

BIBLIOTECA DI SCIENZE STATISTICHE

SERVIZIO BIBLIOTECARIO NAZIONALE

BID 871807 BID 88609  
ACQ. 882102 INV. 88609  
COLL. 5-Coll. WP. 14/2002

**Regression and ratio  
estimators on two occasion**

G. Diana, C. Tommasi, M. Zuin

**2002.14**

**Dipartimento di Scienze Statistiche  
Università degli Studi  
Via C. Battisti 241-243  
35121 Padova**

**Ottobre 2002**

No. 111 / 1981  
 ANNO 1981  
 CLASSE 111

LA BIBLIOTECA

DELLA STABITIONE

DELLA UNIVERSITA'

1981

Department of Science

Library

Year 1981

Class 111

1981

# REGRESSION AND RATIO ESTIMATORS ON TWO OCCASION

Giancarlo Diana

Chiara Tommasi

Marino Zuin

# REGRESSION AND RATIO ESTIMATORS ON TWO OCCASION

Dr. G. S. RAO, Department of Statistics, University of Madras, India



# REGRESSION AND RATIO ESTIMATORS ON TWO OCCASION

## 1 Introduction

In two occasion sampling scheme the regression estimator is usually used for estimating a population mean. On the other hand, Kumar and Dayal (1998) consider the ratio estimator too. Specifically, they show that, for a fixed sample size, the regression estimator is more efficient but the ratio estimator is "cheaper". In this work ratio and regression estimator are compared through their efficiencies, but for fixed costs instead of for fixed sample size. The details are given in Section 2.

Let  $U = \{1, \dots, i, \dots, N\}$  be a finite population where the interest variable  $Y$  is observed on two subsequent occasions. Let  $Y_{1i}$  and  $Y_{2i}$  be the values of  $Y$  on the first and second occasion respectively. A sampling scheme on two occasions is based on two steps:

I step: a simple random sample without replacement (SRSWR) of  $n$  units is selected from  $N$  units, on the first occasion;

II step: out of these  $n$  units,  $m$  are retained for the second occasion (matched portion) and a new SRSWR of  $u = n - m$  units is selected from  $N - n$  units (unmatched portion). In this way the sample size is constant on both occasions.

Let  $y_{1i}$  and  $y_{2i}$  be the sampling realisations on the first and second occasion respectively. The following quantities can be considered

$$\bar{Y}_h = \frac{1}{N} \sum_{i=1}^N Y_{hi}; \quad S_{Y_h}^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_{hi} - \bar{Y}_h)^2; \quad CV(Y_h) = \frac{S_{Y_h}}{\bar{Y}_h} \quad \text{with } h = 1, 2$$

$$S_{Y_1 Y_2} = \frac{1}{N-1} \sum_{i=1}^N (Y_{1i} - \bar{Y}_1)(Y_{2i} - \bar{Y}_2); \quad \rho = \frac{S_{Y_1 Y_2}}{S_{Y_1} S_{Y_2}}$$

$$\bar{y}_h = \frac{1}{n} \sum_{i=1}^n y_{hi}; \quad \bar{y}_{hm} = \frac{1}{m} \sum_{i=1}^m y_{hi}; \quad s_{y_h}^2 = \frac{1}{m-1} \sum_{i=1}^m (y_{hi} - \bar{y}_{hm})^2 \quad \text{with } h = 1, 2$$

# REGRESSION AND RATIO ESTIMATORS ON TWO OCCASION

## 1 Introduction

In two occasion sampling scheme the regression estimator is usually used for estimating a population mean. On the other hand, Kumar and Gupta (1998) consider the ratio estimator for, specifically, they show that for a fixed sample size, the regression estimator is more efficient than the ratio estimator. However, in the work ratio and regression estimator are compared through first efficiency but for fixed cost instead of fixed sample size. The details are given in section 2.

Let  $U = \{1, 2, \dots, N\}$  be a finite population where the interest variable  $Y$  is observed on two successive occasions. Let  $y_i$  and  $y_i'$  be the values of  $Y$  on the first and second occasion respectively. A sampling scheme on two occasions is used on two steps.

