Journal of Econometrics 41 (1989) 267–278. North-Holland

ESTIMATION OF REGRESSION COEFFICIENTS WITH THE HELP OF CHARACTERISTIC FUNCTIONS*

Kees VAN MONTFORT and Ab MOOIJAART

Leiden University, 2312 KM Leiden, The Netherlands

Jan DE LEEUW

University of California, Los Angeles, CA 90024, USA

Received September 1987, final version received August 1988

In this paper some easily applicable estimators of the regression coefficient in a simple regression model with errors in variables are derived. Usually one uses sample moments to estimate the regression coefficient. However, we use empirical characteristic functions in several points. Besides theoretical derivations a small study with generated data is discussed. This study compares our method with estimating methods which use higher-order moments.

1. Introduction

In this paper we discuss regression with errors in variables. We shall treat the bivariate case only, although generalizations to the multivariate case are obvious from our treatment. The observed random variables x and y satisfy the following relations:

$$y = \alpha + \beta\xi + \varepsilon, \qquad x = \xi + \delta,$$

where ξ, ε, and δ are unobserved random variables which are independent of each other, $\mathrm{E}\varepsilon = 0$, and $\mathrm{E}\delta = 0$. We have assumed nothing about the distribution of the variables, except that it is assumed that the means of the error variables are zero, while ξ, ε, and δ are independent of each other.

Expressing the first- and second-order moments of x and y, if they exist, in model parameters, we generate five equations. These five equations contain six unknowns. The regression coefficient β is not identified from these first- and second-order moments of x and y. Van Montfort et al. (1987) discuss in detail the identifiability of the regression coefficient β in our model. Using third-order moments of x and y, they identify β. Furthermore they derive an optimal estimator in the class of consistent β estimators which are functions of the

*The authors are grateful to the associate editor for his constructive and helpful comments.

moments of x and y up to order three. Reiersøl (1941), Geary (1942), and Pakes (1979) also use third- and fourth-order moments for estimating the regression coefficient β.

Reiersøl (1950) shows that β is identifiable if there exists a nonzero (finite or infinite) cumulant κ_{rs} of the joint distribution of x and y, with $r \geq 1$, $s \geq 1$, and either r or s but not both equal to one. So we have the result that β is identifiable if ξ is not normally distributed. In this paper we assume that ξ is not normally distributed, and we estimate β with the use of *empirical characteristic functions*. Of course a link exists between characteristic functions and higher-order moments. If $c(.)$ is a characteristic function of the arbitrary stochastic variable v, then the moments $m_r = \mathrm{E}v^r$ $(r = 1, 2, \ldots)$ of v satisfy the following relation:

$$m_r = (-i)^r \{ \partial c(t) / \partial t^r \} |_{t=0}.$$

From the formula mentioned above follows that the moments of a stochastic variable, if they exist, represent the information of the corresponding characteristic function in the point $t = 0$.

In this paper we identify β with the use of the values of the characteristic functions of x and y in more than one point. So we don't use the information of the characteristic functions of x and y in the point $t = 0$ only. We also use the information of the characteristic functions of x and y in other points on the real line.

Our estimation method has two advantages over methods which use higher-order moments. In the first place we can apply our method for a larger class of distribution functions of x and y. This is because every distribution function has a characteristic function, even if the higher-order moments do not exist. For instance, the Cauchy distribution has no second- and higher-order moments, while the characteristic function exists. In the second place, for small sample sizes higher-order moments estimation methods get bad estimations. This is a corollary of the fact that higher-order sample moments have very big variances for small sample sizes.

In the past Darmois (1940) already used empirical characteristic functions for estimating coefficients in some simple statistical models. For instance, he estimates the regression coefficients in the classical multiple-regression model $(y = \alpha + \beta_1 x_1 + \beta_2 x_2)$. Some others also use empirical characteristic functions for estimating β in simple regression models with errors in the variables. Neyman (1951) gives a consistent estimator when ξ is not normally distributed. Assuming the model is identifiable, Wolfowitz (1952) gives an estimating method for β. This method is often reasonable if the distributions of δ and ε belong to a known small finite-dimensional class, for instance when they are normal. Rubin (1956) gives an estimate of β when the errors δ and ε are normal and ξ is not normal. This estimate converges to β with probability one. Assuming ξ is not normally distributed, Spiegelman (1979) gives a class of estimators of β which are asymptotically normal with mean β and variance proportional to $n^{-1/2}$, under weak assumptions. He shows how to choose a good estimate of β from this class.

