RATIO ESTIMATORS USING CHARACTERISTICS OF POISSON DISTRIBUTION WITH APPLICATION TO EARTHQUAKE DATA

Gamze Özel

Abstract

Natural populations in biology, genetics, education, engineering, insurance, marketing, seismology, social science, and survival analysis are extremely large; consequently, sampling methods have to be conducted for characterizing those populations. Ratio estimators are commonly used to obtain more efficient estimates for the population mean if the study variable is highly correlated with the auxiliary variable. It is well known that the use of the population information of auxiliary variable x improves the precision of the estimate(s) of the parameter(s). Ratio estimators are based on a sample whose distribution is not considered. However, there are situations in which Poisson distributed population may be appropriate. This paper proposes generalized class of ratio estimators from Poisson distributed population. The mean square error (MSE) equations of proposed estimators are compared in application with usual ratio estimator. By these comparisons, we find that ratio estimators using Poisson distribution characteristics as auxiliary variable information is better than usual ratio estimators. The conditions are also found that proposed estimators are more efficient. The findings are supported by numerical illustration with earthquake data of Turkey.

Key words: Ratio-type estimators; Simple random sampling; Mean square error; Poisson distribution; Efficiency.

JEL Code: C02, C60, C80

Introduction

Simple random sampling (SRS) from a finite population has attracted much of the researchers and practitioners working in surveys. Ratio estimators are commonly used in the SRS to obtain more efficient estimates for the population mean if the study variable is highly correlated with the auxiliary variable. It is well known that the use of the population information of auxiliary variable x improves the precision of the estimate(s) of the parameter(s) in the SRS. Several authors including Sisodia and Dwivedi (1981), Upadhyaya and Singh (1999), Kadilar and Cingi (2004), Gupta and Shabbir (2007, 2008), Koyuncu and Kadilar (2009), Singh and Vishwakarma (2010), Shabbir and Gupta (2011), obtained a large number of improved ratio estimators/classes of estimators for the population mean $\overline{Y}$ of the study variable y using auxiliary variable information in the SRS. The problem of estimating the population mean or total in the presence of an auxiliary variable has been widely discussed in the SRS without considering the distribution. However, the Poisson distribution is generally used for the natural populations to express the probability of a given number of rare events and there has been no effort devoted to the development of ratio estimators for a Poisson distributed population. Non-existence of the ratio estimators for the Poisson distributed sample obstacles usage of them in sampling theory itself and its applications (Ozel and Inal, 2008). The aim of this study is to derive new ratio estimators for the population mean from a Poisson distributed population. We also examine the behavior of the estimators of mean for the ratio estimators in the SRS. The earthquake data is used for the numerical example since earthquakes are rare events and generally follows a Poisson distribution (Ozel, 2011a).

1 Suggested Estimators for the Poisson Distributed Population

Consider a finite population $U=(u_1,u_2,...,u_N)$ consisting of N identifiable and distinct units. Let y and x, respectively, be the study and auxiliary variables associated with each unit u_j $(j=1,2,...,N)$ of the population. Assume that X's are known units and Y's are unknown units for all the population. Suppose that a sample of size n is selected according to the SRS. The nature of the sampling distribution depends on the nature of the population from which the random sample is drawn. Let us assume that the parent population has a Poisson distribution. This means that the random samples which are drawn from a Poisson distributed population follow also a Poisson distribution. Then, let us select the observations (y_i,x_i), $i=1,2,....,n$ from a Poisson distributed population. By this way, we suggest that the following a generalized class of ratio estimators for the population mean of the study variable from Poisson distributed population as $\eta=\frac{\overline{y}_{po}}{a\overline{x}_{po}+b}(a\overline{X}+b)=\hat{R}_{po}\overline{X}_{po}$ where $a\overline{x}_{po}+b\neq 0$, $\hat{R}_{po}=\frac{\overline{y}_{po}}{a\overline{x}_{po}+b}$, $\overline{X}_{po}=a\overline{X}+b$. Here, $\overline{y}_{po}=\sum_{i=1}^{n}y_i/n$, $\overline{x}_{po}=\sum_{i=1}^{n}x_i/n$ are the sample means of the study and auxiliary variables from Poisson distributed population, respectively. Note that (a,b) are either constants or function of known parameters of the population such as $a=1$ or C_x, $\beta_1(x)$, $\beta_2(x)$, $\rho_{x_{po}y_{po}}$, and D_x. As mentioned before, the index of dispersion is used for the first time in sampling theory for the auxiliary variable information. Let the auxiliary variable x has a Poisson distribution with parameter $\lambda_1 > 0$, then $\overline{X}$, S_x, C_x, D_x, $\beta_1(x)$, and $\beta_2(x)$ of the auxiliary variable x are given by $\overline{X} = \lambda_1$, $S_x = \sqrt{\lambda_1}$, $C_x = \frac{S_X}{\overline{X}} = \frac{1}{\sqrt{\lambda_1}}$, $D_x = \frac{\mu_{02}}{\overline{X}} = \frac{S_x^2}{\overline{X}} = \frac{\lambda_1}{\lambda_1} = 1$, $\beta_1(x) = \frac{\mu_{03}}{\mu_{02}^{3/2}} = \frac{\lambda_1}{\lambda_1^{3/2}} = \frac{1}{\sqrt{\lambda_1}}$, $\beta_2(x) = \frac{\mu_{04}}{\mu_{02}^2} = \frac{\lambda_1(1+3\lambda_1)}{\lambda_1^2} = \frac{(1+3\lambda_1)}{\lambda_1}$ respectively, where, $\mu_{02} = N^{-1}\sum_{i=1}^{N}(x_i - \overline{X})^2 = \lambda_1$, $\mu_{03} = N^{-1}\sum_{i=1}^{N}(x_i - \overline{X})^3 = \lambda_1$, and $\mu_{04} = N^{-1}\sum_{i=1}^{N}(x_i - \overline{X})^4 = \lambda_1(1+3\lambda_1)$ for the Poisson distributed population. Let the study variable y has a Poisson distribution with parameter λ_2 and let the auxiliary variable has a Poisson distribution with paramater λ_1, then the coefficient of correlation between the study variable y and auxiliary variable x is obtained by the trivariate reduction method (Ozel, 2011b). A bivariate Poisson distribution of y and x is generated by setting $x_i = m_i + z_i$ and $y_i = w_i + z_i$, $i = 1, 2, ..., n$. Assuming that the parameters of m, w, and z are γ_1, γ_2, and γ_3, the coefficient of the correlation between y and x equals