I step: a simple random sample without replacement (SRSWR) of  $n$  units is selected from  $U$  units on the first occasion.  
 II step: out of these  $n$  units,  $w$  are retained for the second occasion (retained portion) and a new SRSWR of  $n-w$  units is selected from  $U-w$  units (unretained portion). In this way the sample size is constant on both occasions. Let  $w_1$  and  $w_2$  be the sampling fractions on the first and second occasion respectively. The following quantities can be considered

$$T = \frac{1}{N} \sum_{i=1}^N y_i, \quad S^2 = \frac{1}{N} \sum_{i=1}^N (y_i - T)^2, \quad C^2 = \frac{S^2}{T^2} \quad \text{with } h = 1$$

$$S_w^2 = \frac{1}{N-w} \sum_{i=1}^N (y_i - T)^2, \quad C_w^2 = \frac{S_w^2}{T^2}$$

$$T' = \frac{1}{N} \sum_{i=1}^N y_i', \quad S'^2 = \frac{1}{N} \sum_{i=1}^N (y_i' - T')^2, \quad C'^2 = \frac{S'^2}{T'^2} \quad \text{with } h = 1$$

$$s_{y_1 y_2} = \frac{1}{m-1} \sum_{i=1}^m (y_{1i} - \bar{y}_{1m})(y_{2i} - \bar{y}_{2m}); \quad \bar{y}_{2u} = \frac{1}{u} \sum_{i=1}^u y_{2i} \quad \text{and} \quad V(\bar{y}_{2u}) = \frac{S_{y_2}^2}{u}.$$

In the above expression of  $V(\bar{y}_{2u})$  we assume that the finite population correction can be ignored. From now on we will use this approximation.

In order to estimate the population mean of  $Y$  on the second occasion,  $\bar{Y}_2$ , the following estimators can be used

$$\bar{y}'_{Re} = \frac{u\bar{y}_{2u} + k\bar{y}_{Re}}{u+k} \quad \text{with} \quad V(\bar{y}'_{Re}) = \frac{S_{y_2}^2(n-u\rho^2)}{(n^2-u^2\rho^2)}$$

$$\bar{y}'_{Ra} = \frac{u\bar{y}_{2u} + k'\bar{y}_{Ra}}{u+k'} \quad \text{with} \quad V(\bar{y}'_{Ra}) = \frac{S_{y_2}^2(n-ut)}{(n^2-u^2t)}$$

where

$$t = 2\rho \frac{CV(Y_1)}{CV(Y_2)} - \left[ \frac{CV(Y_1)}{CV(Y_2)} \right]^2; \quad \frac{1}{k} = \frac{1-\rho^2}{m} + \frac{\rho^2}{n}; \quad \frac{1}{k'} = \frac{1-t}{m} + \frac{t}{n}$$

and  $\bar{y}_{Re}$  and  $\bar{y}_{Ra}$  are the regression and ratio estimators in double sampling when the auxiliary variable is  $Y$  measured on the first occasion. More specifically,

$$\bar{y}_{Re} = \bar{y}_{2m} + b(\bar{y}_1 - \bar{y}_{1m}) \quad \text{with} \quad b = \frac{S_{y_1 y_2}}{S_{y_1}^2}$$

and

$$\bar{y}_{Ra} = \bar{y}_1 \frac{\bar{y}_{2m}}{\bar{y}_{1m}}.$$

Kumar and Dayal (1998) compare  $\bar{y}'_{Re}$  and  $\bar{y}'_{Ra}$  for fixed  $n$  assuming  $CV(Y_1) = CV(Y_2)$  (for a justification of this hypothesis see Zuin, 2002). Specifically, they minimize  $V(\bar{y}'_{Re})$  and  $V(\bar{y}'_{Ra})$  with respect to  $m$  and  $u$ , getting

$$V_{\min}(\bar{y}'_{Re}) = \frac{S_{y_2}^2}{2n} (1 + \sqrt{1-\rho^2}) \quad \text{and} \quad V_{\min}(\bar{y}'_{Ra}) = \frac{S_{y_2}^2}{2n} (1 + \sqrt{1-t}).$$

Since  $V_{\min}(\bar{y}'_{Re}) < V_{\min}(\bar{y}'_{Ra})$  regression estimator is more efficient than ratio estimator. On the other hand,  $\bar{y}'_{Ra}$  is better than  $\bar{y}'_{Re}$  from an economic point of view, since  $m_{Ra} < m_{Re}$  and it is usually more expensive to get information about the unmatched portion than the matched one.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{and} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

