APPLICATIONES MATHEMATICAE
27,3 (2000), pp. 335–342

A. CZAPKIEWICZ and A. L. DAWIDOWICZ (Kraków)

THE TWO-DIMENSIONAL LINEAR RELATION IN THE ERRORS-IN-VARIABLES MODEL WITH REPLICATION OF ONE VARIABLE

Abstract. We present a two-dimensional linear regression model where both variables are subject to error. We discuss a model where one variable of each pair of observables is repeated. We suggest two methods to construct consistent estimators: the maximum likelihood method and the method which applies variance components theory. We study asymptotic properties of these estimators. We prove that the asymptotic variances of the estimators of regression slopes for both methods are comparable.

1. Introduction. A problem sometimes encountered in data analysis is to find a relation between two or more variables. In this paper we discuss the two-dimensional case, where both observables are not measured precisely. Thus let us consider the model

$$X_i = s_i + \varepsilon_i, \qquad Y_i = as_i + b + \delta_i, \qquad i = 1, \ldots, n, \tag{1}$$

where the disturbance errors ε_i and δ_i are independent random variables, with mean and variance equal to zero and $\sigma_\varepsilon^2, \sigma_\delta^2$, respectively. We assume s_i to be an unknown constant. This case is known in the literature as a functional model (Kendall and Stuart 1979). It is well known (Reiersol 1950) that this model, with errors having normal distributions with unknown variances, is nonidentifiable. To overcome this difficulty we need an additional assumption, for example, that the distribution of errors is nonnormal or that either one error variance is known or the ratio of the variances are known. Another approach to construct consistent estimators of regression slopes in model (1) is repeating the random variables X_i, Y_i m_i times (Cox 1976,

2000 *Mathematics Subject Classification*: Primary 62J05; Secondary 62F10, 62J12.
Key words and phrases: linear regression, consistent estimator.

Dolby 1976, Bunke and Bunke 1989). In this case we have

$$(2)\qquad X_{ij} = s_i + \varepsilon_{ij}, \quad Y_{ij} = as_i + b + \delta_{ij}, \quad i = 1, \ldots, n,\ j = 1, \ldots, m_i.$$

In this paper we consider a particular case of the model with replications. We will prove that repeating only one variable, for example Y_i, enables us to construct consistent estimators of the unknown parameters of the linear relation.

We discuss the model

$$(3)\qquad X_i = s_i + \varepsilon_i, \quad Y_{ij} = as_i + b + \delta_{ij}, \quad i = 1, \ldots, n,\ j = 1, \ldots, m.$$

The variables X_i, Y_{ij} are observables, the variables s_i are unknown constants and ε_i, δ_{ij} are assumed to have independent normal distribution with mean zero and unknown variances σ_ε^2 and σ_δ^2:

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2), \qquad \delta_{ij} \sim N(0, \sigma_\delta^2).$$

For constructing consistent estimators of the unknown parameters we present two methods: the maximum likelihood method and a method (Czapkiewicz 1999) based on variance components theory. We compare these two methods by comparing the mean squared errors.

2. Maximum likelihood method

2.1. *Methodology.* We can express the observations X_i, Y_{ij} in (3) as

$$z_i = [X_i, Y_{i1}, \ldots, Y_{im}]', \qquad i = 1, \ldots, n.$$

The independent random vectors z_i have means depending on i:

$$\mu_i = [s_i, as_i + b, \ldots, as_i + b]'$$

and a common $(m+1) \times (m+1)$ covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \ldots & 0 \\ 0 & \sigma_\delta^2 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & \sigma_\delta^2 \end{bmatrix}.$$

The log-likelihood function has the form

$$L(\theta) = \text{const} - n \ln \sigma_\varepsilon - nm \ln \sigma_\delta - \frac{1}{2}\Big[\sum_{i=1}^{n} \frac{(X_i - s_i)^2}{\sigma_\varepsilon^2} + \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{(Y_{ij} - as_i - b)^2}{\sigma_\delta^2}\Big]$$

where $L(\theta) = L(a, s_1, \ldots, s_n, b, \sigma_\varepsilon, \sigma_\delta)$. Solving the log-likelihood equations is not easy. Cox (1976) gives the solutions for model (2) where both X_i and Y_i are repeated m times. When we assume that $X_{ij} = X_i$ for each j in Cox' model, we can use his solutions for our purposes.