$$\rho_{x_{Po}y_{po}} = \frac{\operatorname{cov}(x,y)}{\sqrt{S_x^2 S_y^2}} = \frac{[(\gamma_1+\gamma_3)(\gamma_2+\gamma_3)+\gamma_3] - (\gamma_1+\gamma_3)(\gamma_2+\gamma_3)}{\sqrt{(\gamma_1+\gamma_3)(\gamma_2+\gamma_3)}} = \frac{\gamma_3}{\sqrt{(\gamma_1+\gamma_3)(\gamma_2+\gamma_3)}} \qquad (1)$$

where $E(x,y) = [(\gamma_1+\gamma_3)(\gamma_2+\gamma_3)+\gamma_3]$, $E(x) = (\gamma_1+\gamma_3)$, $E(y) = (\gamma_2+\gamma_3)$, $S_x^2 = \lambda_1 = (\gamma_1+\gamma_3)$ and $S_y^2 = \lambda_2 = (\gamma_2+\gamma_3)$. An obvious properties of $\rho_{x_{po}y_{po}}$ is that the correlation is restricted to be strictly positive since γ_1, γ_2, and $\gamma_3 > 0$. Since we select the observations (y_i, x_i), $i = 1, 2, ..., n$ from a Poisson distributed population with parameters λ_1 and λ_2, then we get $\overline{x}_{po} = \sum_{i=1}^{n}(m_i + z_i)/n$ and $\overline{y}_{po} = \sum_{i=1}^{n}(w_i + z_i)/n$. The covariance of $\overline{x}_{po}$ and $\overline{y}_{po}$ is

$$\operatorname{cov}(\overline{x}_{po}, \overline{y}_{po}) = \left[\frac{(\gamma_1+\gamma_3)}{n}\frac{(\gamma_2+\gamma_3)}{n} + \frac{\gamma_3}{n}\right] - \frac{(\gamma_1+\gamma_3)}{n}\frac{(\gamma_2+\gamma_3)}{n} = \frac{\gamma_3}{n} \qquad (2)$$

where $E(\overline{x}_{po}) = \frac{\lambda_1}{n} = \frac{(\gamma_1+\gamma_3)}{n}$ and $E(\overline{y}_{po}) = \frac{\lambda_2}{n} = \frac{(\gamma_2+\gamma_3)}{n}$. The MSE of the proposed estimator can be found using Taylor series method defined as

$$g(\bar{x},\bar{y}) \cong g(\bar{X},\bar{Y}) + \left.\frac{\partial g(c,d)}{\partial c}\right|_{\bar{X},\bar{Y}} (\bar{x}-\bar{X}) + \left.\frac{\partial g(c,d)}{\partial d}\right|_{\bar{X},\bar{Y}} (\bar{y}-\bar{Y}) \tag{3}$$

where $g(\bar{x},\bar{y}) = \hat{R}_{po}$ and $g(\bar{X},\bar{Y}) = R_{po}$. Eq. (3) can be applied to the generalized class of ratio estimators in order to obtain the MSE equation and we have

$$\mathrm{MSE}(\hat{R}_{po}) \cong \frac{1}{(a\bar{X}+b)^2}\left[a^2R_{po}^2 V(\bar{x}_{po}) - 2aR_{po}\,\mathrm{cov}(\bar{x}_{po},\bar{y}_{po}) + V(\bar{y}_{po})\right] \tag{4}$$

$$\cong \frac{1}{n(a\bar{X}+b)^2}\left[a^2R_{po}^2(\gamma_1+\gamma_3) - 2aR_{po}\gamma_3 + (\gamma_2+\gamma_3)\right]$$

where $R_{po} = \dfrac{\bar{Y}}{a\bar{X}+b}$. Thus, we obtain the MSE of the proposed estimator as

$$\mathrm{MSE}(\eta) \cong (a\bar{X}+b)^2\mathrm{MSE}(\hat{R}_{po}) \cong n^{-1}\left[a^2R_{po}^2(\gamma_1+\gamma_3) - 2aR_{po}\gamma_3 + (\gamma_2+\gamma_3)\right]$$

Since $(\gamma_1+\gamma_3) = \lambda_1$, $(\gamma_2+\gamma_3) = \lambda_2$, and $\gamma_3 = \rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}$, we have

$$\mathrm{MSE}(\eta) \cong n^{-1}\left[a^2R_{po}^2\lambda_1 - 2aR_{po}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2} + \lambda_2\right]. \tag{5}$$

3 Efficiency Comparisons

The ratio estimators presented in Table 1 will be compared with each other according to their MSE equations in the theory.

