Iraqi Journal of Statistical Sciences

http://stats.uomosul.edu.iq

Generalized Multivariate Mixture Ratio Estimators for the Population Mean in Multi-Phase Sampling using Multi-Auxiliary Characteristics

Emmanuel F. Ologunleko[1*] Peter I. Ogunyinka[2] Oluwabunmi E. Ologunleko[3] Ademola Sodipo[3]

[1]Department of Statistics, Olabisi Onabanjo University, Ago Iwoye, Nigeria [2]Independent Statistician, Hull, United Kingdom,[3]University of Ibadan, Ibadan, Nigeria

Article information

Article history:
Received: March 19, 2025
Revised: May 21, 2025
Accepted July 8, 2025
Available online: June 1, 2026

Keywords:
Multi-phase sampling, Population mean estimation, Generalized mixture ratio estimators, Auxiliary variables, Auxiliary attributes

Correspondence:
Emmanuel F. Ologunleko[1*]
ologunlekoemmanuel@gmail.com,
Peter I. Ogunyinka
ogunyinka.peter@gmail.com
Oluwabunmi E. Ologunleko
okikiolaoluwabunmi@gmail.com

Abstract

The integration of auxiliary information into survey sampling has continued to attract significant attention in improving the efficiency of estimators. Traditional ratio and regression estimators, though effective, often show limitations when confronted with multiple auxiliary variables and complex sampling designs. This study proposed a class of *Generalized Multivariate Mixture Ratio Estimators* for estimating the population mean in multi-phase sampling design with multi-auxiliary characteristics. The proposed estimators extend beyond conventional single-variable approaches by combining information from several auxiliary variables and auxiliary attributes. The theoretical properties of the estimator are derived, including the Mean Square Error expressions. Theoretical comparative analysis confirmed that the proposed estimators achieved notable gains in efficiency relative to the reviewed estimators. Simulation studies further confirmed the efficiency of the proposed estimators across varying sample sizes (asymptotically), correlation structures, and distributional conditions. Overall, the generalized multivariate mixture ratio estimators confirmed to be more efficient in population mean estimation under multi-phase sampling design.

DOI: 10.33899/iqjoss.v23i1.63527 , ©Authors, 2026, College of Computer Science and Mathematics University of Mosul.
This is an open access article under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The challenge of accurately estimating the population mean remains a central concern in sampling theory, prompting researchers to explore various approaches that harness several auxiliary variables to improve estimator precision. It is widely recognized among survey statisticians that incorporating auxiliary information can significantly enhance the accuracy of estimates for the study variable. Over time, numerous estimators including ratio, regression, and product types have been developed and refined based on this principle (Ogunyinka, 2019). One of the foundational methods in this area is the classical ratio estimator introduced by Cochran (1940), which uses known data about an auxiliary variable to estimate the population mean of a study variable. Subsequent research by Rao (1957), Tripathi (1970), Srivenkataramana (1980), and Srivenkataramana and Tracy (1979), later expanded bys Chaudhuri and Roy (1994), further advanced the application of auxiliary information in ratio estimation. Olkin (1958) extended these ideas by making use of multiple highly correlated auxiliary variables, showing that this approach produced more accurate results compared to using either no auxiliary information or a single auxiliary variable. Building on Tripathi's (1970) research, Samuiddin and Hanif (2007) categorized the use of auxiliary information into three situations: Firstly, the Full Information Case (FIC), where complete data on all additional variables is available; Secondly, No Information Case (NIC), where such information is absent; and lastly is the Partial Information Case (PIC), where only some auxiliary information is accessible. All these classifications offer flexibility in applying auxiliary data under varying conditions. In addition to continuous auxiliary variables, a dichotomous population characteristic indicating the presence or absence of a particular attribute can also serve as a useful auxiliary variable when it is strongly correlated with the study variable. This type of auxiliary information, often referred to as an **auxiliary attribute**, has been shown to improve estimation accuracy in several studies (Bahl & Tuteja, 1991; Jhajj et al., 2006; Rajesh et al., 2007; Hanif et al., 2009; Moeen et al., 2012; Ogunyinka & Sodipo, 2017). Despite these advancements, Swain (2000) noted that one of the major challenges in multipurpose surveys is the simultaneous estimation of population means for multiple variables. In order to address this, Tripathi and Khattree (1989) suggested techniques for using many additional data under simple random sampling to determine the means of several study variables. Furthermore, Tripathi (1989) later extended this work to the context of two-phase sampling. Building on these contributions, Ahmad et al. (2009, 2010) introduced classes of ratio estimators for two-phase sampling with multiple auxiliary variables, accommodating scenarios corresponding to FIC, NIC, and PIC. More recently, Ogunyinka et al. (2019) proposed a generalized class of ratio-cum-product estimators in two-phase sampling, further enhancing estimation efficiency through the use of multiple auxiliary variables. Moreover, Moeen et al. (2012) established that an estimator combining an auxiliary variable (quantitative) with an auxiliary attribute (qualitative) can produce a more accurate estimator, commonly referred to as a *mixture estimator*. Consequently, there is a clear need to develop generalized multivariate ratio estimators that uses both types of auxiliary characteristics for the estimation of population mean in multi-phase sampling design when there exist both the full and no information situations. In line with this, the present study aims to propose generalized multivariate mixture ratio estimators that effectively use multiple auxiliary characteristics in a multi-phase sampling design to estimate the population mean. The subsequent sections will outline the multi-phase sampling design and present essential properties that form the basis for constructing the proposed estimators.

2. Methodology

Estimating the population mean in multi-phase sampling using several auxiliary characteristics

This study considered a population size of N units, having $Y_1, Y_2, \ldots, Y_p$ study variables, with $X_1, X_2, \ldots, X_t$ auxiliary variables and $P_1, P_2, \ldots, P_q$ auxiliary attributes. The n_h and n_k $(n_k < n_h)$ are the sample sizes of the h^{th} and k^{th} phases, respectively. The $\bar{x}_{(h)i}$ and $\bar{x}_{(k)i},$ $p_{(h)f}$ and $p_{(k)f}$ are the i^{th} and f^{th} auxiliary variables and auxiliary attributes from the h^{th} and k^{th} phases, respectively. Let ρ_{yx_i} indicate the population correlation coefficient of $Y$and X_i, and let C_{x_i} mean coefficient of variation of i^{th} auxiliary variables. Also, the following hold $\theta_h = \left(\frac{1}{n_h} - \frac{1}{N}\right)$ and $\theta_k = \left(\frac{1}{n_k} - \frac{1}{N}\right)$ such that $n_1 > n_2$ and $\theta_1 < \theta_2$. Similarly,

$$\left.\begin{aligned} \bar{y}_h &= \bar{Y} + \bar{e}_{y_h}, \\ \bar{x}_{(h)_i} &= \bar{X}_i + \bar{e}_{x_{(h)_i}} \text{ or } \bar{e}_{x_{(h)_i}} = \left(\bar{x}_{(h)_i} - \bar{X}_i\right), \text{and} \\ p_{(h)_f} &= P_f + \bar{e}_{\tau_{(h)f.}} \end{aligned}\right\} \quad (1)$$

In general, the sampling errors are expected to be very negligible values. This study confirmed that:

$$E\left(\bar{e}_{x_{(h)_i}}\right) = E\left(\bar{e}_{\tau_{(h)_i}}\right) = E\left(\bar{e}_{y_h}\right) = 0,$$

where E_h and E_k indicates the expectations of errors of hth and kth phase sampling, respectively.