In the above derivation of (10), we assume that the true population  
 intercept can be ignored from now on we will use this approximation.  
 In order to estimate the population mean of  $Y$ , on the second occasion, the  
 following estimator can be used

$$\hat{Y}_1 = \bar{y}_1 + \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} (\bar{X} - \bar{x}_1) \quad \text{with} \quad V(\hat{Y}_1) = \frac{1}{n} \left[ \frac{S_y^2}{n} + \frac{(\bar{X} - \bar{x}_1)^2}{n} \left( \frac{S_x^2}{(\bar{x}_2 - \bar{x}_1)^2} \right) \right]$$

$$V(\hat{Y}_1) = \frac{1}{n} \left[ \frac{S_y^2}{n} + \frac{(\bar{X} - \bar{x}_1)^2}{n} \left( \frac{S_x^2}{(\bar{x}_2 - \bar{x}_1)^2} \right) \right]$$

and  $\bar{x}_1$  and  $\bar{x}_2$  are the regression and ratio estimators double sampling  
 when the auxiliary variable  $X$  measured on the first occasion. More  
 specifically

$$\bar{x}_1 = \bar{y}_1 - b_1(\bar{Y} - \bar{y}_1) \quad \text{with} \quad b_1 = \frac{S_{xy}}{S_x^2}$$

and

$$\bar{x}_2 = \frac{\bar{y}_2 - \bar{y}_1}{\bar{X} - \bar{x}_1} (\bar{X} - \bar{x}_1) + \bar{x}_1$$

Kumar and Dey (1998) compare  $\hat{Y}_1$  and  $\hat{Y}_2$  for fixed  $n$  assuming  
 $CV(X) = CV(Y)$  for a justification of the hypothesis the ratio estimator is generally  
 they minimize  $V(\hat{Y}_1)$  and  $V(\hat{Y}_2)$  with respect to  $w$  and  $w'$  getting

$$w = \frac{1}{2} \left[ 1 + \frac{S_x^2}{S_y^2} \right] \quad \text{and} \quad w' = \frac{1}{2} \left[ 1 + \frac{S_x^2}{S_y^2} \right]$$

Since  $V(\hat{Y}_1) < V(\hat{Y}_2)$  regression estimator is more efficient than ratio  
 estimator. On the other hand,  $\hat{Y}_1$  is better than  $\hat{Y}_2$  from an economic point of  
 view since  $w < w'$  and it is usually more expensive to get information about  
 the unmeasured portion than the measured one.



## 2 Efficiency comparison for fixed costs

In practical setting, it is very common to have cost limitations. Thus, in this section we compare ratio and regression estimators through their efficiencies under a cost constraint instead of for a fixed sample size.

- **Theorem:** ratio estimator is never more efficient than regression estimator, under the linear cost function,

$$C = C_0 + C_m m + C_u u, \quad (1)$$

where  $C_0$  is a fixed cost while  $C_m$  and  $C_u$  ( $C_m < C_u$ ) denote the unit costs for the retained and replaced units, respectively.

### **Proof**

Let us denote the sampling cost as  $C^* = C - C_0$ . Of course we are assuming that  $C^*$  is fixed to some suitable value. Since  $m = n - u$ , from (1) we have that

$$u = a - bn \quad \text{where} \quad a = \frac{C^*}{C_u - C_m} \quad \text{and} \quad b = \frac{C_m}{C_u - C_m}. \quad (2)$$

Assuming  $CV(Y_1) = CV(Y_2)$ , as Kumar and Dayal (1998),  $V(\bar{y}'_{Re})$  and  $V(\bar{y}'_{Ra})$  can be written as

$$V(\bar{y}'_{Re}) = \frac{S_{Y_2}^2 [n - (a - bn)\rho^2]}{n^2 - (a - bn)^2 \rho^2} \quad \text{and} \quad V(\bar{y}'_{Ra}) = \frac{S_{Y_2}^2 [n - (a - bn)(2\rho - 1)]}{n^2 - (a - bn)^2 (2\rho - 1)}. \quad (3)$$

In what follows we prove that  $V(\bar{y}'_{Re}) - V(\bar{y}'_{Ra}) < 0$  for any value of  $n > 0$  and  $0 \leq u \leq n$ , thus

$$\left. \begin{array}{l} u \geq 0 \Rightarrow a - bn \geq 0 \Rightarrow n \leq \frac{a}{b} \\ u \leq n \Rightarrow a - bn \leq n \Rightarrow n \geq \frac{a}{b+1} \end{array} \right\} \Rightarrow \frac{a}{b+1} \leq n \leq \frac{a}{b}. \quad (4)$$

From (3) we have

$$V(\bar{y}'_{Re}) - V(\bar{y}'_{Ra}) = \frac{n S_{Y_2}^2 (\rho - 1)^2 [b(b+1)n^2 - a(2b+1)n + a^2]}{[(b^2 \rho^2 - 1)n^2 - 2ab\rho^2 n + a^2 \rho^2] \{[(2\rho - 1)b^2 - 1]n^2 - 2ab(2\rho - 1)n + (2\rho - 1)a^2\}} \quad (5)$$

