Optimum ratio estimators for the population proportion—CMMSE-10

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>International Journal of Computer Mathematics</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>GCOM-2010-0717-B.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>07-Feb-2011</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Muñoz, Juan Francisco; Universidad de Granada Alvarez, Encarnación; Universidad de Granada Arcos, Antonio; Universidad de Granada Rueda, María del Mar; Universidad de Granada Gonzalez, Silvia; Universidad de Jaen, Estadística y Investigación Operativa Santiago, Agustín; Universidad de Guerrero</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Auxiliary information, ratio type estimator, survey sampling, population proportion estimation, simple random sample</td>
</tr>
</tbody>
</table>
RESEARCH ARTICLE

Optimum ratio estimators for the population proportion-CMMSE-10

Juan Francisco Muñoz a, Encarnación Álvarez b, Antonio Arcos c, María del Mar Rueda d, Silvia González e and A. Santiago f

a,b Department of Quantitative Methods for Economy and Enterprise, University of Granada, Spain ; c,d Department of Statistics and Operational Research, University of Granada, Spain ; e Department of Statistics and Operational Research, University of Jaén, Spain ; f Faculty of Mathematics, University of Guerrero, Mexico

The problem of the estimation of a population proportion using auxiliary information has been recently studied by Rueda et al. (2011), which proposed several ratio estimators of the population proportion and studied some theoretical properties. In this paper, we define a new ratio estimator based on a linear combination of two ratio estimators defined by Rueda et al. (2011). The variance of the new estimator is calculated and it is used to obtain the optimum value into the linear combination in the sense of minimal variance. Theoretical and empirical studies show that the suggested ratio estimator performs better than alternative estimators.

Keywords: Auxiliary information, ratio type estimator, proportion estimation, finite population, simple random sampling

AMS Subject Classification: 62D05

1. Introduction

In the presence of auxiliary information, there exist many design-based approaches (see [4], [6], [2]) to improve the precision of estimators in comparison to customary methods, which do not involve auxiliary information. However, techniques involving auxiliary information have been discussed for quantitative variables, and the extension to the estimation of a population proportion requires further investigation. For example, one should be aware of the risks when confidence intervals are constructed for a population proportion, since limits outside [0,1] could be achieved.

We consider the scenario of a finite population \( U = \{1, \ldots, N\} \) containing \( N \) units. Let \( A_1, \ldots, A_N \) denote the values of an attribute of interest \( A \), where \( A_i = 1 \) if \( i \)th unit possesses the attribute \( A \) and \( A_i = 0 \) otherwise. Let \( B \) denote an auxiliary attribute associated with \( A \) and values given by \( B_1, \ldots, B_N \). We also assume that a sample \( s \), of size \( n \), is selected from \( U \) according to the well known simple random sampling without replacement (SRSWOR).

The aim is to estimate the population proportion of individuals that posses the attribute \( A \), i.e. \( P_A = N^{-1} \sum_{i=1}^{N} A_i \). Assuming a finite population, the naive
estimator of \( P_A \), which makes no use of the auxiliary information, is given by
\[
\hat{p}_A = n^{-1} \sum_{i \in S} A_i.
\]

We assume that the population proportion of individuals that posses the attribute \( B \), \( P_B = N^{-1} \sum_{i=1}^N B_i \), is known from a census or estimated without error.

Rueda et al (2011) defined the following ratio estimator for \( P_A \):
\[
\hat{p}_r = \hat{R}P_B,
\]
where \( \hat{R} = \hat{p}_A/\hat{p}_B \) is an estimator of the population ratio \( R = P_A/P_B \) and \( \hat{p}_B = n^{-1} \sum_{i \in S} B_i \) is the sample proportion of individuals that posses the auxiliary attribute \( B \).