2.1 Reviewed Estimators

Ahmad *et al.* (2009) developed the estimated population mean of a generalised ratio estimator in multi-phase sampling where the population data is provided for all the supplementary variables in the estimator component with pth multi-auxiliary variables. The estimator is termed Full Information Case (FIC) estimator and presented as:

$$Z_1 = \left[\bar{y}_{(k)1} \prod_{i=1}^{q} \left(\frac{\bar{X}_i}{\bar{x}_{(k)i}}\right)^{\alpha_{i1}} * \ \bar{y}_{(k)2} \prod_{i=1}^{q} \left(\frac{\bar{X}_i}{\bar{x}_{(k)i}}\right)^{\alpha_{i2}} * \ldots * \bar{y}_{(k)p} \prod_{i=1}^{q} \left(\frac{\bar{X}_i}{\bar{x}_{(k)i}}\right)^{\alpha_{ip}}\right] \quad (2)$$

The corresponding minimized Mean Square Error (MSE_{min}) is presented as

$$MSE(Z_1)_{min} \cong \theta_k \bar{Y}^2 C_y^2 \left(1 - \rho^2_{y.\underset{\sim}{x}_q}\right)_{p\times p} \quad (3)$$

Furthermore, Ahmad *et al* (2009) developed a generalised ratio estimator in multi-phase sampling for estimating the population mean when the population data is not provided for all the additional variables. The estimator is termed No Information Case (NIC) estimator and presented as:

$$Z_2 = \left[\bar{y}_{(k)1} \prod_{i=1}^{q} \left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}}\right)^{\alpha_{i1}} * \ \bar{y}_{(k)2} \prod_{i=1}^{q} \left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}}\right)^{\alpha_{i2}} * \dots * \bar{y}_{(k)p} \prod_{i=1}^{q} \left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}}\right)^{\alpha_{ip}}\right] \quad (4)$$

The corresponding MSE_{min} is presented as

$$MSE(Z_2)_{min} \cong \bar{Y}^2 C_y^2 \left[\theta_k \left(1 - \rho^2_{y_i\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q}\right) + \theta_h \left(\rho^2_{y_i\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q}\right)\right]_{p\times p} \quad (5)$$

2.2 Proposed Generalized Multivariate Mixture Ratio ($GMMR_{at}$) Estimators with Multi-Auxiliary Characteristics in Multi-phase Sampling

The proposed $GMMR_{at}$ estimators for the population mean with multi-auxiliary characteristics in multi-phase sampling in full information case using q total number of multiple auxiliary attributes and auxiliary variables are presented in equation (2.2.1). The estimator is within the range $\{(1,2,\dots,t),(t+1,\ t+2,\dots,q)\}$ of multi-auxiliary characteristics, such that auxiliary characteristics within the range $\{i = (1,2,\dots,t)\}$ and $\{f = (t+1, t+2,\dots,q)\}$ are multi-auxiliary variables and multi-auxiliary attributes, respectively, which satisfy conditions of ratio estimation method.

The proposed estimated population mean for the FIC is given as:

$$Z_{3(1\times p)} = \left[\bar{y}_{(k)1} \prod_{i=1}^{t} \left(\frac{\bar{X}_i}{\bar{x}_{(k)i}}\right)^{\alpha_{i1}} \prod_{f=t+1}^{q} \left(\frac{P_f}{p_{(k)f}}\right)^{\beta_{f1}} \ \bar{y}_{(k)2} \prod_{i=1}^{q} \left(\frac{\bar{X}_i}{\bar{x}_{(k)i}}\right)^{\alpha_{i2}} \prod_{i=1}^{q} \left(\frac{P_f}{p_{(k)f}}\right)^{\beta_{f2}} \dots \ \bar{y}_{(k)p} \prod_{i=1}^{q} \left(\frac{\bar{X}_i}{\bar{x}_{(k)i}}\right)^{\alpha_{ip}} \prod_{i=1}^{q} \left(\frac{P_f}{p_{(k)f}}\right)^{\beta_{fp}}\right]. \quad (6)$$

Substitute equation (2.1) into equation (2.2.1)

$$Z_{3(1\times p)} = \left[\left(\left(\bar{Y}_1 + \bar{e}_{y(k)1}\right) \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(k)i}}{\bar{X}_i}\right)^{-\alpha_{i1}} \prod_{f=t+1}^{q} \left(1 + \frac{\bar{e}_{\tau(k)f}}{P_f}\right)^{-\beta_{f1}}\right) * \left(\left(\bar{Y}_2 + \bar{e}_{y(k)2}\right) \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(k)i}}{\bar{X}_i}\right)^{-\alpha_{i2}} \prod_{f=t+1}^{q} \left(1 + \frac{\bar{e}_{\tau(k)f}}{P_f}\right)^{-\beta_{f2}}\right) * \dots * \left(\left(\bar{Y}_p + \bar{e}_{y(k)p}\right) \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(k)i}}{\bar{X}_i}\right)^{-\alpha_{ip}} \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(k)i}}{\bar{X}_i}\right)^{-\alpha_{ip}}\right)\right]$$

Apply Taylor's series and expand up to the first order of degree to obtain,

$$Z_{3(1\times p)} = \left\{\left[\bar{Y}_1 + \bar{e}_{y(k)1} - \bar{Y} \sum_{i=1}^{t} \alpha_{i1} \frac{\bar{e}_{x(k)i}}{\bar{X}_i} - \bar{Y} \sum_{f=t+1}^{q} \beta_{f1} \frac{\bar{e}_{\tau(k)f}}{P_f}\right] * \left[\bar{Y}_2 + \bar{e}_{y(k)2} - \bar{Y} \sum_{i=1}^{t} \alpha_{i2} \frac{\bar{e}_{x(k)i}}{\bar{X}_i} - \bar{Y} \sum_{f=t+1}^{q} \beta_{f2} \frac{\bar{e}_{\tau(k)f}}{P_f}\right] * \dots * \left[\bar{Y}_p + \bar{e}_{y(k)p} - \bar{Y} \sum_{i=1}^{t} \alpha_{ip} \frac{\bar{e}_{x(k)i}}{\bar{X}_i} - \bar{Y} \sum_{f=t+1}^{q} \beta_{fp} \frac{\bar{e}_{\tau(k)f}}{P_f}\right]\right\}$$

