ZASTOSOWANIA MATEMATYKI
V (1960)

Z. PAWŁOWSKI (Warszawa)

ON THE TOTAL DIFFERENTIAL METHOD AND ITS EFFICIENCY IN THE CASE OF A LINEAR REGRESSION(¹)

I. Introduction. The paper is divided in two parts. The first presents shortly a method of estimating regression coefficients when the regression function is assumed to be linear. The second part is devoted to the discussion of the efficiency of this method as compared with the efficiency of the classical method of estimating linear regression coefficients. The method which is discussed here was sometimes used in the past and some authors called it the method of finite first differences (because first differences are used in the process of estimation instead of the values of original, i. e. observed, "dependent" and "independent" variables), but this term was never generally accepted.

Instead of the *first finite differences method* the name of the *total differential method* is proposed. The reason is that — as will be shown later in section II — the estimation procedure by this method reduces to the estimation of the parameters of the regression function by using the observed values of the total differential of the dependent variable Y_t when the independent variables $X_1, X_2, \ldots, X_h$ change by the (observed) amounts $\Delta X_1, \Delta X_2, \ldots, \Delta X_h$.

The total differential method can be used when statistical data are time series data, i. e. when the observations are ordered in time. Although it has been generally held that the method is especially suitable when there is a trend in the data, nobody — as far I know — has investigated under which conditions this method is more efficient than R. A. Fisher's classical theory of regression. The present investigation brings some new light on this matter and gives a formula which can be used for the comparison of the efficiency of the two methods.

Although it deals with first finite differences of observed variables, the total differential method should not be identified with the *variate-difference* method explored by O. Anderson [1] and G. Tintner [3].

(¹) The main ideas of the paper were presented at the Seminar of the Statistical Institute of Uppsala University (Sweden) on May 23rd 1958.

The assumptions underlying the two methods and their aims are different. The statistician uses the variate-difference method in order to get some information about the importance of the random element affecting the time series he deals with and in order to find the degree of the polynomial by which the non-random component of his time series can be represented. He can also try to eliminate the spurious correlation between two variables due to an influence of time on those two variables (see "Student" [2] and Tintner [3]). On the other hand, the total differential method is devised for the estimation of regression parameters when the regression function is given, i. e. when its functional form and all its independent variables are specified.

II. A short outline of the total differential method in the case of a linear regression. Let us assume that Y_t is such a random variable that its mean value is linearly dependent on the h following variables (which are not random in the sense of the calculus of probability): $x_{1t}, x_{2t}, \ldots, x_{ht}$. We have therefore

$$E(Y_t) = \sum_{i=1}^{h} a_i x_{it} + a_0, \tag{1}$$

where a_i $(i = 0, 1, \ldots, h)$ are unknown constant parameters which have to be estimated. Because of random fluctuations we shall have for consecutive values of time parameter t the following relations relating observed values of Y_t and of x_{it}'s:

$$y_t = \sum_{i=1}^{h} a_i x_{it} + a_0 + Z , \tag{2}$$

where Z_t is a random variable which is assumed to form a stochastic process, stationary to the second degree (see Wold [4], page 156) with zero mean and finite variance σ_z^2.

Let us assume further that our statistical data are time series referring to n different time points (periods) and that they can be represented by the matrix

$$[y_t, x_{it}] = \begin{bmatrix} y_1 & x_{11} & x_{21} & \cdots & x_{h1} \\ y_2 & x_{12} & x_{22} & \cdots & x_{h2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ y_n & x_{1n} & x_{2n} & \cdots & x_{hn} \end{bmatrix}. \tag{3}$$

In order to estimate the unknown parameters a_i we introduce the first finite differences of y_t and x_{it} defined by the equations

$$\Delta x_{it} = x_{i,t+1} - x_{i,t}, \tag{4}$$

$$\Delta y_t = y_{t+1} - y_t. \tag{5}$$

Transforming the original sample data into first differences defined by (4) and (5) we get another matrix of statistical data:

$$[\Delta y_t, \Delta x_{it}] = \begin{bmatrix} \Delta y_1 & \Delta x_{11} & \ldots & \Delta x_{h1} \\ \Delta y_2 & \Delta x_{12} & \ldots & \Delta x_{h2} \\ \ldots & \ldots & \ldots & \ldots \\ \Delta y_{n-1} & \Delta x_{1,n-1} & \ldots & \Delta x_{h,n-1} \end{bmatrix}. \tag{6}$$

