International Journal of Analysis and Applications

Performance Comparison of Three Ratio Estimators of the Population Ratio in Simple Random Sampling Without Replacement

Nuntida Ounrittichai[1], Patsaporn Utha[1], Boonyarit Choopradit[2], Saowapa Chaipitak[1,*]

[1]*Department of Statistics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand*

[2]*Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathumthani 12120, Thailand*

**Corresponding author:* fscispc@ku.ac.th

ABSTRACT. This study aims to compare the efficacy of three ratio estimators for estimating the population ratio in simple random sampling without replacement (SRSWOR). The estimators under consideration are a customary ratio estimator $(\widehat{R_1})$, a ratio estimator based on a transformed mean estimator $(\widehat{R_2})$ introduced by Onyeka et al. [1], and a regression-type estimator $(\widehat{R_3})$ proposed by Onyeka et al. [2]. We assess the performance of these estimators across three distributions (bivariate normal, bivariate Poisson log-normal, and bivariate Cauchy) while varying both correlation coefficients and sample sizes, utilizing Mean Square Error (MSE) and Percent Relative Efficiency (PRE) as evaluation criteria. The results indicate that for a bivariate normal distribution, the $\widehat{R_1}$ and $\widehat{R_2}$ estimators consistently outperformed the $\widehat{R_3}$ estimator across all sample sizes and correlation coefficients. The $\widehat{R_2}$ estimator demonstrated superiority with very small sample sizes, while $\widehat{R_1}$ exhibited better performance in small sample sizes. The $\widehat{R_2}$ estimator remained reliable for moderately sized samples, demonstrating consistent efficiency. In large samples, $\widehat{R_2}$ maintained its performance advantage, except in weak correlation coefficients, where $\widehat{R_1}$ proved superior. For a bivariate Poisson log-normal distribution, both $\widehat{R_2}$ and $\widehat{R_3}$ performed significantly better than $\widehat{R_1}$ for very small sample sizes, irrespective of correlation direction and strength. For moderately sized samples, $\widehat{R_2}$ and $\widehat{R_3}$ consistently excelled, with $\widehat{R_2}$ leading in cases with positive correlation coefficients. For large sample sizes with negative correlation coefficients, both $\widehat{R_2}$ and $\widehat{R_3}$ were comparable effective and significantly better than $\widehat{R_1}$. Conversely, with positive correlation coefficients, the $\widehat{R_1}$ estimator significantly outperformed both $\widehat{R_2}$ and $\widehat{R_3}$. In a bivariate Cauchy distribution, the $\widehat{R_1}$ estimator demonstrated notable and consistent superiority over the $\widehat{R_2}$ and $\widehat{R_3}$ estimators across all sample sizes and correlation coefficients.

Received May 30, 2024.

2020 Mathematics Subject Classification. 62D05.

Key words and phrases. bivariate normal; bivariate Poisson log-normal; bivariate Cauchy; population ratio.

https://doi.org/10.28924/2291-8639-22-2024-121

ISSN: 2291-8639

1. Introduction

Estimating population mean, ratio, variance, and proportion is crucial in sample surveys. The population ratio (R) is a significant parameter that plays a vital role in statistical analysis. It represents the ratio between the population means of two variables: the study variable (Y) and the auxiliary variable (X). This parameter holds significance in various scenarios, such as estimating the yield-to- plantation-area ratio, the expense-to-income ratio, the employment-to-population ratio, and more. To estimate the population ratio, statisticians often rely on a widely recognized and frequently used estimator known as the ratio estimator (Cochran, [3]). Over the past several years, many researchers have proposed alternative methods for estimating the ratio of two population means. The works of Rao [4], Singh [5], [6], Rao and Pereira [7], Tripathi [8], Ray and Singh [9], Upadhyaya and Singh [10], Upadhyaya et al. [11], Singh [12], [13], [14], Srivastava et al. [15], Okafor and Arnab [16], Khare [17], Okafor [18], Singh et al. [19], Prasad et al. [20], and Singh and Karpe [21] have contributed significantly to the development of these alternative estimators. In 2000, Upadhyaya et al. [22] presented an estimator based on the transformation of auxiliary variables. This estimator has several advantages, including cost-effectiveness, as it is derived from existing auxiliary variables. Singh and Karpe [21] investigated the intricate problem of estimating the ratio and product of two population means. This endeavor was conducted while considering supplementary information derived from an auxiliary variable within the context of prevailing measurement errors. In pursuit of more refined estimations, they introduced novel estimators tailored to this complex scenario. The empirical findings, illuminating in nature, demonstrably establish the proposed estimator's heightened efficacy compared to its precursors within the existing scholarly literature. In 2013, Onyeka et al. [1] proposed six novel estimators of the population ratio in the simple random sampling (SRS) scheme. These estimators utilize a variable transformation technique applied to the auxiliary variable. The empirical illustration accompanying the findings demonstrates that certain estimators proposed in the study exhibit significantly improved efficiency compared to the customary ratio estimator, especially for the given dataset of Johnson [23]. In the scholarly work presented by Onyeka et al. [2], a novel category of estimators is advanced to address estimating the population ratio involving the mean values of two distinct variables. This estimation endeavor was conducted within the SRS scheme, and notably, it employed a variable transformation technique applied to an auxiliary variable. The empirical outcomes of this investigation compellingly demonstrate that several estimators within the proposed class exhibit considerable enhancements in efficiency compared to the conventional ratio estimator, as applied to the specific dataset under consideration.