Representing the result in matrix will give

$$Z_{3(1\times p)} = \left\{\left[\left(\bar{Y}_1 + \bar{e}_{y(k)1}\right)\left(\bar{Y}_2 + \bar{e}_{y(k)2}\right) \cdots \left(\bar{Y}_p + \bar{e}_{y(k)p}\right)\right]_{(1\times p)} - \right.$$

$$\left[\bar{e}_{x_{(k)1}}\ \bar{e}_{x_{(k)2}} \dots \bar{e}_{x_{(k)p}}\right]_{1\times t}\left[\frac{\bar{Y}_j}{\bar{X}_i}\alpha_{ij}\right]_{t\times P} - \left[\bar{e}_{\tau_{(k)1}}\bar{e}_{\tau_{(k)2}} \dots \bar{e}_{\tau_{(k)f}}\right]_{1\times q}\left[\frac{\bar{Y}_j}{P_f}\beta_{fj}\right]_{q\times P}\right\}$$

$$Z_{3(1\times p)} = \left\{\left[\bar{Y}_1\ \bar{Y}_2\ \dots \bar{Y}_p\right]_{(1\times p)} + \left[\bar{e}_{y_{(k)1}}\ \ \bar{e}_{y_{(k)2}}\ \dots\ \bar{e}_{y_{(k)p}}\right]_{(1\times p)} - \left(\left[\bar{e}_{x_{(k)1}}\ \bar{e}_{x_{(k)2}} \dots \bar{e}_{x_{(k)p}}\right]_{1\times t}\right.\right.$$

$$\left.\left[\frac{\bar{Y}_j}{\bar{X}_i}\alpha_{ij}\right]_{t\times P}\right) - \left(\left[\bar{e}_{\tau_{(k)1}}\bar{e}_{\tau_{(k)2}} \dots \bar{e}_{\tau_{(k)f}}\right]_{1\times q}\left[\frac{\bar{Y}_j}{P_i}\beta_{fj}\right]_{q\times P}\right)\right\}$$

Let

$\left[\bar{Y}_1\ \bar{Y}_2\ \dots \bar{Y}_p\right]_{(1\times p)} = \bar{Y}_{(1\times p)}$,

$\left[\bar{e}_{y_{(k)1}}\ \ \bar{e}_{y_{(k)2}}\ \dots\ \bar{e}_{y_{(k)p}}\right]_{(1\times p)} = \bar{D}_{y(1\times p)}$,

$\left[\bar{e}_{x_{(k)1}}\ \bar{e}_{x_{(k)2}} \dots \bar{e}_{x_{(k)p}}\right]_{1\times t} = \bar{D}_{x(1\times t)}$,

$\left[\bar{e}_{\tau_{(k)1}}\bar{e}_{\tau_{(k)2}} \dots \bar{e}_{\tau_{(k)f}}\right]_{1\times q} = \bar{D}_{\tau(1\times q)}$,

$\left[\frac{\bar{Y}_j}{\bar{X}_i}\alpha_{ij}\right]_{t\times P} = A_{(t\times p)}$, and $\left[\frac{\bar{Y}_j}{P_f}\beta_{fj}\right]_{q\times P} = B_{(q\times p)}$.

Then,

$Z_{3(1\times p)} = \bar{Y}_{(1\times p)} + \bar{D}_{y(1\times p)} - \bar{D}_{x(1\times t)}A_{(t\times p)} - \bar{D}_{\tau(1\times q)}B_{(q\times p)}$, and

$Z_{3(1\times p)} - \bar{Y}_{(1\times p)} = \bar{D}_{y(1\times p)} - \bar{D}_{x(1\times t)}A_{(t\times p)} - \bar{D}_{\tau(1\times q)}B_{(q\times p)}$.

Let $\sum$ denote matrix representation, and using information associated to the auxiliary characteristics from $h-th$ and $k-th$ phases, the corresponding MSE$_{min}$ of $Z_{3(1\times p)}$ could be reported as

$$\sum\nolimits_{Z_3(p\times p)} = E_h E_{k/h}\left[\left(Z_{3(1\times p)} - \bar{Y}_{(1\times p)}\right)'\left(Z_{3(1\times p)} - \bar{Y}_{(1\times p)}\right)\right]$$

or

$$\begin{aligned}\sum\nolimits_{Z_{3(p\times p)}} &= \theta_k \sum\nolimits_{y(p\times p)} - A'(\theta_h-\theta_k)\sum\nolimits'_{yx(t\times p)} - B'(\theta_h-\theta_k)\sum\nolimits'_{yx(\times p)} + \\ &(\theta_h-\theta_k)\sum\nolimits_{yx(p\times t)} A - A'(\theta_k-\theta_h)\sum\nolimits_{x(t\times t)} A + B'(\theta_h-\theta_k)\sum\nolimits'_{x\tau(q\times t)} A \\ &- (\theta_h-\theta_k)\sum\nolimits_{y\tau(p\times q)} B + A'(\theta_h-\theta_k)\sum\nolimits_{x\tau(t\times q)} B \\ &+ B'(\theta_k-\theta_h)\sum\nolimits_{\tau(q\times q)} B \end{aligned} \tag{7}$$

Given that there exist $\sum_{x(t\times t)}^{-1}$ and $\sum_{\tau(q\times q)}^{-1}$, the best values of A and B that minimize the variance covariance matrix of the MSE of $\sum_{Z_3(p\times p)}$ is given as

$$A_{(t\times p)} = \left[\sum\nolimits_{x(t\times t)}^{-1}\ \ \sum\nolimits_{yx(t\times p)}^{1}\right] \ and\ B_{(q\times p)} = \left[\sum\nolimits_{z(q\times q)}^{-1}\ \ \sum\nolimits_{y\tau(q\times p)}^{1}\right]\ and\ (2.2.3)$$

$$A'_{(p\times t)} = \left[\sum\nolimits_{yx(p\times t)}\ \ \sum\nolimits_{x(t\times t)}^{-1}\right]\ and\ B'_{(p\times q)} = \left[\sum\nolimits_{y\tau(q\times q)}\ \ \sum\nolimits_{\tau(q\times q)}^{-1}\right]. \tag{8}$$