As is easily seen, if (1) and (2) hold true, we also have

$$\Delta y_t = \sum_{i=1}^{h} a_i \Delta x_{it} + \Delta z_t, \tag{7}$$

where Δz_t is a residual term with zero mean and variance $\sigma^2_{\Delta z}$ for all t. Equation (7) is a basic one for the estimation of unknown regression parameters by the total differential method. In view of (2), equation (7) can be interpreted as the total differential of y_t with respect to $x_{1t}, x_{2t}, \ldots \ldots, x_{ht}$ plus a residual term Δz_t. It is more convenient to regard (7) just as a total differential because such an interpretation permits a generalization of the total differential method also for non-linear cases, as will be shown in the last section of the paper.

In the linear case which we are discussing here the estimation procedure has two steps. In the first, the coefficients $a_1, a_2, \ldots, a_h$ are estimated and in the second the constant term a_0 is obtained on the grounds of the results obtained in the first step. In order to obtain estimators for $a_1, a_2, \ldots, a_h$ we apply the least squares method to minimize the expression

$$S_1 = \sum_{t=1}^{n-1} \Big(\Delta y_t - \sum_{i=1}^{h} a_i \Delta x_{it}\Big)^2$$

with respect to $a_1, a_2, \ldots, a_h$. Let $a_1, a_2, \ldots, a_h$ denote the estimates of the unknown coefficients obtained by that device. Obviously these estimators are obtained by solving with respect to $a_1, a_2, \dots, a_h$ the following system of linear equations:

$$\begin{aligned}
\sum_{i=1}^{h}\sum_{t=1}^{n-1} a_i \Delta x_{1t}\Delta x_{it} &= \sum_{t=1}^{n-1}\Delta x_{1t}\Delta y_t,\\
\sum_{i=1}^{h}\sum_{t=1}^{n-1} a_i \Delta x_{2t}\Delta x_{it} &= \sum_{t=1}^{n-1}\Delta x_{2t}\Delta y_t,\\
&\dots\dots\dots\dots\\
\sum_{i=1}^{h}\sum_{t=1}^{n-1} a_i \Delta x_{ht}\Delta x_{it} &= \sum_{t=1}^{n-1}\Delta x_{ht}\Delta y_t,
\end{aligned}$$

Consequently, the estimator a_i of the parameter α_i $(i = 1, 2, \dots, h)$ is given by the ratio of the two determinants

$$a_i = \frac{D_i}{D}, \quad D \neq 0, \tag{8}$$

where D is defined as

$$D = \begin{vmatrix}
\sum_{t=1}^{n-1}\Delta x_{1t}\Delta x_{1t} & \sum_{t=1}^{n-1}\Delta x_{1t}\Delta x_{2t} & \dots & \sum_{t=1}^{n-1}\Delta x_{1t}\Delta x_{ht}\\
\sum_{t=1}^{n-1}\Delta x_{2t}\Delta x_{1t} & \sum_{t=1}^{n-1}\Delta x_{2t}\Delta x_{2t} & \dots & \sum_{t=1}^{n-1}\Delta x_{2t}\Delta x_{ht}\\
\dots & \dots & \dots & \dots\\
\sum_{t=1}^{n-1}\Delta x_{ht}\Delta x_{1t} & \sum_{t=1}^{n-1}\Delta x_{ht}\Delta x_{2t} & \dots & \sum_{t=1}^{n-1}\Delta x_{ht}\Delta x_{ht}
\end{vmatrix}, \tag{9}$$

while D_i is the determinant obtained from the determinant D when its i-th column is replaced by the vector

$$\begin{bmatrix}
\sum_{t=1}^{n-1}\Delta x_{1t}\Delta y_t\\
\sum_{t=1}^{n-1}\Delta x_{2t}\Delta y_t\\
\dots\dots\\
\sum_{t=1}^{n-1}\Delta x_{ht}\Delta y_t
\end{bmatrix}.$$

In the next section the estimators $a_1, a_2, \dots, a_h$ will be shown to be unbiased and in section IV the proof of their consistency will also be given. In the same section we shall also deal with the problem of standard error of the estimator for a regression coefficient.