In the second section of this paper we approximate the logarithm of the characteristic function of ξ by an algebraic polynomial of order three. With the help of this approximated characteristic function in several points we derive some easily applicable estimators of the regression coefficient β in our model. In section 3 we use algebraic polynomials of order four, five, six, etc. to deduce estimators of β. Furthermore, in section 4, we consider the characteristic functions of our variables in points which tend to zero. In this case the logarithms of the characteristic functions of x and y are 'approximately' polynomials. It follows that the estimators of the second and third section are 'approximately' functions of bivariate cumulants of x and y.

Using our estimating method, in section 5 we discuss some examples with generated data. Furthermore, we compare our method with estimation methods which use higher-order moments. Finally, in the appendix we discuss some elementary properties of empirical characteristic functions.

2. Estimators of the regression coefficient β with the use of empirical characteristic functions

In this section we derive some simple estimators of the regression coefficient β with the use of empirical characteristic functions. In the first place we define the simultaneous characteristic function $c(s, t)$ of y and x and the characteristic functions $e(t)$, $a(t)$, and $b(t)$ of ξ, $\varepsilon + \alpha$, and δ. For every s and t on the real line, we get

$$\ln c(s, t) = \ln a(s) + \ln b(t) + \ln e(\beta s + t). \tag{2}$$

From (2) it follows that

$$\ln c(s, t) - \ln c(s, 0) - \ln c(0, t) = \ln e(\beta s + t) - \ln e(\beta s) - \ln e(t). \tag{2'}$$

The left-hand side and the right-hand side of (2) are multiple-valued functions. The single-valued interpretation of (2) is true for the real part, but for the imaginary part we have equality modulo 2π. Thus, for instance, we have for eq. (2)

$$\operatorname{Im}[\ln c(s, t)] = \operatorname{Im}[\ln a(s) + \ln b(t) + \ln c(\beta s + t)] + 2\pi K,$$

with K an arbitrary integer. This is a bit inconvenient to work with numerically. Therefore, we assume that the logarithms of the characteristic functions in system (2) are real-valued. This assumption causes no loss of generality. If the logarithms of the characteristic functions $a(.)$, $b(.)$, $c(.,.)$, or $e(.)$ are not real, we can replace them by their real parts, for instance $\mathrm{Re}[\ln c(s,t)] = \ln\{(\mathrm{Re}\, c(s,t))^2 + (\mathrm{Im}\, c(s,t))^2\}^{1/2}$. In that case the derivations of this paper do not change and we can derive analogous results.

Now we want to derive an expression for β with the use of (2′). We have two possibilities:

- choosing an explicit expression for the density function of ξ,
- approximating the characteristic function $e(.)$ of ξ by a 'simple approximation function'.

Because we don't want to make assumptions about the distribution of ξ, we choose the second possibility. In this paper we approximate $\ln e(t)$ by an algebraic polynomial function.

We prefer the algebraic polynomial approximation to the approximation by trigonometric polynomials. Theorem 16.5 of Powell (1981) shows that for every continuous function $e(.)$ on an interval $[a, b]$ the trigonometric polynomial approximation of order r has a bigger 'approximation fault' than the approximation by algebraic polynomials of order r. Furthermore approximation by splines leads to technical complications. If $e(.)$ is a real-valued function on the interval $[-1,1]$ that is k times continuously differentiable, then the corollary on p. 197 of Powell (1981) says that

$$\min_{p_r \in AP_r} \|e - p_r\|_\infty \le \left\{ \left((r-k)!(\pi/2)^k / r! \right) \|e^{(k)}\|_\infty \right\},$$

where AP_r is the class of all algebraic polynomials of order r and $k = 0, \dots, r$. Defining $e(.)$ on another interval, we can derive a similar result. So on every interval $[a, b]$ we can approximate $e(.)$ as closely as we want.