Hence, Substitute equations (2.2.3) and (2.2.4) into equation (2.3.2) to obtain the minimized MSE (MSE_{min}) as

$$\begin{aligned}\sum\nolimits_{Z_{3(p\times p)}} &\simeq \theta_k \sum\nolimits_{y(p\times p)} - \left(\sum\nolimits_{yx(p\times t)}\ \sum\nolimits_{x(t\times t)}^{-1}\ \theta_k \sum\nolimits'_{yx(t\times p)} - \sum\nolimits_{y\tau(p\times q)}\right. \\ &\left.\sum\nolimits_{\tau(q\times q)}^{-1}\ \theta_k \sum\nolimits_{y\tau(q\times p)}^{-1}\right) - \left(\theta_k \sum\nolimits_{yx(p\times t)}\ \sum\nolimits_{x(t\times t)}^{-1}\ \sum\nolimits'_{yx(t\times p)}\right) + \\ &\left(\sum\nolimits_{yx(p\times t)}\ \sum\nolimits_{yx(p\times t)}\ \sum\nolimits_{x(t\times t)}^{-1}\ \theta_k \sum\nolimits_{x(t\times t)}\ \sum\nolimits_{x(t\times t)}^{-1}\ \sum\nolimits'_{yx(t\times p)}\right) + \\ &\left(\sum\nolimits_{\tau(q\times q)}^{-1}\ \theta_k\theta_k \sum\nolimits'_{x\tau(t\times q)}\ \sum\nolimits_{x(t\times t)}^{-1} \sum\nolimits'_{yx(t\times p)}\right) - \left(\theta_k \sum\nolimits_{y\tau(p\times q)}\right. \\ &\left.\sum\nolimits_{\tau(q\times q)}^{-1}\ \sum\nolimits'_{y\tau(q\times q)}\ \sum\nolimits'_{y\tau(q\times q)}\right) + \left(\sum\nolimits_{yx(p\times t)}\ \sum\nolimits_{x(t\times t)}^{-1}\right.\end{aligned}$$

$$\theta_k \sum_{x\tau(t\times q)} \sum_{\tau(q\times q)}^{-1} \sum_{y\tau(q\times q)}^{\prime} \Big) + \Big(\sum_{y\tau(p\times q)} \sum_{\tau(q\times q)}^{-1} \theta_k \sum_{\tau(q\times q)} \sum_{\tau(q\times q)}^{-1} \sum_{y\tau(q\times q)}^{\prime} \Big).$$

Further simplification will give

$$\sum_{Z_{3(p\times p)}} \simeq \theta_k \sum_{y(p\times p)} - \left(\theta_k \sum_{yx(p\times t)} \sum_{x(t\times t)}^{-1} \sum_{yx(t\times p)}^{\prime} \right) - \left(\theta_k \sum_{y\tau(p\times q)} \sum_{\tau(q\times q)}^{-1} \sum_{y\tau(q\times q)}^{\prime} \right)$$

or

$$\sum_{Z_{3(p\times p)}} \simeq \left[\theta_k \sum_{y(p\times p)} - \left(\theta_k \sum_{yx(p\times t)} \sum_{x(t\times t)}^{-1} \sum_{yx(t\times p)}^{\prime} \right) - \left(\theta_k \sum_{y\tau(p\times q)} \sum_{\tau(q\times q)}^{-1} \sum_{y\tau(q\times q)}^{\prime} \right) \right]_{p\times p.}$$

Thus, the corresponding MSE_{min} is given as

$$\sum_{Z_{3(p\times p)}} \simeq \theta_k \bar{Y}^2 C_y^2 \left[1 - \rho^2_{y_i \cdot (\underset{\sim}{x}, \underset{\sim}{\tau})_q} \right]_{p\times p.} \tag{9}$$

2.4 Proposed Generalized Multivariate Mixture Ratio Estimators for the Population Mean with Multi-Auxiliary Characteristics in Multi-Phase Sampling in No information case (NIC).

The proposed GMMR_{at} estimators for the population mean with multi-auxiliary characteristics in multi-phase sampling for NIC using q total number of multiple auxiliary attributes and auxiliary variables. The estimator is within the range $\{(1,2,\dots,t),(t+1,\ t+2,\dots,q)\}$ of multi-auxiliary characteristics, such that auxiliary characteristics within the range $\{i=(1,2,\dots,t)\}$ and $\{f=(t+1,t+2,\dots,q)\}$ are multi-auxiliary variables and multi-auxiliary attributes, respectively, which satisfy the condition of using ratio estimation method. The proposed NIC estimator is given as:

$$Z_{4(1\times p)} = \left[\left(\bar{y}_{(k)1} \prod_{i=1}^{t} \left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}} \right)^{\alpha_{i1}} \prod_{f=t+1}^{q} \left(\frac{p_{(h)f}}{p_{(k)f}} \right)^{\beta_{f1}} \right) * \left(\bar{y}_{(k)2} \prod_{i=1}^{t} \left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}} \right)^{\alpha_{i2}} \prod_{f=t+1}^{q} \left(\frac{p_{(h)f}}{p_{(k)f}} \right)^{\beta_{f2}} \right) * \dots * \left(\bar{y}_{(k)p} \prod_{i=1}^{t} \left(\frac{\bar{x}_{(h)i}}{\bar{x}_{(k)i}} \right)^{\alpha_{ip}} \prod_{f=t+1}^{q} \left(\frac{p_{(h)f}}{p_{(k)f}} \right)^{\beta_{fp}} \right) \right]. \tag{10}$$

Substitute equation (2.1) to (2.3.1) to obtain

$$Z_{4(1\times p)} = \Bigg[(\bar{Y}_1 + \bar{e}_{y(k)1}) \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(h)i}}{\bar{X}_i}\right) \left(1 + \frac{\bar{X}_i}{\bar{e}_{x(k)i}}\right)^{\alpha_{i1}} \prod_{f=t+1}^{q} \left(1 + \frac{\bar{e}_{\tau(h)f}}{P_f}\right) \left(1 + \frac{P_f}{\bar{e}_{\tau(k)f}}\right)^{\beta_{f1}} * (\bar{Y}_2 + \bar{e}_{y(k)2}) \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(h)i}}{\bar{X}_i}\right) \left(1 + \frac{\bar{X}_i}{\bar{e}_{x(k)i}}\right)^{\alpha_{i2}} \prod_{f=t+1}^{q} \left(1 + \frac{\bar{e}_{\tau(k)f}}{P_f}\right) \left(1 + \frac{P_f}{\bar{e}_{\tau(k)f}}\right)^{\beta_{f2}} *** (\bar{Y}_p + \bar{e}_{y(k)p}) \prod_{i=1}^{t} \left(1 + \frac{\bar{e}_{x(h)i}}{\bar{X}_i}\right) \left(1 + \frac{\bar{X}_i}{\bar{e}_{x(k)i}}\right)^{\alpha_{ip}} \prod_{f=t+1}^{q} \left(1 + \frac{\bar{e}_{\tau(h)f}}{P_f}\right) \left(1 + \frac{P_f}{\bar{e}_{\tau(k)f}}\right)^{\beta_{fp}} \Bigg].$$