To write estimators, set

$$s_{yy} = \sum_{i=1}^{n}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_{i.})^2/(nm), \qquad b_{yy} = \sum_{i=1}^{n}(\overline{Y}_{i.} - \overline{Y})^2/n,$$

$$b_{xx} = \sum_{i=1}^{n}(X_i - \overline{X})^2/n, \qquad b_{xy} = \sum_{i=1}^{n}(X_i - \overline{X})(\overline{Y}_{i.} - \overline{Y})/n$$

and

$$B(a) = b_{yy} - 2ab_{xy} + a^2 b_{xx}.$$

Solving the likelihood equations we get estimators in terms of a:

$$\begin{aligned}\widehat{b} &= \overline{Y} - a\overline{X},\\ \widehat{\sigma}_\varepsilon^2 &= s_{yy} + (b_{yy} - ab_{xy})^2/B(a),\\ \widehat{\sigma}_\delta^2 &= s_{xx} + (ab_{xx} - b_{xy})^2/B(a),\\ \widehat{s}_i &= \big((ab_{xx} - b_{xy})(\overline{Y}_{i.} - \overline{Y} + a\overline{X}) + (b_{yy} - ab_{xy})X_i\big)\,/B(a).\end{aligned} \tag{4}$$

But to get an estimator of a we must solve an equation of the fourth degree in a:

$$-s_{yy}(ab_{xx} - b_{yx})B(a) - (b_{yy} - ab_{xy})(ab_{xx} - b_{xy})(b_{yy} - a^2 b_{xx}) = 0. \tag{5}$$

When $m > 2$ we solve (5) numerically and then check whether the absolute maximum has been found.

2.2. *Asymptotic behaviour of maximum likelihood estimators.* In this section we look for the asymptotic properties of maximum likelihood estimators in the model discussed in the previous section. The random vectors z_i are independent, with normal but not identical distribution. The expectations of their distributions depend on i. The number of unknown parameters which we estimate increases with n.

Assume that s_i, $i = 1, \ldots, n$, belong to a bounded set as n tends to infinity and the following two limits exist:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} s_i \quad\text{and}\quad \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} s_i^2.$$

Then we can prove:

LEMMA 1. *When $n \to \infty$ and $m \to \infty$, the solutions of the likelihood equations give strongly consistent estimators of the unknown parameters $a, b, \sigma_\delta, \sigma_\varepsilon$. For sufficiently large n and m, the variance of the vector*

$$[\widehat{a} - a, \widehat{s}_1 - s_1, \ldots, (\widehat{s}_n - s_n)(\widehat{b} - b), \widehat{\sigma}_\varepsilon - \sigma_\varepsilon, \widehat{\sigma}_\delta - \sigma_\delta] \tag{6}$$

can be approximated by

$$\left[-E\left(\frac{\partial^2}{\partial\xi\partial\phi}L(\theta)\right)\right]^{-1} \tag{7}$$

where ξ, ϕ belong to the set of unknown parameters.

This lemma may be proved by a method analogous to that described in Lehmann's monograph (1983, p. 404, Th. 4.1). We thus have the following asymptotic variances of unknown regression slopes:

THEOREM 1. *When n and m are large, the asymptotic variances of $\widehat{a}$ and $\widehat{b}$*, $\operatorname{avar}(\widehat{a})$ *and* $\operatorname{avar}(\widehat{b})$, *are*

$$\operatorname{avar}(\widehat{a}) = \frac{ma^2\sigma_\varepsilon^2 + \sigma_\delta^2}{m\sum_{i=1}^n (s_i - \bar{s})^2}, \tag{8}$$

$$\operatorname{avar}(\widehat{b}) = \frac{ma^2\sigma_\varepsilon^2 + \sigma_\delta^2}{mn} \cdot \frac{\sum_{i=1}^n s_i^2}{\sum_{i=1}^n (s_i - \bar{s})^2}. \tag{9}$$

