimator, heteroscedastic, MLE, Taylor approximation

1. 서론

표본조사에서 모수 추정의 정확성 향상을 위한 많은 이론과 방법이 개발되었고 최근에는 비형태추정량(ratio-type estimator; RT)에 대한 연구가 활발히 이루어지고 있다. 비형태추정량은 비추정량의 특징인 계산의 편리성, 사용의 용이성을 유지하면서도 대표본의 경우 회귀추정량의 장점인 정확성과 정밀성을 갖는 것으로 알려져 있다. 특히 자료 형태가 비모형을 만족하지 않고 절편이 있는 모형의 경우 추정에 편향이 발생하는 비추정량의 단점을 보완할 수 있어 실제 자료 분석에 매우 유용하게 사용될 수 있다.

Singh과 Vishwakarma (2007)는 이중추출법에서 한 개의 보조변수가 있을 때 양의 비례 관계 시 사용하는 지수적 비추정량(exponential ratio estimator)과 음의 관계 시 사용하는 지수적 곱추정량(exponential product estimator)을 연구하였다. Koyuncu와 Kadilar (2009)에서는 두 개의 보조변수가 있을 때 양의 비례 관계 시 사용하는 비추정량(ratio estimator), 음의 비례 관계 시 사용하는 곱추정량(product estimator) 그리고 비추정량과 곱추정량의 곱으로 얻어지는 비곱추정량(ratio cum product estimator)을 제안하였으며 Singh 등 (2010)에서는 이중추출법에서 캘리브레이션 기법을 통해 얻어지는 비추정량과 곱추정량 그리고 회귀추정량(regression estimator)을 연구하였다. Tailor 등 (2014)과 Lone 등 (2015)은 이중추출법에서 비형태지수추정량(ratio type exponential estimator)과 곱형태지수추정량(product type exponential estimator) 그리고 일반화 비곱형태지수추정량(generalized ratio-cum-product type exponential estimator)을 제안하였고, Tailor 등 (2015)은 이중추출법에서

This research was supported by the Hankuk University of Foreign Studies research fund (2018).

[1]Corresponding author: Department of Statistics, Hankuk University of Foreign Studies, 81, Oedae-ro, Mohyeon-eup, Cheoin-gu, Yongin-si, Gyeonggi-do 17035, Korea. E-mail: keyshin@hufs.ac.kr

일반화 비곱형태추정량(generalized ratio-cum-product type estimator)을 연구하였다. 또한 Etuk 등 (2016), Haq 등 (2017)과 Clement와 Enang (2017)은 비추정량의 정확성 향상을 위한 새로운 추정량을 제안하였으며 Oh와 Shin (2017)은 Taylor 근사식을 이용한 비형태추정량과 모형에서 오차 분산에 포함된 모수를 최대가능도추정량(maximum likelihood estimator; MLE)을 이용하여 추정하는 비형태추정량을 제안하였다.

하지만 비추정량과 비형태추정량은 편향이 있는 것으로 알려져 있다. 이러한 편향은 특히 소표본에서 추정의 정확성에 큰 영향을 주고 있어 이를 해결하기 위한 연구가 수행되었다. Singh와 Singh (1998)은 비추정량에서 편향을 줄이기 위한 연구를 실시하였고 Swain (2014)은 한 개의 보조변수가 있을 때 제곱근 변환을 활용한 비형태추정량을 연구하였다. 그리고 Sharma와 Singh (2014)은 두 개의 보조변수가 있을 때 새로운 비형태추정량을 제안하였고 Khan과 Singh (2015)은 새로운 Chain Ratio-Type Estimator를 제안하는 등 보조정보를 이용하여 비형태추정량의 편향과 mean squared error (MSE)를 줄이기 위한 많은 연구가 이루어졌다.

본 연구에서는 절편이 있는 회귀모형에서 편향을 제거하기 위해 Oh와 Shin (2017)이 제안한 비형태추정량에 편향 제거를 위한 모수를 추가하고 Taylor 2차 근사를 이용하여 추가된 모수를 추정하는 방법을 사용한 일반화 비형태추정량(generalized ratio-type estimator; GRT)을 제안하였으며 모의실험을 통하여 제안된 추정량의 우수성을 확인하였다.

본 논문의 구성은 다음과 같다. 먼저 2절에서는 회귀추정량과 비형태추정량 그리고 비형태추정량의 편향과 MSE를 살펴보았으며 등분산성을 만족하지 않는 회귀추정량도 살펴보았다. 3절에서는 편향을 제거하기 위한 새로운 일반화 비형태추정량을 제안하였다. 4절에서는 제안된 추정량과 기존의 비형태추정량의 성능을 비교하기 위한 모의실험이 수행되었으며 마지막으로 5절에 결론이 있다.

2. 회귀추정량과 비형태추정량

2.1. 비형태추정량

보조변수가 있는 표본조사에서 추정량의 정확성 향상을 위해 일반화회귀추정량(generalized regression estimator)이 사용된다. 이 중에서 보조변수가 하나인 단순회귀모형에서 보조변수의 모평균 $\bar{X}$가 알려진 경우의 회귀추정량은 다음과 같다 (linear regression estimator; LR).

$$\hat{\bar{Y}}_{\mathrm{LR}} = \hat{\beta}_0 + \hat{\beta}_1\bar{X} \tag{2.1}$$

또한 절편이 '0'이고 보조변수 값이 x_i일 때, 오차의 분산이 $x_i\sigma^2$인 비모형에서는 $\hat{\bar{Y}}_R = \bar{y}(\bar{X}/\bar{x})$인 비추정량이 사용된다. 최근 많은 논문에서 비추정량을 일반화한 비형태추정량이 연구되고 있으며 비형태추정량은 다음과 같다.

$$\hat{\bar{Y}}_{\mathrm{RT}} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha}, \tag{2.2}$$

여기서 Oh와 Shin (2017)은 비형태추정량을 Taylor 1차 근사를 이용하여 식 (2.1)과 (2.2)인 회귀추정량과 비형태추정량의 관계를 연구하였다. 즉

$$\hat{\bar{Y}}_{\mathrm{RT}} \approx \bar{y}\left(1+\alpha\left(\frac{\bar{X}}{\bar{x}}-1\right)\right) = \bar{y}(1-\alpha) + \frac{\bar{y}}{\bar{x}}\alpha\bar{X} \tag{2.3}$$

가 되므로 $\hat{\beta}_0 = \bar{y}(1-\hat{\alpha})$, $\hat{\beta}_1 = (\bar{y}/\bar{x})\hat{\alpha}$라 하면 근사적으로 두 식이 같아지게 된다. 또한 두 식을 동시에 만족하면서 MSE를 최소로 하는 α추정량은 $\hat{\alpha} = \hat{\beta}_1/\hat{R}$, $\hat{R} = \bar{y}/\bar{x}$으로 얻어진다. 이러한 결과는 Tailor

등 (2015)에서도 확인할 수 있다. 그러나 Oh와 Shin (2017)의 모의실험 결과를 살펴보면 $\hat{\alpha} = \hat{\beta}_1/\hat{R}$, $\hat{R}$을 사용할 때 소표본에서 편향이 크게 발생하는 것을 확인할 수 있다.