Let \( A^c \) and \( B^c \) denote the complementary attributes of \( A \) and \( B \), and consider the population two-way table given by
\[
\begin{array}{c|cc}
B & B^c \\
A & N_1 & N_2 \\
A^c & N_{11} & N_{12} \\
\hline
N_{21} & N_{22} & N_2
\end{array}
\]
(2)

where \( N_1 = \sum_{i=1}^N A_i \) is the number of units in the population that posses the attribute \( A \), \( N_2 \) is the number of units in the population that not to posses the attribute \( A \), etc. Analogously, \( N_{11} \) is the number of units in the population that simultaneously posses the attributes \( A \) and \( B \), \( N_{12} \) is the number of units in the population that simultaneously posses the attributes \( A \) and \( B^c \), etc. Classification (2) can be also defined at the sample level as
\[
\begin{array}{c|cc}
B & B^c \\
A & n_1 & n_2 \\
A^c & n_{11} & n_{12} \\
\hline
n_{21} & n_{22} & n
\end{array}
\]
(3)

The estimator \( \hat{p}_r \) is a biased estimator of \( P_A \) and the asymptotic variance of \( \hat{p}_r \) is given by
\[
AV(\hat{p}_r) = \frac{N - n}{(N - 1)n} \left( P_A Q_A + R^2 P_B Q_B - 2R\phi\sqrt{P_A Q_A P_B Q_B} \right),
\]
(4)
where
\[
\phi = \frac{N_{11}N_{22} - N_{12}N_{21}}{\sqrt{N_1N_2N_1N_2}}
\]
is the Cramer’s \( V \) coefficient ([1]) based on the two-way classification (2).

We observe that the customary estimator \( \hat{p}_A \) can be also obtained as \( \hat{p}_A = 1 - \hat{q}_A \), where \( \hat{q}_A = n^{-1} \sum_{i \in S} A^c_i \), hence \( \hat{p}_A \) has the same performance in the estimation of \( P_A \) then the performance of \( \hat{q}_A \) in the estimation of \( Q_A \). However, this property is not satisfied by \( \hat{p}_r \), i.e. it can be easily seen that \( \hat{p}_r \neq 1 - \hat{q}_r \), where \( \hat{q}_r = \hat{R}_c Q_B \) and \( \hat{R}_c = (\hat{q}_A/\hat{q}_B) \). For this reason, Rueda et al. (2011) defined the ratio estimator
\[
\hat{p}_{r,q} = 1 - \hat{q}_r \text{ for } P_A \text{ and showed that } AV(\hat{p}_r) < AV(\hat{p}_{r,q}) \text{ when } P_A < P_B.
2. The optimum ratio estimator

In this section, we define a new ratio type estimator using a linear combination of the ratio estimators \( \hat{p}_r \) and \( \hat{p}_{r,q} \) previously defined. The choice of the optimum weight value into the linear combination is achieved by minimizing the variance. Finally, some interesting theoretical properties are also obtained.

The new ratio type estimator is

\[
\hat{p}_{r,w} = w\hat{p}_r + (1 - w)\hat{p}_{r,q}. \tag{5}
\]

**Theorem 2.1**

The optimum value for \( w \) in the sense of minimum variance into the class of estimators \( \hat{p}_{r,w} \) is

\[
w_{opt} = \frac{AV(\hat{p}_{r,q}) - cov(\hat{p}_r, \hat{p}_{r,q})}{AV(\hat{p}_r) + AV(\hat{p}_{r,q}) - 2cov(\hat{p}_r, \hat{p}_{r,q})}. \tag{6}
\]

**Proof**

Next, we determine the optimum value of \( w \) by minimizing the variance of \( \hat{p}_{r,w} \).

The asymptotic variance of \( \hat{p}_{r,w} \) is given by

\[
AV(\hat{p}_{r,w}) = AV(w\hat{p}_r + (1 - w)\hat{p}_{r,q}) =
\]

\[
= w^2AV(\hat{p}_r) + (1 - w)^2AV(\hat{p}_{r,q}) + 2w(1 - w)cov(\hat{p}_r, \hat{p}_{r,q}).
\]

By denoting \( V_1 = AV(\hat{p}_r) \), \( V_2 = V(\hat{p}_{r,q}) \) and \( C = cov(\hat{p}_r, \hat{p}_{r,q}) \), the variance of \( \hat{p}_{r,w} \) can be expressed as

\[
AV(\hat{p}_{r,w}) = w^2V_1 + (1 - w)^2V_2 + 2w(1 - w)C.
\]

The first derivative of \( AV(\hat{p}_{r,w}) \) with respect to \( w \) is

\[
\frac{\partial AV(\hat{p}_{r,w})}{\partial w} = 2wV_1 - 2(1 - w)V_2 + 2(1 - 2w)C = 0;
\]

\[
w_{opt} = \frac{V_2 - C}{V_1 + V_2 - 2C}.
\]