Prior research has centered on developing estimators for population ratios and assessing their efficacy within the context of their respective inquiries. However, these investigations have been restricted to specific populations to circumvent challenging comparisons. As such, these estimators have yet to be evaluated against those from other studies concerning data distribution. This study underscores the significance of selecting estimators attuned to population distribution, correlation magnitude and direction, and sample size. Such selection is paramount in obtaining precise inferences and predictions in various statistical scenarios. Consequently, our focus is on evaluating the efficacy of three established estimators detailed in Cochran [3], Onyeka et al. [1], and Onyeka et al. [2]. This examination involves exploring the performance of these three estimators under three distinct population distributions: the bivariate normal distribution, the bivariate Poisson log-normal distribution, and the bivariate Cauchy distribution. The effect of the magnitude and direction of the correlation coefficient between two variables on the ability to estimate population ratios of these estimates was studied.

The rest of the article is organised as follows. Section 2 provides comprehensive information regarding the three ratio estimators. Section 3 outlines the research methodology. Section 4 presents simulation results. Finally, Section 5 provides the conclusion.

2. Three Estimators for the Population Ratio

The population ratio (R) can be computed utilizing the $R = \frac{\mu_y}{\mu_x}$ formula, where $\mu_y(\mu_x)$ denotes the population mean of variable *Y(X)*, respectively, that is, $\mu_y = \frac{1}{N}\sum_{i=1}^{N} Y_i$ and $\mu_x = \frac{1}{N}\sum_{i=1}^{N} X_i$. Assume a sample of size n is selected from a population of size N under the SRSWOR scheme. Three estimators are considered for estimating the population ratio, detailed as follows.

2.1 Ratio estimator

The ratio estimator $(\widehat{R_1})$, described in Cochran's work from 1977 [3], is widely used and calculated using Equation (1)

$$\widehat{R_1} = \bar{y}/\bar{x} \tag{1}$$

where $\bar{y}(\bar{x})$ denotes the sample mean of the variables $Y(X)$, respectively, that is, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i$. The Mean Square Error (MSE) of $\widehat{R_1}$ is defined as:

$$MSE(\widehat{R_1}) = \frac{1-f}{n\mu_x^2}\sum_{i=1}^{n}\frac{(y_i - \widehat{R_1}x_i)^2}{n-1} = \frac{1-f}{n\mu_x^2}(\sigma_y^2 + R^2\sigma_x^2 + 2R\sigma_{yx}).$$

Here, $f = n/N$ represents a sampling fraction, $\sigma_y^2(\sigma_x^2)$ represents the population variance of the variables $Y(X)$, while σ_{yx} represents the population covariance between the variables $Y(X)$.

2.2 Simple ratio estimator based on the transformed sample mean

Onyeka et al. [1] proposed a simple ratio estimator based on the transformed sample mean $(\widehat{R_2})$, which is calculated using Equation (2)

$$\widehat{R_2} = \bar{y}/\bar{x}^* \tag{2}$$

where $\bar{x}^*$ represents a transformed sample mean estimator that is associated with the variable transformation and $\bar{x}^* = (1+\pi)\mu_x - \pi\bar{x},\ \pi = \frac{n}{N-n}$. The transformed auxiliary variable x_i^* is calculated based on $x_i^* = \frac{N\mu_x - nx_i}{N-n}$, i = 1, 2, …, N. The MSE of $\widehat{R_2}$ can be determined by

$$MSE(\widehat{R_2}) = \frac{1-f}{n\mu_x^2}(\sigma_y^2 + \pi^2R^2\sigma_x^2 + 2\pi R\sigma_{yx}).$$

2.3 Regression-type estimator based on both the sample mean and transformed sample mean

Onyeka et al. [2] proposed a regression-type estimator $(\widehat{R_3})$ that uses of both the sample mean and transformed sample mean in their 2014 publication. This estimator can be conveniently calculated using Equation (3).