2.2. 비형태추정량의 편향과 MSE

이 절에서는 비형태추정량인 식 (2.2)의 편향과 MSE를 살펴보았다. 먼저 오차를 $e_0 = (\bar{y} - \bar{Y})/\bar{Y}$, $e_1 = (\bar{x} - \bar{X})/\bar{X}$라 하면 $\bar{y} = \bar{Y}(1 + e_0)$, $\bar{x} = \bar{X}(1 + e_1)$이 된다. 이러한 오차 가정은 비형태추정량과 관련된 연구에서 흔히 사용하는 것으로 자세한 내용은 Singh와 Vishwakarma (2007)와 Sharma와 Singh (2014)을 참조하기 바란다. 따라서 $E(e_0) = E(e_1) = 0$이고 또한 다음을 만족하게 된다.

$$\mathrm{Var}(e_0) = \frac{1}{\bar{Y}^2}\frac{S_y^2}{n}(1-f), \qquad \mathrm{Var}(e_1) = \frac{1}{\bar{X}^2}\frac{S_x^2}{n}(1-f), \qquad \mathrm{Cov}(e_0, e_1) = \frac{1}{\bar{X}\bar{Y}}\frac{S_{xy}}{n}(1-f),$$

여기서 $f = n/N$, $S_x^2 = \sum_{i=1}^N (x_i - \bar{X})^2/(N-1)$, $S_y^2 = \sum_{i=1}^N (y_i - \bar{Y})^2/(N-1)$, $S_{xy} = \sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y})/(N-1)$이다. 또한 비형태추정량의 Taylor 2차 근사를 이용하면 다음의 결과를 얻는다.

$$\begin{aligned}
\hat{\bar{Y}}_{\mathrm{RT}} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha} &= \bar{Y}(1+e_0)(1+e_1)^{-\alpha} \\
&\approx \bar{Y}(1+e_0)\left((1-\alpha e_1) + \frac{\alpha(\alpha+1)}{2}e_1^2\right) \\
&= \bar{Y}\left(1 + e_0 - \alpha e_1 - \alpha e_0 e_1 + \frac{\alpha(\alpha+1)}{2}e_1^2 + \frac{\alpha(\alpha+1)}{2}e_0 e_1^2\right).
\end{aligned}$$

따라서 근사적인 편향과 MSE는 다음과 같이 얻어진다.

$$\begin{aligned}
\mathrm{Bias}\left(\hat{\bar{Y}}_{\mathrm{RT}}\right) &\approx \bar{Y}E\left(e_0 - \alpha e_1 - \alpha e_0 e_1 + \frac{\alpha(\alpha+1)}{2}e_1^2 + \frac{\alpha(\alpha+1)}{2}e_0 e_1^2\right) \\
&= \bar{Y}\left(-\alpha\frac{1}{\bar{X}\bar{Y}}\frac{S_{xy}}{n}(1-f) + \frac{\alpha(\alpha+1)}{2}\frac{1}{\bar{X}^2}\frac{S_x^2}{n}(1-f)\right), \\
\mathrm{MSE}\left(\hat{\bar{Y}}_{\mathrm{RT}}\right) &\approx \bar{Y}^2 E\left(e_0 - \alpha e_1 - \alpha e_0 e_1 + \frac{\alpha(\alpha+1)}{2}e_1^2 + \frac{\alpha(\alpha+1)}{2}e_0 e_1^2\right)^2 \\
&\approx \bar{Y}^2 E\left(e_0 - \alpha e_1 - \alpha e_0 e_1\right)^2 \\
&= \bar{Y}^2\left(\frac{1}{\bar{Y}^2}\frac{S_y^2}{n}(1-f) - 2\alpha\frac{1}{\bar{X}\bar{Y}}\frac{S_{xy}}{n}(1-f) + \alpha^2\frac{1}{\bar{X}^2}\frac{S_x^2}{n}(1-f)\right).
\end{aligned}$$

이제 MSE를 최소로 하는 α를 구하면 $\alpha = (S_{xy}/S_x^2)/(\bar{Y}/\bar{X})$가 되는 것을 확인할 수 있다. 따라서 α의 추정량은 $\hat{\alpha} = (s_{xy}/s_x^2)/(\bar{y}/\bar{x}) = \hat{\beta}_1/\hat{R}$이고 여기서 $\hat{R} = \bar{y}/\bar{x}$, $s_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2/(n-1)$, $s_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})/(n-1)$이므로 식 (2.3)에서 얻어진 값과 일치하게 된다. 또한 최소제곱추정량으로 얻어진 절편추정량은 $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$이므로 추정된 $\hat{\alpha}$을 사용한 결과와 일치하게 된다. 물론 이 추정량의 경우 분산이 등분산성을 만족한다는 가정 하에서 얻어진 결과이다. 결론적으로 비형태추정량은 편향이 존재한다.

2.3. 등분산성을 만족하지 않을 때의 회귀추정량

회귀모형식이 다음과 같을 경우 식 (2.1)의 회귀추정량은 최적이 아니다.