Table 1 Some members of the class of ratio estimator for t and η

Usual Ratio Estimators	a	b	Ratio Estimators	a	b
$t_1 = \bar{y}\dfrac{\bar{X}}{\bar{x}}$	1	0	$\eta_1 = \bar{y}_{po}\dfrac{\bar{X}_{po}}{\bar{x}_{po}}$	1	0
$t_2 = \bar{y}\dfrac{\bar{X}+C_x}{\bar{x}+C_x}$	1	C_x	$\eta_2 = \bar{y}_{po}\dfrac{\bar{X}_{po}+C_x}{\bar{x}_{po}+C_x}$	1	C_x
$t_3 = \bar{y}\dfrac{\bar{X}+\beta_1(x)}{\bar{x}+\beta_1(x)}$	1	$\beta_1(x)$	$\eta_3 = \bar{y}_{po}\dfrac{\bar{X}_{po}+\beta_1(x)}{\bar{x}_{po}+\beta_1(x)}$	1	$\beta_1(x)$
$t_4 = \bar{y}\dfrac{\bar{X}+\beta_2(x)}{\bar{x}+\beta_2(x)}$	1	$\beta_2(x)$	$\eta_4 = \bar{y}_{po}\dfrac{\bar{X}_{po}+\beta_2(x)}{\bar{x}_{po}+\beta_2(x)}$	1	$\beta_2(x)$
$t_5 = \bar{y}\dfrac{\bar{X}+\rho_{yx}}{\bar{x}+\rho_{yx}}$	1	ρ_{xy}	$\eta_5 = \bar{y}_{po}\dfrac{\bar{X}_{po}+\rho_{x_{po}y_{po}}}{\bar{x}_{po}+\rho_{x_{po}y_{po}}}$	1	$\rho_{x_{po}y_{po}}$
$t_{PR1} = \bar{y}\dfrac{\bar{X}+D_x}{\bar{x}+D_x}$	1	D_x	$\eta_6 = \bar{y}_{po}\dfrac{\bar{X}_{po}+D_x}{\bar{x}_{po}+D_x}$	1	D_x
$t_6 = \bar{y}\dfrac{\bar{X}C_x+\beta_1(x)}{\bar{x}C_x+\beta_1(x)}$	C_x	$\beta_1(x)$	$\eta_7 = \bar{y}_{po}\dfrac{\bar{X}_{po}C_x+\beta_1(x)}{\bar{x}_{po}C_x+\beta_1(x)}$	C_x	$\beta_1(x)$