$$\widehat{R_3} = \frac{\bar{y}}{\frac{1}{4}(\bar{x}-\bar{x}^*+\mu_x)+\frac{3}{4}(\bar{x}^*-\bar{x}+\mu_x)} \tag{3}$$

The MSE of $\widehat{R_3}$ is defined as:

$$MSE(\widehat{R_3}) = \frac{1-f}{n\mu_x^2}\left(\sigma_y^2 + \frac{1}{4}(1+\pi)^2R^2\sigma_x^2 + (1+\pi)R\sigma_{yx}\right).$$

3. Methodology

The research methodology can be outlined as follows:

3.1 Simulate population data.

The present study utilized the R programming language version 3.3.3 to generate data. The simulated population consisted of 1,000 individuals and comprised two variables: the study variable (Y) and an auxiliary variable (X). The study population was characterized by bivariate normal distribution, bivariate Poisson log-normal distribution, and bivariate Cauchy distribution. Both variables were assigned a population mean and variance of 200 and one, respectively. The population correlation coefficient (ρ) between the two variables was designated as either positive (indicating a similar relationship) or negative (indicating an opposite relationship). The correlation strength was defined at three levels: weak (0.25), moderate (0.50), and strong (0.75).

3.2 Calculate the population ratio value.

3.3 Random 1,000 samples from the population.

A sample of size n was randomly drawn using SRSWOR with each sample size repeated 1,000 times. The sample sizes were classified into four categories: very small (10), small (15), moderate (30), and large (100).

3.4 Calculate the population ratio estimates.

Calculate the ratio estimates obtained by the three ratio estimators: $\widehat{R_1}, \widehat{R_2}$ and $\widehat{R_3}$, as represented in Equations (1) to (3), respectively.

3.5 Calculate the estimated MSE and PRE.

This study uses the MSE and the Prediction Error (PRE) as criteria to evaluate the performance of different estimators and compare them. A reduced MSE value and an elevated PRE indicate an estimator's superior performance. To estimate the MSE of parameter estimates, we computed the sum of squared error estimates for every 1,000 iterations. Then, we divided the sum by the total number of iterations, also 1,000. Accordingly, the formula for estimating the MSE of parameter estimates is formulated as follows:

$$MSE(\hat{\theta}) = \frac{1}{1,000} \sum_{i=1}^{1,000} \left(\hat{\theta}_i - \theta\right)^2$$

where $\hat{\theta}_i$ represents the parameter estimate of the i-th iteration, and θ represents the parameter value. Furthermore, the PRE of the estimator $\hat{\theta}_i$ over $\hat{\theta}_j$ is computed using the formula:

$$PRE(\hat{\theta}_i) = MSE(\hat{\theta}_j)/MSE(\hat{\theta}_i).$$

4. Simulation Results

A simulation was conducted on a population comprising two variables, Y and X. The study involved the application of three distributions: the bivariate normal distribution, the bivariate Poisson log-normal distribution, and the bivariate Cauchy distribution. The correlation coefficients were varied at ± 0.25, ± 0.50, and ± 0.75 during the analysis. Samples of sizes 10, 15, 30, and 100 were randomly chosen using the SRSWOR scheme. The results of the estimated MSE and PRE of the estimators, compared to the classical ratio estimator $\widehat{R_1}$, considering populations following bivariate normal, bivariate Poisson log-normal, and bivariate Cauchy distributions, are presented in Tables 1 to 3, respectively.

In Table 1, it is evident that in the context of a bivariate normal distribution, when analysing scenarios with negative correlation coefficients, the $\widehat{R_2}$ estimator consistently outperforms the $\widehat{R_1}$ estimator regarding estimated MSE and PRE. This superior performance is observed across a wide range of sample sizes and correlation coefficients, with the exception being scenarios with weak correlation between variables ($\rho = -0.25$) in small (n = 15) to very small (n = 10) sample sizes, where the $\widehat{R_1}$ estimator exhibited better results. The estimated MSE indicates a discernible trend where the $\widehat{R_1}$ and $\widehat{R_2}$ estimators consistently outperform $\widehat{R_3}$, irrespective of sample size or correlation coefficient. In terms of PRE, the PRE estimates reveal an interesting pattern in which the $\widehat{R_2}$ estimator consistently displays a heightened level of efficiency relative to the others. Moreover, this efficiency tends to increase as the sample size increases. This observation underscores the critical role of sample size in the choice of estimator, particularly when considering the population correlation coefficient. For instance, when confronted with larger sample sizes, the $\widehat{R_2}$ estimator emerges as a favourable choice. Turning our attention to scenarios characterized by positive correlation coefficients, it remains evident that the $\widehat{R_2}$ estimator consistently outperforms $\widehat{R_1}$ regarding estimated MSE and PRE across nearly all sample sizes and correlation coefficients. Once again, the MSE estimates reinforce the superior performance of estimators $\widehat{R_1}$ and $\widehat{R_2}$ while concurrently highlighting the inferior accuracy of the $\widehat{R_3}$ estimator across all sample sizes and correlation coefficients. This underscores the limited accuracy of the estimator $\widehat{R_3}$ within this analytical framework. Additionally, the PRE estimates consistently depict $\widehat{R_2}$ as the more efficient estimator, with efficiency generally increasing as both the sample size and correlation coefficient rise. Among the three estimators, the $\widehat{R_3}$ estimator consistently holds the lowest position in terms of PRE, indicating that it is the least efficient estimator in this context. Consequently, the choice of estimator remains contingent upon the interplay between sample size and correlation coefficient. However, mostly $\widehat{R_2}$ emerges as the most dependable and efficient estimators within the framework of a bivariate normal distribution, except for instances where there is a weak correlation ($\rho = 0.25$) in both small (n = 15) and large (n = 100) sample sizes, and situations involving a moderate correlation ($\rho = 0.50$) specifically in large sample sizes.