The goal is to study the sign of the difference  $V(\bar{y}'_{Re}) - V(\bar{y}'_{Ra})$ . It is straightforward to prove that the numerator of expression (5) is negative for any values of  $n$  in

## 2 Efficiency comparison for fixed costs

In practical setting, it is very common to have a fixed cost. For example, in the above we suppose that a fixed cost is incurred for each observation. This fixed cost is not included in the regression equation, but it is included in the total cost function.

The regression line estimator is still the same, but the total cost function is now given by

$$C(x) = C_0 + C_1x + C_2x^2 \quad (2)$$

where  $C_0$  is a fixed cost,  $C_1$  and  $C_2$  are the regression coefficients. The regression line estimator is still the same, but the total cost function is now given by

Let us denote the sampling cost as  $C(x) = C_0 + C_1x + C_2x^2$ . Of course we are assuming that

$C_0$  is fixed to some value. Since  $x = x - \bar{x} + \bar{x}$ , we have that

$$C(x) = C_0 + C_1(\bar{x} + (x - \bar{x})) + C_2(\bar{x} + (x - \bar{x}))^2 \quad (3)$$

Assuming  $C_0$ ,  $C_1$ ,  $C_2$  are known and fixed,  $C(x)$  and  $C(x)$  can be written as

$$C(x) = C_0 + C_1\bar{x} + C_2\bar{x}^2 + 2C_2\bar{x}(x - \bar{x}) + C_2(x - \bar{x})^2 \quad (4)$$

In what follows we prove that  $C(x) < C(x)$  for any value of  $x$  and  $\bar{x}$  if and only if

$$C_2 < 0 \quad (5)$$

from (5) we have

$$C(x) < C(x) \quad (6)$$

the goal is to study the sign of the difference  $C(x) - C(x)$ . It is straightforward to prove that the numerator of expression (6) is negative for any value of  $x$  and

$[a(b+1)^{-1}, ab^{-1}]$ . On the other hand for studying the sign of the denominator we have to consider two cases,  $0 < \rho \leq 0,5$  and  $0,5 < \rho < 1$ . For computational reasons, let us factorise the first term of the denominator as  $(n - n_1)(n - n_2)$  where  $n_1 = a\rho(b\rho + 1)^{-1}$  and  $n_2 = a\rho(b\rho - 1)^{-1}$ .

**a)**  $0 < \rho \leq 0,5$ . We have three possibilities:

$$1) \rho^2 < b^{-2} \Rightarrow n_2 < n_1 < \frac{a}{b+1};$$

$$2) \rho^2 > b^{-2} \Rightarrow n_1 < \frac{a}{b+1} < \frac{a}{b} < n_2;$$

$$3) \rho^2 = b^{-2} \Rightarrow \text{the first factor becomes } -2ab^{-1}n + a^2b^{-2}.$$

In all the cases the first factor is negative in  $[a(b+1)^{-1}, ab^{-1}]$ . On the other hand, the second factor of the denominator is always negative, thus the denominator is positive when  $0 < \rho \leq 0,5$ .

**b)**  $0,5 < \rho < 1$ . The second term of the denominator in (5) can be factorised as

$$(n - n_3)(n - n_4) \text{ where } n_3 = a\sqrt{2\rho - 1}(b\sqrt{2\rho - 1} + 1)^{-1} \text{ and } n_4 = a\sqrt{2\rho - 1}(b\sqrt{2\rho - 1} - 1)^{-1}$$

and we have three possibilities:

$$1) 2\rho - 1 < \rho^2 < b^{-2} \Rightarrow n_2 < n_1 < \frac{a}{b+1} \quad \text{and} \quad n_4 < n_3 < \frac{a}{b+1};$$

$$2) 2\rho - 1 < b^{-2} < \rho^2 \Rightarrow n_1 < \frac{a}{b+1} < \frac{a}{b} < n_2 \quad \text{and} \quad n_4 < n_3 < \frac{a}{b+1};$$

$$3) b^{-2} < 2\rho - 1 < \rho^2 \Rightarrow n_1 < \frac{a}{b+1} < \frac{a}{b} < n_2 \quad \text{and} \quad n_3 < \frac{a}{b+1} < \frac{a}{b} < n_4.$$

In all the cases as well as when  $\rho^2 = b^{-2}$  or  $2\rho - 1 = b^{-2}$ , both the first and the second factor are negative for any  $n$  in  $[a(b+1)^{-1}, ab^{-1}]$ , thus the denominator is positive when  $0,5 < \rho < 1$ .