Now we approximate $\ln e(.)$ by the algebraic polynomial $p_3(.)$ of order three:

$$p_3(t) = a_1 t + a_2 t^2 + a_3 t^3. \tag{3}$$

With the help of approximation (3) we can approximate the right-hand side of eq. (2′) by the polynomial

$$P_3(s,t) = a_2 2\beta st + a_3\left(3\beta^2 s^2 t + 3\beta s t^2\right). \tag{4}$$

From (2′) and (4) it follows that, for $s, t \in R$,

$$\ln c(s,t) - \ln c(s,0) - \ln c(0,t)$$
$$\approx P_3(s,t) = a_2 2\beta st + a_3\left(3\beta^2 s^2 t + 3\beta s t^2\right). \tag{5}$$

If we choose in (5) for (s, t) the values (s_1, t_1), (s_2, t_2), and (s_3, t_3), then we get the system

$$C \approx Mv, \tag{6}$$

with

$$C = \left(C(s_1, t_1), C(s_2, t_2), C(s_3, t_3)\right)^T,$$

$$C(s, t) = \ln c(s, t) - \ln c(s, 0) - \ln c(0, t),$$

$$M = \begin{bmatrix} 2s_1t_1 & 3s_1t_1^2 & 3s_1^2t_1 \\ 2s_2t_2 & 3s_2t_2^2 & 3s_2^2t_2 \\ 2s_3t_3 & 3s_3t_2^2 & 3s_3^2t_2 \end{bmatrix},$$

$$v = \left(a_2\beta \quad a_3\beta \quad a_3\beta^2\right)^T.$$

If $\det M = 18\, s_1s_2s_3t_1t_2t_3\,(s_3t_2 + s_2t_1 + s_1t_3 - s_1t_2 - s_2t_3 - s_3t_1) \neq 0$, then the matrix M is nonsingular and the vector v is identified by system (6). Because β is the quotient of the third and second component of the vector v, assuming $a_3 \neq 0$, β is also identified.

In system (6) we have to choose the values of s_1, s_2, s_3, t_1, t_2, and t_3 in such a way that the matrix M is nonsingular. Furthermore, we have to choose s_1, s_2, s_3, t_1, t_2, and t_3 in the neighbourhood of zero. If we approximate $\ln e(t)$, $\ln e(\beta s)$, and $\ln e(\beta s + t)$ by polynomials with the help of small values of s_i and t_i $(i = 1, \ldots, 3)$, for instance $s_i, t_i \in [-0.1; 0.1]$, then the higher-order terms are negligible. Finally, we can estimate $C(s, t)$ consistently by the estimator $\hat{C}(s, t) = \ln \hat{c}(s, t) - \ln \hat{c}(s, 0) - \ln \hat{c}(0, t)$, where $\hat{c}(s, t)$ is the simultaneous empirical characteristic function corresponding to $c(s, t)$. So with the help of (6) we can estimate β. We can even find a set of estimators of β by choosing several sets of points (s_1, t_1), (s_2, t_2), and (s_3, t_3). Generally the above-mentioned estimators are not consistent.

Remark 1. If ξ is normally distributed, then $\operatorname{Re} \ln e(t) = -(t^2\sigma_\xi)/2$ and $a_3 = 0$. In that special case it follows from system (6) that the second and the third component of the vector v are zero, and our method cannot be used to estimate β.

Remark 2. Assuming the polynomial approximations to be perfect, estimators can be derived for the asymptotic variances of our estimators. With the use of the delta method we can express the asymptotic variances of our estimators in terms of asymptotic variances and covariances of the simultaneous empirical characteristic function of x and y in some points. The formulas for the asymptotic variances and covariances of simultaneous empirical characteristic functions in several points are given in the appendix. Because the expressions for the asymptotic variances of our estimators are very awkward, we don't present them in this paper.

3. Approximating by polynomials of order four and higher

Instead of approximating $\ln e(t)$ by a polynomial of order three, we can approximate $\ln e(t)$ by polynomials of order four and higher. Let $\ln e(.)$ be approximated by the following polynomial of order four:

$$p_4(t) = a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4.$$

Now we find

$$P_4(s, t) = a_2 2\beta st + a_3(3\beta^2 s^2 t + 3\beta st^2) + a_4(4\beta^3 s^3 t + 6\beta^2 s^2 t^2 + 4\beta st^3).$$