		MSE			PRE		
n	ρ	$\widehat{R_1}$	$\widehat{R_2}$	$\widehat{R_3}$	$\widehat{R_1}$	$\widehat{R_2}$	$\widehat{R_3}$
10	-0.75	0.00004	0.00003	0.94733	100	133.3333	0.0042
	-0.50	0.00006	0.00004	0.94744	100	150.0000	0.0063
	-0.25	0.00002	0.00004	0.94746	100	50.0000	0.0021
	0.25	0.00007	0.00006	0.94732	100	116.6667	0.0074
	0.50	0.00003	0.00002	0.94734	100	150.0000	0.0032
	0.75	0.00003	0.00001	0.94751	100	300.0000	0.0032
15	-0.75	0.00005	0.00009	0.94746	100	55.5556	0.0053
	-0.50	0.00004	0.00008	0.94744	100	50.0000	0.0042
	-0.25	0.00002	0.00009	0.94747	100	22.2222	0.0021
	0.25	0.00001	0.00006	0.94744	100	16.6667	0.0011
	0.50	0.00003	0.00008	0.94748	100	37.5000	0.0032
	0.75	0.00008	0.00005	0.94743	100	160.0000	0.0084
30	-0.75	0.00007	0.00004	0.94749	100	175.0000	0.0074
	-0.50	0.00006	0.00005	0.94745	100	120.0000	0.0063
	-0.25	0.00008	0.00003	0.94742	100	266.6667	0.0084
	0.25	0.00005	0.00009	0.94746	100	55.5556	0.0053
	0.50	0.00006	0.00003	0.94741	100	200.0000	0.0063
	0.75	0.00008	0.00007	0.94744	100	114.2857	0.0084
100	-0.75	0.00009	0.00005	0.94746	100	180.0000	0.0095
	-0.50	0.00008	0.00003	0.94747	100	266.6667	0.0084
	-0.25	0.00003	0.00008	0.94748	100	37.5000	0.0032
	0.25	0.00005	0.00009	0.94745	100	55.5556	0.0053
	0.50	0.00007	0.00005	0.94745	100	140.0000	0.0074
	0.75	0.00006	0.00001	0.94743	100	600.0000	0.0063

Table 1 MSE and PRE estimates of the estimators $\widehat{R_1}, \widehat{R_2}$, and $\widehat{R_3}$ over the $\widehat{R_1}$ estimator when the population has a bivariate normal distribution with different sample sizes (n) and different correlation coefficients (ρ)