Proof. To show the formula for $\operatorname{avar}(\widehat{a})$, let us calculate $\partial L/\partial\xi\partial\phi$ where $\xi, \phi \in \{a, s_1, \ldots, s_n, b, \sigma_\varepsilon, \sigma_\delta\}$. The matrix (7) has the form

$$\Theta_n^{-1} = \begin{bmatrix} \frac{m}{\sigma_\delta^2}\sum s_i^2 & \frac{ma}{\sigma_\delta^2}s' & \frac{m}{\sigma_\delta^2}\sum s_i & 0 & 0 \\ \frac{ma}{\sigma_\delta^2}s & \frac{ma^2\sigma_\varepsilon^2+\sigma_\delta^2}{\sigma_\varepsilon^2\sigma_\delta^2}I_n & \frac{am}{\sigma_\delta^2}1_n & 0 & 0 \\ \frac{m}{\sigma_\delta^2}\sum s_i & \frac{am}{\sigma_\delta^2}1_n' & \frac{mn}{\sigma_\delta^2} & 0 & 0 \\ 0 & \ldots & 0 & \frac{2n}{\sigma_\varepsilon^2} & 0 \\ 0 & \ldots & 0 & 0 & \frac{2nm}{\sigma_\delta^2} \end{bmatrix}^{-1}$$

where $s = (s_1, \ldots, s_n)'$ and 1_n is the n-dimensional vector of ones.

Let us partition Θ_n as

$$\Theta_n = \begin{bmatrix} \frac{m}{\sigma_\delta^2}\sum s_i^2 & w' \\ w & M \end{bmatrix}.$$

The element in Θ_n^{-1} which is the required asymptotic variance of $\widehat{a}$ can be obtained by a standard result on the inverse of a partition matrix:

$$\operatorname{avar}(\widehat{a}) = \left(\frac{m}{\sigma_\delta^2}\sum s_i^2 - w'M^{-1}w\right)^{-1}.$$

But

$$M^{-1} = \begin{bmatrix} Q^{-1} & 0 \\ 0 & T^{-1} \end{bmatrix}$$

where

$$Q^{-1} = \begin{bmatrix} \frac{ma^2\sigma_\varepsilon^2+\sigma_\delta^2}{\sigma_\varepsilon^2\sigma_\delta^2}I_n & \frac{am}{\sigma_\delta^2}1_n \\ \frac{am}{\sigma_\delta^2}1_n' & \frac{mn}{\sigma_\delta^2} \end{bmatrix}^{-1} \quad \text{and} \quad T^{-1} = \begin{bmatrix} \frac{2n}{\sigma_\varepsilon^2} & 0 \\ 0 & \frac{2nm}{\sigma_\delta^2} \end{bmatrix}^{-1}$$

so

$$\operatorname{avar}(\widehat{a}) = \left(\frac{m\sum s_i^2}{\sigma_\delta} - q'Q^{-1}q - t'T^{-1}t\right)^{-1} \tag{10}$$

where $q' = \left[\frac{ma}{\sigma_\delta^2}s', \frac{m}{\sigma_\delta^2}\sum s_i\right]$ and $t' = [0, 0]$.

Taking into account that $t'T^{-1}t = 0$ and inserting the calculated value of $q'Q^{-1}q$ into expression (10) we obtain the asymptotic variance of the estimator $\widehat{a}$.

To obtain the asymptotic variance of the estimator $\widehat{b}$, $\operatorname{avar}(\widehat{b})$, we repeat the previous argument for the matrix of $\partial L/\partial\xi\partial\phi$ where ξ, ϕ are taken in the order $\{b, s_1, \ldots, s_n, a, \sigma_\varepsilon, \sigma_\delta\}$. ∎

3. Variance components estimation method. In this section we present another method of estimating unknown parameters in model (3). This method (Czapkiewicz 1999) is based on some properties of a linear model with two variance components. We discuss the model

$$X_i = s_i + \varepsilon_i, \quad Y_{ij} = as_i + b + \delta_{ij}, \quad i = 1, \ldots, n, \ j = 1, \ldots, m.$$

If we substitute s_i in the last formula we obtain

$$Y_{ij} = aX_i + b + \delta_{ij} - a\varepsilon_i. \tag{11}$$

Replacement of the distribution of (X_i, Y_{ij}) by the conditional distribution of Y_{ij} with respect to X_i enables us to use a different model (treating X_i as a constant) to estimate the same parameters $a, b, \sigma_\delta, \sigma_\varepsilon$ as in model (3). The technique of variance components can be applied for this purpose.