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \overset{\mathrm{iid}}{\sim} N\left(0, x_i^{\nu}\sigma^2\right), \; 0 \le \nu, \tag{2.4}$$

여기서 만약 ν가 알려져 있다면 알려진 ν와 가중최소제곱추정법(weighted least squares estimation; WLSE)을 이용하여 β_0, β_1을 추정할 수 있다. 즉 $\beta = (\beta_0, \beta_1)^T$라 하면

$$\hat{\beta}_{\rm WLSE} = \left(\hat{\beta}_{0\rm WLSE}, \hat{\beta}_{1\rm WLSE}\right)^T = \left(X^T W^{-1} X\right)^{-1} \left(X^T W^{-1} y\right)$$

로 추정할 수 있으며 여기서 $W = {\rm diag}\{x_1, x_2, \ldots, x_n\}$인 대각행렬이다. 따라서 평균 추정량은 다음과 같다.

$$\hat{\bar{Y}}_{\rm WLSE} = \hat{\beta}_{0\rm WLSE} + \hat{\beta}_{1\rm WLSE}\bar{X} \tag{2.5}$$

또한 ν가 알려져 있지 않다면 MLE를 이용하여 $\beta = (\beta_0, \beta_1)^T$를 추정할 수 있다. 예를 들면 오차가 정규분포를 따를 경우 다음의 로그-우도함수를 이용하여 모수를 추정할 수 있다.

$$l\left(y|\beta_0, \beta_1, \nu, \sigma^2\right) = -\sum_{i=1}^{n} \log\left(x_i^{\frac{\nu}{2}}\right) - \frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left\{\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{x_i^{\nu}}\right\}. \tag{2.6}$$

이때 얻어진 추정량을 $\hat{\beta}_{\rm MLE} = (\hat{\beta}_{0\rm MLE}, \hat{\beta}_{1\rm MLE})^T$라 하면 회귀추정량은 다음과 같이 얻어진다(linear regression using MLE; MLR).

$$\hat{\bar{Y}}_{\rm MLR} = \hat{\beta}_{0\rm MLE} + \hat{\beta}_{1\rm MLE}\bar{X}, \tag{2.7}$$

여기서 2.2절에서 얻어진 편향과 MSE 결과는 식 (2.5) 또는 (2.7)에서 정의된 추정량을 사용한 경우에서 얻어진 결과와는 다르게 된다. 특히 Oh와 Shin (2017)의 결과를 보면 절편이 있는 회귀모형에서 MLE를 이용한 비형태추정량은 회귀추정량에 비해 편향이 커지고 따라서 MSE도 커지는 것을 확인할 수 있다. 하지만 식 (2.1), (2.5), 그리고 (2.7)과 같은 회귀추정량은 모두 같은 형태이므로 비형태추정량을 회귀추정량 형태로 근사하여 추정된 모수를 적용하면 편향과 MSE를 줄일 수 있게 된다.

3. 제안된 일반화 비형태추정량

식 (2.5), (2.7)과 같이 등분산성을 만족하지 않을 때 WLSE와 MLE를 이용한 경우에는 $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$와 같은 성질을 만족하지 않으며 또한 절편이 있는 회귀모형의 비형태추정량은 편향이 크게 발생하고 있다. 이에 본 연구에서는 식 (2.3)의 Taylor 1차 근사식을 확장한 2차 근사식을 사용하여 편향을 줄일 수 있는 새로운 추정량을 제안하였다. 이를 위해 식 (2.2)에 모수 c를 추가한 다음의 비형태추정량을 고려하자.

$$\hat{\bar{Y}}_{\rm RT} = \bar{y}(1+c)\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha}. \tag{3.1}$$

Taylor 2차 근사식을 사용하면

$$\begin{aligned}\hat{\bar{Y}}_{\rm RT} &\approx \bar{y}(1+c)\left(1 + \alpha\left(\frac{\bar{X}}{\bar{x}} - 1\right) + \frac{\alpha(\alpha-1)}{2}\left(\frac{\bar{X}}{\bar{x}} - 1\right)^2\right) \\ &= \bar{y}(1+c)\left[(1-\alpha) + \frac{\alpha(\alpha-1)}{2}\left(\frac{\bar{X}}{\bar{x}} - 1\right)^2\right] + \bar{y}(1+c)\frac{\alpha}{\bar{x}}\bar{X}\end{aligned}$$

이 된다. 이제

$$\hat{\beta}_0 = \bar{y}(1+\hat{c})\left[(1-\hat{\alpha}) + \frac{\hat{\alpha}(\hat{\alpha}-1)}{2}\left(\frac{\bar{X}}{\bar{x}} - 1\right)^2\right], \qquad \hat{\beta}_1 = \bar{y}(1+\hat{c})\frac{\hat{\alpha}}{\bar{x}} \tag{3.2}$$

을 만족하는 $\hat{\alpha}$, $\hat{c}$을 찾으면 편향을 줄이면서 모수 $0 \leq \nu$을 고려한 비형태추정량을 만들 수 있다. 즉 $\hat{\beta}_0$ 대신에 $\hat{\beta}_{0\text{MLE}}$ 그리고 $\hat{\beta}_1$ 대신에 $\hat{\beta}_{1\text{MLE}}$ 등을 이용한다면 $0 \leq \nu$인 경우에도 사용할 수 있게 된다. 이제 식 (3.2)를 정리하면

$$c = \frac{1}{\bar{y}}\left(\hat{\beta}_0 + \hat{\beta}_1\bar{x} - \bar{y} - \hat{\beta}_1\bar{x}\frac{(\hat{\alpha}-1)}{2}\left(\frac{\bar{X}}{\bar{x}} - 1\right)^2\right)$$

이 되고 여기에 $\hat{\alpha} = \hat{\alpha}_{\text{MLE}} = \hat{\beta}_{1\text{MLE}}(\bar{x}/\bar{y})$값을 대입하면 c의 추정값이 얻어진다. 즉

$$\hat{c}^{(1)} = \frac{1}{\bar{y}}\left(\hat{\beta}_{0\text{MLE}} + \hat{\beta}_{1\text{MLE}}\bar{x} - \bar{y} - \hat{\beta}_{1\text{MLE}}\bar{x}\frac{(\hat{\alpha}_{\text{MLE}}-1)}{2}\left(\frac{\bar{X}}{\bar{x}} - 1\right)^2\right) \tag{3.3}$$

가 된다. 또한 식 (3.3)에서 구한 $\hat{c}^{(1)}$을 이용하여

$$\hat{\alpha}_{\text{ADJ}} = \hat{\beta}_{1\text{MLE}}\left(\frac{\bar{x}}{\bar{y}}\right)\Big/\left(1+\hat{c}^{(1)}\right) = \frac{\hat{\alpha}_{\text{MLE}}}{1+\hat{c}^{(1)}}$$