Apply the Taylor's series of expansion, expand and simplify up to the first order of degree, and further representing the result in matrix will yield

$$Z_{4(1\times p)} = \Big\{ \left[\bar{Y}_1 \;\; \bar{Y}_2 \;\; \dots \bar{Y}_p\right]_{(1\times p)} + \left[\bar{e}_{y_{(k)1}} \;\; \bar{e}_{y_{(k)2}} \;\; \dots \;\; \bar{e}_{y_{(k)p}}\right]_{(1\times p)} + \left[\left(\bar{e}_{x_{(h)1}} - \bar{e}_{x_{(k)1}}\right)\left(\bar{e}_{x_{(h)2}} - \bar{e}_{x_{(k)2}}\right)\cdots\left(\bar{e}_{x_{(h)p}} - \bar{e}_{x_{(k)p}}\right)\right]_{1\times t}\left[\frac{\bar{Y}_j}{\bar{X}_i}\alpha_{ij}\right]_{t\times P} + \left[\left(\bar{e}_{\tau_{(h)1}} - \bar{e}_{\tau_{(k)1}}\right)\;\left(\bar{e}_{\tau_{(h)2}} - \bar{e}_{\tau_{(k)2}}\right)\cdots\left(\bar{e}_{\tau_{(h)p}} - \bar{e}_{\tau_{(k)p}}\right)\right]_{1\times q}\left[\frac{\bar{Y}_j}{P_f}\beta_{fj}\right]_{q\times P}\Big\}.$$

Let

$$\left[\bar{Y}_1 \;\; \bar{Y}_2 \;\; \dots \bar{Y}_p\right]_{(1\times p)} = \bar{Y}_{(1\times p)},$$

$$\left[\bar{e}_{y_{(k)1}} \;\; \bar{e}_{y_{(k)2}} \;\; \dots \;\; \bar{e}_{y_{(k)p}}\right]_{(1\times p)} = \bar{D}_{y(1\times p)},$$

$$\left[\left(\bar{e}_{x_{(h)1}} - \bar{e}_{x_{(k)1}}\right)\left(\bar{e}_{x_{(h)2}} - \bar{e}_{x_{(k)2}}\right)\cdots\left(\bar{e}_{x_{(h)p}} - \bar{e}_{x_{(k)p}}\right)\right]_{1\times t} = \bar{D}_{x(1\times t)},$$

$$\left[\left(\bar{e}_{\tau_{(h)1}} - \bar{e}_{\tau_{(k)1}}\right)\;\left(\bar{e}_{\tau_{(h)2}} - \bar{e}_{\tau_{(k)2}}\right)\cdots\left(\bar{e}_{\tau_{(h)p}} - \bar{e}_{\tau_{(k)p}}\right)\right]_{1\times q} = \bar{D}_{\tau(1\times q)},$$

$$\left[\frac{\bar{Y}_j}{\bar{X}_i}\alpha_{ij}\right]_{t\times P} = A_{(t\times p)}, \qquad and \;\left[\frac{\bar{Y}_j}{P_f}\beta_{fj}\right]_{q\times P} = B_{(q\times p)}.$$

Then,

$Z_{4(1\times p)} = \bar{Y}_{(1\times p)} + \bar{D}_{y(1\times p)} + \bar{D}_{x(1\times t)}A_{(t\times p)} + \bar{D}_{\tau(1\times q)}B_{(q\times p)}$, and

$Z_{4(1\times p)} - \bar{Y}_{(1\times p)} = \bar{D}_{y(1\times p)} + \bar{D}_{x(1\times t)}A_{(t\times p)} + \bar{D}_{\tau(1\times q)}B_{(q\times p)}$.

Let $\sum$ denote matrix representation. Using information associated to the auxiliary characteristics from $h-th$ and $k-th$ phases, the corresponding MSE$_{min}$ of $Z_{4(1\times p)}$ could be reported as

$$\sum\nolimits_{Z_4(p\times p)} = E_h E_{k/h}\left[\left(Z_{4(1\times p)} - \bar{Y}_{(1\times p)}\right)'\left(Z_{4(1\times p)} - \bar{Y}_{(1\times p)}\right)\right]$$

or

$$\sum\nolimits_{Z_4(p\times p)} = \theta_k \sum\nolimits_{y(p\times p)} + A'(\theta_h - \theta_k)\sum\nolimits'_{yx(t\times p)} + B'(\theta_h - \theta_k)\sum\nolimits'_{yx(\times p)} + (\theta_h - \theta_k)\sum\nolimits_{yx(p\times t)} A + A'(\theta_k - \theta_h)\sum\nolimits_{x(t\times t)} A + B'(\theta_h - \theta_k)\sum\nolimits'_{x\tau(q\times t)} A + (\theta_h - \theta_k)\sum\nolimits_{y\tau(p\times q)} B + A'(\theta_h - \theta_k)\sum\nolimits_{x\tau(t\times q)} B + B'(\theta_k - \theta_h)\sum\nolimits_{\tau(q\times q)} B \tag{11}$$

Given that there exists $\sum_{x(t\times t)}^{-1}$ and $\sum_{\tau(q\times q)}^{-1}$, the best values of A and B that minimize the variance covariance matrix of the MSE of $\sum_{Z_4(p\times p)}$ is given as

$$A_{(t\times p)} = \left[\sum\nolimits_{x(t\times t)}^{-1} \;\; \sum\nolimits_{yx(t\times p)}^{1}\right] \; and \; B_{(q\times p)} = \left[\sum\nolimits_{\tau(q\times q)}^{-1} \;\; \sum\nolimits_{y\tau(q\times p)}^{1}\right] \; and \tag{12}$$

$$A'_{(p\times t)} = \left[\sum\nolimits_{yx(p\times t)} \;\; \sum\nolimits_{x(t\times t)}^{-1}\right] \; and \; B'_{(p\times q)} = \left[\sum\nolimits_{y\tau(q\times q)} \;\; \sum\nolimits_{\tau(q\times q)}^{-1}\right]. \tag{13}$$