$t_7 = \bar{y}\frac{\bar{X}C_x + \beta_2(x)}{\bar{x}C_x + \beta_2(x)}$	C_x	$\beta_2(x)$	$\eta_8 = \bar{y}_{po}\frac{\bar{X}_{po}C_x + \beta_2(x)}{\bar{x}_{po}C_x + \beta_2(x)}$	C_x	$\beta_2(x)$
$t_8 = \bar{y}\frac{\bar{X}C_x + \rho_{xy}}{\bar{x}C_x + \rho_{xy}}$	C_x	ρ_{xy}	$\eta_9 = \bar{y}_{po}\frac{\bar{X}_{po}C_x + \rho_{x_{po}y_{po}}}{\bar{x}_{po}C_x + \rho_{x_{po}y_{po}}}$	C_x	$\rho_{x_{po}y_{po}}$
$t_{PR2} = \bar{y}\frac{\bar{X}C_x + D_x}{\bar{x}C_x + D_x}$	C_x	D_x	$\eta_{10} = \bar{y}_{po}\frac{\bar{X}_{po}C_x + D_x}{\bar{x}_{po}C_x + D_x}$	C_x	D_x
$t_9 = \bar{y}\frac{\bar{X}\beta_1(x) + C_x}{\bar{x}\beta_1(x) + C_x}$	$\beta_1(x)$	C_x	$\eta_{11} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_1(x) + C_x}{\bar{x}_{po}\beta_1(x) + C_x}$	$\beta_1(x)$	C_x
$t_{10} = \bar{y}\frac{\bar{X}\beta_1(x) + \beta_2(x)}{\bar{x}\beta_1(x) + \beta_2(x)}$	$\beta_1(x)$	$\beta_2(x)$	$\eta_{12} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_1(x) + \beta_2(x)}{\bar{x}_{po}\beta_1(x) + \beta_2(x)}$	$\beta_1(x)$	$\beta_2(x)$
$t_{11} = \bar{y}\frac{\bar{X}\beta_1(x) + \rho_{xy}}{\bar{x}\beta_1(x) + \rho_{xy}}$	$\beta_1(x)$	ρ_{xy}	$\eta_{13} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_1(x) + \rho_{x_{po}y_{po}}}{\bar{x}_{po}\beta_1(x) + \rho_{x_{po}y_{po}}}$	$\beta_1(x)$	$\rho_{x_{po}y_{po}}$
$t_{PR3} = \bar{y}\frac{\bar{X}\beta_1(x) + D_x}{\bar{x}\beta_1(x) + D_x}$	$\beta_1(x)$	D_x	$\eta_{14} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_1(x) + D_x}{\bar{x}_{po}\beta_1(x) + D_x}$	$\beta_1(x)$	D_x
$t_{12} = \bar{y}\frac{\bar{X}\beta_2(x) + C_x}{\bar{x}\beta_2(x) + C_x}$	$\beta_2(x)$	C_x	$\eta_{15} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_2(x) + C_x}{\bar{x}_{po}\beta_2(x) + C_x}$	$\beta_2(x)$	C_x
$t_{13} = \bar{y}\frac{\bar{X}\beta_2(x) + \beta_1(x)}{\bar{x}\beta_2(x) + \beta_1(x)}$	$\beta_2(x)$	$\beta_1(x)$	$\eta_{16} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_2(x) + \beta_1(x)}{\bar{x}_{po}\beta_2(x) + \beta_1(x)}$	$\beta_2(x)$	$\beta_1(x)$
$t_{14} = \bar{y}\frac{\bar{X}\beta_2(x) + \rho_{xy}}{\bar{x}\beta_2(x) + \rho_{xy}}$	$\beta_2(x)$	ρ_{xy}	$\eta_{17} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_2(x) + \rho_{x_{po}y_{po}}}{\bar{x}_{po}\beta_2(x) + \rho_{x_{po}y_{po}}}$	$\beta_2(x)$	$\rho_{x_{po}y_{po}}$
$t_{PR4} = \bar{y}\frac{\bar{X}\beta_2(x) + D_x}{\bar{x}\beta_2(x) + D_x}$	$\beta_2(x)$	D_x	$\eta_{18} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_2(x) + D_x}{\bar{x}_{po}\beta_2(x) + D_x}$	$\beta_2(x)$	D_x
$t_{15} = \bar{y}\frac{\bar{X}\rho_{xy} + C_x}{\bar{x}\rho_{xy} + C_x}$	ρ_{xy}	C_x	$\eta_{19} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}} + C_x}{\bar{x}_{po}\rho_{x_{po}y_{po}} + C_x}$	$\rho_{x_{po}y_{po}}$	C_x
$t_{16} = \bar{y}\frac{\bar{X}\rho_{xy} + \beta_1(x)}{\bar{x}\rho_{xy} + \beta_1(x)}$	ρ_{xy}	$\beta_1(x)$	$\eta_{20} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}} + \beta_1(x)}{\bar{x}_{po}\rho_{x_{po}y_{po}} + \beta_1(x)}$	$\rho_{x_{po}y_{po}}$	$\beta_1(x)$
$t_{17} = \bar{y}\frac{\bar{X}\rho_{xy} + \beta_2(x)}{\bar{x}\rho_{xy} + \beta_2(x)}$	ρ_{xy}	$\beta_2(x)$	$\eta_{21} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}} + \beta_2(x)}{\bar{x}_{p}\rho_{x_{po}y_{po}} + \beta_2(x)}$	$\rho_{x_{po}y_{po}}$	$\beta_2(x)$
$t_{PR5} = \bar{y}\frac{\bar{X}\rho_{xy} + D_x}{\bar{x}\rho_{xy} + D_x}$	ρ_{xy}	D_x	$\eta_{22} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}} + D_x}{\bar{x}_{po}\rho_{x_{po}y_{po}} + D_x}$	$\rho_{x_{po}y_{po}}$	D_x
$t_{PR6} = \bar{y}\frac{\bar{X}D_x + C_x}{\bar{x}D_x + C_x}$	D_x	C_x	$\eta_{23} = \bar{y}_{po}\frac{\bar{X}_{po}D_x + C_x}{\bar{x}_{po}D_x + C_x}$	D_x	C_x
$t_{PR7} = \bar{y}\frac{\bar{X}D_x + \beta_1(x)}{\bar{x}D_x + \beta_1(x)}$	D_x	$\beta_1(x)$	$\eta_{24} = \bar{y}_{po}\frac{\bar{X}_{po}D_x + \beta_1(x)}{\bar{x}_{po}D_x + \beta_1(x)}$	D_x	$\beta_1(x)$
$t_{PR8} = \bar{y}\frac{\bar{X}D_x + \beta_2(x)}{\bar{x}D_x + \beta_2(x)}$	D_x	$\beta_2(x)$	$\eta_{25} = \bar{y}_{po}\frac{\bar{X}_{po}D_x + \beta_2(x)}{\bar{x}_{po}D_x + \beta_2(x)}$	D_x	$\beta_2(x)$
$t_{PR9} = \bar{y}\frac{\bar{X}D_x + \rho_{xy}}{\bar{x}D_x + \rho_{xy}}$	D_x	ρ_{xy}	$\eta_{26} = \bar{y}_{po}\frac{\bar{X}_{po}D_x + \rho_{x_{po}y_{po}}}{\bar{x}_{po}D_x + \rho_{x_{po}y_{po}}}$	D_x	$\rho_{x_{po}y_{po}}$

We first compare η_i, $i = 1, ..., 6$, for $a = 1$ with η_j, $j = 7,...,26$, for $a \neq 1$ in Table 1 to obtain the efficiency comparison as follows:

$$MSE(\eta_i) > MSE(\eta_j) . \quad (6)$$

Using Eq. (5), we can write

$$n^{-1}\left[R_{po}^2\lambda_1 - 2R_{po}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2} + \lambda_2\right] > n^{-1}\left[a^2R_{po}^2\lambda_1 - 2aR_{po}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2} + \lambda_2\right] .$$

From this inequality, we have $\left[R_{po}\lambda_1(1-a^2) - 2\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}(1+a)\right] > 0$ where $(1-a^2) > 0$. Then, we have $\left[R_{po}\lambda_1 - 2\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}\frac{(1+a)}{(1-a)(1+a)}\right] > 0$. It is written as $\left[(1-a)R_{po}\lambda_1 > 2\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}\right]$. Hence, the efficiency condition for Eq. (6) is found as $\frac{(1-a)R_{po}}{2\rho_{x_{po}y_{po}}}\frac{\lambda_1}{\sqrt{\lambda_1\lambda_2}} > 0$. This condition is always satisfied since R_{po}, λ_1, λ_2 are always positive when $a < 1$, $\rho_{x_{po}y_{po}} > 0$ and $\gamma_1, \gamma_2, \gamma_3 > 0$. Hence, we can infer that the proposed ratio estimators η_j, $j = 7,...,26$ are more efficient than the estimators η_i, $i = 1,...,6$ using the auxiliary variable information.