가 얻어진다. 최종적으로 $\hat{c}_{\text{ADJ}}$는 식 (3.3)의 $\hat{\alpha}_{\text{MLE}}$에 $\hat{\alpha}_{\text{ADJ}}$를 대입하여 다음의 식으로 얻어진다.

$$\hat{c}_{\text{ADJ}} = \frac{1}{\bar{y}}\left(\hat{\beta}_{0\text{MLE}} + \hat{\beta}_{1\text{MLE}}\bar{x} - \bar{y} - \hat{\beta}_{1\text{MLE}}\bar{x}\frac{(\hat{\alpha}_{\text{ADJ}}-1)}{2}\left(\frac{\bar{X}}{\bar{x}} - 1\right)^2\right) \tag{3.4}$$

또한 최종적으로 $\hat{\alpha}_{\text{ADJ}}$는 다음의 식으로 얻어진다.

$$\hat{\alpha}_{\text{ADJ}} = \hat{\beta}_{1\text{MLE}}\frac{\bar{x}}{\bar{y}}/\left(1+\hat{c}_{\text{ADJ}}\right) = \frac{\hat{\alpha}_{\text{MLE}}}{1+\hat{c}_{\text{ADJ}}}. \tag{3.5}$$

결론적으로 식 (3.4)와 (3.5)를 이용하여 얻어진 제안된 일반화 비형태추정량은 다음과 같다.

$$\hat{\bar{Y}}_{\text{GRT}} = \bar{y}\left(1+\hat{c}_{\text{ADJ}}\right)\left(\frac{\bar{X}}{\bar{x}}\right)^{\hat{\alpha}_{\text{ADJ}}} = \bar{y}\left(1+\hat{c}_{\text{ADJ}}\right)\left(\frac{\bar{X}}{\bar{x}}\right)^{\frac{\hat{\alpha}_{\text{MLE}}}{1+\hat{c}_{\text{ADJ}}}}. \tag{3.6}$$

4. 모의실험

4.1. 자료 생성 과정

이 절에서는 회귀추정량, 비형태추정량 그리고 본 논문에서 제안한 일반화 비형태추정량의 추정의 효율을 비교하기 위한 모의실험을 실시하였다. 기본적으로 자료 생성 과정은 Oh와 Shin (2017)의 방법을 사용하였으며 모집단 자료는 다음 식 (2.4)의 모형에서 생성되었다.

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \overset{\text{iid}}{\sim} N\left(0, x_i^{\nu}\sigma^2\right), \ 0 \leq \nu, \tag{4.1}$$

여기서 보조자료 x_i는 로그-정규분포 LN$(0,4)$에서 추출하였다. 이는 흔히 표본조사에서 사용되는 보조변수 x가 오른쪽으로 꼬리가 긴 분포를 따르기 때문이다. 다음으로 $\sigma = 5$를 사용하였고 오차의 분포는 정규분포를 사용하였다. 이는 2절에서 설명하였듯이 MLE를 이용하여 $\hat{\beta}$을 추정할 때 오차를 정규분포로 가정했기 때문이다. 다음으로 모수 $\beta_0 = 0, 7$ 그리고 $\beta_1 = 5$를 사용하여 절편이 있는 경우와 없는 경우를 비교하였으며 오차에 모함된 모수 ν가 증가함에 따른 결과의 변화를 확인하기 위해 $\nu = 0, 0.5, 1, 1.3$을 사용하였다. 최종적으로 식 (4.1)을 이용하여 50,000개의 모집단 자료가 생성되었으며 이 자료에서 단순 랜덤추출로 표본 수 $n = 50, 100, 200, 500$개를 추출하였다. 추출된 표본 자료를 이용하여 3절에서 설명한 다음의 5개 추정량을 이용하여 모평균을 추정하였다. 사용된 추정량의 정의는 다음과 같다.

- M1. 회귀추정량

  주어진 보조변수의 모평균 $\bar{X}$와 절편이 있는 회귀모형을 이용하여 절편과 기울기 β_0, β_1을 최소제곱 추정법으로 추정하여 $\hat{\beta}_{0\mathrm{OLS}}$, $\hat{\beta}_{1\mathrm{OLS}}$를 얻으며 이를 이용하여 관심변수 y의 모평균을 추정한다. 사용된 회귀추정량 공식은 식 (2.1)의 $\hat{\bar{Y}}_{\mathrm{LR}} = \hat{\beta}_{0\mathrm{OLS}} + \hat{\beta}_{1\mathrm{OLS}}\bar{X}$이다.

- M2. 비형태추정량

  주어진 보조변수의 모평균 $\bar{X}$와 표본에서 얻어진 표본평균 $\bar{x}, \bar{y}$ 그리고 M1에서 얻어진 $\hat{\beta}_{1\mathrm{OLS}}$를 이용하여 관심변수 y의 모평균을 추정한다. 즉 비형태추정량은 식 (2.2)의 $\hat{\bar{Y}}_{\mathrm{RT}} = \bar{y}(\bar{X}/\bar{x})^{\hat{\alpha}_{\mathrm{OLS}}}$이다. 여기서 $\hat{\alpha}_{\mathrm{OLS}} = \hat{\beta}_{1\mathrm{OLS}}/\hat{R}_1$이고, $\hat{R}_1 = \bar{y}/\bar{x}$이다.

- M3. MLE를 이용한 비형태추정량

  Oh와 Shin (2017)에서 사용한 방법으로 모수 ν와 β_0, β_1을 MLE를 이용하여 동시에 추정하며 얻어진 추정량을 비형태추정량 수식에 적용한 $\hat{\bar{Y}}_{\mathrm{MRT}} = \bar{y}(\bar{X}/\bar{x})^{\hat{\alpha}_{\mathrm{MLE}}}$이다. 여기서 $\hat{\alpha}_{\mathrm{MLE}} = \hat{\beta}_{1\mathrm{MLE}}/\hat{R}_1$, $\hat{R}_1 = \bar{y}/\bar{x}$이고 $\hat{\beta}_{1\mathrm{MLE}}$는 식 (2.6)에서 추정된다.