Furthermore, the three estimators have been subject to evaluation across varying sample sizes and distinct correlation coefficients within the context of a bivariate Poisson log-normal distribution, as presented in Table 2. It is noteworthy that, particularly when exploring scenarios featuring a negative correlation coefficient, the MSE estimates of the $\widehat{R_2}$ and $\widehat{R_3}$ estimators consistently exhibit superior performance when compared to the $\widehat{R_1}$ estimator. This phenomenon holds across the entire spectrum of sample sizes and correlation coefficients emphasizing instances where ρ = –0.75 or when n = 30 and 100. Additionally, the PRE values exhibit variability contingent upon the sample size and the correlation coefficient. Conversely, in scenarios characterized by ρ = –0.75 and larger sample sizes, $\widehat{R_2}$ and $\widehat{R_3}$ consistently outperform $\widehat{R_1}$. These findings collectively imply that, as a general trend, $\widehat{R_2}$ and $\widehat{R_3}$ tend to serve as the most accurate and efficient estimators within this context. However, it is imperative to exercise prudence in selecting the appropriate estimator, as the choice should be contingent upon the specific interplay of sample size and correlation coefficient. Shifting our attention to positive correlation coefficients, it is worth mentioning that when the correlation coefficient is 0.75, the $\widehat{R_1}$ estimator outperforms $\widehat{R_2}$ and $\widehat{R_3}$ in terms of MSE and PRE across various sample sizes, except for scenarios with very small sample sizes (n = 10). Within the context of a correlation coefficient of 0.5, an empirical examination reveals that the $\widehat{R_1}$ estimator demonstrates superior performance when compared to $\widehat{R_2}$ and $\widehat{R_3}$ under circumstances characterized by very small sample sizes (n = 10) and considerably large sample sizes (n = 100). When, when sample sizes are small (n = 15) or moderate (n = 30), the $\widehat{R_2}$ and $\widehat{R_3}$ estimators are more effective. However, when ρ is 0.25, $\widehat{R_1}$ consistently outperforms both $\widehat{R_2}$ and $\widehat{R_3}$ in situations with both small sample sizes (n = 15) and large sample sizes (n = 100). Interestingly, the $\widehat{R_2}$ and $\widehat{R_3}$ estimators perform better when faced with very small samples (n = 10) or moderate samples (n = 30).

		MSE			PRE		
n	ρ	$\widehat{R_1}$	$\widehat{R_2}$	$\widehat{R_3}$	$\widehat{R_1}$	$\widehat{R_2}$	$\widehat{R_3}$
10	-0.75	1.4697	1.0013	1.0000	100	146.78	146.97
	-0.50	0.8673	0.9992	1.0000	100	86.80	86.73
	-0.25	3.4834	0.8552	0.9476	100	407.32	367.60
	0.25	1.1944	1.0034	1.0001	100	119.04	119.43
	0.50	0.0018	0.993	0.9998	100	0.18	0.18
	0.75	2.5756	1.0027	1.0001	100	256.87	257.53
15	-0.75	0.0074	1.0062	1.0002	100	0.74	0.74
	-0.50	13.2038	1.0285	1.0008	100	1,283.79	1,319.32
	-0.25	3.2057	0.9871	0.9996	100	324.76	320.70
	0.25	0.0381	0.9946	0.9999	100	3.83	3.81
	0.50	1.4759	0.8648	0.9962	100	170.66	148.15
	0.75	0.0377	1.7822	1.0183	100	2.12	3.70
30	-0.75	4.0998	1.0116	1.0003	100	405.28	409.86
	-0.50	3.602	1.0181	1.0005	100	353.80	360.02
	-0.25	2.7775	1.0204	1.0006	100	272.20	277.58
	0.25	4.1961	1.0062	1.0002	100	417.02	419.53
	0.50	2,011.27	0.9859	0.9996	100	204,003.14	201,207.18
	0.75	0.0109	1.0072	1.0002	100	1.08	1.09
100	-0.75	577.9088	1.0148	1.0004	100	56,948.05	57,767.77
	-0.50	1.4657	1.004	1.0001	100	145.99	146.56
	-0.25	5.4544	0.952	0.9986	100	572.94	546.20
	0.25	0.2101	0.9827	0.9995	100	21.38	21.02
	0.50	0.4307	1.0363	1.0011	100	41.56	43.02
	0.75	0.2473	1.0112	1.0003	100	24.46	24.72

Table 2 MSE and PRE estimates of the estimators $\widehat{R_1}, \widehat{R_2}$, and $\widehat{R_3}$ over the $\widehat{R_1}$ estimator when the population has a bivariate Poisson log-normal distribution with different sample sizes (n) and different correlation coefficients (ρ)

n	ρ	MSE $\widehat{R_1}$	MSE $\widehat{R_2}$	MSE $\widehat{R_3}$	PRE $\widehat{R_1}$	PRE $\widehat{R_2}$	PRE $\widehat{R_3}$
10	-0.75	0.4331	195.7600	57.1220	100	0.22	0.76
	-0.50	0.8710	1549.0000	169.7100	100	0.06	0.51
	-0.25	0.0013	1171.4400	25.1660	100	0.00	0.01
	0.25	0.1148	57.9840	12.5130	100	0.20	0.92
	0.50	0.1896	85.5070	11.7720	100	0.22	1.61
	0.75	0.5615	1373.1000	141.0700	100	0.04	0.40
15	-0.75	0.0875	400.8000	36.4510	100	0.02	0.24
	-0.50	0.1532	166857.0000	15.6550	100	0.00	0.98
	-0.25	0.3928	226.6600	39.8050	100	0.17	0.99
	0.25	0.1555	1610.2000	56.4380	100	0.01	0.28
	0.50	0.1438	197.6900	30.0410	100	0.07	0.48
	0.75	0.1251	414.9800	29.7730	100	0.03	0.42
30	-0.75	0.1913	299.7000	42.0570	100	0.06	0.45
	-0.50	0.1275	183.6800	44.0310	100	0.07	0.29
	-0.25	0.2122	9.5413	18.8010	100	2.22	1.13
	0.25	0.0544	85.0080	52.4910	100	0.06	0.10
	0.50	0.0007	69.3170	32.1760	100	0.00	0.00
	0.75	0.0160	79.3030	40.2830	100	0.02	0.04
100	-0.75	0.0147	20.8391	32.1110	100	0.07	0.05
	-0.50	0.0232	20.8542	48.0810	100	0.11	0.05
	-0.25	0.0279	34.4529	49.7480	100	0.08	0.06
	0.25	0.2032	14.4208	42.8850	100	1.41	0.47
	0.50	0.0448	17.5957	33.1070	100	0.25	0.14
	0.75	0.0461	18.2860	27.8550	100	0.25	0.17

