UIT Report Vol. 13, No. 2, pp. 513~517, 1982　　연구논문집 제13권 제 2 호 pp. 513~517, 1982

# 層化集落抽出에서의 不偏比推定量에 關한 硏究

金 元 卿
經 濟 學 科
(1982. 6. 30 접수)

〈要　約〉

母數의 推定量으로서 標本資料와 補助資料를 함께 利用하는 比推定量은 標本統計量보다 推定量의 精度를 높일 수 있으나 一般的으로 不偏推定量이 되지 못한다.

本 硏究는 層化集落抽出인 경우의 不偏比推定量을 얻기 위해서 單純任意抽出에서의 「Hartely—Ross」의 不偏比推定量을 擴張하여 不偏分離比推定量과 不偏結合比推定量을 導出하였고 또한 各各의 分散을 구하여 有效性을 比較 分析하였다.

# A Study on Ratio Estimators in Stratified Cluster Sampling

Kim Won Kyung
Dept. of Economics
(Received June 30, 1982)

〈**Abstract**〉

Unbiased ratio estimators in stratified clusler sampling are presented. The two unbiased ratio estimators, separate unbiased ratio estimator and combined unbiased ratio estimator, are obtained by extension of the Hartely-Ross unbiased ratio estimator and their variances are developed by using the moments of bivariate sample cumulants. The relative efficiencies of these estimators with respect to the biased ratio estimators are investigated.

## I. 序　論

理論統計의 가장 重要한 目的의 하나는 母集團으로부터 얻은 標本을 통해서 母集團 特性値인 母數를 보다 正確하게 推定하는 것이다. 標本으로 母數를 推定하는 경우에는 必然的으로 誤差가 發生하게 되는데 이러한 誤差는 標本抽出方法을 달리함으로써 줄일 수도 있지만 推定方法을 달리함으로서도 줄일 수 있다. 母數의 推定方法으로는 標本으로만 얻어지는 標本統計量을 母數의 推定量으로 삼는 경우와 補助資料를 함께 使用하여 얻어지는 比推定量을 母數의 推定量으로 삼는 경우가 있는데 後者의 推定量은 追加的으로 주어지는 補助資料가 標本資料와 正相關이 높은 경우에 前者에 비해서 推定量의 精度가 높아진다. 補助資料로는 흔히 前期에 調査된 資料를 使用하므로 대체로 標本資料와는 正相關이 높아 比推定量은 有效推定量이라 할 수 있다. 그러나 比推定量은 一般的으로 不偏推定量이 되지 못하는 결점을 가지고 있다. 물론 그 偏倚는 대부분의 경우 그러 크지않아 實際標本設計에서는 별로 큰 問題가 없으나 層化된 母集團에서 層數가 크고 各 層으로부터 標本을 적게 抽出하는 경우 그 偏倚는 상당히 커져 이러한 標本設計에서는 不偏比推定量이 매우 바람직스럽다.

不偏比推定量은 처음으로 「Hartely-Ross」에 의해서 提示되었다. 單純任意抽出인 경우에 母平均에 대한 「Hartely-Ross」의 不偏比推定量 $\bar{y}'$은 다음과

같다.(1)

(1) $\bar{y}'=\bar{r}\bar{X}+\dfrac{N-1}{N}\dfrac{1}{n-1}(\bar{y}-\bar{r}\bar{x})$

여기서 $\bar{r}=\dfrac{1}{n}\sum_{i=1}^{n}r_i=\dfrac{1}{n}\sum_{i=1}^{n}\dfrac{y_i}{x_i}$

$\bar{x}, \bar{y}$은 $x_i, y_i$의 標本平均

$\bar{X}$는 $x$變量의 주어진 母平均

$N, n$은 母集團 및 標本크기

이 推定量의 分散은 「Goodman—Hartely」가 二變量累積積率(Moment of bivariate sample cumulants)을 이용하여 구하였다.(2)

層化抽出인 경우에 不偏比推定量은 「Pascual」이 「Hartely-Ross」의 不偏比推定量을 擴張해서 分離比推定量과 結合比推定量을 提示하였고 각각의 效率性을 比較分析하였다.(3)

本 硏究에서는 「Hartely-Ross」의 不偏比推定量을 層化集落抽出(Stratified Cluster Sampling)인 경우로 擴張하여 不偏分離比推定量과 不偏結合比推定量을 導出하고 각각의 分散을 구하여 그 效率性을 比較分析하고자 한다.