5. Numerical Illustration

In the study, we consider the earthquake data of Turkey for the numerical comparisons of the proposed and other ratio estimators in the SRS. We consider mainshocks that occured in Turkey between 1900 and 2011 having surface wave magnitudes $M_S \geq 5.0$, their foreshocks within five days with $M_S \geq 3.0$ and aftershocks within one month with $M_S \geq 4.0$. In this area, 120 mainshocks with surface magnitude $M_S \geq 5.0$ have occured between 1900 and 2011. The population consists of the destructive earthquakes. In the population data set the number of aftershocks is a study variable and the number of foreshocks is an auxiliary variable. The MSE values of usual ratio estimators $t_r, r = 1,...,17$ and $t_{PR1},...,t_{PR9}$ are obtained from Eq. (4) without considering the distributions of the study and auxiliary variables. The summary statistics for the population are given. Then, the MSE values of the proposed estimators $(\eta_1,...,\eta_{26})$ are computed from Eq. (5) with considering the distributions of the study and auxiliary variables. Several studies modeled earthquakes as a Poisson distribution (Ozel, 2011a, b). To obtain the distribution of these variables, we fit Poisson distribution to the earthquake dataset. The Poisson distribution provided an adequate fit with p-value <0.01 and chi-square value $(\chi^2 = 0.964)$ for the goodness of fit test. This means that the Poisson distribution with parameter $\lambda_1 = 6.925$ (year) fits the probability function of the auxiliary variable and the expected number of foreshocks of a main shock approximately equals to seven per year. After obtaining the frequency distribution of aftershocks and goodness of fit test ($\chi^2 = 0.964$, p-value= <0.01), it is seen that the study variable has a Poisson distribution with parameter $\lambda_2 = 10.216$. The summary statistics for the Poisson distributed population are given. To obtain $\rho_{x_{po}y_{po}}$ for the Poisson distributed data, Turkey is divided into three main neotectonic domains based on the neotectonic zones of Turkey. The foreshocks in Turkey are separated according to these neotectonic zones. By this way, the parameters γ_1, γ_2, and γ_3 are obtained. According to the goodness of fit test, it is seen that the Poisson distribution fits the number of shocks for area Region 1 with parameter $\gamma_1 = 4.813$ $(\chi^2 = 0.048, \text{p-value} = 0.043)$, with parameter $\gamma_2 = 8.104$ $(\chi^2 = 0.014,$ $\text{p-value} = 0.032)$ for Region 2, and $\gamma_3 = 2.112$ $(\chi^2 = 0.013, \text{p-value} = 0.025)$ for Region 3. Then, the correlation between the study variable and auxiliary variable is positive ($\rho_{x_{po}y_{po}} = 0.712$) and it can said that the number of foreshocks is related to the number of aftershocks. Therefore, the ratio estimators can be used for the estimation of the population mean in the SRS. The MSE values of the usual ratio estimators $t_r, r = 1,...,17$ and $t_{PR1},...,t_{PR9}$ are obtained and the proposed mean estimators $(\eta_1,...,\eta_{26})$ are computed using SRS and Table 2.

Table 2 R_{po} values and MSE equations for the ratio estimators of the Poisson distributed population