Table 3 MSE and PRE estimates of the estimators $\widehat{R_1}, \widehat{R_2}$, and $\widehat{R_3}$ over the $\widehat{R_1}$ estimator when the population has a bivariate Cauchy distribution with different sample sizes (n) and different correlation coefficients (ρ)

Finally, an empirical analysis of populations has been conducted to examine bivariate Cauchy distributions. This analysis reveals clear patterns regarding the characteristics of MSE and PRE displayed in Table 3. For negative correlation coefficients, the $\widehat{R_1}$ estimator exhibits superior performance across all sample sizes. Notably, when dealing with relatively small to moderate sample sizes (specifically, sample sizes of 10, 15, and 30), the $\widehat{R_3}$ estimator exhibits a statistically significant performance advantage over $\widehat{R_2}$. In contrast, this trend is reversed for larger sample sizes, wherein the $\widehat{R_2}$ estimator exhibits superior performance compared to $\widehat{R_3}$. This pattern signifies a size-dependent relationship between the two estimators, contingent upon the magnitude of the sample size. Shifting focus to scenarios characterized by positive correlation coefficient values, the $\widehat{R_1}$ estimator once again surfaces as the most adept choice, displaying comparable performance levels. This tandem outperforms estimators $\widehat{R_2}$ and $\widehat{R_3}$, with statistical significance underpinning the observed discrepancies. However, mirroring the dynamics observed in positive ρ values, the $\widehat{R_3}$ estimator surpasses the $\widehat{R_2}$ estimator in scenarios characterized by relatively small to moderate sample sizes. This phenomenon highlights the sensitivity of estimator performance to sample size, consistent with prior findings. Ultimately, in the endeavor to focus our scholarly scrutiny upon the evaluation of PRE, it becomes evident that the resultant findings bear a remarkable semblance to the antecedently elucidated outcomes delineated within the purview of MSE.

For convenience, Table 4 summarizes the most effective estimator according to population distribution, sample size, and correlation coefficient.

n	ρ	Distribution		
		Bivariate Normal	Bivariate Poisson Log- Normal	Bivariate Cauchy
10	-0.75	$\widehat{R_2}$	$\widehat{R_2}$, $\widehat{R_3}$	$\widehat{R_1}$
	-0.50	$\widehat{R_2}$	$\widehat{R_1}$	$\widehat{R_1}$
	-0.25	$\widehat{R_1}$	$\widehat{R_2}$ (1), $\widehat{R_3}$ (2)	$\widehat{R_1}$
	0.25	$\widehat{R_2}$	$\widehat{R_2}$, $\widehat{R_3}$	$\widehat{R_1}$
	0.50	$\widehat{R_2}$	$\widehat{R_1}$	$\widehat{R_1}$
	0.75	$\widehat{R_2}$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
15	-0.75	$\widehat{R_1}$	$\widehat{R_1}$	$\widehat{R_1}$
	-0.50	$\widehat{R_1}$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	-0.25	$\widehat{R_1}$	$\widehat{R_2}$ (1), $\widehat{R_3}$ (2)	$\widehat{R_1}$
	0.25	$\widehat{R_1}$	$\widehat{R_1}$	$\widehat{R_1}$
	0.50	$\widehat{R_1}$	$\widehat{R_2}$ (1), $\widehat{R_3}$ (2)	$\widehat{R_1}$
	0.75	$\widehat{R_2}$	$\widehat{R_1}$	$\widehat{R_1}$
30	-0.75	$\hat{R}_2$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	-0.50	$\hat{R}_2$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	-0.25	$\hat{R}_2$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	0.25	$\widehat{R_1}$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	0.50	$\hat{R}_2$	$\widehat{R_2}$ (1), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	0.75	$\hat{R}_2$	$\widehat{R_1}$	$\widehat{R_1}$
100	-0.75	$\widehat{R_2}$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	-0.50	$\widehat{R_2}$	$\widehat{R_2}$ (2), $\widehat{R_3}$ (1)	$\widehat{R_1}$
	-0.25	$\widehat{R_1}$	$\widehat{R_2}$ (1), $\widehat{R_3}$ (2)	$\widehat{R_1}$
	0.25	$\widehat{R_1}$	$\widehat{R_1}$	$\widehat{R_1}$
	0.50	$\widehat{R_2}$	$\widehat{R_1}$	$\widehat{R_1}$
	0.75	$\widehat{R_2}$	$\widehat{R_1}$	$\widehat{R_1}$