層化集落抽出方法은 層化抽出과 集落抽出의 特性을 結合한 標本抽出方法으로서 이는 標本設計를 함에 있어 設計의 융통성을 가질 수 있을 뿐만 아니라 調査費用과 時間이 경감되는 등의 長點때문에 現實的으로도 흔히 使用되는 抽出方法이다. 이러한 實質的 適用性을 감안하면 層化集落抽出에서 不偏性, 有效性 등 바람직한 推定量의 性質을 갖춘 母數의 推定量을 구하다는 것은 매우 所望스럽다. 本 硏究는 이와같은 目的에서 層化集落抽出에서의 不偏比推定量을 導出, 分析하고자 한다.

## Ⅱ. 不偏比推定量

### 1. 不偏分離比推定量

分離比推定量은 層化된 母集團으로부터 各 層에서의 比推定量을 구하여 이를 層別로 合해서 얻은 推定量이다.

이제 $L$개로 層化된 크기 $N_h$의 母集團의 各 層이 다시 $M_h$개의 集落으로 이루어져 있고 이들 集落으로부터 $m_h$개의 集落을 標本으로 抽出한다고 하자. 또한 抽出된 各 集落으로부터 $x$, $y$ 각각의 變量에 대해 단 하나씩의 標本을 얻어 이들 比를 취하고 다시 위와같은 獨立的인 標本抽出을 $k$번 계속한다고 하자.

그러면 各 層에서의 母平均에 대한 比推定量 $\bar{y}_h$는 다음과 같다.

(2) $\bar{y}_h=\bar{r}_h\bar{X}_h$

여기서 $\bar{r}_h=\dfrac{1}{k}\sum_{i=1}^{k}r_{hi}=\dfrac{1}{k}\sum_{i=1}^{k}\dfrac{y_{hi}}{x_{hi}}$

그러나 이 推定量은 不偏推定量이 되지 못한다. 왜냐하면

(3) $\text{Cov}(r_{hi}, x_{hi})=E[r_{hi}-E(r_{hi})][x_{hi}-\bar{X}_h]$
$\qquad=\bar{Y}_h-\bar{X}_hE(r_{hi})$

$E(r_{hi})=E(\bar{r}_h)$이므로

(4) $E(\bar{y}_h)=E(\bar{r}_h\bar{X}_h)=\bar{Y}_h-\text{Cov}(r_{hi}, x_{hi})$

따라서 $\bar{y}_h$는 $-\text{Cov}(r_{hi}, x_{hi})$의 偏倚를 가진다. 그런데 $\text{Cov}(r_{hi}, x_{hi})$의 不偏推定量은

(5) $\dfrac{M_h-1}{M_h}\dfrac{1}{k-1}\sum_{i=1}^{k}(r_{hi}-\bar{r}_h)(x_{hi}-\bar{x}_h)$

$\qquad=\dfrac{M_h-1}{M_h}\dfrac{k}{k-1}(\bar{y}_h-\bar{r}_h\bar{x}_h)$

이므로 $h$層에서의 母平均에 대한 不偏比推定量 $\bar{y}_h'$는 위의 偏倚를 (2)式에 더해 줌으로서 얻을 수 있다.

(6) $\bar{y}_h'=\bar{r}_h\bar{X}_h+\dfrac{M_h-1}{M_h}\dfrac{k}{k-1}(\bar{y}_h-\bar{r}_h\bar{x}_h)$

결국 母平均에 대한 不偏分離比推定量 $\bar{y}_s'$는 다음과 같이 導出된다.

(7) $\bar{y}_s'=\sum_{h=1}^{L}\dfrac{N_h}{N}\bar{y}_h'$

여기서 $N$은 母集團크기$(N=\sum_{h=1}^{L}N_h)$

$N_h$은 $k$層의 母集團크기$(N_h=\sum_{i=1}^{M_h}N_{hi})$

不偏分離比推定量의 分散은 「Goodman-Hartely」가 單純任意抽出인 경우에 二變量累積積率을 利用하여 分散을 구한 方法을 適用하면 얻어질 수 있다.(4)

(1) Hartely, H.O. and Ross, A., "Unbiased Ratio Estimate," *Nature*, Vol. 174(1954), pp. 270—271.