The second derivative is

\[
\frac{\partial^2 AV(\hat{p}_{r,w})}{\partial^2 w} = 2V_1 + 2V_2 - 4C = 2(V_1 + V_2 - 2C) = 2AV(\hat{p}_r - \hat{p}_{r,q}) > 0,
\]

and we conclude that \( w_{opt} \) really minimizes \( AV(\hat{p}_{r,w}) \).
Therefore, the optimum ratio estimator in the sense of minimum variance into the class (5) is

$$\hat{p}_{r,OPT} = w_{opt}\hat{p}_r + (1 - w_{opt})\hat{p}_{r,q}.$$  

In practice, $$\hat{p}_{r,OPT}$$ could be unknown, since $$w_{opt}$$ depends on population variances, which are generally unknown. In this situation, we can use the estimator

$$\hat{p}_{r,opt} = \hat{w}_{opt}\hat{p}_r + (1 - \hat{w}_{opt})\hat{p}_{r,q}, \quad (7)$$

where

$$\hat{w}_{opt} = \frac{\hat{V}(\hat{p}_{r,q}) - \hat{\text{cov}}(\hat{p}_r, \hat{p}_{r,q})}{\hat{V}(\hat{p}_r) + \hat{V}(\hat{p}_{r,q}) - 2\hat{\text{cov}}(\hat{p}_r, \hat{p}_{r,q})}. \quad (8)$$

Following Särndal et al. (1992) pg 372, the variance of $$\hat{p}_{r,w}$$ can be expressed as

$$AV(\hat{p}_{r,w}) = (V_1 + V_2 - 2C)\left(\frac{V_1 V_2 - C^2}{V_1 + V_2 - 2C}\right)^2 + \frac{V_1 V_2 - C^2}{V_1 + V_2 - 2C},$$

and we can deduce that the variance of the optimum estimator is

$$AV(\hat{p}_{r,OPT}) = \frac{V_1 V_2 - C^2}{V_1 + V_2 - 2C}.$$  

An estimator of the variance of the optimum estimator can be obtained as

$$\hat{V}(\hat{p}_{r,opt}) = \frac{\hat{V}(\hat{p}_r)\hat{V}(\hat{p}_{r,q}) - \hat{\text{cov}}^2(\hat{p}_r, \hat{p}_{r,q})}{\hat{V}(\hat{p}_r) + \hat{V}(\hat{p}_{r,q}) - 2\hat{\text{cov}}(\hat{p}_r, \hat{p}_{r,q})}.$$  

We observe that $$AV(\hat{p}_{r,OPT})$$ depends on the covariance $$C = \text{cov}(\hat{p}_r, \hat{p}_{r,q})$$. Theorem 2 gives an expression for $$C$$.

**Theorem 2.2** The covariance between the ratio estimators $$\hat{p}_r$$ and $$\hat{p}_{r,q}$$ is

$$\text{cov}(\hat{p}_r, \hat{p}_{r,q}) = \frac{N - n}{N - 1} \frac{1}{n} \left( PAQ_A + RR_cP_BP_B - (R + R_c)\phi \sqrt{PAQ_A P_BQ_B} \right),$$

where $$R_c = Q_A/Q_B$$ is the population ratio of the complementary proportions of the attributes $$A$$ and $$B$$.

**Proof**

Using Taylor series (see Särndal et al. 1992, pg 178), $$\hat{R}$$ can be expressed as

$$\hat{R} \equiv R + \frac{1}{nF_B} \sum_{i \in s} (A_i - RB_i) = R + \frac{1}{F_B} (\hat{p}_A - R\hat{p}_B),$$

and similarly

$$\hat{R}_c \equiv R_c + \frac{1}{Q_B} (\hat{q}_A - R_c\hat{q}_B).$$
Using the previous expressions we obtain

\[ C = \text{cov}(\hat{p}_r, 1 - \hat{q}_r) = -\text{cov}(\hat{R}_B, \hat{R}_c Q_B) = \]

\[ -P_B Q_B \text{cov} \left( R + \frac{1}{P_B} (\hat{p}_A - R\hat{P}_B), R_c + \frac{1}{Q_B} (\hat{q}_A - R_c\hat{q}_B) \right) = \]

\[ = V(\hat{p}_A) + R R_c V(\hat{p}_B) - (R + R_c) \text{cov}(\hat{p}_A, \hat{p}_B) = \]

\[ = \frac{N - n}{N - 1} \left( P_A Q_A + R R_c P_B Q_B - (R + R_c) \phi \sqrt{P_A Q_A P_B Q_B} \right). \]