Note: The number within parentheses indicates the order of an appropriate estimator.

Table 4 The most effective estimator categorized according to population distribution, sample size, and correlation coefficient

5. Conclusion

In this study, we conducted a simulation to compare three estimators for estimating the population ratio in SRSWOR. The three estimators are the traditional ratio estimator $\widehat{R_1}$ defined by Cochran [3]; the simple ratio estimator $\widehat{R_2}$, which is based on a transformed mean estimator introduced by Onyeka et al. [1]; and the regression-type estimator $\widehat{R_3}$ proposed by Onyeka et al.

[2]. Our research aimed to determine the effectiveness of these estimators across different sample sizes and correlation coefficients for populations following bivariate normal, bivariate Poisson log-normal, and bivariate Cauchy distributions. The study revealed that when the population follows a bivariate normal distribution, the ratio estimators $\widehat{R_1}$ and $\widehat{R_2}$ consistently outperformed the regression-type estimator $\widehat{R_3}$ across all sample sizes and correlation coefficients. In situations with very small sample sizes (n = 10), and for most correlation coefficients examined, the performance of the $\widehat{R_2}$ estimator was significantly superior to that of $\widehat{R_1}$. In small sample sizes (n = 15), the $\widehat{R_1}$ estimator outperformed $\widehat{R_2}$ for most correlation coefficients. In moderately sized samples (n = 30), the $\widehat{R_2}$ estimator outperformed $\widehat{R_1}$ for most correlation coefficients. Finally, with large sample sizes (n = 100) and moderate to strong correlation coefficients in both positive and negative directions (ρ = ±0.75, ±0.50), the $\widehat{R_2}$ estimator demonstrated significantly better performance than $\widehat{R_1}$, however $\widehat{R_1}$ outperformed $\widehat{R_2}$ significantly in cases of weak correlation coefficients (ρ = ±0.25).

In cases where the underlying population follows a bivariate Poisson log-normal distribution, for very small sample sizes (n = 10), the $\widehat{R_2}$ and $\widehat{R_3}$ estimators showed equal and significantly better performance than $\widehat{R_1}$ for weak and strong correlation coefficients, regardless of their directions. When strong correlation coefficients in both positive and negative directions (ρ = ±0.75) were present for small sample sizes (n = 15), the $\widehat{R_1}$ estimator outperformed $\widehat{R_2}$ and $\widehat{R_3}$. However, in most cases with weak to moderate correlation coefficients (ρ = ±0.25, ±0.50), the $\widehat{R_2}$ and $\widehat{R_3}$ estimators were equally effective and significantly better than the $\widehat{R_1}$ estimator. For moderately sized samples (n = 30), the $\widehat{R_2}$ and $\widehat{R_3}$ estimators were equally effective and significantly superior to $\widehat{R_1}$ for the majority of examined correlation coefficients. In the cases of large sample sizes (n = 100) with negative correlation coefficients, both $\widehat{R_2}$ and $\widehat{R_3}$ were equally effective and significantly better than $\widehat{R_1}$. However, with positive correlation coefficients, the $\widehat{R_1}$ estimator significantly outperformed both the $\widehat{R_2}$ and $\widehat{R_3}$ estimators. In populations with a bivariate Cauchy distribution, the $\widehat{R_1}$ estimator shows robust performance compared to the $\widehat{R_2}$ and $\widehat{R_3}$ estimators, across all sample sizes and correlation coefficients examined.

The authors suggest further studies on populations with different distributions, such as bivariate gamma distributions. The performance of the estimators can also be compared to other distribution conditions, such as coefficient of variation, kurtosis, skewness, and additional ratio estimators, as well as by increasing the correlation coefficient level. For example, exploring scenarios where the correlation coefficient approaches one or zero or investigating ratio estimators derived from alternative sampling methods like stratified or successive sampling could be valuable areas for future research.