(2) 分散의 誘導는 다음을 參照.
Goodman, L.A. and Hartely, H.O., "The Precision of Unbiused Ratio Type Estimators," *Journal of the American Statistical Association*, Vol. 53 (1958). pp. 491—508. Robson. D.S., "Application of Multivariate Polykays to the Theory of Unbiased Ratio Type Estimation," *Journal of the American Statistical Association*, Vol. 52(1957), pp. 511—522.

(3) Pascual, J.N., "Unbiased Ratio Estimators in Stratified Sampling," *Journal of the American Statistical Association*, Vol. 56(1961), pp. 70—87.

(4) 二變量累積積率에 관해서는 다음을 參照.
Kendall, M.G. and Staurt, A., The Advanced Theory of Statistics, Vol. 1, 3rd. ed., Harfner Pulshing Company, New York, 1969. pp. 235—236.

(8) $\mathrm{Var}(\bar{y}_s') = \sum_{h=1}^{L} \frac{N_h^2}{N^2} \mathrm{Var}(\bar{y}_h')$

(9) $\mathrm{Var}(\bar{y}_h') = \frac{1}{k} [\bar{X}_h^2 \mathrm{var}(\bar{r}_h) + 2\bar{X}_h \mathrm{Cov}(\bar{r}_h, c) + \mathrm{Var}(c)]$

여기서 $c = \frac{1}{k-1} \sum_{i=1}^{k} (r_{hi} - \bar{r}_h)(x_{hi} - \bar{x}_h)$

윗 式으로부터 $\mathrm{Cov}(\bar{r}_h, c)$과 $\mathrm{Var}(c)$를 구하여 代入하면 다음과 같다.

(10) $\mathrm{Var}(\bar{y}_s') \doteq \sum_{h=1}^{L} \frac{1}{N^2} \frac{1}{k} [\sigma^2_{y_{hi}} - \bar{R}_h^2 \sigma^2_{x_{hi}} - 2\bar{R}_h \mathrm{Cov}(x_{hi}, y_{hi}) + \frac{1}{k-1} [\sigma^2_{r_{hi}} \sigma^2_{x_{hi}} + \mathrm{Cov}(r_{hi}, x_{hi})]]$

$= \sum_{h=1}^{L} \frac{1}{N^2} \frac{1}{k} [\mathrm{Var}(y_{hi} - \bar{R}_h x_{hi})] + \frac{1}{k-1} [\sigma^2_{r_{hi}} \sigma^2_{x_{hi}} + \mathrm{cov}(r_{hi}, x_{hi})]$

여기서 $\bar{R}_h = E(r_{hi})$

$$\sigma^2_{y_{hi}} = \frac{1}{M_h - 1} \sum_{i=1}^{M_h} (y_{hi} - \bar{Y}_h)^2$$

$$\sigma^2_{x_{hi}} = \frac{1}{M_h - 1} \sum_{i=n}^{M_h} (x_{hi} - \bar{X}_h)^2$$

$$\sigma^2_{r_{hi}} = \frac{1}{M_h - 1} \sum_{i=1}^{M_h} (r_{hi} - \bar{R}_h)^2$$

## 2. 不偏結合比推定量

層化된 母集團으로부터 比推定量을 얻을 수 있는 또 하나의 方法은 結合比推定量으로서 이는 各 層을 하나로 結合해서 얻는 方法이다. 즉 層化集落抽出에서의 母平均에 대한 結合比推定量 $\bar{y}_c$는 다음과 같이 얻어진다.

(11) $\bar{y}_c = \frac{\bar{y}_{sc}}{\bar{x}_{sc}} \bar{X} = \bar{r}_{sc} \bar{X}$

여기서 $\bar{y}_{sc} = \frac{1}{N} \sum_{h=1}^{L} \frac{M_h}{m_h} \sum_{i=1}^{m_h} \frac{N_{hi}}{n_{hi}} \sum_{j=1}^{n_{hi}} y_{hij}$

$$\bar{x}_{sc} = \frac{1}{N} \sum_{h=1}^{L} \frac{M_h}{m_h} \sum_{i=1}^{m_h} \frac{N_{hi}}{n_{hi}} \sum_{j=1}^{n_{hi}} x_{hij}$$

이제 不偏結合比推定量을 얻기 위해 다음과 같은 標本抽出을 고려해 보자.