- M4. 제안된 일반화 비형태추정량

  본 논문에서 제안한 방법으로 식 (2.6)에서 추정된 $\hat{\beta}_{0\mathrm{MLE}}, \hat{\beta}_{1\mathrm{MLE}}$를 사용하며 편향을 제거하기 위해 모수 c를 추가한 일반화 비형태추정량으로 식 (3.6)인 $\hat{\bar{Y}}_{\mathrm{GRT}} = \bar{y}(1+\hat{c}_{\mathrm{ADJ}})(\bar{X}/\bar{x})^{\hat{\alpha}_{\mathrm{ADJ}}}$이다. 여기서 $\hat{c}_{\mathrm{ADJ}}$는 식 (3.4)에서 추정되며, $\hat{\alpha}_{\mathrm{ADJ}}$는 식 (3.5)에서 추정된다.

- M5. MLE를 이용한 회귀추정량

  Oh와 Shin (2017)에서 사용한 방법으로 식 (2.7)인 $\hat{\bar{Y}}_{\mathrm{MLR}} = \hat{\beta}_{0\mathrm{MLE}} + \hat{\beta}_{1\mathrm{MLE}}\bar{X}$이다. 여기서 $\hat{\beta}_{0\mathrm{MLE}}$, $\hat{\beta}_{1\mathrm{MLE}}$은 식 (2.6)에서 추정된다.

본 연구에서 사용한 비교 통계량은 편향(bias), 절대편향(absolute bias), 제곱근 MSE(root mean squared error; RMSE)이며 각 비교 통계량의 정의는 다음과 같다.

$$\mathrm{Bias} = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\bar{Y}}^{(r)} - \bar{Y}^{(r)}\right),$$

$$\text{Absolute bias} = \frac{1}{R}\sum_{r=1}^{R}\left|\hat{\bar{Y}}^{(r)} - \bar{Y}^{(r)}\right|,$$

$$\mathrm{RMSE} = \left\{\frac{1}{R}\sum_{r=1}^{R}\left(\hat{\bar{Y}}^{(r)} - \bar{Y}^{(r)}\right)^2\right\}^{\frac{1}{2}},$$

여기서 반복수 R = 5,000을 사용하였으며 각 반복마다 생성된 모집단은 같은 분포를 따르지만 다른 seed를 사용했기 때문에 각각의 모집단 자료는 달라진다. 이에 r번째 모집단의 모평균을 $\bar{Y}^{(r)}$이라 표시하였다.

4.2. 모의실험 결과

모의실험 결과에는 표본수가 증가함에 따라 오차 분산의 모수 ν와 비형태추정량의 모수인 α의 추정량의 특징을 살펴보기 위해 $\hat{\nu}, \hat{\alpha}_{\mathrm{OLS}}, \hat{\alpha}_{\mathrm{MLE}}, \hat{\alpha}_{\mathrm{ADJ}}$ 결과를 정리하였으며, 편향을 줄이기 위해 추가한 모수 c의 추정량의 특징을 확인하기 위해 $\hat{c}_{\mathrm{ADJ}}$ 결과도 정리하였다. 또한 비교통계량을 기준으로 각 방법에서 얻어진 추정량의 효율성이 비교되었다.

Table 4.1. Parameter estimation

Sample size	β_0	ν	Parameter				
			$\hat{\nu}$	$\hat{\alpha}_{\text{OLS}}$	$\hat{\alpha}_{\text{MLE}}$	$\hat{\alpha}_{\text{ADJ}}$	$\hat{c}_{\text{ADJ}}$
50	0	0.0	0.0350	1.0004	1.0004	1.0017	−0.0003
		0.5	0.4412	1.0007	1.0007	1.0016	−0.0001
		1.0	0.8945	1.0005	1.0001	1.0004	0.0002
		1.3	1.1692	0.9973	0.9998	0.9996	0.0008
	7	0.0	0.0335	0.8007	0.8007	0.7876	0.0183
		0.5	0.4409	0.8011	0.8009	0.7878	0.0183
		1.0	0.8946	0.8005	0.8002	0.7868	0.0188
		1.3	1.1647	0.7989	0.7989	0.7856	0.0186
100	0	0.0	0.0243	1.0005	1.0005	1.0008	0.0001
		0.5	0.4550	1.0008	1.0007	1.0012	−0.0003
		1.0	0.9198	1.0003	1.0000	0.9999	0.0002
		1.3	1.1940	0.9992	1.0000	1.0001	0.0001
	7	0.0	0.0249	0.8158	0.8158	0.8041	0.0160
		0.5	0.4571	0.8157	0.8155	0.8029	0.0174
		1.0	0.9197	0.8153	0.8154	0.8030	0.0172
		1.3	1.1981	0.8153	0.8149	0.8024	0.0173
200	0	0.0	0.0173	1.0001	1.0001	1.0002	−0.0001
		0.5	0.4671	0.9998	0.9998	1.0001	−0.0002
		1.0	0.9338	0.9999	1.0001	1.0002	−0.0002
		1.3	1.2104	0.9993	1.0001	1.0001	0.0001
	7	0.0	0.0174	0.8249	0.8249	0.8161	0.0118
		0.5	0.4662	0.8266	0.8266	0.8181	0.0114
		1.0	0.9329	0.8259	0.8253	0.8165	0.0118
		1.3	1.2124	0.8229	0.8257	0.8167	0.0120
500	0	0.0	0.0117	1.0001	1.0001	1.0002	0.0001
		0.5	0.4708	0.9999	1.0000	1.0000	−0.0001
		1.0	0.9417	1.0003	1.0001	1.0001	0.0000
		1.3	1.2219	0.9996	1.0001	1.0000	0.0001
	7	0.0	0.0116	0.8330	0.8330	0.8292	0.0048
		0.5	0.4720	0.8331	0.8331	0.8293	0.0048
		1.0	0.9417	0.8329	0.8331	0.8294	0.0046
		1.3	1.2210	0.8336	0.8334	0.8297	0.0047

OLS = ordinary least squares; MLE = maximum likelihood estimator; ADJ = adjusted.