An estimator of the covariance \( \text{cov}(\hat{p}_r, \hat{p}_{r,q}) \) is

\[ \hat{\text{cov}}(\hat{p}_r, \hat{p}_{r,q}) = \frac{1 - f}{n - 1} \left( \hat{p}_{\hat{p}A} + \hat{R} \hat{R}_c \hat{p}_B \hat{q}_B - (\hat{R} + \hat{R}_c) \hat{\phi} \sqrt{\hat{p}_{\hat{p}A} \hat{p}_{\hat{p}B} \hat{q}_{\hat{q}B}} \right), \]

where

\[ \hat{\phi} = \frac{n_{11}n_{22} - n_{12}n_{21}}{\sqrt{n_{11}n_{22} - n_{12}^2}}. \]

**Theorem 2.3**

The optimum weight \( w_{opt} \) in expression (6) can be expressed as

\[ w_{opt} = \frac{R_c - \beta}{R_c - R}, \]

where

\[ \beta = \frac{\text{cov}(\hat{p}_A, \hat{p}_B)}{V(\hat{p}_B)}. \]

**Proof**

Knowing that

\[ V_1 = V(\hat{p}_A) + R^2 V(\hat{p}_B) - 2R \text{cov}(\hat{p}_A, \hat{p}_B), \]

\[ V_2 = V(\hat{q}_A) + R^2 V(\hat{q}_B) - 2R \text{cov}(\hat{q}_A, \hat{q}_B) = \]

\[ = V(\hat{p}_A) + R^2 V(\hat{p}_B) - 2R \text{cov}(\hat{p}_A, \hat{p}_B) \]
and

\[ C = V(\hat{p}_A) + R R_c V(\hat{p}_B) - (R + R_c) \text{cov}(\hat{p}_A, \hat{p}_B), \]

the numerator and the denominator of \( w_{opt} \) in (6) are given by

\[ V_2 - C = V(\hat{p}_B)(R_c^2 - R R_c) - \text{cov}(\hat{p}_A, \hat{p}_B)[2R_c - (R - R_c)] = \]
\[ = V(\hat{p}_B)R_c(R_c - R) - \text{cov}(\hat{p}_A, \hat{p}_B)(R_c - R) = \]
\[ = (R_c - R)[V(\hat{p}_B)R_c - \text{cov}(\hat{p}_A, \hat{p}_B)] \]

and

\[ V_1 + V_2 - 2C = V(\hat{p}_B)(R^2 + R_c^2 - 2R R_c) - \text{cov}(\hat{p}_A, \hat{p}_B)[2R + 2R_c - 2(R + R_c)] = \]
\[ = V(\hat{p}_B)(R_c - R)^2 \]

By replacing these expressions in (6) we obtain

\[ w_{opt} = \frac{V_2 - C}{V_1 + V_2 - 2C} = \frac{(R_c - R)[V(\hat{p}_B)R_c - \text{cov}(\hat{p}_A, \hat{p}_B)]}{V(\hat{p}_B)(R_c - R)^2} = \]
\[ = \frac{V(\hat{p}_B)R_c - \text{cov}(\hat{p}_A, \hat{p}_B)}{(R_c - R)V(\hat{p}_B)} = \frac{R_c - \beta}{R_c - R} \]

Following theorem 2.3, the estimated optimum weight \( \hat{w}_{opt} \) given by (8) can be calculated as

\[ \hat{w}_{opt} = \frac{\hat{R}_c - \hat{\beta}}{\hat{R}_c - \hat{R}}, \tag{9} \]

where

\[ \hat{\beta} = \frac{\text{cov}(\hat{p}_A, \hat{p}_B)}{V(\hat{p}_B)}. \]

From expression (9) we conclude that \( \hat{w}_{opt} = 1 \), that is \( \hat{p}_r:opt = \hat{p}_r \), if \( \hat{\beta} = \hat{R} \), whereas \( \hat{w}_{opt} = 0 \), that is \( \hat{p}_r:opt = \hat{p}_{r,q} \), if \( \hat{\beta} = \hat{R_c} \). In other words, the ratio estimator \( \hat{p}_r \) has a larger weight into the optimum estimator \( \hat{p}_{r,opt} \) as \( \hat{\beta} \) is closer to \( \hat{R} \). On the
other hand, the ratio estimator $\hat{p}_{r,q}$ has a larger weight into the optimum estimator $\hat{p}_{r,opt}$ as $\beta$ is closer to $\tilde{R}_c$.