Acknowledgments: This work was supported by the International SciKU Branding (ISB), Faculty of Science, Kasetsart University, and the Department of Statistics, Faculty of Science, Kasetsart University, Thailand.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] A. C. Onyeka, V. U. Nlebedim, and C. H. Izunobi, Estimation of Population Ratio in Simple Random Sampling using Variable Transformation, Glob. J. Sci. Front. Res. 13 (2013), 57–65.

[2] A.C. Onyeka, V.U. Nlebedim, C.H. Izunobi, A Class of Estimators for Population Ratio in Simple Random Sampling Using Variable Transformation, Open J. Stat. 04 (2014), 284–291. https://doi.org/10.4236/ojs.2014.44029.

[3] W.G. Cochran, Sampling Techniques, 3rd ed., John Wiley & Sons, New York, 1977.

[4] J.N.K. Rao, Double Ratio Estimate in Forest Surveys, J. Indian Soc. Agric. Stat. 9 (1957), 191–204.

[5] M.P. Singh, On the Estimation of Ratio and Product of the Population Parameters, Sankhyā Ser. B. 27 (1965), 321–328. http://www.jstor.org/stable/25051553.

[6] M.P. Singh, Ratio Cum Product Method of Estimation, Metrika 12 (1967), 34–42. https://doi.org/10.1007/bf02613481.

[7] J.N.K. Rao, N.P. Pereira, On Double Ratio Estimators, Sankhyā Ser. A. 30 (1968), 83–90. http://www.jstor.org/stable/25049511.

[8] T.P. Tripathi, A General Class of Estimators for Population Ratio, Sankhyā Ser. C. 42 (1980), 63–75.

[9] S.K. Ray, R.K. Singh, Some Estimators for the Ratio and Product of Population Parameters, J. Indian Soc. Agric. Stat. 37 (1985), 1–10.

[10] L.N. Upadhyaya, H.P. Singh, A Class of Estimators using Auxiliary Information for Estimating Ratio of Two Finite Population Means, Gujarat Stat. Rev. 12 (1985), 7–16.

[11] L.N. Upadhyaya, H.P. Singh, J.W.E. Vos, On the Estimation of Population Means and Ratios Using Supplementary Information, Stat. Neerl. 39 (1985), 309–318. https://doi.org/10.1111/j.1467-9574.1985.tb01149.x.

[12] R.K. Singh, On Estimating Ratio and Product of Population Parameters, Calcutta Stat. Assoc. Bull. 31 (1982), 69–76. https://doi.org/10.1177/0008068319820106.

[13] H.P. Singh, Estimation of Ratio, Product and Mean using Auxiliary Information in Sample Surveys, Aligarh J. Stat. 6 (1986), 32–44.

[14] H. P. Singh, On the Estimation of Ratio and Product of Two Finite Population Means, Proc. Nat. Acad. Sci. India. Sec. A, 58 (1988), 399–403.

[15] S.R. Srivastava, B.B. Khare, S.R. Srivastava, On Generalized Chain Estimator for Ratio and Product of Two Population Means using Auxiliary Characters, Assam Stat. Rev. 2 (1988), 21–29.

[16] F.C. Okafor, R. Arnab, Some Strategies of Two-Stage Sampling for Estimating Population Ratios Over two Occasions, Aust. J. Stat. 29 (1987), 128–142. https://doi.org/10.1111/j.1467-842x.1987.tb00729.x.

[17] B. B. Khare, On Generalized Class of Estimators for Ratio of Two Population Means using Multiauxiliary Characters, Aligarh J. Stat. 11 (1991), 81–90.

[18] F.C. Okafor, The Theory and Application of Sampling over Two Occasions for the Estimation of Current Population Ratio, Statistica, 52 (1992), 137–147. https://doi.org/10.6092/issn.1973-2201/897.

[19] V.K. Singh, H.P. Singh, H.P. Singh, D. Shukla, A General Class of Chain Estimators for Ratio and Product of Two Means of a Finite Population, Comm. Stat. – Theory Methods 23 (1994), 1341–1355. https://doi.org/10.1080/03610929408831325.

[20] B. Prasad, R.S. Singh, H.P. Singh, Some Chain Ratio-type Estimators for Ratio of Two Population Means using Two Auxiliary Characters in Two Phase Sampling, Metron, 54 (1996), 95–113.

[21] H.P. Singh, N. Karpe, On the Estimation of Ratio and Product of Two Population Means using Supplementary Information in Presence, Statistica, 69 (2009), 27–47. https://doi.org/10.6092/issn.1973-2201/3546.

[22] L.N. Upadhyaya, G.N. Singh, H.P. Singh, Use of Transformed Auxiliary Variable in the Estimation of Population Ratio in Sample Survey, Stat. Transit. 4 (2000), 1019–1027.

[23] J. Johnston, Econometric Methods. 2nd ed. McGraw-Hill Book Company, New York, 1972.