$L$개로 層化된 母集團의 각 層으로부터 任意로 단 하나의 集落을 抽出하고 $N_{hi}$ 크기의 각 集落으로부터는 $n_{hi}$개의 標本을 抽出하여 이들 標本으로 $\bar{y}_{sc}$, $\bar{x}_{sc}$의 값을 구한다. 抽出된 標本을 다시 되넣고 위와같은 抽出方法으로 또 $\bar{y}_{sc}$, $\bar{x}_{sc}$를 얻는다. 이렇게 獨立的인 標本抽出을 $k$번 계속하여 각각 $\bar{y}_{sc}$, $\bar{x}_{sc}$의 값을 얻는다고 하고 $i$번째 標本으로부터 얻은 $\bar{y}_{sc}$, $\bar{x}_{sc}$를 다음과 같이 표시하자.

$$(\bar{y}_{sc})_i = \bar{y}_i, \quad i = 1, 2, \cdots, k$$
$$(\bar{x}_{sc})_i = \bar{x}_i, \quad i = 1, 2, \cdots, k$$

그러면 결국

$$\bar{y}_{sc} = \frac{1}{k} \sum_{i=1}^{k} y_i$$

$$\bar{x}_{sc} = \frac{1}{k} \sum_{i=1}^{k} \bar{x}_i$$

가 될 것이다.

이들 標本값으로부터 結合比推定量 $\bar{r}_{sc}$는 다음과 같이 얻을 수 있다.

(12) $\bar{r}_{sc} = \frac{1}{k} \sum_{i=1}^{k} \frac{\bar{y}_i}{\bar{x}_i} = \frac{1}{k} \sum_{i=1}^{k} \bar{r}_i$

그러면

(13) $\mathrm{Cov}(\bar{r}_i, \bar{x}_i) = \bar{Y} - \bar{X} E(\bar{r}_i)$

$E(\bar{r}_i)$은 $E(\bar{r}_{sc})$과 같으므로

(14) $E(\bar{y}_c) = \bar{X} E(\bar{r}_{sc}) = \bar{Y} - \mathrm{Cov}(\bar{r}_i, \bar{x}_i)$

따라서 結合比推定量의 偏倚는 $-\mathrm{Cov}(\bar{r}_i, \bar{x}_i)$가 된다. 그런데 $\mathrm{Cov}(\bar{r}_i, \bar{x}_i)$의 不偏推定量은

$$\frac{N-1}{N} \frac{1}{k-1} \sum_{i=1}^{k} (\bar{r}_i - \bar{r}_{sc})(\bar{x}_i - \bar{x}_{sc})$$

(15) $= \frac{N-1}{N} \frac{k}{k-1} (\bar{y}_{sc} - \bar{x}_{sc} \bar{r}_{sc})$

이므로 母平均에 대한 結合比推定量 $\bar{y}_c'$를 다음과 같이 定義하면

(16) $\bar{y}_c' = \bar{r}_{sc} \bar{X} + \frac{N-1}{N} \frac{k}{k-1} (\bar{y}_{sc} - \bar{r}_{sc} \bar{x}_{sc})$

이는 분명히 不偏推定量이 된다.

不偏結合比推定量의 分散도 二變量累積積率에 의해서 얻어질 수 있는데 이를 구해보면 결국 不偏分離比推定量의 分散에서 $\bar{R}_h$ 대신에 $\bar{R}$를 代入한 것이다.

(17) $\mathrm{Var}(\bar{y}_c') = \frac{1}{k} \left\{ \sigma^2_{\bar{y}_i} + \bar{R}^2 \sigma^2_{\bar{x}_i} - 2\bar{R} \mathrm{Cov}(\bar{y}_i, \bar{x}_i) + \frac{1}{k-1} \sigma^2_{\bar{x}_i} \sigma^2_{\bar{r}_i} + \mathrm{Cov}(\bar{r}_i, \bar{x}_i) \right\}$

여기서 $\bar{R} = E(\bar{r}_i)$

$$\sigma^2_{\bar{y}_i} = \frac{1}{k-1} \sum_{i=1}^{k} (y_i - \bar{y}_{sc})^2$$

$$\sigma^2_{\bar{x}_i} = \frac{1}{k-1} \sum_{i=1}^{k} (\bar{x}_i - \bar{x}_{sc})^2$$

$$\sigma^2_{\bar{r}_i} = \frac{1}{k-1} \sum_{i=1}^{k} (\bar{r}_i - \bar{r}_{sc})^2$$