Ratio Estimators	R_{po}	MSE Equations
$\eta_1 = \bar{y}_{po}\dfrac{\bar{X}_{po}}{\bar{x}_{po}}$	$\dfrac{\bar{Y}}{\bar{X}}$	$MSE(\eta_1) \cong \dfrac{1}{n}\left[\dfrac{\bar{Y}^2}{\bar{X}^2}\lambda_1 - 2\dfrac{\bar{Y}}{\bar{X}}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_2 = \bar{y}_{po}\dfrac{\bar{X}_{po}+C_x}{\bar{x}_{po}+C_x}$	$\dfrac{\bar{Y}}{\bar{X}+C_x}$	$MSE(\eta_2) \cong \dfrac{1}{n}\left[\dfrac{\bar{Y}^2}{(\bar{X}+C_x)^2}\lambda_1 - 2\dfrac{\bar{Y}}{\bar{X}+C_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_3 = \bar{y}_{po}\dfrac{\bar{X}_{po}+\beta_1(x)}{\bar{x}_{po}+\beta_1(x)}$	$\dfrac{\bar{Y}}{\bar{X}+\beta_1(x)}$	$MSE(\eta_3) \cong \dfrac{1}{n}\left[\dfrac{\bar{Y}^2}{(\bar{X}+\beta_1(x))^2}\lambda_1 - 2\dfrac{\bar{Y}}{\bar{X}+\beta_1(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_4 = \bar{y}_{po}\dfrac{\bar{X}_{po}+\beta_2(x)}{\bar{x}_{po}+\beta_2(x)}$	$\dfrac{\bar{Y}}{\bar{X}+\beta_2(x)}$	$MSE(\eta_4) \cong \dfrac{1}{n}\left[\dfrac{\bar{Y}^2}{(\bar{X}+\beta_2(x))^2}\lambda_1 - 2\dfrac{\bar{Y}}{\bar{X}+\beta_2(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_5 = \bar{y}_{po}\dfrac{\bar{X}_{po}+\rho_{x_{po}y_{po}}}{\bar{x}_{po}+\rho_{x_{po}y_{po}}}$	$\dfrac{\bar{Y}}{\bar{X}+\rho_{x_{po}y_{po}}}$	$MSE(\eta_5) \cong \dfrac{1}{n}\left[\dfrac{\bar{Y}^2}{(\bar{X}+\rho_{x_{po}y_{po}})^2}\lambda_1 - 2\dfrac{\bar{Y}}{\bar{X}+\rho_{x_{po}y_{po}}}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_6 = \bar{y}_{po}\dfrac{\bar{X}_{po}+D_x}{\bar{x}_{po}+D_x}$	$\dfrac{\bar{Y}}{\bar{X}+D_x}$	$MSE(\eta_6) \cong \dfrac{1}{n}\left[\dfrac{\bar{Y}^2}{(\bar{X}+D_x)^2}\lambda_1 - 2\dfrac{\bar{Y}}{\bar{X}+D_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_7 = \bar{y}_{po}\dfrac{\bar{X}_{po}C_x+\beta_1(x)}{\bar{x}_{po}C_x+\beta_1(x)}$	$\dfrac{\bar{Y}}{\bar{X}C_x+\beta_1(x)}$	$MSE(\eta_7) \cong \dfrac{1}{n}\left[\dfrac{C_x^2\bar{Y}^2}{(\bar{X}C_x+\beta_1(x))^2}\lambda_1 - 2\dfrac{C_x\bar{Y}}{\bar{X}C_x+\beta_1(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_8 = \bar{y}_{po}\dfrac{\bar{X}_{po}C_x+\beta_2(x)}{\bar{x}_{po}C_x+\beta_2(x)}$	$\dfrac{\bar{Y}}{\bar{X}C_x+\beta_2(x)}$	$MSE(\eta_8) \cong \dfrac{1}{n}\left[\dfrac{C_x^2\bar{Y}^2}{(\bar{X}C_x+\beta_2(x))^2}\lambda_1 - 2\dfrac{C_x\bar{Y}}{\bar{X}C_x+\beta_2(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_9 = \bar{y}_{po}\dfrac{\bar{X}_{po}C_x+\rho_{x_{po}y_{po}}}{\bar{x}_{po}C_x+\rho_{x_{po}y_{po}}}$	$\dfrac{\bar{Y}}{\bar{X}C_x+\rho_{x_{po}y_{po}}}$	$MSE(\eta_9) \cong \dfrac{1}{n}\left[\dfrac{C_x^2\bar{Y}^2}{(\bar{X}C_x+\rho_{x_{po}y_{po}})^2}\lambda_1 - 2\dfrac{C_x\bar{Y}}{\bar{X}C_x+\rho_{x_{po}y_{po}}}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{10} = \bar{y}_{po}\dfrac{\bar{X}_{po}C_x+D_x}{\bar{x}_{po}C_x+D_x}$	$\dfrac{\bar{Y}}{\bar{X}C_x+D_x}$	$MSE(\eta_{10}) \cong \dfrac{1}{n}\left[\dfrac{C_x^2\bar{Y}^2}{(\bar{X}C_x+D_x)^2}\lambda_1 - 2\dfrac{C_x\bar{Y}}{\bar{X}C_x+D_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{11} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_1(x)+C_x}{\bar{x}_{po}\beta_1(x)+C_x}$	$\dfrac{\bar{Y}}{\bar{X}\beta_1(x)+C_x}$	$MSE(\eta_{11}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_1(x))^2\bar{Y}^2}{(\bar{X}\beta_1(x)+C_x)^2}\lambda_1 - 2\dfrac{\beta_1(x)\bar{Y}}{\bar{X}\beta_1(x)+C_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{12} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_1(x)+\beta_2(x)}{\bar{x}_{po}\beta_1(x)+\beta_2(x)}$	$\dfrac{\bar{Y}}{\bar{X}\beta_1(x)+\beta_2(x)}$	$MSE(\eta_{12}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_1(x))^2\bar{Y}^2}{(\bar{X}\beta_1(x)+\beta_2(x))^2}\lambda_1 - 2\dfrac{\beta_1(x)\bar{Y}}{\bar{X}\beta_1(x)+\beta_2(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{13} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_1(x)+\rho_{x_{po}y_{po}}}{\bar{x}_{po}\beta_1(x)+\rho_{x_{po}y_{po}}}$	$\dfrac{\bar{Y}}{\bar{X}\beta_1(x)+\rho_{x_{po}y_{po}}}$	$MSE(\eta_{13}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_1(x))^2\bar{Y}^2}{(\bar{X}\beta_1(x)+\rho_{x_{po}y_{po}})^2}\lambda_1 - 2\dfrac{\beta_1(x)\bar{Y}}{\bar{X}\beta_1(x)+\rho_{x_{po}y_{po}}}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{14} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_1(x)+D_x}{\bar{x}_{po}\beta_1(x)+D_x}$	$\dfrac{\bar{Y}}{\bar{X}\beta_1(x)+D_x}$	$MSE(\eta_{14}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_1(x))^2\bar{Y}^2}{(\bar{X}\beta_1(x)+D_x)^2}\lambda_1 - 2\dfrac{\beta_1(x)\bar{Y}}{\bar{X}\beta_1(x)+D_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{15} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_2(x)+C_x}{\bar{x}_{po}\beta_2(x)+C_x}$	$\dfrac{\bar{Y}}{\bar{X}\beta_2(x)+C_x}$	$MSE(\eta_{15}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_2(x))^2\bar{Y}^2}{(\bar{X}\beta_2(x)+C_x)^2}\lambda_1 - 2\dfrac{\beta_2(x)\bar{Y}}{\bar{X}\beta_2(x)+C_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{16} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_2(x)+\beta_1(x)}{\bar{x}_{po}\beta_2(x)+\beta_1(x)}$	$\dfrac{\bar{Y}}{\bar{X}\beta_2(x)+\beta_1(x)}$	$MSE(\eta_{16}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_2(x))^2\bar{Y}^2}{(\bar{X}\beta_2(x)+\beta_1(x))^2}\lambda_1 - 2\dfrac{\beta_2(x)\bar{Y}}{\bar{X}\beta_2(x)+\beta_1(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{17} = \bar{y}_{po}\dfrac{\bar{X}_{po}\beta_2(x)+\rho_{x_{po}y_{po}}}{\bar{x}_{po}\beta_2(x)+\rho_{x_{po}y_{po}}}$	$\dfrac{\bar{Y}}{\bar{X}\beta_2(x)+\rho_{x_{po}y_{po}}}$	$MSE(\eta_{17}) \cong \dfrac{1}{n}\left[\dfrac{(\beta_2(x))^2\bar{Y}^2}{(\bar{X}\beta_2(x)+\rho_{x_{po}y_{po}})^2}\lambda_1 - 2\dfrac{\beta_2(x)\bar{Y}}{\bar{X}\beta_2(x)+\rho_{x_{po}y_{po}}}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$

