

## A regression approach to estimation in two-phase sampling using two auxiliary variables

J. Sahoo, L. N. Sahoo\* and S. Mohanty

Department of Statistics, Orissa University of Agriculture and Technology, Bhubaneswar 751 003, India

\*Department of Statistics, Utkal University, Bhubaneswar 751 004, India

Using two-phase sampling mechanism, a regression-type estimator in the presence of the available knowledge on second auxiliary variable  $z$  is considered, when the population mean of the main auxiliary variable  $x$  is unknown. The suggested estimator is found to be more efficient than the usual two-phase regression estimator and the regression-type estimators suggested by Kiregyera.

### Two-phase sampling set up

CONSIDER a finite population  $U = \{1, 2, \dots, k, \dots, N\}$ . Let  $y$  and  $x$  be the variable for study and auxiliary variable, taking values  $y_k$  and  $x_k$  respectively for the  $k$ th unit. When the two variables are strongly related but no information is available on the population mean  $\bar{X}$  of  $x$ , we seek to estimate the population mean  $\bar{Y}$  of  $y$  from a sample  $r$ , obtained through a two-phase selection. Allowing simple random sampling (WOR) design in each phase, the two-phase sampling scheme will be as follows:

(a) The first phase sample  $s (s \subset U)$  of fixed size  $n$  is drawn to observe only  $x$  in order to furnish a good estimate of  $\bar{X}$ .

(b) Given  $s$ , the second-phase sample  $r (r \subset s)$  of fixed size  $m$  is drawn to observe  $y$  only.

$$\text{Let } \bar{x}_r = \frac{1}{m} \sum_{k \in r} x_k, \bar{y}_r = \frac{1}{m} \sum_{k \in r} y_k$$

$$\text{and } \bar{x}_s = \frac{1}{n} \sum_{k \in s} x_k.$$

The two-phase sampling regression estimator in this case will be

$$t_{RG} = \bar{y}_r + b_{yx}^{(r)} (\bar{x}_s - \bar{x}_r), \quad (1)$$

where  $b_{yx}^{(r)}$  is the sample regression coefficient of  $y$  on  $x$  computed using data on  $r$ . The mean square error of  $t_{RG}$  to a first order of approximation is

$$\text{MSE}(t_{RG}) = S_y^2 [f_r (1 - \rho_{yx}^2) + f_s \rho_{yx}^2], \quad (2)$$

$$\text{where } f_r = \left( \frac{1}{m} - \frac{1}{N} \right), f_s = \left( \frac{1}{n} - \frac{1}{N} \right),$$

$S_y^2 = \frac{1}{N-1} \sum_{k=1}^N (y_k - \bar{Y})^2$ , and  $\rho_{yx}$  is the correlation coefficient between  $y$  and  $x$ .

### Use of a second auxiliary variable

Sometimes even if  $\bar{X}$  is unknown, information on a cheaply ascertainable variable  $z$ , closely related to  $x$  but compared to  $x$  remotely related to  $y$ , is available on all units of the population. Chand<sup>1</sup>, Kiregyera<sup>2,3</sup>, Sri venkataramana and Tracy<sup>4</sup>, Srivastava *et al*<sup>5,6</sup> gave brief discussions of this type of situation.

Chand<sup>1</sup> suggested a chain ratio-type estimator

$$t_1 = \bar{y}_r \frac{\bar{x}_s \bar{Z}}{\bar{x}_r \bar{z}_s},$$

where  $\bar{Z}$  is the known population mean of  $z$  and  $\bar{z}_s = \frac{1}{n} \sum_{k \in s} z_k$ . He obtained  $t_1$  from the two-phase sampling ratio estimator  $t_R = \bar{y}_r \frac{\bar{x}_s}{\bar{x}_r}$  simply replacing  $\bar{x}_s$  by  $\frac{\bar{x}_s}{\bar{z}_s} \bar{Z}$ .