In the same way as in the previous section we can derive

$$C \approx Mv, \tag{7}$$

with

$$C = (C(s_1, t_1), \dots, C(s_6, t_6))^T,$$

$$M = \begin{bmatrix} 2s_1t_1 & 3s_1t_1^2 & 3s_1^2t_1 & 4s_1t_1^3 & 6s_1^2t_1^2 & 4s_1^3t_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2s_6t_6 & 3s_6t_6^2 & 3s_6^2t_6 & 4s_6t_6^3 & 6s_6^2t_6^2 & 4s_6^3t_6 \end{bmatrix},$$

$$v = (a_2\beta, a_3\beta, a_3\beta^2, a_4\beta, a_4\beta^2, a_4\beta^3)^T$$

$$= (v_{1,1}, v_{1,2}, v_{2,1}, v_{1,3}, v_{2,2}, v_{3,1})^T.$$

$\beta \approx v_{2,1}/v_{1,2} \approx v_{2,2}/v_{1,3} \approx v_{3,1}/v_{2,2} \approx (v_{3,1}/v_{1,3})^{1/2}$. We define the estimator $(\boldsymbol{v}_{1,1}, \boldsymbol{v}_{1,2}, \boldsymbol{v}_{2,1}, \boldsymbol{v}_{1,3}, \boldsymbol{v}_{2,2}, \boldsymbol{v}_{3,1})^T \equiv M^{-1}\hat{C}$ of $(v_{1,1}, v_{1,2}, v_{2,1}, v_{1,3}, v_{2,2}, v_{3,1})^T$, with $\hat{C}$ a consistent estimator of C. Now $\boldsymbol{v}_{2,1}/\boldsymbol{v}_{1,2}$, $\boldsymbol{v}_{2,2}/\boldsymbol{v}_{1,3}$, $\boldsymbol{v}_{3,1}/\boldsymbol{v}_{2,2}$, and $(\boldsymbol{v}_{3,1}/\boldsymbol{v}_{1,3})^{1/2}$ are four estimators of β. We can prove that the following estimator $\boldsymbol{\beta}$ has asymptotic minimal variance in the class of consistent estimators of β which are a linear combination of the β estimators $\boldsymbol{v}_{2,1}/\boldsymbol{v}_{1,2}$, $\boldsymbol{v}_{2,2}/\boldsymbol{v}_{1,3}$, $\boldsymbol{v}_{3,1}/\boldsymbol{v}_{2,2}$, and $(\boldsymbol{v}_{3,1}/\boldsymbol{v}_{1,3})^{1/2}$:

$$\boldsymbol{\beta} \equiv (u^T V^{-1}\boldsymbol{t})/(u^T V^{-1} u),$$

with

$$t^T = \left(v_{2,1}/v_{1,2}, v_{2,2}/v_{1,3}, v_{3,1}/v_{2,2}, \left(v_{3,1}/v_{1,3} \right)^{1/2} \right),$$

and V the asymptotic covariance matrix of t, while u has four elements equal to one.

Proof. $\boldsymbol{\beta}$ is of the form $d^T t$. In order for $\boldsymbol{\beta}$ to be consistent, we must have $d^T u = 1$, because each component of t is a consistent estimator of β (at least as far as the approximation goes). Now it follows $d^T = (u^T V^{-1})/(u^T V^{-1} u) + h^T$, with h a (4×1) vector that satisfies the condition $h^T u = 0$. Using this expression and the delta method, we find $\operatorname{var} \boldsymbol{\beta} = d^T V d = (u^T V^{-1} u)^{-1} + h^T V h$. Since V is positive definite, this is minimal for $h^T = (0, 0, 0, 0)$. Q.E.D.

Approximating $\ln e(t)$ by polynomials of order five and higher we can derive analogously 'optimal' estimators. Due to the approximations used, the estimators so found are not consistent in a strict sense; they are only approximately consistent. The approximation can be improved, however, by applying higher-degree polynomials.

Remark. In the previous section we could also find a set of estimators for β by choosing several sets of points (s_1, t_1), (s_2, t_2), and (s_3, t_3). Similar to the proof stated above it is possible to derive from such a set of β estimators an 'optimal' combination.

4. The case that s and t tend to zero

In this section we use the method of the preceding sections. However, we choose 'very small' values for $|s|$ and $|t|$. If s and t tend to zero, then the characteristic functions of x and y are 'approximately' algebraic polynomials.

Let us approximate $\ln c(s, t)$ by a bivariate polynomial of degree r. Thus,

$$\ln c(s, t) = \sum_{g=0}^{r} \sum_{j=0}^{r-g} \left(b_{gj}((g+j)!/(g!j!)) s^g t^j \right).$$

b_{gj} $(g = 0, \ldots, r,\ j = 0, \ldots, r-g)$, multiplied by $((g+j)!(-i)^{g+j})$, is a bivariate cumulant of x and y. It follows that

$$C(s, t) = \sum_{g=1}^{r} \sum_{j=1}^{r-g} \left(b_{gj}((g+j)!/(g!j!)) s^g t^j \right). \tag{9}$$

Furthermore, we approximate the logarithm of the characteristic function $e(t)$ of $\boldsymbol{\xi}$ by a polynomial of order r $(r \geq 1)$.