The estimation of α_0 is done in the second step of the estimation procedure. The constant term α_0 is estimated under the assumption that $\alpha_1 = a_1, \alpha_2 = a_2, \ldots, \alpha_h = a_h$, which leads to the minimization of the expression

$$S_2 = \sum_{t=1}^{n}\Big(y_t - \sum_{i=1}^{h} a_i x_{it} - \alpha_0\Big)^2$$

with respect to α_0, so that we get the following estimator of α_0:

$$a_0 = \frac{1}{n}\Big[\sum_{t=1}^{n} y_t - \sum_{t=1}^{n}\sum_{i=1}^{h} a_i x_{it}\Big]. \tag{10}$$

EXAMPLE 1. We assume that $E(Y_t) = \alpha_1 x_t + \alpha_0$. We are to estimate α_1 and α_0 on the basis of the following statistical data (table 1).

TABLE 1

Time period	y_t	x_t	Δy_t	Δx_t
1	1,07	1,05	0,01	0,02
2	1,08	1,07	0,00	0,03
3	1,08	1,10	0,03	0,02
4	1,11	1,12	0,05	0,03
5	1,16	1,15	—	—

The first step gives us the estimate of α_1:

$$a_1 = \frac{\sum_{t=1}^{4} \Delta x_t \Delta y_t}{\sum_{t=1}^{4} (\Delta x_t)^2} = \frac{0,0023}{0,0026} = 0,88.$$

Substituting 0,88 for α_1 we assume that $E(Y_t) = 0,88 x_t + \alpha_0$. Minimizing the expression

$$S = \sum_{t=1}^{5} (y_t - 0,88 x_t - \alpha_0)^2$$

we get $a_0 = 0,13$. In the following sections it will be shown that also a_0 is an unbiased and consistent estimator.

III. The problem of unbiasedness of the total differential method estimators. In this section we shall investigate the problem of unbiasedness of estimators for regression coefficients in the case of a linear relationship when these estimators are obtained by the total differential method. As we shall see, such estimators are unbiased as long as our assumptions about the regression function, i. e. about the independent variables of this function, are correct. On the other hand, these estimators will generally be biased if we omit in the regression function one or more explanatory variables the first differences of which are correlated with the first finite differences of the dependent variable, or with the first finite differences of other independent variables of the regression function.

We shall prove our statements under the assumption that the x_i's, i. e. the explanatory (independent) variables of the regression function are not random — in accordance with the Gauss-Fisher theory of regression [2].

THEOREM 1. *If Y_t is such a random variable that*

$$E(Y_t) = \sum_{i=1}^{h} a_i x_{it} + a_0 \tag{11}$$

and therefore

$$E(\Delta Y_t) = \sum_{i=1}^{h} a_i \Delta x_{it}, \tag{11'}$$

then the estimators $a_0, a_1, \ldots, a_h$ of the parameters $a_0, a_1, \ldots, a_h$, obtained by the total differential method, are unbiased.

Proof. We shall prove first that the estimators a_i for $i = 1, 2, \ldots, h$ are unbiased and afterwards we shall show the unbiasedness of the estimator of the constant term a_0. As was shown in section 1 (see formula (8)), the total differential method estimator of a_i is given by

$$a_i = \frac{D_i}{D}.$$

Expanding the determinant D_i with respect to the i-th column, we get

$$a_i = \frac{A_{1i}\sum_{t=1}^{n-1} \Delta x_{1t}\Delta y_t + A_{2i}\sum_{t=1}^{n-1} \Delta x_{2t}\Delta y_t + \ldots + A_{hi}\sum_{t=1}^{n-1} \Delta x_{ht}\Delta y_t}{D},$$

where A_{ji} is the cofactor of the determinant D_i (and also of D) when the i-th column and the j-th row of this determinant are cancelled. Taking

([2]) See H. Wold [4], p. 205-207.