Substitute equation (2.3.3)and (2.3.4) into equation (2.3.2) to obtain the MSE as

$$\sum\nolimits_{Z_4(p\times p)} \simeq \theta_k \sum\nolimits_{y(p\times p)} + \sum\nolimits_{yx(p\times t)} \sum\nolimits_{x(t\times t)}^{-1} (\theta_h - \theta_k)\sum\nolimits'_{yx(t\times p)} + \sum\nolimits_{y\tau(p\times q)} \sum\nolimits_{\tau(q\times q)}^{-1} (\theta_h - \theta_k)\sum\nolimits'_{y\tau(q\times p)} + (\theta_h - \theta_k)\sum\nolimits_{yx(p\times t)} \sum\nolimits_{x(t\times t)}^{-1} \sum\nolimits'_{yx(t\times p)} + \sum\nolimits_{yx(p\times t)} \sum\nolimits_{x(t\times t)}^{-1} (\theta_h - \theta_k)\sum\nolimits_{x(t\times t)} \sum\nolimits_{x(t\times t)}^{-1} \sum\nolimits'_{yx(t\times p)} + \sum\nolimits_{\tau(q\times q)}^{-1} (\theta_h - \theta_k)\sum\nolimits'_{x\tau(t\times q)} \sum\nolimits_{x(t\times t)}^{-1} \sum\nolimits'_{yx(t\times p)} + (\theta_h - \theta_k)\sum\nolimits_{y\tau(p\times q)} \sum\nolimits_{\tau(q\times q)}^{-1}$$

$$\sum\nolimits'_{y\tau(q\times q)} + \sum\nolimits_{yx(p\times t)} \sum\nolimits^{-1}_{x(t\times t)} (\theta_h - \theta_k) \sum\nolimits_{x\tau(t\times q)} \sum\nolimits^{-1}_{\tau(q\times q)} \sum\nolimits'_{y\tau(q\times q)} + \sum\nolimits_{y\tau(p\times q)} \sum\nolimits^{-1}_{\tau(q\times q)} (\theta_k - \theta_h) \sum\nolimits_{\tau(q\times q)} \sum\nolimits^{-1}_{\tau(q\times q)} \sum\nolimits'_{y\tau(q\times q)}$$

Further simplification gives

$$\sum\nolimits_{Z_4(p\times p)} \simeq \Big[\theta_k \sum\nolimits_{y(p\times p)} - (\theta_k - \theta_h) \sum\nolimits_{yx(p\times t)} \sum\nolimits^{-1}_{x(t\times t)} \sum\nolimits'_{yx(t\times p)} - (\theta_h - \theta_k) \sum\nolimits_{y\tau(p\times q)} \sum\nolimits^{-1}_{\tau(q\times q)} \sum\nolimits'_{y\tau(q\times q)} \Big]_{p\times p,}$$

and the corresponding MSE_{min} is given as

$$\sum\nolimits_{Z_4(p\times p)} \simeq \bar{Y}^2 C_y^2 \left[\theta_k \left(1 - \rho^2_{y_i\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q}\right) + \theta_h \left(\rho^2_{y_i\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q}\right)\right]_{p\times p.} \tag{14}$$

3. Result

3.1 Theoretical Analysis of the Reviewed Estimators and proposed Estimators

This section presents a theoretical and empirical comparison between the reviewed estimators and the proposed estimator. In the theoretical analysis, efficiency is determined by comparing the minimum MSE of the competing estimators. For instance, if $MSE(\bar{y}_{st})_{min} - MSE(\bar{y}_{\delta})_{min} < 0$, this implies that the estimator $\bar{y}_{st}$ achieves a lower error and is therefore more efficient than $\bar{y}_{\delta}$. Conversely, if the inequality does not hold, the estimator $\bar{y}_{\delta}$ is preferred.

3.1.1 The proposed FIC estimator versus Ahmad *et al.* (2009) FIC estimator

This comparison is obtained with the expression

$$MSE(Z_3)_{min} - MSE(Z_1)_{min}$$

If $\rho^2_{y\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q} > \rho^2_{y\cdot\underset{\sim}{x}_q}$, then the expression gives

$-\theta_k \overline{Y}^2 C_y^2 \left(\rho^2_{y\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q} - \rho^2_{y\cdot\underset{\sim}{x}_q}\right) < 0$. This implies that $MSE(Z_4)_{min} - MSE(Z_1)_{min} < 0$. Hence, estimator Z_3 is more efficient than estimator Z_1.

3.1.2. The proposed FIC estimator versus Ahmad *et al.* (2009) NIC estimator.

This comparison is obtained with the expression

$$MSE(Z_3)_{min} - MSE(Z_2)_{min}$$

If $\rho^2_{y\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q} > \rho^2_{y\cdot\underset{\sim}{x}_q}$, then the expression gives

$-\overline{Y}^2 C_y^2 \left(\theta_k \rho^2_{y\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q} + \theta_h \rho^2_{y\cdot\underset{\sim}{x}_q}\right) + \theta_k \overline{Y}^2 C_y^2 \rho^2_{y\cdot\underset{\sim}{x}_q} < 0$. This implies that $MSE(Z_3)_{min} - MSE(Z_2)_{min} < 0$. Hence, Z_3 is more efficient than Z_2.

3.1.3. The proposed NIC estimator with Ahmad *et al.* (2009) NIC estimator.

This comparison is obtained with the expression

$$MSE(Z_4)_{min} - MSE(Z_2)_{min}$$

If $\left(\rho^2_{y\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q} > \rho^2_{y\cdot\underset{\sim}{x}_q}\right)$ and $(\theta_k - \theta_h) > 0$, then the expression is given

$\overline{Y}^2 C_y^2 \left(-(\theta_k - \theta_h) \rho^2_{y\cdot(\underset{\sim}{x},\underset{\sim}{\tau})_q} + (\theta_k - \theta_h) \rho^2_{y\cdot\underset{\sim}{x}_q}\right) < 0$. This implies that $MSE(Z_4)_{min} - MSE(Z_2)_{min} < 0$. Hence, Z_4 is more efficient than Z_2.

3.2 Empirical Comparison of all the Estimators

The empirical analyses used data simulation technique. The distribution of the population sizes and the phase-wise samples sizes are presented in table 1.0.

Table 1. Distribution of population size and other phase-wise sample sizes over the sixteen asymptotically simulated populations.