Substituting (9) into (7), we get the following system:

$$Mb = Mv. \tag{10}$$

Matrix M and vector v are defined in (7); replacing v_{gj} by b_{gj} in vector v, we get the vector b. Because the matrix M is nonsingular, from system (10) it follows now:

$$b = v \;\Rightarrow\; v_{gj} = b_{gj} \quad \text{for every} \quad g = 1,\ldots, r,\; j = 1,\ldots, r-g. \tag{11}$$

So in this case our estimators are functions of bivariate cumulants of x and y asymptotically. In particular, for $r = 3$ asymptotically it follows that our estimator is the consistent moment estimator β_2 of Pal (1980) [$\beta_2 = m_{21}/m_{12}$; m_{21} and m_{12} are consistent estimators of $\mathrm{E}(y - \mathrm{E}y)^2(x - \mathrm{E}x)$ and $\mathrm{E}(y - \mathrm{E}y)(x - \mathrm{E}x)^2$].

5. An example with generated data

In this example we discuss, by means of some generated data, several estimation methods. Because sample size may play an important role with respect to the accuracy of the estimates, both sample sizes of 50 and 200 are used. For the method discussed in this paper, the degree of the polynomials varies from three to five. We analyzed the simple regression model as given in the introduction, with $\alpha = 0$ and $\beta = 1$. Further the variables ε and δ come from a standard normal distribution. Four different types of distributions are defined for the ξ variable. These four distributions are all chi-square distributions, with degrees of freedom one, two, five, and ten. These distributions differ, among other things, in their degree of skewness. The higher the number of degrees of freedom, the more symmetric the distribution is. To make the variances of the x and y variables equal in all situations, the variances of the ξ variables are set equal to one.

We use the estimation method discussed in this paper. The simultaneous characteristic functions of x and y are computed in the arbitrary points $(-0.211, 0.183)$, $(-0.111, -0.123)$, and $(-0.083, -0.143)$ for approximating polynomials of degree three, in the arbitrary points $(-0.211, 0.183)$, $(-0.111, -0.123)$, $(-0.083, -0.143)$, $(0.189, 0.067)$, $(-0.029, 0.113)$, and $(0.023, -0.057)$ for approximating polynomials of degree four, and in the arbitrary points $(-0.211, 0.183)$, $(-0.111, -0.123)$, $(-0.083, -0.143)$, $(0.189, 0.067)$, $(-0.029, 0.113)$, $(0.023, -0.057)$, $(0.101, 0.087)$, $(0.147, -0.091)$, $(0.037, 0.121)$, and $(0.133, -0.051)$ for approximating polynomials of degree five. Two other methods are used. These methods are given by Van Montfort

Table 1

Empirical characteristic function method, number of replications = 100 [means and standard deviations (in parentheses) of estimates $\hat{\beta}$].

	n = 50			n = 200		
	Degree			Degree		
DF	3	4	5	3	4	5
1	3.42 (15.27)	1.29 (1.05)	0.95 (0.09)	1.24 (0.64)	1.02 (0.26)	1.05 (0.11)
2	1.03 (7.74)	1.05 (0.58)	0.97 (0.20)	1.54 (3.47)	1.07 (0.46)	1.31 (0.09)
5	−0.34 (18.83)	1.47 (1.37)	1.00 (0.18)	2.94 (9.42)	1.18 (1.27)	1.03 (0.15)
10	3.23 (18.19)	1.19 (0.85)	1.03 (0.16)	2.29 (23.18)	1.23 (0.97)	1.06 (0.15)

et al. (1987) and Aigner et al. (1984). The basic idea of these methods is that estimation of the parameters is carried out not by using first- and second-order information from the data, but also third- and/or fourth-order information. This means that in those methods also third and fourth moments of the variables play a role. The estimator of Van Montfort et al. has asymptotically minimal variance in the class of consistent estimators which are functions of moments up to order three. The Aigner estimator is a function of moments up to order four.

For the results of the empirical characteristic function method see table 1, and the results of the moments method are given in table 2. These tables consist of the means and the standard deviations of the estimates over the 100 replications.