the mathematical expectation of a_i we have, since the x_i's are assumed to be fixed,

$$E(a_i) = \frac{A_{1i}\sum_t \Delta x_{1t}E(\Delta y_t) + A_{2i}\sum_t \Delta x_{2t}E(\Delta y_t) + \ldots + A_{hi}\sum_t \Delta x_{ht}E(\Delta y_t)}{D}$$

$$= \frac{A_{1i}\sum_t \Delta x_{1t}\sum_{j=1}^{h}\alpha_j\Delta x_{jt} + A_{2i}\sum_t \Delta x_{2t}\sum_{j=1}^{h}\alpha_j\Delta x_{jt} + \ldots + A_{hi}\sum_t \Delta x_{ht}\sum_{j=1}^{h}\alpha_j\Delta x_{jt}}{D}$$

because of (11′). Rearranging the terms in the numerator we can write

$$(12) \qquad E(a_i) = \frac{\alpha_1\sum_{j=1}^{h}A_{ji}\Delta x_{jt}\Delta x_{1t} + \alpha_2\sum_{j=1}^{h}A_{ji}\Delta x_{jt}\Delta x_{2t} + \ldots + \alpha_h\sum_{j=1}^{h}A_{ji}\Delta x_{jt}\Delta x_{ht}}{D}.$$

But, by a well-known theorem of linear algebra, when $i \neq r$, we have the equality

$$\sum_{j=1}^{h} A_{ji}\Delta x_{jt}\Delta x_{rt} = 0,$$

while for $i = r$ we have

$$\sum_{j=1}^{h} A_{ji}\Delta x_{jt}\Delta x_{rt} = D.$$

The numerator of formula (12) therefore equals $\alpha_i \cdot D$, whence $E(a_i) = \alpha_i$, i. e. a_i ~~is an~~ unbiased estimator of α_i. As the index i was chosen freely, the same relation holds for estimators $a_1, a_2, \ldots a_h$. We have shown thus that $a_1, a_2, \ldots, a_h$ are unbiased estimators of the parameters $\alpha_1, \alpha_2, \ldots, \alpha_h$.

We come now to the proof for a_0. From formula (10) we infer that the total differential method estimator of the constant term α_0 is

$$a_0 = \frac{1}{n}\Big[\sum_{t=1}^{n} y_t - \sum_{t=1}^{n}\sum_{i=1}^{h} a_i x_{it}\Big].$$

Therefore

$$E(a_0) = \frac{1}{n}\Big[\sum_{t=1}^{n} E(y_t) - \sum_{t=1}^{n}\sum_{i=1}^{h} x_{it}E(a_i)\Big]$$

because of the assumption that the x_i's are fixed. Using (11) we get

$$E(a_0) = \frac{1}{n}\Big[\sum_{t=1}^{n}\Big(\sum_{i=1}^{h}\alpha_i x_{it} + \alpha_0\Big) - \sum_{t=1}^{n}\sum_{i=1}^{h}\alpha_i x_{it}\Big] = \alpha_0,$$

and a_0 is an unbiased estimator of the constant term α_0.

The estimators $a_0, a_1, \ldots, a_h$ will be biased, however, when our assumptions about the form of the regression function are not correct. Below we shall give an example in which it will be assumed that the specification of independent variables entering the regression function is wrong, i. e. that we have omitted in the regression function a variable V_t such that the mathematical expectation of Y_t depends on V_t. In order to avoid tedious algebra we shall take the simplest case.

Let us suppose that our assumption about a regression function is

$$E(Y_t) = \alpha_1 X_t + \alpha_0, \tag{13}$$

while the "true" regression function has the form

$$E(Y_t) = \alpha_1 X_t + \alpha_2 V_t + \alpha_0, \tag{13'}$$

where $\alpha_2 \neq 0$. In this case the estimators a_1 and a_0 of the parameters α_1 and α_0 will no longer be unbiased. In fact, we have

$$E(a_1) = E\left(\frac{\sum\limits_{t=1}^{n-1} \Delta x_t \Delta y_t}{\sum\limits_{t=1}^{n-1} (\Delta x_t)^2}\right) = \frac{\sum\limits_{t=1}^{n-1} \Delta x_t E(\Delta y_t)}{\sum\limits_{t=1}^{n-1} (\Delta x_t)^2} \tag{14}$$