Mohanty<sup>7</sup> considered regression-cum-ratio estimators

$$t_2 = \frac{\bar{y}_r + b_{yx}^{(r)} (\bar{x}_s - \bar{x}_r)}{\bar{z}_s} \bar{Z}$$

$$\text{and } t_3 = \frac{\bar{y}_r + b_{yx}^{(r)} (\bar{x}_s - \bar{x}_r)}{\bar{z}_r} \bar{Z}$$

where  $\bar{z}_r = \frac{1}{m} \sum_{k \in r} z_k$ .

Modifying  $t_1$ , Kiregyera<sup>2</sup> obtained a ratio-to-regression estimator

$$t_4 = \frac{\bar{y}_r}{\bar{x}_r} [\bar{x}_s + b_{xz}^{(s)} (\bar{Z} - \bar{z}_s)]$$

which is found to be more suitable in this context, where  $b_{xz}^{(s)}$  is the sample regression coefficient of  $x$  on  $z$  based on  $s$ . Kiregyera<sup>3</sup> also extended this formulation to develop a ratio-in-regression estimator

$$t_5 = \bar{y}_r + b_{yx}^{(r)} \left( \frac{\bar{x}_s}{\bar{z}_s} \bar{Z} - \bar{x}_r \right),$$

and a regression-in-regression estimator

$$t_6 = \bar{y}_r + b_{yx}^{(r)} [(\bar{x}_s - \bar{x}_r) - b_{xz}^{(s)} (\bar{z}_s - \bar{Z})].$$

Table 1. Description of populations

Pop No	Source	units (N)	y	x	z
1	Srivastava <i>et al.</i> <sup>1</sup>	21 plants	Pods/plant	no. of primary branches	seeds/pod
2	Sahoo and Swain <sup>8</sup>	50 plants	yield/plant	no. of tillers	percentage of sterility
3	Tripathi <sup>9</sup>	225 persons	persons in service	educated persons	size of the household

The mean square error of  $t_6$ , to the first order of approximation, is

$$MSE(t_6) = S_y^2 [f_r(1 - \rho_{yx}^2) + f_s(\rho_{yx}^2 + \rho_{yx}^2 \rho_{xz}^2 - 2\rho_{yx}\rho_{yz}\rho_{xz})], \quad (3)$$

where  $\rho_{yz}$  and  $\rho_{xz}$  have their usual meanings.

It may be seen that  $t_6$  would be better than  $t_{RG}$ , provided

$$\rho_{yz} > \frac{1}{2} \rho_{yx} \rho_{xz}. \quad (4)$$

But, when  $z$  is remotely related to  $y$  as supposed by Kiregyera, this condition cannot be easily realizable in practice. Therefore,  $t_6$  may not be effectively used in many situations.

*An alternative approach*

Using covariates  $x$  and  $z$  simultaneously, we may consider an unbiased estimator,

$$\mathcal{Y} = \bar{y}_r + d_1(\bar{x}_s - \bar{x}_r) + d_2(\bar{z}_s - \bar{Z}),$$

where  $d_1$  and  $d_2$  are constants, which in particular may be random variables converging in probability to  $D_1$  and  $D_2$  respectively, with  $n \rightarrow N$ . The optimum values of  $d_1$  and  $d_2$  are obtained by minimizing the variance of  $\mathcal{Y}$  given by

$$V(\mathcal{Y}) = (f_r - f_s)(S_y^2 + d_1^2 S_x^2 + 2d_1 S_y S_x \rho_{yx}) + f_s(S_y^2 + d_2^2 S_z^2 + 2d_2 S_y S_z \rho_{yz}), \quad (5)$$

where  $S_x^2 = \frac{1}{N-1} \sum_{k=1}^N (x_k - \bar{X})^2$  and  $S_z^2 = \frac{1}{N-1} \sum_{k=1}^N (z_k - \bar{Z})^2$ .

clearly,  $V(\mathcal{Y})$  is minimized for

$$d_1 = -\beta_{yx} \quad \text{and} \quad d_2 = -\beta_{yz},$$

where  $\beta_{yx}$  and  $\beta_{yz}$  are the regression coefficients of  $y$  on  $x$  and  $y$  on  $z$  respectively.