우선 모수 ν, α의 추정 결과인 Table 4.1를 살펴보면 기본적으로 $0 \leq \nu$를 가정하고 있으므로 $\nu = 0$인 경우에는 과대추정의 결과를 얻었으며, 이 경우를 제외하고는 전체적으로 과소추정의 결과를 얻었다. 또한 α 추정 결과를 살펴보면 ν의 크기와는 상관없이 $\hat{\alpha}_{\text{OLS}}$와 $\hat{\alpha}_{\text{MLE}}$는 절편이 '0'인 경우 1에 가깝게 추정되었으며, 절편이 '7'인 경우에는 표본이 커지면서 0.8에서 0.83으로 증가하며 1 보다 작은 값으로 추정되었다. 또한 $\hat{\alpha}_{\text{ADJ}}$는 절편이 '0'인 경우 $\hat{\alpha}_{\text{OLS}}$, $\hat{\alpha}_{\text{MLE}}$와 유사하게 1에 가깝게 추정된 반면 절편이 '7'인 경우에는 $\hat{\alpha}_{\text{OLS}}$와 $\hat{\alpha}_{\text{MLE}}$보다 작은 수로 추정되었다. $\hat{c}_{\text{ADJ}}$은 절편이 '0'인 경우에는 0에 가깝게 추정되었으며, 절편이 '7'인 경우에는 0보다 크게 추정되었다.

Table 4.2에서 Table 4.4에는 비교통계량 결과를 수록하였으며 이 결과에서 M1과 M5는 회귀추정량

Table 4.2. Simulation results for bias

Sample size	β_0	ν	Bias				
			M1	M2	M3	M4	M5
50	0	0.0	0.0174	0.0286	0.0276	0.0438	0.0186
		0.5	0.0127	0.0310	0.0327	0.0402	0.0123
		1.0	−0.0433	−0.0137	−0.0492	−0.0402	−0.0538
		1.3	0.0476	0.1134	0.0561	0.0546	0.0419
	7	0.0	−0.0038	−1.0887	−1.0899	−0.5522	−0.0066
		0.5	0.0129	−1.0698	−1.0720	−0.5373	−0.0002
		1.0	0.0247	−1.0549	−1.0764	−0.5277	0.0043
		1.3	0.1287	−0.9555	−1.0184	−0.4750	0.1035
100	0	0.0	−0.0073	−0.0063	−0.0056	−0.0009	−0.0068
		0.5	−0.0176	−0.0122	−0.0138	−0.0116	−0.0159
		1.0	−0.0042	0.0047	−0.0023	0.0028	0.0013
		1.3	0.0454	0.0754	0.0304	0.0359	0.0326
	7	0.0	−0.0004	−0.6200	−0.6202	−0.1375	0.0018
		0.5	−0.0148	−0.6409	−0.6462	−0.1334	−0.0206
		1.0	−0.0025	−0.6171	−0.6162	−0.1081	0.0026
		1.3	0.0669	−0.5435	−0.5822	−0.0752	0.0377
200	0	0.0	−0.0026	−0.0020	−0.0017	−0.0024	−0.0026
		0.5	−0.0008	−0.0002	−0.0010	−0.0058	−0.0059
		1.0	−0.0177	−0.0151	−0.0201	−0.0241	−0.0244
		1.3	−0.0043	0.0013	−0.0125	−0.0161	−0.0163
	7	0.0	0.0026	−0.3586	−0.3586	0.0100	0.0028
		0.5	−0.0112	−0.3722	−0.3720	−0.0079	−0.0092
		1.0	−0.0051	−0.3585	−0.3611	0.0102	−0.0030
		1.3	0.0016	−0.3504	−0.3487	0.0324	0.0257
500	0	0.0	−0.0012	−0.0012	−0.0013	−0.0020	−0.0020
		0.5	0.0042	0.0042	0.0045	0.0018	0.0017
		1.0	−0.0064	−0.0056	−0.0078	−0.0082	−0.0083
		1.3	−0.0120	−0.0079	0.0052	0.0057	0.0056
	7	0.0	−0.0011	−0.1693	−0.1694	0.0059	−0.0002
		0.5	−0.0029	−0.1722	−0.1722	0.0028	−0.0021
		1.0	−0.0135	−0.1821	−0.1818	−0.0118	−0.0164
		1.3	−0.0018	−0.1668	−0.1703	0.0025	−0.0020

결과이다. 따라서 자료 생성 과정에서 회귀모형이 사용되었고 각 모형에 맞는 추정량을 사용하기 때문에 M1과 M5는 M2–M4인 비형태추정량에 비해 전체적으로 우수한 결과를 준다. 즉 $\nu = 0$인 경우는 M1이 우수하고 $\nu > 0$인 경우는 M5가 우수하다. 또한 절편이 없고 $\nu = 1$인 경우가 비모형이므로 M2 또는 M3가 우수한 결과를 주는 경우도 있다. 이는 본 연구에서 제안한 추정량의 경우 $\nu \neq 0, 1$이고 절편이 있는 경우에 최적화되도록 만들어진 추정량이기 때문이다. 따라서 향후 비교에서는 비형태추정량, 특히 M3와 M4의 비교에 초점을 맞추었다.

먼저 비교통계량 중 Bias 결과인 Table 4.2에서 절편이 ‘0’인 경우를 살펴보면, 표본이 가장 작은 50에서 M1과 M5는 큰 차이를 보이지 않으며 전체적으로 우수한 결과를 준다.

비형태추정량인 M2-M4를 비교하면 M2는 $\nu = 1$일 때 가장 좋은 결과를 보이고 있으나 $\nu = 1.3$인 경우 매우 큰 편향을 보이고 있다. 본 논문에서 제안한 일반화 비형태추정량인 M4와 Oh와 Shin (2017)에서