$$= \frac{\sum\limits_{t=1}^{n-1} [\Delta x_t(\alpha_1 \Delta x_t + \alpha_2 \Delta v_t)]}{\sum\limits_{t=1}^{n-1} (\Delta x_t)^2} = \alpha_1 + \alpha_2 \frac{\sum\limits_{t=1}^{n-1} \Delta x_t \Delta v_t}{\sum\limits_{t=1}^{n-1} (\Delta x_t)^2},$$

which proves that a_1 will be biased unless the sum $\sum\limits_t \Delta x_t \Delta v_t$ equals zero. By using the argument which will be fully explained in the next section, it can be shown that this sum is equal to the expression

$$2\operatorname{cov}(x_t, v_t) - \operatorname{cov}(x_t, v_{t+1}) - \operatorname{cov}(x_{t+1}, v_t) + O(n^{-1}).$$

This result is different from the well-known result obtained by H. Wold (see [4], chapter XII, theorem 3) for the case of the classical regression analysis. He proved namely that the estimator a_1 can be unbiased in spite of the specification error simply if the correlation coefficient between x_t and v_t is zero. That statement, however, does not hold true in our case of the total differential method. On the other hand, it can be seen from formula (14) that the bias of a_1 is generally not serious when the coefficient α_2 is near zero.

When relation (13′) holds true instead of (13), the estimator a_0 will also be biased. This can be shown by simple calculations involving some algebraic transformations. We have now

$$(15)\qquad E(a_0) = \frac{1}{n}\Big\{E\Big[\sum_{t=1}^{n} y_t - a_1\sum_{t=1}^{n} x_t\Big]\Big\}$$

$$= \frac{1}{n}\sum_{t=1}^{n}(\alpha_1 x_t + \alpha_2 v_t + \alpha_0) - \frac{1}{n}\sum_{t=1}^{n} x_t E(a_1)$$

$$= \frac{\alpha_1}{n}\sum_{t=1}^{n} x_t + \frac{\alpha_2}{n}\sum_{t=1}^{n} v_t + \alpha_0 - \frac{1}{n}\sum_{t=1}^{n} x_t\left[\alpha_1 + \alpha_2\frac{\sum\limits_{t=1}^{n-1}\Delta x_t\Delta v_t}{\sum\limits_{t=1}^{n-1}(\Delta x_t)^2}\right]$$

$$= \alpha_0 + \frac{\alpha_2}{n}\left(\sum_{t=1}^{n} v_t - \sum_{t=1}^{n} x_t\frac{\sum\limits_{t=1}^{n-1}\Delta x_t\Delta v_t}{\sum\limits_{t=1}^{n-1}(\Delta x_t)^2}\right),$$

which shows the bias of the estimator a_0 in the case of a specification error in the assumed form of the regression function.

IV. Standard errors of estimators in the case of one independent variable. In this section we shall derive the formula for the variance of the total differential method estimator a_1 of the coefficient α_1 in the regression function $E(Y_t) = \alpha_1 X_t + \alpha_0$. This will permit us later to compare the efficiency of the method discussed with the efficiency of the classical method of regression analysis and to find cases when the new method is better, i. e. has smaller standard errors of estimates. In this section we shall show also that the two total differential method estimators of α_1 and α_0 are consistent. First we find the variance of a_1.

THEOREM 2. *Let us assume that $Y_t = \alpha_1 X_t + \alpha_0 + Z_t$ where X_t is a non-random variable and Z_t is a random variable such that $E(Z_t) = 0$ and $D^2(Z_t) = \sigma_z^2 < \infty$ for all t, so that Z_t forms a stochastic process, stationary to the second degree. Then the variance of a_1, the total differential method estimator of α_0 defined by (12), is given by*

$$(16)\qquad D^2(a_1) = \frac{\sigma_{\Delta Z}^2\sum\limits_{t=1}^{n-1}(\Delta x_t)^2 + 2\sigma_{\Delta Z}^2\sum\limits_{v=1}^{n-2}\sum\limits_{t=1}^{n-1-v}\Delta x_t\Delta x_{t+v}\varrho_v^*}{\Big[\sum\limits_{t=1}^{n-1}(\Delta x_t)^2\Big]^2}$$

where the ϱ_v^*'s are autocorrelation coefficients of ΔZ_t:

$$(17)\qquad \varrho_v^* = \frac{E(\Delta Z_t\Delta Z_{t+1})}{\sigma_{\Delta Z}^2}.$$