$\eta_{18} = \bar{y}_{po}\frac{\bar{X}_{po}\beta_2(x)+D_x}{\bar{x}_{po}\beta_2(x)+D_x}$	$\frac{\bar{Y}}{\bar{X}\beta_2(x)+D_x}$	$MSE(\eta_{18}) \cong \frac{1}{n}\left[\frac{(\beta_2(x))^2\bar{Y}^2}{(\bar{X}\beta_2(x)+D_x)^2}\lambda_1 - 2\frac{\beta_2(x)\bar{Y}}{\bar{X}\beta_2(x)+D_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{19} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}}+C_x}{\bar{x}_{po}\rho_{x_{po}y_{po}}+C_x}$	$\frac{\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+C_x}$	$MSE(\eta_{19}) \cong \frac{1}{n}\left[\frac{\rho^2_{x_{po}y_{po}}\bar{Y}^2}{(\bar{X}\rho_{x_{po}y_{po}}+C_x)^2}\lambda_1 - 2\frac{\rho_{x_{po}y_{po}}\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+C_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{20} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}}+\beta_1(x)}{\bar{x}_{po}\rho_{x_{po}y_{po}}+\beta_1(x)}$	$\frac{\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+\beta_1(x)}$	$MSE(\eta_{20}) \cong \frac{1}{n}\left[\frac{\rho^2_{x_{po}y_{po}}\bar{Y}^2}{(\bar{X}\rho_{x_{po}y_{po}}+\beta_1(x))^2}\lambda_1 - 2\frac{\rho_{x_{po}y_{po}}\bar{Y}-}{\bar{X}\rho_{x_{po}y_{po}}+\beta_1(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{21} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}}+\beta_2(x)}{\bar{x}_{po}\rho_{x_{po}y_{po}}+\beta_2(x)}$	$\frac{\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+\beta_2(x)}$	$MSE(\eta_{21}) \cong \frac{1}{n}\left[\frac{\rho^2_{x_{po}y_{po}}\bar{Y}^2}{(\bar{X}\rho_{x_{po}y_{po}}+\beta_2(x))^2}\lambda_1 - 2\frac{\rho_{x_{po}y_{po}}\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+\beta_2(x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{22} = \bar{y}_{po}\frac{\bar{X}_{po}\rho_{x_{po}y_{po}}+D_x}{\bar{x}_{po}\rho_{x_{po}y_{po}}+D_x}$	$\frac{\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+D_x}$	$MSE(\eta_{22}) \cong \frac{1}{n}\left[\frac{\rho^2_{x_{po}y_{po}}\bar{Y}^2}{(\bar{X}\rho_{x_{po}y_{po}}+D_x)^2}\lambda_1 - 2\frac{\rho_{x_{po}y_{po}}\bar{Y}}{\bar{X}\rho_{x_{po}y_{po}}+D_x}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{23} = \bar{y}_{po}\frac{\bar{X}_{po}D_x+C_x}{\bar{x}_{po}D_x+C_x}$	$\frac{\bar{Y}}{\bar{X}D_x+C_x}$	$MSE(\eta_{23}) \cong \frac{1}{n}\left[\frac{D_x^2\bar{Y}^2}{(\bar{X}D_x+C_x)^2}\lambda_1 - 2\frac{D_x\bar{Y}}{(\bar{X}D_x+C_x)}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{24} = \bar{y}_{po}\frac{\bar{X}_{po}D_x+\beta_1(x)}{\bar{x}_{po}D_x+\beta_1(x)}$	$\frac{\bar{Y}}{\bar{X}D_x+\beta_1(x)}$	$MSE(\eta_{24}) \cong \frac{1}{n}\left[\frac{D_x^2\bar{Y}^2}{(\bar{X}D_x+\beta_1(x))^2}\lambda_1 - 2\frac{D_x\bar{Y}}{(\bar{X}D_x+\beta_1(x))}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{25} = \bar{y}_{po}\frac{\bar{X}_{po}D_x+\beta_2(x)}{\bar{x}_{po}D_x+\beta_2(x)}$	$\frac{\bar{Y}}{\bar{X}D_x+\beta_2(x)}$	$MSE(\eta_{25}) \cong \frac{1}{n}\left[\frac{D_x^2\bar{Y}^2}{(\bar{X}D_x+\beta_2(x))^2}\lambda_1 - 2\frac{D_x\bar{Y}}{(\bar{X}D_x+\beta_2(x))}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$
$\eta_{26} = \bar{y}_{po}\frac{\bar{X}_{po}D_x+\rho_{x_{po}y_{po}}}{\bar{x}_{po}D_x+\rho_{x_{po}y_{po}}}$	$\frac{\bar{Y}}{\bar{X}D_x+\rho_{x_{po}y_{po}}}$	$MSE(\eta_{26}) \cong \frac{1}{n}\left[\frac{D_x^2\bar{Y}^2}{(\bar{X}D_x+\rho_{x_{po}y_{po}})^2}\lambda_1 - 2\frac{D_x\bar{Y}}{(\bar{X}D_x+\rho_{x_{po}y_{po}})}\rho_{x_{po}y_{po}}\sqrt{\lambda_1\lambda_2}+\lambda_2\right]$