그런데

$\bar{y}_{sc}=\frac{1}{k}\sum_{i=1}^{k}\bar{y}_i$, $\mathrm{Var}(\bar{y}_{sc})=\frac{1}{k}\mathrm{Var}(\bar{y}_i)$ 이므로

(17)式은 다음과 같이 整理된다.

$$(18)\ \mathrm{Var}(\bar{y}_c')=\mathrm{Var}(\bar{y}_{sc}-\bar{R}\bar{x}_{sc})+\frac{1}{k(k-1)}(\sigma^2_{\bar{x}_i}\sigma^2_{\bar{y}_i}+\mathrm{cov}(\bar{y}_i,\bar{x}_i))$$

## Ⅲ. 有效性 比較

앞서 導出한 各 比推定量의 精度를 測定하기 위해 먼저 不偏結合比推定量의 分散과 偏倚結合比推定量의 그것을 比較하여 보자.

母平均에 대한 偏倚結合比推定量은 (11)式이고 이의 分散은 다음과 같다.[5]

$$(19)\ \mathrm{Var}(\bar{y}_c)=\frac{1}{N^2}\sum_{h=1}^{L}\frac{M_h^2}{m_h}\frac{M_h-m_h}{M_h}S_{1h}^2+\frac{1}{N^2}\sum_{h=1}^{L}\frac{M_h}{m_h}\sum_{i=1}^{M_h}\frac{N_{hi}^2}{n_{hi}}\frac{N_{hi}-n_h}{N_{hi}}S_{2hi}^2$$

여기서 $S_{1h}^2=S_{1hY}+Q^2S_{1hX}-2QS_{1hXY}$

$$S^2_{1hY}=\frac{1}{M_h-1}\sum_{i=1}^{M_h}(Y_{hi}-\bar{Y}_h)^2$$

$$S^2_{1hX}=\frac{1}{M_h-1}\sum_{i=1}^{M_h}(X_{hi}-\bar{X}_h)^2$$

$$S_{1hXY}=\frac{1}{M_h-1}\sum_{i=1}^{M_h}(X_{hi}-\bar{X}_h)(Y_{hi}-\bar{Y}_h)$$

$$Q=\frac{\bar{Y}}{\bar{X}}$$

$$S_{2hi}^2=S_{2hiY}+Q^2S_{2hiX}-2RS_{2hiXY}$$

$$S_{2hiY}=\frac{1}{N_{hi}-1}\sum_{j=1}^{N_{hi}}(Y_{hij}-Y_{hi})^2$$

$$S^2_{2hiX}=\frac{1}{N_{hi}-1}\sum_{j=1}^{N_{hi}}(X_{hij}-X_{hi})^2$$

$$S_{2hiXY}=\frac{1}{N_{hi}-1}\sum_{j=1}^{N_{hi}}(X_{hij}-X_{hi})(Y_{hij}-Y_{hi})$$

$$X_{hi}=\frac{1}{N_{hi}}X_{hi},\quad Y_{hi}=\frac{1}{N_{hi}}Y_{hi}$$

이 分散값은 $S_{1h}^2$, $S_{2hi}^2$의 값에 따라 決定되나 $m_h=\bar{m}$, $n_{hi}=\bar{n}$일 경우 그것은 주로 $S_{1h}^2$값에 의존하게 된다[6] 더우기 $M_h \gg m_h$이면 (19)式은 다음과 같이 되어 두번째 項이 무시된다.

$$(20)\ \mathrm{Var}(\bar{y}_c)\doteqdot\frac{1}{N^2}\sum_{h=1}^{L}\left[\left(\frac{M_h}{m_h}\right)^2\left(m_hS_{1h}^2+\frac{m_h}{M_h}\sum\frac{N_{hi}}{n_{hi}}\frac{N_{hi}-n_{hi}}{N_{hi}}\frac{S_{2hi}^2}{n_{hi}}\right)\right]$$

$$\doteqdot\frac{1}{N^2}\sum_{h=1}^{L}M_h^2\frac{S_{1h}^2}{m_h}$$

그런데 $\mathrm{Var}(\bar{y}_{sc})=\frac{1}{N^2}\sum_{h=1}^{L}\frac{M_h^2}{m_h}\mathrm{Var}(y_{hi})$이므로 위는 결국 다음과 같다.

$$(21)\quad \mathrm{Var}(\bar{y}_c)\doteqdot\mathrm{Var}(\bar{y}_{sc}-Q\bar{x}_{sc})$$

반면에 不偏結合比推定量의 分散은 (17)式이나 分散의 比較가 可能하도록 $k$값이 커서 (17)式의 두번째 項이 무시될 수 있다면 그 分散은

$$(22)\quad \mathrm{Var}(\bar{y}_c')\doteqdot\mathrm{Var}(\bar{y}_{sc}-\bar{R}\bar{x}_{sc})$$

가 된다.