Proof. From (13) we can write the variance of a_1 as

$$D^2(a_1) = D^2\left(\frac{\sum_{t=1}^{n-1} \Delta y_t \Delta x_t}{\sum_{t=1}^{n-1} (\Delta x_t)^2}\right) = D^2\left(\frac{\sum_{t=1}^{n-1} \Delta Z_t \Delta x_t}{\sum_{t=1}^{n-1} (\Delta x_t)^2}\right),$$

but as X_t's are not random variables we can put the non-random term before the operator D^2 and write

$$D^2(a_1) = \frac{\sum_{t=1}^{n-1} \sigma^2_{\Delta Z_t}(\Delta x_t)^2 + 2\sum_{v=1}^{n-2}\sum_{t=1}^{n-1-v} \Delta x_t \Delta x_{t+v} \sigma_{\Delta Z_t}\sigma_{\Delta Z_{t+v}} \varrho_v^*}{[\sum_{t=1}^{n-1} (\Delta x_t)^2]^2}.$$

Because the variance of Z_t is assumed to be the same for all t the same must hold true for the variance of ΔZ_t, and therefore putting in the formula above $\sigma^2_{\Delta Z_t} = \sigma^2_{\Delta Z_{t+v}} = \sigma^2_{\Delta Z}$ we get (16), which completes the proof.

Expression (16) can be written in a simpler way if we introduce into it the coefficients $\varkappa_v$ defined as follows:

$$\varkappa_v = \frac{\sum_{t=1}^{n-1-v} \Delta x_t \Delta x_{t+v}}{\sum_{t=1}^{n-1} (\Delta x_t)^2}. \tag{18}$$

Inserting (18) into (16) we get

$$D^2(a_1) = \frac{\sigma^2_{\Delta Z}}{\sum_{t=1}^{n-1} (\Delta x_t)^2}\left(1 + 2\sum_{v=1}^{n-2} \varkappa_v \varrho_v^*\right). \tag{19}$$

Computing the variance of a_1 in practice we shall find two difficulties which, however, appear also when the classical regression formula is used. The first of these difficulties is that we usually do not know the value of the variance of ΔZ_t. This difficulty can be overcome by computing the sample variance of ΔZ_t and using it as an approximate value of the true parameter. The second, similar difficulty arises because we do not know the autoregression coefficients of ΔZ_t. Again, we can compute instead the sample values of those coefficients. It is interesting to note that we usually need only a few first coefficients ϱ_v^* because in general absolute values of $\varkappa_v$ will decrease rapidly to zero so that for large v products $\varkappa_v \varrho_v^*$ will only little differ from zero. The practical impossibility of computing the exact value of $D^2(a_1)$ cannot be considered, though, as a weak point of the total differential method because the same situation usually holds when one uses classical methods of estimation.

It is easy to show that both a_1 and a_0 are consistent estimators of α_1 and α_0 respectively. In fact, we shall prove below the following

THEOREM 3. *Under the assumptions stated in theorem 2 and assuming that either the series* $A_n = \sum_{v=1}^{n-2} \varkappa_v \varrho_v^*$ *is convergent when* $n \to \infty$ *or that*

$$\lim_{n\to\infty} \frac{A_n}{n} = 0, \tag{20}$$

the statistic a_1 *defined by* (12) *is a consistent estimator for* α_1 *and the statistic* a_0 *defined by* (22) *is a consistent estimator for* α_0.

Proof. As is well known, it follows from Tchebyschev's inequality that an estimator t of a parameter Θ is consistent when it is unbiased and when its variance tends to zero as the sample size increases to infinity. In our case we know from section 3 that a_1 and a_0 are unbiased, and thus we need only to prove that their variances tend to zero as $n \to \infty$.