Discussion. From table 1 we see that the estimates of β become better for increasing degree of the polynomial. In particular for degree five the estimates and the corresponding standard errors behave rather well. It also seems that there is not a big difference between the case $n = 50$ and $n = 200$ for degree five of the polynomial. So we suggest that a polynomial of degree five will fit the empirical characteristic function quite well. Further it holds that the four types of distribution do not have a great influence on the estimates and the standard errors.

In table 2 we see that the difference in sample size plays an important role; sample sizes of 200 will give better estimates. In the case of fitting by third-order moments, estimates are better for the most skew-distributed variables, i.e., whenever $DF = 1$. This is quite obvious, because in the case of symmetric distributions, third-order moments are zero and so they play no

Table 2
Moment method, number of replications = 100 [means and standard deviations (in parentheses) of estimates $\hat{\beta}$].

	$n = 50$		$n = 200$	
	Order of moments		Order of moments	
DF	3	4	3	4
1	0.84 (1.15)	3.19 (17.70)	1.02 (0.17)	1.06 (0.62)
2	1.50 (2.30)	1.20 (6.21)	1.05 (0.24)	1.52 (8.10)
5	1.30 (5.20)	0.90 (2.27)	1.07 (0.33)	0.80 (3.27)
10	0.96 (4.52)	− 7.87 (93.30)	0.80 (0.96)	1.66 (10.15)

role in estimating the parameters. In this latter case fourth-order moments could give better estimates. However, as we see from the table this is not the case. The reason for this is that the standard errors of fourth-order moments are very large, in particular with small sample sizes. So if we want to use fourth-order moments sample sizes should be larger. The most interesting result from tables 1 and 2 is the comparison of the two tables. From this small study (sample size $n = 50$ and $n = 200$) it is quite obvious that our method of characteristic functions, with polynomial approximations of order four and five, gives better results than the methods of moments. The superiority of our method of characteristic functions crucially depends on the degree of the approximating polynomial. Approximating polynomials of order three causes worse estimators than approximating polynomials of order four and five. Finally we mention that our estimators with empirical characteristic functions are biased asymptotically, while the estimators with sample moments are consistent.

Appendix

In this appendix we discuss briefly some elementary properties of empirical characteristic functions. Csorgo (1981), Feuerverger and Mureika (1977), Feuerverger and McDunnough (1981a, b), Kent (1975), and Ramachandran (1967) discuss in detail empirical characteristic functions and their properties.

Suppose $c_n(t) = n^{-1} \sum_{j=1}^{n} \exp(itx_j)$ is the empirical characteristic function corresponding to an arbitrary characteristic function $c(t)$. Feuerverger and

Mureika (1977) prove that, for fixed $T < \infty$, the convergence $\sup_{|t|<T}|c_n(t) - c(t)| \to 0$ is almost surely as $n \to \infty$. Furthermore, the empirical characteristic function process $c_n(.)$ is seen to be an average of n independent processes of the type $\exp(itx)$. By means of the central limit theorem it follows asymptotically that $y_n(t) = n^{1/2}(c_n(t) - c(t))$ is normally distributed, has mean zero, and the covariance structure

$$\operatorname{cov}(\operatorname{Re} y_n(s), \operatorname{Re} y_n(t)) = \tfrac{1}{2}[\operatorname{Re} c(s+t) + \operatorname{Re} c(s-t)] - \operatorname{Re} c(s) \operatorname{Re} c(t),$$

$$\operatorname{cov}(\operatorname{Re} y_n(s), \operatorname{Im} y_n(t)) = \tfrac{1}{2}[\operatorname{Im} c(s+t) + \operatorname{Im} c(s-t)] - \operatorname{Re} c(s) \operatorname{Im} c(t),$$

$$\operatorname{cov}(\operatorname{Im} y_n(s), \operatorname{Im} y_n(t)) = \tfrac{1}{2}[-\operatorname{Re} c(s+t) + \operatorname{Re} c(s-t)] - \operatorname{Im} c(s) \operatorname{Im} c(t).$$