Table 4.3. Simulation results for absolute bias

Sample size	β_0	ν	Absolute bias				
			M1	M2	M3	M4	M5
50	0	0.0	0.6106	0.5981	0.6089	0.6642	0.6297
		0.5	0.9250	0.9357	0.9195	0.8999	0.8824
		1.0	1.9035	1.9900	1.8004	1.8061	1.7974
		1.3	3.1915	3.3752	2.9014	2.8410	2.8301
	7	0.0	0.6047	1.3151	1.3190	1.1397	0.6292
		0.5	0.9110	1.4616	1.4677	1.2862	0.8812
		1.0	1.9364	2.2656	2.1756	2.0993	1.8455
		1.3	3.2683	3.4964	3.1923	3.0833	2.8712
100	0	0.0	0.4032	0.3744	0.3774	0.4223	0.4199
		0.5	0.5936	0.5862	0.5821	0.5549	0.5543
		1.0	1.2506	1.2726	1.1865	1.2108	1.2111
		1.3	2.2500	2.3340	2.0760	1.9975	1.9952
	7	0.0	0.3979	0.7914	0.7915	0.6070	0.4160
		0.5	0.5920	0.9111	0.9178	0.7137	0.5589
		1.0	1.2453	1.4625	1.4038	1.3097	1.1978
		1.3	2.2527	2.4032	2.2202	2.1109	2.0195
200	0	0.0	0.2858	0.2713	0.2714	0.2986	0.2984
		0.5	0.3926	0.3842	0.3857	0.3620	0.3628
		1.0	0.8383	0.8454	0.7992	0.8148	0.8155
		1.3	1.5608	1.5913	1.4749	1.4088	1.4086
	7	0.0	0.2901	0.4954	0.4954	0.3554	0.3038
		0.5	0.3884	0.5767	0.5780	0.4010	0.3590
		1.0	0.8451	0.9541	0.9273	0.8577	0.8333
		1.3	1.5576	1.6191	1.5125	1.4103	1.3943
500	0	0.0	0.1767	0.1717	0.1718	0.1850	0.1850
		0.5	0.2357	0.2319	0.2327	0.2110	0.2111
		1.0	0.5121	0.5131	0.4984	0.5170	0.5171
		1.3	0.9597	0.9680	0.9244	0.8818	0.8817
	7	0.0	0.1777	0.2677	0.2677	0.1945	0.1860
		0.5	0.2385	0.3150	0.3154	0.2264	0.2176
		1.0	0.5096	0.5549	0.5403	0.5134	0.5095
		1.3	0.9728	1.0032	0.9590	0.8961	0.8933

제안한 비형태추정량인 M3를 비교해보면 $\nu < 1$일 때는 M3가 M4보다 우수하고 $\nu \geq 1$일 때는 M4가 M3보다 우수한 결과를 보여주고 있다. 그러나 절편이 '0'인 경우의 편향은 크지 않기 때문에 추정에는 큰 영향을 주지 않는다. 반면 절편이 '7'일 때에는 ν의 변화와는 관계없이 M4의 결과가 M2, M3의 결과보다 확실히 우수한 것을 확인할 수 있으며 따라서 본 연구에서 제안한 일반화 비형태추정량이 편향을 크게 줄인다고 판단된다.

다음으로 표본 개수가 100 이상인 경우의 편향 결과에서 절편이 '0'일 경우의 결과를 보면 M4가 M2, M3 보다 전체적으로 좋거나 유사하다. 따라서 2차 근사식이 만족되면 매우 우수한 결과를 얻을 수 있게 되며 M5 결과와 유사하게 된다. 또한 절편이 '7'일 경우 M4와 M2, M3를 비교했을 때 M4가 매우 우수하며 결과의 차이가 명확하다.

절대 편향 결과인 Table 4.3에서는 M1과 M5가 우수한 결과를 준다. 즉 절편이 '0'이고 $\nu = 0$인 경우

Table 4.4. Simulation results for RMSE

Sample size	β_0	ν	RMSE				
			M1	M2	M3	M4	M5
50	0	0.0	0.7749	0.8287	0.8531	0.9094	0.7997
		0.5	1.2014	1.2774	1.2356	1.2232	1.1660
		1.0	2.4869	2.7163	2.3468	2.3513	2.3322
		1.3	4.1464	4.4987	3.7106	3.6306	3.6078
	7	0.0	0.7753	1.9828	1.9952	1.8672	0.8059
		0.5	1.1854	2.1805	2.1878	2.0505	1.1630
		1.0	2.5491	3.0879	2.9390	2.8702	2.3994
		1.3	4.2512	4.5975	4.1007	4.0074	3.6757
100	0	0.0	0.5082	0.4753	0.4803	0.5344	0.5290
		0.5	0.7549	0.7525	0.7436	0.7174	0.7136
		1.0	1.6097	1.6615	1.5344	1.5627	1.5630
		1.3	2.8850	3.0574	2.6164	2.5219	2.5184
	7	0.0	0.5001	1.1492	1.1514	0.9687	0.5239
		0.5	0.7564	1.2828	1.2962	1.0816	0.7236
		1.0	1.5978	1.9291	1.8339	1.7266	1.5334
		1.3	2.8900	3.1008	2.8116	2.6868	2.5512
200	0	0.0	0.3594	0.3422	0.3423	0.3760	0.3755
		0.5	0.4970	0.4875	0.4899	0.4641	0.4646
		1.0	1.0679	1.0868	1.0197	1.0416	1.0425
		1.3	1.9907	2.0448	1.8560	1.7771	1.7770
	7	0.0	0.3633	0.6989	0.6992	0.5053	0.3804
		0.5	0.4876	0.8005	0.8032	0.5649	0.4552
		1.0	1.0775	1.2454	1.2000	1.0964	1.0586
		1.3	1.9919	2.0915	1.9127	1.7824	1.7553
500	0	0.0	0.2215	0.2154	0.2155	0.2318	0.2319
		0.5	0.2956	0.2909	0.2923	0.2660	0.2660
		1.0	0.6473	0.6500	0.6301	0.6537	0.6539
		1.3	1.2158	1.2319	1.162	1.1096	1.1096
	7	0.0	0.2229	0.3755	0.3753	0.2484	0.2342
		0.5	0.2995	0.4339	0.4347	0.2960	0.2755
		1.0	0.6458	0.7141	0.6937	0.6503	0.6454
		1.3	1.2280	1.2699	1.2087	1.1206	1.1163

RMSE = root mean squared error.

는 M1이 그리고 절편이 '7'이고 $\nu > 0$인 경우는 M5 결과가 우수하다. 이제 비형태추정량인 M2, M3, M4 결과를 살펴보면 절편이 '0'이고 $\nu = 0, 1$인 경우를 제외하면 모든 경우에서 M4가 우수한 결과를 준다. 특히 절편이 '7'인 경우에는 매우 우수한 결과를 준다. 마지막으로 RMSE 결과인 Table 4.4를 보면 Table 4.3 결과에서 얻어진 우수성 결과와 매우 유사하다.

따라서 회귀모형을 만족하거나 비모형을 만족할 경우에는 이 모형에 최적화되어 있는 추정량을 사용하면 된다. 또한 흔히 ν가 알려져 있지 않기 때문에 ν를 추정한 후 회귀모형을 사용할 경우에는 M5를 사용하면 된다. 그러나 비형태추정량을 사용할 경우에는 특히 절편이 '0'이 아닌 경우에는 M4를 사용함으로써 매우 우수한 결과를 얻을 수 있을 것이다. 물론 절편이 '0'이고 $\nu = 0$ 또는 $\nu = 1$로 알려진 경우에는 M3를 사용할 수 있다.