**Theorem 2.4**

The asymptotic variance of the optimum ratio estimator $\hat{p}_{r,OPT}$ can be calculated as

$$AV(\hat{p}_{r,OPT}) = V(\hat{p}_A)(1 - \phi^2).$$

**Proof**

The asymptotic variance of $\hat{p}_{r,OPT}$ is

$$AV(\hat{p}_{r,OPT}) = \frac{V_1V_2 - C}{V_1 + V_2 - 2C},$$

where the denominator, as seen in proof of theorem 2.3, can be obtained as

$$V_1 + V_2 - 2C = V(\hat{p}_B)(R - R_c)^2.$$

Next, we obtain the numerator of $AV(\hat{p}_{r,OPT})$. For the sake of simplicity, we denote $V_A = V(\hat{p}_A)$, $V_B = V(\hat{p}_B)$ and $C_{AB} = cov(\hat{p}_A, \hat{p}_B)$. We had that

$$V_1 = V_A + R^2V_B - 2Rc_{AB},$$

$$V_2 = V_A + R^2cV_B - 2Rc_{AB}$$

and

$$C = V_A + RRcV_B - (R + Rc)C_{AB}.$$

First,

$$V_1V_2 = V_A^2 + R^2c^2V_AV_B - 2RcV_Ac_{AB} + R^2V_AV_B + R^2c^2V_B^2$$

$$-2R^2cV_Bc_{AB} - 2RV_Ac_{AB} - 2R^2c^2V_Bc_{AB} + 4RRc^2c_{AB}.$$

The square of the covariance can be expressed as

$$C^2 = V_A^2 + R^2c^2V_B^2 + (R + Rc)^2c_{AB}^2 + 2V_ARcV_B$$

$$-2(R + Rc)V_Ac_{AB} - 2RRc(R + Rc)V_Bc_{AB} =$$

$$= V_A^2 + R^2c^2V_B^2 + R^2c_{AB}^2 + R^2c_{AB}^2 + 2RRc^2c_{AB} + 2V_ARcV_B.$$
\[-2RVA_{CAB} - 2RcVACAB - 2R^2RCVCAB - 2RR_c^2VCAB.\]

Then, the numerator of \(AV(\hat{p}_{r,OPT})\) is

\[V_1V_2 - C^2 = V_AV_B(R_c^2 - 2RR_c + R^2) - C_{AB}^2(R^2 + R_c^2 - 2RR_c) =\]

\[= (V_AV_B - C_{AB}^2)(R - R_c)^2.\]

The variance of \(\hat{p}_{r,OPT}\) can be also obtained as

\[AV(\hat{p}_{r,OPT}) = \frac{V_AV_B - C_{AB}^2}{V_B} = \frac{V(\hat{p}_A)V(\hat{p}_B) - \text{cov}(\hat{p}_A, \hat{p}_B)^2}{V(\hat{p}_B)}.\] (10)

Replacing \(V(\hat{p}_A), V(\hat{p}_B)\) and \(\text{cov}(\hat{p}_A, \hat{p}_B)\) in (10) by their respective expressions under SRSWOR we obtain

\[AV(\hat{p}_{r,OPT}) = \frac{N - n - 1}{N - 1}\left[ \frac{P_AQ_AP_BQ_B - \phi^2P_AQ_AP_BQ_B}{P_BQ_B} \right] =\]

\[= \frac{N - n - 1}{N - 1}P_AQ_A(1 - \phi^2) = V(\hat{p}_A)(1 - \phi^2).\]

Theoretical comparison between the ratio estimator \(\hat{p}_{r,OPT}\) and the simple expansion estimator \(\hat{p}_A\) is fairly simple using theorem 2.4. In fact, \(\hat{p}_{r,OPT}\) is more efficient than \(\hat{p}_A\), since \(1 - \phi^2 \leq 1\), and both estimators has the same performance when \(\phi^2 = 0\).