We obtain a model

$$Y = X\beta + U_1\Phi_1 + U_2\Phi_2,$$

where β is a vector of unknown parameters a and b,

$$Y = [y_1, \ldots, y_n]', \quad y_i = [Y_{i1}, \ldots, Y_{im}],$$

$$X = \begin{bmatrix} X_1 1_m & 1_m \\ \vdots & \vdots \\ X_n 1_m & 1_m \end{bmatrix},$$

the matrix U_1 is

$$U_1 = I_n \otimes 1_m,$$

whereas U_2 is the $nm \times nm$ unit matrix. The vectors Φ_1, Φ_2 are

$$\Phi_1 = [\gamma_1, \ldots, \gamma_n]', \quad \Phi_2 = [\delta_{11}, \ldots, \delta_{1m}, \ldots, \delta_{n1}, \ldots, \delta_{nm}]'$$

where $\gamma_i = -a\varepsilon_i$. The variance components are

$$\sigma_1^2 = a^2\sigma_\varepsilon^2, \quad \sigma_2^2 = \sigma_\delta^2. \tag{12}$$

First we recall the following result:

Theorem 2. *The uniformly best*, *invariant unbiased estimators of* σ_1 *and* σ_2 *are*

$$\widetilde{\sigma}_1 = \frac{nm-2}{m^2(n-2)(m-1)n} Y'MVMY - \frac{1}{mn(m-1)} Y'MY, \tag{13}$$

$$\widetilde{\sigma}_2 = \frac{1}{n(m-1)} Y'MY - \frac{1}{mn(m-1)} Y'MVMY. \tag{14}$$

The estimator of $\widetilde{\beta} = [\widetilde{a}, \widetilde{b}]'$ *has the form*

$$\widetilde{\beta} = (X'\widetilde{Z}^{-1}X)^{-1}X'\widetilde{Z}^{-1}Y \tag{15}$$

where $\widetilde{Z} = \widetilde{\sigma}_1 V + \widetilde{\sigma}_2 I_{mn}$.

The proof of this theorem is given in Czapkiewicz (1999). Now, we prove the following theorem:

THEOREM 3. *The estimators of unknown parameters, based on variance components theory, have the following properties:*

(i) *The estimator defined in* (15) *does not depend on the values of* $\widetilde{\sigma}_1$ *and* $\widetilde{\sigma}_2$.

(ii) *The estimator* $\widetilde{\beta}$ *has a normal distribution with expectation* β *and covariance matrix*

$$(m\sigma_1^2 + \sigma_2^2)(X'X)^{-1} = (ma^2\sigma_\varepsilon^2 + \sigma_\delta^2)(X'X)^{-1}. \tag{16}$$

(iii) *The estimator* $\widetilde{\beta}$ *is unbiased with minimal covariance matrix in the class of linear unbiased estimators, the estimator* $\widetilde{\sigma}_\delta$ *is the uniformly best unbiased estimator of* σ_δ, *and the estimator*

$$\widetilde{\sigma}_\varepsilon = \sqrt{\widetilde{\sigma}_2^2/\widetilde{a}^2}$$

is weakly consistent.

Proof. A simple calculation shows that for every p and q we have

$$X'(pV + qI_{mn}) = (mp + q)X'. \tag{17}$$

(i) From (17) we have

$$X'\widetilde{Z} = X'(\widetilde{\sigma}_1^2 V + \widetilde{\sigma}_2^2 I_{mn}) = (m\widetilde{\sigma}_1^2 + \widetilde{\sigma}_2^2)X',$$

so

$$X' = (m\widetilde{\sigma}_1^2 + \widetilde{\sigma}_2^2)X'(\widetilde{Z})^{-1}$$

and

$$\widetilde{\beta} = (m\widetilde{\sigma}_1^2 + \widetilde{\sigma}_2^2)(X'X)^{-1}\frac{1}{m\widetilde{\sigma}_1^2 + \widetilde{\sigma}_2^2}X'Y = (X'X)^{-1}X'Y.$$

(ii) If we assume that Y has a normal distribution, then the estimator $\widetilde{\beta}$, which is a linear function of Y, also has a normal distribution. The expectation of $\widetilde{\beta}$ is

$$E((X'X)^{-1}X'Y) = (X'X)^{-1}X'E(Y) = (X'X)^{-1}X'X\beta = \beta.$$

The covariance matrix of $\widetilde{\beta}$, $\mathrm{Var}(\widetilde{\beta})$, is

$$\begin{aligned}\mathrm{Var}(\widetilde{\beta}) &= E((X'X)^{-1}X'YY'X(X'X)^{-1}) - \beta\beta' \\ &= (X'X)^{-1}X'E(YY')X(X'X)^{-1} - \beta\beta'.\end{aligned}$$