Usually,  $\beta_{yx}$  and  $\beta_{yz}$  are unknown and can be estimated by  $b_{yx}^{(r)}$  and  $b_{yz}^{(r)}$  respectively, computed using data on  $r$  yielding a regression-type estimator

$$t_{TRG} = \bar{y}_r + b_{yx}^{(r)}(\bar{x}_s - \bar{x}_r) + b_{yz}^{(r)}(\bar{Z} - \bar{z}_s).$$

This estimator may be termed as a chain regression-type estimator analogous to the chain ratio-type estimator of Chand<sup>1</sup>. To the first order of approximation

$$MSE(t_{TRG}) = S_y^2 [f_r(1 - \rho_{yx}^2) + f_s(\rho_{yx}^2 - \rho_{yz}^2)] \quad (6)$$

From equations (2), (3) and (6) we see that the suggested estimator  $t_{TRG}$  is more efficient than  $t_{RG}$  and  $t_6$ . Further, it can be checked that the estimators  $t_R, t_1, t_2, t_3, t_4$  and  $t_5$  are always less efficient than  $t_{TRG}$ .

The method of estimation suggested here can be extended to a three or higher phases of sampling using more than two auxiliary variables.

*Numerical illustrations*

For the purpose of illustration we consider three populations as described in Table 1.

Relative efficiencies of different estimators compared to the mean per unit estimator  $\bar{y}_r$ , are presented in Table 2. As expected, the suggested regression-type estimator  $t_{TRG}$  performed better than other estimators under comparison in case of all the three populations.

Table 2. Relative efficiency of different estimators compared to  $\bar{y}_r$ , (in %).

Estimator	Population		
	1 $n=12, m=5$	2 $n=20, m=8$	3 $n=60, m=25$
$t_R$	147	156	117
$t_{RG}$	159	157	151
$t_1$	127	145	116
$t_2$	134	149	149
$t_3$	80	122	136
$t_4$	146	156	118
$t_5$	147	144	111
$t_6$	159	156	151
$t_{TRG}$	160	162	152



1. Chand, L., Unpublished Ph D dissertation, Iowa State University, Ames, Iowa, 1975.
2. Kiregyera, B., *Metrika*, 1980, **27**, 217-223.
3. Kiregyera, B., *Metrika*, 1984, **31**, 215-226.
4. Srivenkataramana, T. and Tracy, D. S., *Biometrika*, 1989, **76**, 818-821.
5. Srivastava, Rani S., Khare, B. B. and Srivastava, S. R., *Assam Stat. Rev.*, 1988, **2**, 21-29.
6. Srivastava, Rani, Khare, B. B. and Srivastava, S. R., *J. Indian Soc. Agric. Stat.*, 1990, **42**, 108-117.
7. Mohanty, S., *J. Indian Stat. Assoc.*, 1967, **5**, 1-14.
8. Sahoo, L. N. and Swain, A. K. P. C., *Sankhyā*, 1980, **C42**, 56-62.
9. Tripathi, T. P., *Sankhyā*, 1980, **C42**, 63-75.

Received 30 January 1993; accepted 26 March 1993

## Electron transfer chemiluminescence of Buckminsterfullerene radical anion and thianthrene cation

Neeraj Gupta and K. S. V. Santhanam

Chemical Physics Group, Tata Institute of Fundamental Research, Bombay 400 005, India

The electron transfer reaction between  $C_{60}$  (Buckminsterfullerene) radical anion and thianthrene radical cation has been carried out in methylene chloride for the first time, liberating an energy of 1.72 eV which is sufficient to form the first triplet state of  $C_{60}$ . This reaction produces a red chemiluminescence. The excited singlet state appears to be formed by a triplet-triplet annihilation reaction.