		T12		T13		T14		T15		T23		T24		T25		T34		T35		T45	
Pop.	Pop. Size	1st	2nd	1st	3rd	1st	4th	1st	5th	2nd	3rd	2nd	4th	2nd	5th	3rd	4th	3rd	5th	4th	5th
1	10000	3333	1111	3333	833	3333	667	3333	556	1667	556	1667	417	1667	333	1111	278	1111	222	833	167
2	9650	3217	1072	3217	804	3217	643	3217	536	1608	536	1608	402	1608	322	1072	268	1072	214	804	161
3	9300	3100	1033	3100	775	3100	620	3100	517	1550	517	1550	388	1550	310	1033	258	1033	207	775	155
4	8950	2983	994	2983	746	2983	597	2983	497	1492	497	1492	373	1492	298	994	249	994	199	746	149
5	8600	2867	956	2867	717	2867	573	2867	478	1433	478	1433	358	1433	287	956	239	956	191	717	143
6	8250	2750	917	2750	688	2750	550	2750	458	1375	458	1375	344	1375	275	917	229	917	183	688	138
7	7900	2633	878	2633	658	2633	527	2633	439	1317	439	1317	329	1317	263	878	219	878	176	658	132
8	7550	2517	839	2517	629	2517	503	2517	419	1258	419	1258	315	1258	252	839	210	839	168	629	126
9	7200	2400	800	2400	600	2400	480	2400	400	1200	400	1200	300	1200	240	800	200	800	160	600	120
10	6850	2283	761	2283	571	2283	457	2283	381	1142	381	1142	285	1142	228	761	190	761	152	571	114
11	6500	2167	722	2167	542	2167	433	2167	361	1083	361	1083	271	1083	217	722	181	722	144	542	108
12	6150	2050	683	2050	513	2050	410	2050	342	1025	342	1025	256	1025	205	683	171	683	137	513	103
13	5800	1933	644	1933	483	1933	387	1933	322	967	322	967	242	967	193	644	161	644	129	483	97
14	5450	1817	606	1817	454	1817	363	1817	303	908	303	908	227	908	182	606	151	6O6	121	454	19
15	5100	1700	567	1700	425	1700	340	1700	283	850	283	850	213	850	170	567	142	567	113	425	85
16	4750	1583	528	1583	396	1583	317	1583	264	792	264	792	198	792	158	528	132	528	106	396	79

Table 2. MSEs of the proposed estimators for first-phase and second-phase samplings, respectively, for the sixteen asymptotically simulated populations.

S/N	Info Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	FIC	1.984	1.710	2.230	2.322	2.366	2.443	2.737	2.737	2.668	3.108	2.863	3.542	3.346	3.782	3.935	3.790
2	NIC	2.211	1.955	2.421	2.562	2.562	2.731	2.727	3.104	3.011	3.423	3.230	3.906	3.717	4.266	4.389	4.267

Table 3. Ranking of the proposed estimators based on the obtained MSEs for the first-phase and second phase samplings, respectively, over the sixteen asymptotically simulated populations

		Populations																	
S/N	Info. Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average of Ranks	Rank of Average
1	FIC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.00	1
2	NIC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2.00	2

Table 4. Ranking of the proposed estimators based on the obtained MSEs for the second-phase and third phase samplings, respectively, over the sixteen asymptotically simulated populations

		Populations																	
S/N	Info. Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average of Ranks	Rank of Average
1	FIC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.00	1
2	NIC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2.00	2

Table 5. Ranking of the proposed estimators based on the obtained MSEs for the fourth-phase and fifth-phase samplings, respectively, over the sixteen asymptotically simulated populations

		Populations																
S/N	Info. Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average of Ranks	Rank of Average
1	FIC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.00	1
2	NIC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2.00	2

Table 6. Ranking of the proposed estimators based on the obtained MSEs for the phase-wise sampling when $N = 10000$ respectively,

		Populations																	
S/N	Info. Case	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average of Ranks	Rank of Average
1	FIC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.00	1
2	NIC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2.00	2

4. Discussion and Conclusions

R statistical programming software was employed in this study to simulate and analyse the simulated data obtained for the empirical analysis in comparing the proposed estimators (FIC and NIC). Simulation was conducted for sixteen populations with respect to ten (10) different pair-wise sampling. Table 1 shows the distribution of the population sizes and sample sizes across the sixteen simulated populations and ten (10) pair-wise sampling. T12 is a two-phase sampling which is interpreted as the pair-wise sampling with first-phase and second-phase sampling which is dependent on the first-phase sample (Hidiroglou, 2001). Similarly, T45 is a multi-phase sampling with the fourth-phase sample paired with the independent fifth-phase sample.

Table 2 revealed that as the population size and the sample size reduce down the sixteen (16) populations, the MSEs increases for the two (2) proposed estimators. This shows that the proposed estimators are asymptotically efficient. The ranking in table 3, table 4, and table 5 show that the FIC estimator is ranked first (1st) and the NIC estimator is ranked least (2nd) in the proposed estimator. This result is in accordance with the research of Ogunyinka (2019) and Kung'u *et al.* (2014) as they suggested that an estimator for which the information about the population is available for all auxiliary variables (FIC) would be more effective to the estimator which the information about the population is not available on all auxiliary variables (NIC).

It was confirmed in Table 6 that the MSEs increase as the number of sampling phase increases. This means that the suggested estimators performs better when the number of sampling phase is small. This is justified, according to the result above, that as the number of sampling phases increases, the sample size (where the auxiliary information is obtained) reduces and the sample size for the subsequent phase also reduces. Consequent to this reduction in sample size, the MSE increases.

5. Conclusion

In the theoretical comparison, the proposed FIC and NIC estimators are more asymptotically efficient than the reviewed ratio estimators with respect to the type of information cases. Furthermore, the proposed FIC estimator is more efficient than the proposed NIC estimator. In general, the proposed generalised multivariate mixture ratio estimators for the population mean in multi-phase sampling design are efficient than the corresponding reviewed estimators.

Acknowledgement

Nil

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Ethical Approval

Ethical approval was not required for this study as it did not involve human participants, personal data.

References

1. Ahmad Z, Ahmad, M. and Hanif, M. 2009. Generalized multi-phase multivariate ratio estimators using multi-auxiliary variables, Proceeding of the Tenth Islamic Countries Conference on Statistical Sciences (ICCS-X), Vol 1, 357-366.

   https://www.researchgate.net/publication/266736494_GENERALIZED_MULTI-PHASE_RATIO_ESTIMATORS_USING_MULTI-

AUXILIARY_VARIABLES/citation/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