이제 (21)式과 (22)式으로부터 分散을 比較하여 보자.

$$(23)\ \mathrm{Var}(\bar{y}_c)-\mathrm{Var}(\bar{y}_c')\doteqdot\mathrm{Var}(\bar{y}_{sc}-Q\bar{x}_{sc})-\mathrm{Var}(\bar{y}_{sc}-\bar{R}\bar{x}_{sc})\doteqdot(Q^2-\bar{R}^2)\mathrm{Var}(\bar{x}_{sc})-2\mathrm{Cov}(\bar{y}_{sc},\ \bar{x}_{sc})(Q-\bar{R})$$

$B$를 $\bar{y}_{sc}$와 $\bar{x}_{sc}$의 回歸係數, 즉

$$B=\frac{\mathrm{Cov}(\bar{y}_{sc},\bar{x}_{sc})}{\mathrm{Var}(\bar{x}_{sc})}$$

라 하면 위는 다음과 같다.

$$(24)\quad \mathrm{Var}(\bar{y}_c)-\mathrm{Var}(\bar{y}_c')=\mathrm{Var}(\bar{x}_{st})[(Q-B)^2-(\bar{R}-B)^2]$$

이 式으로부터 不偏比推定量의 精度가 測定될 수 있는데 결국 $\bar{R}$값이 $Q$값보다 $B$값에 더욱 近似하면 不偏結合比推定量 $\bar{y}_c'$가 偏倚結合比推定量 $\bar{y}_c'$보다 有效하고 그렇지 않으면 後者가 前者보다 分散이 작아지므로 偏倚를 고려치 않는다면 보다 有效하다고 말할 수 있다.[7] 그러나 推定量의 精度를 測定할 경우에는 偏倚가 고려되지 않을 수 없으므로 이 때에는 偏倚의 제곱과 分散의 合인 平均平方誤差(Mean Squared Error)로서 精度를 測定하는 것이 바람직하나 平均平方誤差를 測定手段으로 使用할 경우에는 一般的인 比較條件을 나타내기가 매우 어렵다.

둘째로 不偏分離比推定量과 偏倚分離比推定量 間

---

(5) 分散誘導는 다음을 參照.
Hansen, M.H., Hurwitz, W.N. and Madow, W.G., Sample Survey Method and Theory, Vol. Ⅱ, John Wiley and Sons, New York, 1953, pp. 177—181.

(6) 이는 層化集落抽出인 경우에 母平均의 分散公式에서 成立되므로 결국 比推定量의 分散公式에서도 成立된다.
Yamane, T., Elementary Sampling Theory, Prentice-Hall, Englewood Cliffs, N.J., 1967, p. 319 參照.

(7) 이 條件은 層化抽出에서 「Pascual」이 얻은 條件과 같으나 層化抽出과 層化集落抽出에서 $\bar{R}$, $Q$, $B$값이 서로 달리 定義되어 있다.

의 分散을 比較하여 보자. 이 分散의 比較는 앞의 方法과 똑같은 方法으로 行해 질 수 있다.

分離比推定量의 分散의 近似값은

$$(25)\quad \mathrm{Var}(\bar{y}_s) \doteq \sum_{h=1}^{L} \frac{N_h^2}{N^2}\left[\frac{1}{k}\mathrm{Var}(y_{hi} - Q_h x_{hi})\right]$$

여기서 $Q_h = \dfrac{\bar{Y}_h}{\bar{X}_h}$

이고 不偏分離比推定量의 分散은

$$(26)\quad \mathrm{Var}(\bar{y}_s') \doteq \sum_{h=1}^{L} \frac{N_h^2}{N^2}\left[\frac{1}{k}\mathrm{Var}(y_{hi} - \bar{R}_h x_{hi})\right]$$

이므로 두 分散의 差는 다음과 같다.