Therefore for any  $n$  in  $[a(b+1)^{-1}, ab^{-1}]$ ,  $V(\bar{y}'_{Ra}) \geq V(\bar{y}'_{Re})$ .

### 3 Numerical simulation

Even if, for fixed costs, regression estimators are more efficient than ratio estimators, it may be interesting to have an idea about the efficiency loss when

$(x_1 - \mu)^2$ . On the other hand, for  $x_1 < \mu$ , the sign of the denominator is  
 now to consider two cases:  $x_1 < \mu$  and  $x_1 > \mu$ . For computational  
 reasons, let us consider the first case, the denominator of  $(x_1 - \mu)^2$  where

$$\begin{aligned}
 & \text{denominator} = (x_1 - \mu)^2 \\
 & \text{denominator} = (x_1 - \mu)^2
 \end{aligned}$$

$(x_1 - \mu)^2$  is a positive quantity, and the denominator is positive.

In the case of  $x_1 > \mu$ , the denominator is positive, and the numerator is

also positive, and the denominator is positive.

In the case of  $x_1 < \mu$ , the denominator is positive, and the numerator is

also positive, and the denominator is positive.

In all the cases as well as when  $x_1 = \mu$ , the denominator is positive, and the

second factor is positive, and the denominator is positive.

### 3 Numerical simulation

Even if, for fixed costs, repeated estimators are more efficient than ratio



using the ratio estimator. For this reason we have performed a simulation study. For different values of  $C^*$ ,  $C_u/C_m$  and  $\rho \in (0.5, 1)$  we have computed the following quantities:  $n_{re}^{ott}$  and  $n_{ra}^{ott}$ , that is the optimal values of  $n$  minimizing  $V(\bar{y}'_{Re})$  and  $V(\bar{y}'_{Ra})$ , respectively; the corresponding optimal values for the variances, i.e.  $V^{ott}(\bar{y}'_{Re})$  and  $V^{ott}(\bar{y}'_{Ra})$  (with  $S_{Y_2}^2 = 1000$ ) and finally the efficiency loss rate.

From the simulation results, which are summarized in Table 3.1, it seems that

- i) the precision loss rate for using  $\bar{y}'_{Ra}$  instead of  $\bar{y}'_{Re}$  essentially does not change whatever the fixed value for  $C^*$ ;
- ii) the precision loss rate is always less than 2,7%.

**Remark 1:** in order to ensure a sampling cost not greater than the fixed value, we have rounded the optimal value of  $n$  to the closest integer and the optimal value for  $u$  to its integer part. That must cause an efficiency gain for  $\bar{y}'_{Ra}$ . For instance in Table 3.1 we have six cases.

**Remark 2:** In this simulation study we have considered only  $\rho \in (0.5, 1)$  since regression and ratio estimators are not useful when  $\rho$  is "low".

## 4 Conclusions and practical considerations

With or without cost constraints, regression estimators are more efficient than ratio estimators. The precision loss rate for using ratio estimators is however quite small, thus the simplicity can justify its use.

From a practical point of view, for using  $\bar{y}'_{Re}$  and/or  $\bar{y}'_{Ra}$  we need for  $n_{re}^{ott}$  and  $n_{ra}^{ott}$ , which are given by

$$n_{re}^{ott} = \frac{a\rho[\rho\sqrt{b^2\rho^2-1} + \sqrt{1-\rho^2}]\sqrt{b^2\rho^2-1}}{b^3\rho^4 + b^2\rho^2 - b\rho^2 - 1} \quad \text{and} \quad n_{ra}^{ott} = \frac{a[k(b^2k-1) + \sqrt{k(1-k)(b^2k-1)}]}{b^3k^2 + b^2k - bk - 1}, \quad \text{where } k = 2\rho - 1.$$

It is easy to prove that whenever  $n_{re}^{ott}$  and  $n_{ra}^{ott}$  exist they are greater than the lower bound  $a(a+b)^{-1}$ , whatever the value for  $\rho$ . Furthermore for any  $\rho \in (0.5, 1)$ , if  $2\rho - 1 > b^{-1}$ , i.e.  $\rho > 0.5(1 + b^{-1})$ , then both  $n_{re}^{ott}$  and  $n_{ra}^{ott}$  exist and are less than the upper bound  $ab^{-1}$ , otherwise if  $b^{-2} < 2\rho - 1 \leq b^{-1} < \rho^2$ , then both  $n_{re}^{ott}$  and  $n_{ra}^{ott}$  exist