2. Ahmad Z, Ahmad, M. and Hanif, M. 2010. Generalized multi-phase multivariate ratio estimators in partial informationm case using multi-auxiliary variables, Pakistan Journal of Statistics, 32, 462-470. https://www.researchgate.net/publication/266735501_Generalized_Multi-Phase_Multivariate_Ratio_Estimators_for_Partial_Information_Case_Using_Multi-Auxiliary_Vatiables
3. Bahl, S. and Tetuja, R.K. 1991. Ratio and product type exponential estimators. Journal of Information and Optimization Sciences, 12.1, 159-164. doi: 10.1080/02522667.1991.10699058
4. Chaudhuri, A., & Roy, D. (1994). Model assisted survey sampling strategy in two phases. Metrika, 41, 355-362. https://www.academia.edu/104875014/Model_assisted_survey_sampling_strategy_in_two_phases
5. Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. The journal of agricultural science, 30(2), 262-275. http://dx.doi.org/10.1017/S0021859600048012
6. Hanif, M., Haq I. and Shahbaz, M. Q. 2009. On a new family of estimators using multiple-auxiliary Attributes. World App. Sci. 7.2: 1419-1422. https://www.academia.edu/21226308/Two_Phase_Sampling
7. Hidiroglou, M. A. 2001. Double sampling. Survey Methodology Statistics, Canada 27.2: 143-154.
8. Jhajj, H. S., Sharma, M. K. and Groover, L. K. 2006. Dual of ratio estimators of finite population mean obtained on using linear transformation to auxiliary variables. Journal of the Japan Statistical Society, 36 (1), 107-119. https://cir.nii.ac.jp/crid/1390282680262427136
9. Kung'u J., Chumba G. and Odongo L. 2014. Mixture regression estimators using multi-auxiliary variables and attributes in two-phase sampling. Open Journal of Statistics, 04, 355-366. 10.4236/ojs.2014.45035. https://www.researchgate.net/publication/276497132_Mixture_Regression_Estimators_Using_Multi-Auxiliary_Variables_and_Attributes_in_Two-Phase_Sampling
10. Moeen, M., Shahbaz, Q. and Hanif, M. 2012. Mixture ratio and regression estimators using multi-auxiliary variables and attributes in single phase sampling. Islamic Countries Society of Statistical Sciences, 231. https://www.researchgate.net/publication/286743679
11. Ogunyinka, P. I. and Sodipo, A. A. 2017b. Extension of information case on mixture ratio estimators using multi-auxiliary variables and attributes on two-phase sampling. Transactions of the Nigerian Association of Mathematical Physics, Vol. 4, pp. 237-250. https://www.researchgate.net/publication/317548846_EXTENSION_OF_INFORMATION_CASES_ON_MIXTURE_REGRESSION_ESTIMATORS_USING_MULTI-AUXILIARY_VARIABLES_AND_ATTRIBUTES_IN_TWO-PHASE_SAMPLING
12. Ogunyinka, P. I. 2019. Generalised mixed estimators in two-phase sampling with multi-auxiliary characteristics. An Unpublished Ph.D. Thesis Dissertation, University of Ibadan, Ibadan, Nigeria.
13. Ogunyinka, P. I., Ologunleko, E. F., and Sodipo, A. A. 2019. General class of ratio-cum-product estimators in two-phase sampling using multi-auxiliary variables. Anale Seria Informatica, Vol XVII fasc. 2, 225-233. https://anale-informatica.tibiscus.ro/download/lucrari/17-2-32-Ogunyinka.pdf

14. Olkin, I. 1958. Multivariate ratio estimation for finite populations. Biometrika, 45, 154–165. http://dx.doi.org/10.1093/biomet/45.1-2.154

15. Rajesh, S., Pankaj, C., Nirimala, S. and Florentins, S. 2007. Ratio-Product type exponential estimator for estimating Finite population mean using Information on Auxiliary Attributes. Renaissance High Press, USA. Chapter 1:18-32.

https://www.researchgate.net/publication/308606231_Ratio-Product_Type_Exponential_Estimator_for_Estimating_Finite_Population_Mean_Using_Information_on_Auxiliary_Attribute_Published

16. Rao, J. N. K. (1957). Double ratio estimate in forest surveys. *Journal of the Indian Society of Agricultural Statistics*, *9*, 191-204.

17. Samiuddin M. and Hanif M. 2007. Estimation of population mean in single phase sampling with or without additional information. Pakistan Journal of Statistics, 23, 99-118.

https://www.researchgate.net/publication/237369953

18. Srivenkataramana, T. (1980). A dual to ratio estimator in sample surveys. Biometrika, 67(1), 199-204. http://dx.doi.org/10.1093/biomet/67.1.199

19. Srivenkataramana, T., & Tracy, D. S. (1979). On ratio and product methods of estimation in sampling. Statistica Neerlandica, 33(1), 37-49 https://www.researchgate.net/publication/229650381_On_ratio_and_product_methods_of_estimation_in_sampling

20. Swain, A. K. P. C. 2000. Current developments in survey sampling. Om Prakash, Bhubaneswar, India.

21. Tripathi, T. 1970. Contributions to the sampling theory using multivariate information. Ph.D. Thesis. Department of Statistics, Punjabi University, India. xiv+156pp.

22. Tripathi, T. 1989. Optimum estimation of mean vector for dynamic population. Invited paper in the proceeding of Internation symposium on Optimization and Statistics., held at AMU. Aligarch.

توسط مقدرات نسبة الخليط متعدد المتغيرات المعممة للسكان في أخذ العينات متعدد المراحل باستخدام خصائص متعددة المساعدة

أولوغونليكو [1] * ، بيتر إ. أوغونينكا [2] ، أولووابونمي إ. أولوغونليكو [3] ، أديمولا أ. سوديبو [3]

[1]قسم الإحصاء ، جامعة أولابيسي أونابانجو ، آغو إيوي ، نيجيريا ،[2] إحصائي مستقل ، هال ، المملكة المتحدة ،[3] جامعة إبادان ، إبادان ، نيجيريا.

الخلاصة: وقد استمر إدماج المعلومات المساعدة في أخذ عينات المسح في جذب اهتمام كبير في تحسين كفاءة المقدرين. غالبا ما تظهر مقدرات النسب والانحدار التقليدية، على الرغم من فعاليتها قيودا عند مواجهة متغيرات مساعدة متعددة وتصميمات أخذ العينات المعقدة. اقترحت هذه الدراسة فئة من مقدرات نسبة الخليط متعددة المتغيرات المعممة لتقدير متوسط السكان في تصميم أخذ العينات متعدد المراحل بخصائص متعددة المساعدة. تمتد المقدرات المقترحة إلى ما وراء الأساليب التقليدية أحادية المتغير من خلال الجمع بين المعلومات من العديد من المتغيرات المساعدة والسمات المساعدة. يتم اشتقاق الخصائص النظرية للمقدر ، بما في ذلك تعبيرات الخطأ المربع المتوسط. أكد التحليل المقارن النظري أن المقدرين المقترحين حققوا مكاسب ملحوظة في الكفاءة مقارنة بالمقدرين الذين تمت مراجعتهم. أكدت دراسات المحاكاة كذلك كفاءة المقدرات المقترحة عبر أحجام العينات المختلفة (بشكل مقارب) ، وهياكل الارتباط ، وظروف التوزيع. بشكل عام ، أكدت مقدرات نسبة الخليط متعددة المتغيرات المعممة أنها أكثر كفاءة في تقدير متوسط السكان في إطار تصميم أخذ العينات متعدد المراحل.

الكلمات المفتاحية: أخذ العينات متعدد المراحل ، تقدير متوسط السكان ، مقدرات نسبة الخليط المعممة ، المتغيرات المساعدة ، السمات المساعدة.