$$\mathrm{Var}(\bar{y}_s) - \mathrm{Var}(\bar{y}_s') = \sum_{h=1}^{L} \frac{N_h^2}{N^2}\left[\frac{1}{k}\mathrm{Var}(x_{hi})\right.$$

$$(27)\quad \left.\{(Q_h - B_h)^2 - (\bar{R}_h - B_h)\}\right]$$

여기서 $B_h = \dfrac{\mathrm{Cov}(x_{hi}, y_{hi})}{\mathrm{Var}(x_{hi})}$

이 結果는 分散값이 각 層別로 계산되어 있다는 것 이외에 結合比推定量의 分散比較結果와 다를 것이 없다. 즉 各層別 $\bar{R}_h$ 값이 $Q_h$ 보다 $B_h$ 에 더욱 가까우면 不偏分離比推定量이 偏倚分離比推定量보다 有效한 推定量이 된다.[8]

## Ⅴ. 結 論

標本抽出方法으로 層化集落抽出을 하였을 경우 補助資料 및 標本資料로 얻어지는 比推定量은 分離的인 方法과 結合的인 方法으로 구할 수 있다. 그러나 이렇게 얻어진 比推定量은 一般的으로 不偏推定量이 되지 못하고 特히 層數가 크고 각 層別로 集落을 적게 抽出하는 경우에 그 偏倚는 상당히 커지게 된다.

이에 따라 本 硏究는 層化集落抽出에서의 不偏比推定量을 구하기 위하여 「Hately-Ross」의 不偏比推定量을 擴張함으로서 不偏分離比推定量과 不偏結合比推定量을 導出하고 各各의 分散을 구하여 이를 偏倚比推定量의 그것과 比較하였다. 그 結果結合比推定量의 경우 各 變量 標本값의 比의 平均 $\bar{R}$ 가 平均比인 $Q$ 보다 回歸係數 $B$ 에 더욱 近似하면 不偏結合比推定 이 偏倚結合比推定量보다 有效한 것으로 나타났다. 分離比推定量의 경우도 마찬가지로 各 層別로 위의 條件이 만족되면 不偏分離比推定量이 偏倚分離比推定量보다 有效한 것으로 나타났다.

## 參 考 文 獻


1. Cochran, W.G., Sampling Techniques, 3rd. ed., John Wiley and Sons, New Yoak, 1977.
2. Goodman, L.A. and Hartely, H.O., "The Precision of Unbiased Ratio Type Estimators," *Journal of the American Statistical Association*, Vol. 53(1958), pp. 491—508.
3. Hansen, M.H., Hurwitz, W.N. and Madow, W.G., Sample Survey Method and Theory, Vol. Ⅰ, Ⅱ, John Wiley and Sons, New York, 1953.
4. Hartely, H.O. and Ross, A., "Unhiased Ratio Estimate," *Nature*, Vol. 174(1954), pp. 270—271.
5. Kendall, M.G. and Staurt, A., The Advanced Theory of Statistics, Vol. 1, 3rd. ed., Harfner Pubishing Company, New York, 1969.
6. Pascual, J.N., "Un biaseed Ratio Estimators in Stratified Sampling," *Journal of the American Statistical Association*, Vol. 56(1961), pp. 70—87.
7. Raj, D., Sampling Theory, McGraw-Hill, New York, 1968.
8. Raj. D., "A Note on the Variance of the Ratio Estimate," *Journal of the American Statistical Association* Vol. 59(1964), pp. 895—899.
9. Robson, D.S., "Application of Multivariate PolyKays to the Theory of Unbiased Ratio Type Estimation," *Journal of the American Statistical Association*, Vol. 52(1957), pp. 511—522.
10. Yamane, T., Elementary Samping Theory, Prentice-Hall, Englewoold Cliffs, N.J., 1967.


(8) 實質的으로 $\bar{R}$, $Q$, $B$(또는 $\bar{R}_h$, $Q_h$, $B_h$)의 값을 比較할 때에는 이들의 推定量인 $\bar{r}_{sc}$, $\dfrac{\bar{y}_{sc}}{\bar{x}_{sc}}$, $\dfrac{\mathrm{Cov}(\bar{x}_{sc}, \bar{y}_{sc})}{\mathrm{Var}(\bar{x}_{sc})}$ (또는 各 層에서의 값)으로 計算하게 된다.