Using the data generated from the simulation, we have plotted a simulation study for different values of  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ . We have computed the following quantities: (i) the mean of the optimal values of  $\alpha$  minimizing  $R(\alpha, \beta, \gamma, \delta)$ , (ii) the standard deviation of the optimal values of  $\alpha$ , (iii) the probability of the optimal values of  $\alpha$  being less than  $\alpha_0$ , (iv) the probability of the optimal values of  $\alpha$  being greater than  $\alpha_0$ , and (v) the probability of the optimal values of  $\alpha$  being equal to  $\alpha_0$ .

For the simulation study, we have used a random number generator for the distribution of  $\alpha$  and  $\beta$  and a uniform distribution for  $\gamma$  and  $\delta$ . The distribution of  $\alpha$  is  $N(0, 1)$  and the distribution of  $\beta$  is  $N(0, 1)$ . The distribution of  $\gamma$  is  $U(0, 1)$  and the distribution of  $\delta$  is  $U(0, 1)$ . The simulation study is based on 10,000 replications. The results of the simulation study are presented in Table 1. From Table 1, it is clear that the mean of the optimal values of  $\alpha$  is close to zero, the standard deviation of the optimal values of  $\alpha$  is small, and the probability of the optimal values of  $\alpha$  being less than zero is close to 0.5. The probability of the optimal values of  $\alpha$  being greater than zero is also close to 0.5. The probability of the optimal values of  $\alpha$  being equal to zero is close to 0.

From Table 1, it is clear that the mean of the optimal values of  $\alpha$  is close to zero, the standard deviation of the optimal values of  $\alpha$  is small, and the probability of the optimal values of  $\alpha$  being less than zero is close to 0.5. The probability of the optimal values of  $\alpha$  being greater than zero is also close to 0.5. The probability of the optimal values of  $\alpha$  being equal to zero is close to 0.

## 4. Conclusions and practical considerations

With or without any constraints, the regression estimator is more efficient than the ordinary least squares estimator. The question arises for using this estimator in the regression study. In this simulation study, we have considered only  $\alpha$  and  $\beta$  since the regression and ratio estimator are not used when  $\alpha$  is zero.

From a practical point of view, for using the estimator, we need to know the values of  $\alpha$  and  $\beta$ . The values of  $\alpha$  and  $\beta$  are given by  $\alpha = \frac{1}{2} \sqrt{1 + \frac{1}{\beta^2}}$  and  $\beta = \frac{1}{2} \sqrt{1 + \frac{1}{\alpha^2}}$ . It is easy to prove that whenever  $\alpha$  and  $\beta$  exist they are greater than the lower bound  $\alpha_0$  (whenever the value of  $\beta$  is finite) and  $\beta$  is greater than the upper bound  $\beta_0$  (whenever the value of  $\alpha$  is finite). It is also clear that  $\alpha$  and  $\beta$  exist and are greater than the lower bound  $\alpha_0$  and  $\beta_0$  whenever  $\alpha$  and  $\beta$  exist.