RECENTLY several authors<sup>1-5</sup> reported the fluorescence and phosphorescence properties of fullerenes, especially of  $C_{60}$  and  $C_{70}$  in condensed media. The fluorescence (at room temperature) and phosphorescence (only at 77 K) emissions were observed in the red spectral regions (600-800 nm). These studies have established that the first triplet and singlet states energies of  $C_{60}$  are located at  $E_{T_1} = 1.56$  eV and  $E_{S_1} = 2.00$  eV (refs. 1-3).

A question that stems from these studies is that whether it is possible to populate these excited states through chemical reactions involving fullerenes via electron transfer reactions of their radical ions. From a consideration of the above energies and Marcus theory of electron transfer reactions<sup>6,7</sup> it is evident that if a chemical reaction between fullerene radical anion and a suitable radical cation can be initiated with a negative free energy change equivalent to or greater than  $E_{S_1}$  or  $E_{T_1}$  then it should be possible to produce the excited singlet state either directly or via triplet-triplet annihilation<sup>8-10</sup>. As it has so far not been feasible to produce a stable radical cation of fullerene ( $C_{60}$  or  $C_{70}$ ) a reaction involving it with its own radical anion cannot be considered; stable radical anion of fullerenes has, however, been generated electrolytically in a number of studies<sup>11-14</sup>. In the light of this limitation, one will have

to examine the electron transfer reactions with other stable radical cations which could be generated electrochemically. Thianthrene radical cation is a good choice for this purpose as it can be generated electrochemically and is stable for several minutes<sup>15,16</sup>. So we report here the first case of an electron transfer reaction between fullerene radical anion and thianthrene radical cation to produce a chemiluminescence.

The experiments were conducted using pure  $C_{60}$  (99.5%) (Strem Chemicals, USA) or the carbon soot which contains a mixture of  $C_{60}$  and  $C_{70}$  (2-20% of  $C_{60}/C_{70}$ ) extracted with  $CH_2Cl_2$  (supplied by Strem Chemicals, USA).

The current-voltage curve of solution of  $\sim 1$  mM Buckminsterfullerene ( $C_{60}$ ) in  $CH_2Cl_2$  (SDS Fine Chemicals Ltd., b.p. 40°C) containing 0.1 M tetra-*n*-butyl ammonium fluoborate ( $TBABF_4$ , South Western Analytical Chemicals Company Inc.) showed three successive one electron cathodic peaks at  $E_{pc}^I = -0.45$  V,  $E_{pc}^{II} = -0.88$  V and  $E_{pc}^{III} = -1.25$  V vs. SCE. The significant characteristics of cyclic voltammetry,  $E_p - E_{p/2}$ ,  $E_{pa} - E_{pc}$  and  $i_{pa}/i_{pc}$  are suggestive of a stable reduction product. Upon sweeping the potential towards positive potentials, there was no anodic peak within the potential window of this solvent. With a mixture of  $C_{60}$  and thianthrene (Aldrich Chemical Co., m.p. 154-156°C) in the above medium, an anodic peak due to the oxidation of thianthrene is observed at  $E_{pa}^I = +1.32$  V vs. SCE. At the negative potentials the thianthrene reduction could not be observed within the potential limits available in this medium. Experiments carried out with 1 mg of carbon soot ( $C_{60} + C_{70}$ ), 10 mM thianthrene and 0.1 M TBA  $BF_4$  in  $CH_2Cl_2$  showed the cathodic and anodic peaks as shown in Figure 1. No appreciable shift in peak potentials were observed by the presence of  $C_{70}$  in the medium. Although the original recordings were done with a