To prove the consistency of a_1 let us denote by $m_{(\Delta x)^2}$ the mean value of $(\Delta x_t)^2$'s, i. e. let

$$m_{(\Delta x)^2} = \frac{1}{n-1} \sum_{t=1}^{n-1} (\Delta x_t)^2. \tag{21}$$

Of course as X_t's are finite this mean will be bounded for all n. Using (21) we can rewrite (19) as follows:

$$D^2(a_1) = \frac{\sigma_{\Delta Z}^2}{(n-1)\, m_{(\Delta x)^2}} \Big(1+2\sum_{v=1}^{n-2} \varkappa_v \varrho_v^*\Big)$$

which is easily seen to tend to zero as $n \to \infty$ and therefore a_1 is consistent. Whether (20) will hold true depends on ϱ_v^* and these depend on the type of the stochastic stationary process which the ΔZ_t's represent. In particular (20) will hold true when ΔZ_t form a process of moving averages or the so-called autoregressive process because then, as has been shown by Wold [5], the autoregression coefficients of the process are all zero for all $v > v_0$ where v_0 is a natural finite number. In such cases the series A_n is obviously converging.

Now we shall prove that also a_0 is consistent. Under the assumptions about the regression function the total differential method estimator of a_0 is

$$a_0 = \frac{\sum_{t=1}^{n} y_t}{n} - \frac{a_1}{n} \sum_{t=1}^{n} x_t = \bar{y} - a_1 \bar{x}. \tag{22}$$

Therefore we have

$$D^2(a_0) = D^2(\bar{y}) + D^2(a_1)\bar{x}^2 - 2\bar{x} D(\bar{y}) D(a_1)\, \varrho_{\bar{y}, a_1}.$$

But $D^2(\bar{y}) = \sigma_y^2/n = \sigma_z^2/n$ and therefore it tends to zero as $n \to \infty$. The same has been proved for $D^2(a_1)$ and as $\bar{x}$ is finite we find that the variance of a_0 tends to zero. This shows that a_0 is a consistent estimator of α_0.

By similar arguments it can also be proved that also in a more general case of the function (1) (i. e. when there are more than one independent variables in the regression function which are not random) the total differential method estimators of parameters are unbiased and consistent. The proof is essentially the same but needs much more tedious algebra.

V. The efficiency of the total differential method estimators. The present section is devoted to the comparison of the efficiency of estimators obtained by the total differential method with the efficiency of estimators used in Fisher's theory of regression, which was later generalised by H. Wold (see [5] and [6]) to the case of time series data where consecutive observations are usually not independent in the sense of the calculus of probability.

In order to perform such comparisons we must transform formula (19) in such a way as to get $D^2(a_1)$ in terms of parameters for the original variables X_t and Z_t only, and not in terms of parameters of their first finite differences. For this purpose we shall prove below several simple relations which connect parameters of Z_t and of X_t with those of ΔZ_t and of ΔX_t.

THEOREM 4. *If Z_t represents a stochastic process stationary to the second degree, then*

$$\sigma_{\Delta Z}^2 = 2\sigma_Z^2(1-\varrho_1) \tag{23}$$

where ϱ_1 is the coefficient of autocorrelation between Z_t and Z_{t+1}.

Proof. We can write down the following identities:

$$\begin{aligned}\sigma_{\Delta Z}^2 &= E(Z_{t+1}-Z_t)^2-[E(Z_{t+1}-Z_t)]^2\\ &= E(Z_{t+1}^2)-[E(Z_{t+1})]^2+E(Z_t^2)-[E(Z_t)]^2-2E(Z_{t+1}Z_t)+2E(Z_t)E(Z_{t+1})\\ &= \sigma_{Z_{t+1}}^2+\sigma_{Z_t}^2-2\operatorname{cov}(Z_{t+1},Z_t) = 2\sigma_Z^2-2\sigma_Z^2\varrho_1.\end{aligned}$$

It should be noted that formula (23) is not a new one. It can be derived from a more general expression, obtained first by O. Anderson ([2], page 114), relating variances of higher degree differences to the variance of an original variable, say Z_t.