**TABLE 3.1:** Efficiency comparison for fixed costs between  $\bar{y}'_{Re}$  and  $\bar{y}'_{Ra}$ .

$C^*$	$C_u/C_m$	$\rho$	$n_{Re}^{out}$	$n_{Ra}^{out}$	$V(\bar{y}'_{Re})$	$V(\bar{y}'_{Ra})$	$\frac{V(\bar{y}'_{Ra}) - V(\bar{y}'_{Re})}{V(\bar{y}'_{Re})} \%$	
100	1,2	0,6	95	100	9,773	10,000	2,326	
		0,7	93	94	9,424	9,674	2,659	
		0,8	91	91	8,905	9,057	1,714	
		0,9	89	89	8,125	8,180	0,681	
	1,4	0,7	94	100	9,930	10,000	0,701	
		0,8	88	90	9,598	9,709	1,152	
		0,9	83	83	8,970	9,018	0,533	
	1,6	0,8	96	100	10,025	10,000	-	0,250
		0,9	82	83	9,625	9,678	0,551	
		1,8	0,9	98	100	10,039	10,000	-
	200	1,2	0,6	189	200	4,886	5,000	2,340
			0,7	185	188	4,711	4,837	2,678
0,8			182	183	4,452	4,528	1,691	
0,9			178	178	4,063	4,090	0,681	
1,4		0,7	188	200	4,965	5,000	0,701	
		0,8	176	179	4,799	4,859	1,245	
		0,9	166	167	4,476	4,505	0,637	
1,6		0,8	191	200	4,992	5,000	0,157	
		0,9	165	166	4,817	4,839	0,452	
		1,8	0,9	195	200	5,004	5,000	-
400		1,2	0,6	378	400	2,443	2,500	2,340
			0,7	370	376	2,356	2,419	2,678
	0,8		364	365	2,226	2,264	1,690	
	0,9		356	357	2,031	2,045	0,672	
	1,4	0,7	377	400	2,484	2,500	0,651	
		0,8	353	358	2,401	2,427	1,094	
		0,9	332	333	2,238	2,251	0,582	
	1,6	0,8	382	400	2,496	2,500	0,157	
		0,9	329	332	2,407	2,416	0,352	
		1,8	0,9	390	400	2,502	2,500	-
	800	1,2	0,6	757	800	1,221	1,250	2,341
			0,7	741	751	1,178	1,209	2,679
0,8			728	731	1,113	1,132	1,687	
0,9			712	713	1,016	1,022	0,671	
1,4		0,7	753	800	1,242	1,250	0,677	
		0,8	705	716	1,200	1,214	1,120	
		0,9	664	667	1,119	1,125	0,557	
1,6		0,8	764	800	1,248	1,250	0,157	
		0,9	659	664	1,203	1,208	0,402	
		1,8	0,9	781	800	1,251	1,250	-
1600		1,2	0,6	1.513	1.600	0,611	0,625	2,341
			0,7	1.482	1.502	0,589	0,605	2,679
	0,8		1.455	1.462	0,557	0,566	1,688	
	0,9		1.425	1.427	0,508	0,511	0,671	
	1,4	0,7	1.507	1.600	0,621	0,625	0,689	
		0,8	1.411	1.433	0,600	0,607	1,145	
		0,9	1.328	1.333	0,560	0,563	0,545	
	1,6	0,8	1.528	1.600	0,624	0,625	0,157	
		0,9	1.317	1.328	0,602	0,604	0,339	
		1,8	0,9	1.561	1.600	0,625	0,625	-

TABLE 1. Efficiency comparison for load factor between  $T_{in}$  and  $T_{out}$

Case	$T_{in}$ (K)	$T_{out}$ (K)	$\eta_{eff}$	$\eta_{th}$	$\eta_{c}$	$\eta_{p}$
100	1.00	1.00	0.00	0.00	0.00	0.00
	1.01	1.01	0.00	0.00	0.00	0.00
	1.02	1.02	0.00	0.00	0.00	0.00
	1.03	1.03	0.00	0.00	0.00	0.00
	1.04	1.04	0.00	0.00	0.00	0.00
	1.05	1.05	0.00	0.00	0.00	0.00
	1.06	1.06	0.00	0.00	0.00	0.00
	1.07	1.07	0.00	0.00	0.00	0.00
	1.08	1.08	0.00	0.00	0.00	0.00
	1.09	1.09	0.00	0.00	0.00	0.00
200	2.00	2.00	0.00	0.00	0.00	0.00
	2.01	2.01	0.00	0.00	0.00	0.00
	2.02	2.02	0.00	0.00	0.00	0.00
	2.03	2.03	0.00	0.00	0.00	0.00
	2.04	2.04	0.00	0.00	0.00	0.00
	2.05	2.05	0.00	0.00	0.00	0.00
	2.06	2.06	0.00	0.00	0.00	0.00
	2.07	2.07	0.00	0.00	0.00	0.00
	2.08	2.08	0.00	0.00	0.00	0.00
	2.09	2.09	0.00	0.00	0.00	0.00
300	3.00	3.00	0.00	0.00	0.00	0.00
	3.01	3.01	0.00	0.00	0.00	0.00
	3.02	3.02	0.00	0.00	0.00	0.00
	3.03	3.03	0.00	0.00	0.00	0.00
	3.04	3.04	0.00	0.00	0.00	0.00
	3.05	3.05	0.00	0.00	0.00	0.00
	3.06	3.06	0.00	0.00	0.00	0.00
	3.07	3.07	0.00	0.00	0.00	0.00
	3.08	3.08	0.00	0.00	0.00	0.00
	3.09	3.09	0.00	0.00	0.00	0.00
400	4.00	4.00	0.00	0.00	0.00	0.00
	4.01	4.01	0.00	0.00	0.00	0.00
	4.02	4.02	0.00	0.00	0.00	0.00
	4.03	4.03	0.00	0.00	0.00	0.00
	4.04	4.04	0.00	0.00	0.00	0.00
	4.05	4.05	0.00	0.00	0.00	0.00
	4.06	4.06	0.00	0.00	0.00	0.00
	4.07	4.07	0.00	0.00	0.00	0.00
	4.08	4.08	0.00	0.00	0.00	0.00
	4.09	4.09	0.00	0.00	0.00	0.00
500	5.00	5.00	0.00	0.00	0.00	0.00
	5.01	5.01	0.00	0.00	0.00	0.00
	5.02	5.02	0.00	0.00	0.00	0.00
	5.03	5.03	0.00	0.00	0.00	0.00
	5.04	5.04	0.00	0.00	0.00	0.00
	5.05	5.05	0.00	0.00	0.00	0.00
	5.06	5.06	0.00	0.00	0.00	0.00
	5.07	5.07	0.00	0.00	0.00	0.00
	5.08	5.08	0.00	0.00	0.00	0.00
	5.09	5.09	0.00	0.00	0.00	0.00