Because

$$E(YY') = \sigma_1^2 V + \sigma_2^2 I + (X\beta)(X\beta)',$$

from (17) we obtain

$$\mathrm{Var}(\widetilde{\beta}) = (m\sigma_1^2 + \sigma_2^2)(X'X)^{-1}.$$

(iii) Let us consider another linear unbiased estimator $L'Y$ of β. We wll prove that its covariance matrix is not smaller than the covariance matrix of $(X'X)^{-1}X'Y$ (i.e. the difference between these matrices is non-negative definite). Set

$$A' = L' - (X'X)^{-1}X'.$$

Notice that $E(A'Y) = 0$ and $A'X = 0$. From this and from (17) we have

$$\begin{aligned}E(A'Y((X'X)^{-1}X'Y)') &= A'E(YY')X(X'X)^{-1} \\ &= A'(\sigma_1^2 V + \sigma_2^2 I_{mn})X(X'X)^{-1} \\ &= (m\sigma_1^2 + \sigma_2^2)A'X(X'X)^{-1} = 0.\end{aligned} \tag{18}$$

Now we write $\mathrm{Var}(L'Y)$ as

$$\mathrm{Var}(L'Y) = \mathrm{Var}(L'Y - (X'X)^{-1}X'Y + (X'X)^{-1}X'Y).$$

By (18),

$$\mathrm{Var}(L'Y) = \mathrm{Var}(L'Y - (X'X)^{-1}X'Y) + \mathrm{Var}((X'X)^{-1}X'Y).$$

The first component is a non-negative definite matrix, so we have

$$\mathrm{Var}(L'Y) \geq \mathrm{Var}((X'X)^{-1}X'Y).$$

The properties of $\widetilde{\sigma}_\delta$ follow from Theorem 2 whereas the properties of $\widetilde{\sigma}_\varepsilon$ follow from Słucki's Theorem (see e.g. Bartoszewicz 1989, p. 53, Th. 5.3). ■

Remark. The variances of the estimators $\widetilde{a}$ and $\widetilde{b}$ are

$$\mathrm{var}(\widetilde{a}) = \frac{ma^2\sigma_\varepsilon^2 + \sigma_\delta^2}{m\sum_{i=1}^n (X_i - \overline{X})^2}, \tag{19}$$

$$\mathrm{avar}(\widetilde{b}) = \frac{ma^2\sigma_\varepsilon^2 + \sigma_\delta^2}{mn} \cdot \frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i - \overline{X})^2}. \tag{20}$$

Conclusion. The variances of the estimators of a and b obtained using the maximum likelihood method and the theory of variance components are comparable. The differences between formulas (8), (19) and (9), (20) result from differences in the definitions of the models from which we estimated the same parameter a or b.

References

J. Bartoszewicz (1989), *Lectures in Mathematical Statistics*, PWN, Warszawa (in Polish).

O. Bunke and H. Bunke (1989), *Non-Linear Regression, Functional Relationships, and Robust Methods*, Wiley, New York.

A. Czapkiewicz (1999), *On estimation of parameters in the bivariate linear errors-in-variables model*, Appl. Math. (Warsaw) 25, 401–410.

N. R. Cox (1976), *The linear structural relation for several groups of data*, Biometrika 63, 231–237.

G. R. Dolby (1976), *The ultrastructural relation: A synthesis of the functional and structural relations*, Biometrika 63, 39–50.

W. A. Fuller (1987), *Measurement Error Models*, Wiley, New York.

S. Gnot (1991), *Estimation of Variance Components in Linear Models*, Wyd. Naukowo-Techniczne, Warszawa (in Polish).

M. G. Kendall and A. Stuart (1979), *The Advanced Theory of Statistics*, Vol. 2, Griffin, London.

E. L. Lehmann (1983), *Theory of Point Estimation*, Wiley, New York.

A. Olsen, J. Seely and D. Birkes (1976), *Invariant quadratic unbiased estimation for two variance components*, Ann. Statist. 4, 878–890.

C. R. Rao and J. Kleffe (1988), *Estimation of Variance Components and Applications*, North-Holland Ser. Statist. Probab. 3, North-Holland, Amsterdam.

O. Reiersol (1950), *Identifiability of a linear relation between variables which are subject to error*, Econometrica 18, 575–589.

Anna Czapkiewicz
Faculty of Management
University of Mining and Metallurgy
Gramatyka 10, 30-067 Kraków
E-mail: gzrembie@cyf-kr.edu.pl

Antoni Leon Dawidowicz
Institute of Mathematics
Jagiellonian University
Reymonta 4/510, 30-059 Kraków
E-mail: dawidowi@im.uj.edu.pl

Received on 17.6.1999;
revised version on 25.10.1999