Using theorem 2.4, an estimator of the optimum ratio type estimator variance is

\[\hat{V}(\hat{p}_{r,opt}) = \hat{V}(\hat{p}_A)(1 - \phi^2).\]

3. Simulation study

In this section, the proposed optimum ratio estimator \(\hat{p}_{r,opt}\) is compared numerically with alternative proportion estimators. Simulation studies are based on several simulated populations which cover a wide number of possible scenarios, including small and large proportions, small and large Cramer’s \(V\) coefficients between the attribute of interest and the auxiliary attributes, etc. Simulated populations are briefly described as follows.

A total of 30 populations of \(N = 1000\) units were generated to study the effect of different aspects on the estimators of a population proportion. Populations were generated as a random sample of 1000 units from a Bernoulli distribution with parameter \(p = \{0.1, 0.25, 0.5, 0.75, 0.9\}\), and the attributes of interest were thus achieved with the aforementioned population proportions. Auxiliary attributes were also generated by using the same distribution, but we randomly change a given proportion of values in order to the Cramer’s \(V\) coefficient between the attribute of interest and the auxiliary attribute goes from 0.5 to 0.9. Since \(P_A < P_B\) when \(P_A = 0.25\), we also generated populations with \(P_A = 0.25\) and \(P_A > P_B\), which
allow us to study the effect of the relation between \( P_A \) and \( P_B \) on the various estimators, specially on the estimator \( \hat{p}_r \).

For each of the 30 populations, \( D = 10000 \) samples, with size \( n = 100 \), were selected to compare the various estimators in terms of relative bias (\( RB \)) and relative efficiency (\( RE \)), where

\[
RB = \frac{E[\hat{p}] - P_A}{P_A} ; \quad RE = \frac{MSE[\hat{p}_A]}{MSE[\hat{p}]},
\]

\( \hat{p} \) is a given estimator and the empirical expectation (\( E[\cdot] \)) and the empirical mean square error (\( MSE[\cdot] \)) are given by

\[
E[\hat{p}] = \frac{1}{D} \sum_{d=1}^{D} \hat{p}(d) ; \quad MSE[\hat{p}] = \frac{1}{D} \sum_{d=1}^{D} (\hat{p}(d) - P_A)^2,
\]

\( \hat{p}(d) \) denotes the estimator \( \hat{p} \) calculated at the \( d \)th simulation run. Values of \( RE \) larger than 1 indicate that the estimator \( \hat{p} \) is more efficient than the customary estimator \( \hat{p}_A \), which is considered as the reference estimator in the efficiency studies.

We considered the proposed optimum ratio estimator \( \hat{p}_{r,opt} \), the ratio estimators \( \hat{p}_r \) and \( \hat{p}_{r,e} \) proposed by Rueda et al. (2011) and the difference estimator given by

\[
\hat{p}_d = \hat{p}_A + (P_B - \hat{p}_B).
\]

Values of \( RB \) in this simulation study are within a reasonable range, i.e., they are all less than 1% and are thus omitted. Figures 1 and 2 reports the values of \( RE \) for the various estimators and samples selected under SRSWOR. We observe that the proposed optimum estimator is more efficient than alternative estimators, whereas other estimators such as \( \hat{p}_r \) and \( \hat{p}_d \) can be less efficient than the customary estimator \( \hat{p}_A \). The gain in efficiency of the estimators based upon auxiliary information increases considerably as \( \phi \) is closer to 0.9. For this reason and for clarity in figures, values of \( RE \) are separately plotted according to \( \phi \).

Table 1 reports the proportion of number of times that estimators take values outside the limits [0, 1]. For small as well as for large samples, we observe that the proposed optimum ratio estimator and \( \hat{p}_{r,e} \) always take values within the interval [0, 1], whereas estimators \( \hat{p}_r \) and \( \hat{p}_d \) can give estimates outside [0, 1], with a percentage larger than 10% for small samples.
4. Figures

Figure 1. Values of Relative Efficiency (RE) for the various estimators of \( P_A \). \( \phi \) goes from 0.5 to 0.9 and \( P_A \) takes the values 0.1, 0.5 and 0.9.

Acknowledgments. The authors would like to thank the referees for their many helpful comments and suggestions. This work is partially supported by Ministerio de Educación y Ciencia (contract No. MTM2009-10055)

References

Figure 2. Values of Relative Efficiency (RE) for the various estimators of $P_A$. $\phi$ goes from 0.5 to 0.9 and $P_A$ takes the values 0.25 and 0.75.