We can generalize the formulas above to the simultaneous characteristic function $c(s, t)$ and the corresponding empirical characteristic function $c_n(s, t) = n^{-1}\Sigma_{j=1}^{n} \exp(isy_j + itx_j)$. If $y_n(s, t) = n^{1/2}(c_n(s, t) - c(s, t))$, then it follows asymptotically that

$$\operatorname{cov}(\operatorname{Re} y_n(s_1, t_1), \operatorname{Re} y_n(s_2, t_2))$$
$$= \tfrac{1}{2}[\operatorname{Re} c(s_1 - s_2, t_1 - t_2) + \operatorname{Re} c(s_1 + s_2, t_1 + t_2)]$$
$$- \operatorname{Re} c(s_1, t_1) \operatorname{Re} c(s_2, t_2),$$

$$\operatorname{cov}(\operatorname{Re} y_n(s_1, t_1), \operatorname{Im} y_n(s_2, t_2))$$
$$= \tfrac{1}{2}[\operatorname{Im} c(s_1 - s_2, t_1 - t_2) + \operatorname{Im} c(s_1 + s_2, t_1 + t_2)]$$
$$- \operatorname{Re} c(s_1, t_1) \operatorname{Im} c(s_2, t_2),$$

$$\operatorname{cov}(\operatorname{Im} y_n(s_1, t_1), \operatorname{Im} y_n(s_2, t_2))$$
$$= \tfrac{1}{2}[\operatorname{Re} c(s_1 - s_2, t_1 - t_2) - \operatorname{Re} c(s_1 + s_2, t_1 + t_2)]$$
$$- \operatorname{Im} c(s_1, t_1) \operatorname{Im} c(s_2, t_2).$$

References

Aigner, D.J., C. Hsiao, A. Kapteyn, and T. Wansbeek, 1984, Latent variable models in econometrics, in: Z. Griliches and M.D. Intriligator, eds., Handbook of econometrics (North-Holland, Amsterdam).

Csorgo, S., 1981, Limit behavior of the empirical characteristic function, Annals of Probability 9, 130–144.

Csorgo, S. and F. Revesz, 1981, Strong approximation in probability and statistics, Series in probability and mathematical statistics (Academic Press, New York).

Darmois, G., 1940, Les mathematiques de la psychologie, Memorial des sciences mathematiques fasc. 48 (Gauthier-Villars, Paris).

Feron, R. and C. Fourgeaud, 1952, Publications de l'Institut de Statistique de l'Université de Paris, Vol. 1, fasc. 2, 44–49.

Feuerverger, A. and R.A. Mureika, 1977, The empirical characteristic function and its application, Annals of Statistics 55, 88–97.

Feuerverger, A. and P. McDunnough, 1981a, On some Fourier methods for inference, Journal of the American Statistical Association 76, 379–387.

Feuerverger, A. and P. McDunnough, 1981b, On the efficiency of empirical characteristic function procedures, Journal of the American Statistical Society B 43, 20–27.

Geary, R.C., 1942, Inherent relations between random variables, Proceedings of the Royal Irish Academy A 47, 63–76.

Kent, J.T., 1975, A weak convergence theorem for the empirical characteristic function, Journal of Applied Probability 12, 515–523.

Neyman, J., 1951, Existence of consistent estimates of the directional parameter in a linear structural relation between two variables, Annals of Mathematical Statistics 22, 497–512.

Pakes, A., 1979, Aggregation effects and panel data estimation problems: An empirical examination of the R&D intensity decision, Discussion paper 699 (Harvard Institute of Economic Research, Cambridge, MA).

Pal, M., 1980, Consistent moment estimators of regression coefficients in the presence of errors in variables, Journal of Econometrics 14, 349–364.

Powell, M.J.D., 1981, Approximation theory and methods (Cambridge University Press, Cambridge).

Ramachandran, B., 1967, Advanced theory of characteristic functions (Statistical Publishing Society, London).

Reiersøl, O., 1941, Conference analysis by means of lag moments and other methods of confluence analysis, Econometrica 9, 1–24.

Reiersøl, O., 1950, Identifiability of a linear relation between variables which are subject to error, Econometrica 18, 375–389.

Rubin, H., 1956, Uniform convergence of random functions with applications to statistics, Annals of Mathematical Statistics 27, 200–203.

Spiegelman, C., 1979, On estimating the slope of a straight line when both variables are subject to error, Annals of Statistics 7, 201–206.

Van Montfort, K., J. de Leeuw, and A. Mooijaart, 1987, Regression with errors in variables: Estimators based on third order moments, Statistica Neerlandica 41, 223–239.

Wolfowitz, J., 1952, Consistent estimators of the parameters of a linear structural relation, Skandinavisk Aktuarietidskrift 35, 132–151.