5. 결론

비형태추정량은 편향이 존재하며 특히 절편이 있는 경우의 소표본에서는 추정의 정확성에 큰 영향을 주고 있다. 이를 해결하기 위해 본 논문에서는 Oh와 Shin (2017)에서 연구한 비형태추정량에 새로운 모수를 추가하고 Taylor 2차 근사식을 이용한 일반화 비형태추정량을 제안하였다. 모의실험 결과를 보면 $\nu = 0$일 경우에는 회귀추정량인 M1이 가장 우수한 결과를 보였으며 이는 자료가 등분산성을 만족하기 때문으로 당연한 결과라 판단된다. 또한 $\nu = 1$이고 절편이 없을 때는 자료가 비모형을 만족하기 때문에 M3가 다른 비형태추정량에 비해 전체적으로 우수한 결과를 준다. $0 < \nu < 1$ 그리고 $1 < \nu$인 경우에는 M5가 전체적으로 우수한 결과를 주며 이 또한 당연한 결과이다.그러나 비형태추정량 비교 결과를 살펴보면 전체적으로 본 연구에서 제안한 일반화 비형태추정량인 M4가 기존의 비형태 추정량인 M2, M3보다 우수하고 특히 절편이 있는 경우에는 편향이 크게 줄어드는 것을 확인하였으며 절대편향 및 RMSE 기준에서도 매우 우수한 결과를 주는 것을 확인하였다. 또한 자료의 수가 증가 하면서 M4는 M5와 유사한 값을 갖는다. 따라서 절편이 있는 모형에서 비형태추정량을 사용할 경우에 본 연구에서 제안한 일반화 비형태추정량을 사용한다면 추정의 정확성과 정밀성이 크게 향상될 것으로 판단된다.

References

Clement, E. P. and Enang, E. I. (2017). On the efficiency of ratio estimator over the regression estimator, *Communications in Statistics-Theory and Methods*, **46**, 5357–5367.

Etuk, S. I., Enang, E. I., and Ekpenyong, E. J. (2016). A modified class of ratio estimators for population mean, *Communications in Statistics-Theory and Methods*, **45**, 837–849.

Haq, A., Khan, M., and Hussain, Z. (2017). A new estimator of finite population mean based on the dual use of the auxiliary information, *Communications in Statistics-Theory and Methods*, **46**, 4425–4436

Khan, M. and Singh, R. (2015). Estimation of population mean in chain ratio-type estimator under systematic sampling, *Journal of Probability and Statistics*, **2015**, 248374–248378

Koyuncu, N. and Kadilar, C. (2009). Family of estimators of population mean using two auxiliary variable in stratified random sampling, *Communications in Statistics-Theory and Methods*, **38**, 2938–2417.

Lone, H. A., Tailor, R., and Singh, H. P., (2015), Generalized ratio-cum-product type exponential estimator in stratified random sampling, *Communications in Statistics-Theory and Methods*, **45**, 3302–3309.

Oh, J. T. and Shin, K. I. (2017). A study on the ratio type estimator in survey sampling, *The Korean Association for Survey Research*, **18**, 105–125

Sharma, P. and Singh, R. (2014). Improved ratio type estimator using two auxiliary variables under second order approximation, mathematical *Journal of Interdisciplinary Sciences*, **2**, 179–190

Singh, H. P., Kumar, S., and Kozak, M. (2010). Improved estimation of finite-population mean using sub-sampling to deal with non response in two-phase sampling scheme, *Communications in Statistics-Theory and Methods*, **39**, 791–802

Singh, H. P. and Vishwakarma, G. K. (2007). Modified exponential ratio and product estimators for finite population mean in double sampling, *Austrian Journal of Statistics*, **36**, 217–225

Singh, R. and Singh, H. P. (1998). Almost unbiased ratio and product-type estimators in systematic sampling, *Qüestiió*, **22**, 403–416

Swain, A. (2014). On an improved ratio type estimator of finite population mean in sample surveys, *Investigacion Operacional*, **35**, 49–57

Tailor, R., Chouhan, S., and Kim, J.-M. (2014). Ratio and product type exponential estimators of population mean in double sampling for stratification, *Communications for Statistical Applications and Methods*, **21**, 1–9.

Tailor, R., Lone, H. A., and Pandey, R. (2015). Generalized ratio-cum-product type estimator of finite population mean in double sampling for stratification, *Communications for Statistical Applications and Methods*, **22**, 255–264.

편향 보정 비형태추정량에 관한 연구

오정택[a] · 신기일[a,1]

[a]한국외국어대학교 통계학과

(2018년 5월 28일 접수, 2018년 6월 7일 수정, 2018년 6월 8일 채택)

요 약

표본조사에서는 정확한 모수 추정을 위한 다양한 방법이 개발되었으며 이 중에서 보조정보를 이용한 비추정량 또는 회귀추정량이 흔히 사용된다. 최근 많은 연구가 진행되고 있는 비형태추정량(ratio type estimator)은 비추정량의 단점을 보완하여 추정의 정확성을 향상시키는 것으로 알려져 있다. 그러나 비형태추정량은 편향이 있는 것으로 알려져 있어 이를 해결하기 위한 연구가 활발히 진행되고 있다. 이에 본 연구에서는 편향을 제거하기 위해 비형태추정량에 새로운 모수를 추가한 일반화 비형태추정량(generalized ratio-type estimator)을 제안하였다. 또한 사업체조사와 같이 등분산성을 만족하지 않는 자료에서 추정의 정확성 향상을 위해 모형의 오차에 포함된 분산 모수를 추정하고 제안된 추정량을 적용하는 방법을 제안하였다. 또한 모의실험을 통해 일반화 비형태추정량은 기존의 비추정량에 비해 매우 우수한 결과를 주는 것을 확인하였다.

주요용어: 편향, 불편추정량, 이분산성, 최대가능도추정량, Taylor 근사

이 연구는 2018년 한국외국어대학교 교내연구비 지원을 받아 수행되었음.

[1]교신저자: (17035) 경기도 용인시 처인구 모현읍 외대로 81, 한국외국어대학교 통계학과.
E-mail: keyshin@hufs.ac.kr