but only  $n_{re}^{ott}$  is less than  $ab^{-1}$ , while  $n_{ra}^{ott} \geq ab^{-1}$ . For this reason we have considered only values of  $\rho$  in  $[0.5(1+b^{-1}), 1)$ .

For practical purposes we suggest to round  $n_{re}^{ott}$  and  $n_{ra}^{ott}$  to their closest integers. Of course, the unmatched portion is given by (2) replacing  $n$  with  $n_{re}^{ott}$  or  $n_{ra}^{ott}$ . Specifically, we suggest to use the integer part of such obtained values in order to ensure a sampling cost not greater than the fixed value.

## 5 Bibliography

Kumar R. and Dayal S. (1998), *Regression and ratio estimators in sampling on two occasion*, "Journal of the Indian Statistical Association", 36, pp. 85-89.

Zuin M. (2002), *Il campionamento in due occasioni*, Graduation thesis, Department of Statistics, University of the study of Padova.

### SUMMARY

Recent studies in sampling on two occasion have highlighted the preference for regression estimator in terms of precision and for ratio estimators in terms of economic convenience. These conclusions are valid until no cost constraint is considered. In this paper an efficiency comparison for fixed costs is performed: it is shown that ratio estimator is always worse than regression estimator even if the efficiency loss is often small.

### RIASSUNTO

Recenti studi nell'ambito del campionamento in due occasioni hanno evidenziato la preferibilità dello stimatore per regressione in termini di precisione e dello stimatore per rapporto in termini di convenienza economica. Si trattava, però, di un'analisi di tipo marginale, valida quando la numerosità campionaria è fissata. In questo lavoro si propone un confronto di efficienza tra i due stimatori a parità di costi. Da tale confronto emerge che lo stimatore per rapporto non è mai preferibile anche se la perdita di efficienza è spesso trascurabile.

but only if a less than  $w_1$  while  $w_2$  is considered only values of  $w_1$  in (1).  
 for practical purposes we suggest to round  $w_1$  and  $w_2$  to the nearest integer.  
 Of course, the more rounded position is given by  $w_1$  and  $w_2$  and  
 so naturally we suggest to use the integer part of  $w_1$  and  $w_2$  obtained values in order  
 to avoid a sampling cost not greater than the fixed value.

## Bibliography

James R. Durbin (1973), *Journal of the Royal Statistical Society*, 35, pp. 35-45.  
 John M. (1973), *Il campionamento in due occasioni*, Gioluffo, Trieste.  
 Department of Statistics, University of the study of Padova.

## SUMMARY

Recent studies regarding on two occasions have highlighted the preference for  
 regression estimator in terms of precision and in terms of economic  
 convenience. These conclusions are valid only under certain conditions. In the  
 case of efficiency comparison, fixed cost is neglected. It is shown that this  
 estimator is more precise than regression estimator even if the efficiency is lower  
 than

## RISULTATI

Recenti studi effettuati in due occasioni in merito del campionamento a due  
 preselezioni, hanno evidenziato la preferenza per l'estimatore di regressione in termini di  
 precisione e di convenienza economica. Tali conclusioni sono valide solo in  
 particolari condizioni. In particolare, si è trascurato il costo fisso. Si è  
 dimostrato che l'estimatore di regressione è più preciso dell'estimatore  
 di differenza anche se l'efficienza è inferiore a quella dell'estimatore di  
 differenza.