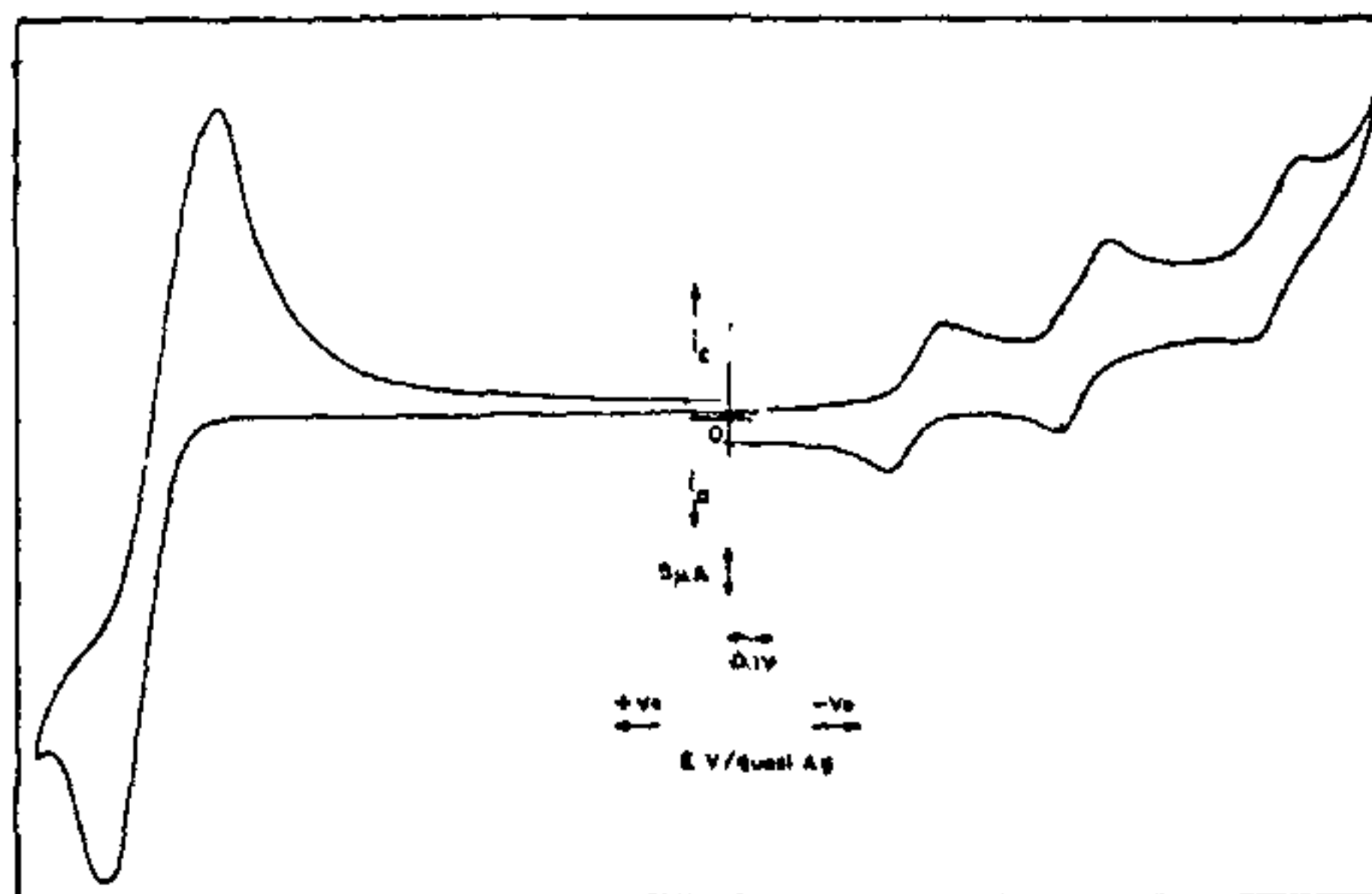


Figure 1. Cyclic voltammograms of the fullerene ( $C_{60} + C_{70}$ ) and 10 mM thianthrene mixed system in methylene chloride containing 0.1 M TBA  $BF_4$ . The area of the platinum wire electrode is 0.014  $cm^2$ . Scan rate = 50 mV/s.