We use the following expression to find the relative efficiency (RE) of ratio estimators using the characteristics of Poisson distribution when compared with the usual ratio estimators. Then, the proposed and usual ratio estimators are compared with respect to their MSE and RE values. We found that (i) the proposed η family ratio estimators using characteristics of Poisson distribution perform better than usual t family usual ratio estimators, (ii) the relative efficiency of the proposed η family ratio estimators are approximately 56 times more than the usual ratio estimators for the Poisson distributed data, (iii) the largest gain in efficiency is observed by using $\beta_2(x)$ and D_x with $\beta_2(x)$ if inter-group comparison of for the proposed estimators is done for the Poisson distributed data., (iv) the MSE value of the proposed t-family ratio estimator using C_x and D_x together is smaller than the other usual t-family ratio estimators, (v) the proposed ratio estimators $(\eta_{23} ..., \eta_{26})$ have the same value of MSE with $\eta_2,...,\eta_5$ since $D_x = 1$. However, if there is a population for different distributed populations, index of dispersion will differ from 1. In such a case $\eta_{23} ..., \eta_{26}$ yield different MSE values from $\eta_2,...,\eta_5$. Thus the class of the proposed ratio estimators is to be preferred to usual ratio estimators for the Poisson distributed population in the SRS.

Conclusion

In this study, first we suggested ratio estimators for the population mean using index of dispersion as an auxiliary variable. Then, we have developed new ratio estimators using characteristics of Poisson distributed auxiliary variable for the population mean in SRS and obtained their MSE equations. Different classes of ratio estimators are also proposed using the auxiliary variable information with considering the distribution of population. By MSE equations and RE values, the MSE values are compared and it is found that the proposed η family estimators are always more efficient than the usual t-family estimators for the Poisson distributed earthquake data. This theoretical result is also supported by a numerical example based on an earthquake data of Turkey. In the forthcoming studies, we hope to develop new estimators for the population mean for the Poisson distributed population using other sampling methods.

References

Gupta S, Shabbir J (2007) On the use of transformed auxiliary variable in estimating population mean. J Stat Plan Inference 137(5): 1606-1611.

Gupta S, Shabbir J (2008) On improvement in estimating the population mean in simple random sampling. Journal of Applied Statistics 35(5): 559-566.

Kadilar C, Cingi H (2004) Ratio estimators in simple random sampling. Applied Mathematics and Computation 151: 893-902.

Koyuncu N, Kadilar C (2009) Ratio and product estimators in stratified random sampling. J Stat Plan Inference 139(8): 2552-2558.

Ozel G, Inal C (2008) The probability function of the compound Poisson process and an application aftershock sequences in Turkey, Environmetrics 19(1): 79-85.

Ozel G (2011a) A bivariate compound Poisson model for the occurrence of foreshock and aftershock sequences in Turkey, Environmetrics 22(7): 847-856.

Ozel G (2011b) On certain properties of a class of bivariate compound Poisson distributions and an application to earthquake data, Revista Colombiana Estadistica, 34(3): 545-566.

Shabbir J, Gupta S (2011) On estimating the finite population mean in simple and stratified random sampling, Communication in Statistics: Theory and Methods 140(10): 2870-2874.

Singh HP, Vishwakarma GK (2010) A general procedure for estimating the population mean in stratified sampling using auxiliary information. Metron 68(1): 47-65

Sisodia BVS, Dwivedi VK (1981) A modified ratio estimator using coefficient of variation of auxiliary variable. Journal of the Indian Society of Agricultural Statistics 33(2): 13-18.

Upadhyaya LN, Singh HP (1999) Use of transformed auxiliary variable in estimating the finite population mean. Biometrical Journal 41(5): 627-636.

Contact

Gamze Özel
Hacettepe University, Department of Statistics, Ankara, Turkey
gamzeozl@hacettepe.edu.tr