THEOREM 5. *If* $Y_t = \sum_{i=1}^{h} a_i x_{it} + a_0 + Z_t$ *where* x_{it} *are not random variables, while* Z_t *is random and represents a stochastic process stationary to the second degree with* $E(Z_t) = 0$ *and* $D^2(Z_t) = \sigma_Z^2 < \infty$, *then the following equalities* (24) *and* (25) *hold true*:

$$\varrho_v^* = \varrho(\Delta Z_t, \Delta Z_{t+v}) = \frac{2\varrho_v + \varrho_{v+1} - \varrho_{v-1}}{2(1-\varrho_1)}, \tag{24}$$

$$\varrho_v = \frac{\operatorname{cov}(Y_t, Y_{t+v})}{\sigma_{Y_t}^2} = \frac{\operatorname{cov}(Z_t, Z_{t+v})}{\sigma_Z^2}. \tag{25}$$

Proof. First, in order to prove (25) we need only to observe that we have

$$D^2(Y_t) = D^2\Big(\sum_{i=1}^{h} a_i x_{it} + a_0 + Z_t\Big) = D^2(Z_t) = \sigma_Z^2.$$

Furthermore we have $Z_t = Y_t - E(Y_t)$ and $Z_{t+v} = Y_{t+v} - E(Y_{t+v})$. Then

$$\operatorname{cov}(Z_t, Z_{t+v}) = E(Z_t Z_{t+v}) = E\big[\big(Y_t - E(Y_t)\big)\big(Y_{t+v} - E(Y_{t+v})\big)\big],$$

which proves relation (25). Now in order to prove (24) we shall transform the covariance of ΔZ_t and of ΔZ_{t+v} in the following way:

$$\operatorname{cov}(\Delta Z_t, \Delta Z_{t+v}) = E(\Delta Z_t \Delta Z_{t+v}) - E\Delta Z_t \cdot E\Delta Z_{t+v},$$

$$\operatorname{cov}(\Delta Z_t, \Delta Z_{t+v}) = E[(Z_{t+v+1} - Z_{t+v})(Z_{t+1} - Z_t)] - \\ - E(Z_{t+1} - Z_t) E(Z_{t+v+1} - Z_{t+v}),$$

which, after the multiplication and rearrangement of terms, gives

$$\operatorname{cov}(\Delta Z_t, \Delta Z_{t+v}) = \operatorname{cov}(Z_{t+v+1}, Z_{t+1}) - \operatorname{cov}(Z_{t+v+1}, Z_t) - \\ - \operatorname{cov}(Z_{t+1}, Z_{t+v}) + \operatorname{cov}(Z_{t+v}, Z_t).$$

Dividing now both sides of the last equation by σ_Z^2 and using formula (23) we finally get

$$\varrho_v^* = \frac{2\varrho_v + \varrho_{v+1} - \varrho_{v-1}}{2(1-\varrho_1)},$$

which completes the proof.

Theorem 5 has a great practical significance. It shows that under some definite conditions we can find such situations that though consecutive observations made, for instance on a variable W_t are strongly correlated, the first differences of them may not be. This can sometimes be useful in practical applications of the total differential method to the estimation of regression parameters.

THEOREM 6. *If we introduce the notation*

$$M_{t+1} = \frac{1}{n-1}\sum_{t=1}^{n-1} x_{t+1}, \qquad M_t = \frac{1}{n-1}\sum_{t=1}^{n-1} x_t,$$

$$S^2_{t+1} = \frac{1}{n-1}\sum_{t=1}^{n-1} (x_{t+1}-M_{t+1})^2, \qquad S^2_t = \frac{1}{n-1}\sum_{t=1}^{n-1} (x_t-M_t)^2,$$

$$r_1 = \frac{\operatorname{cov}(x_{t+1}, x_t)}{S_t S_{t+1}}$$

where the x_t's are observed values of X_t at different time points, then the following equality holds:

$$\sum_{t=1}^{n-1} (\Delta x_t)^2 = (n-1)[S^2_{t+1} + S^2_t - 2r_1 S_t S_{t+1} + (M_{t+1}-M_t)^2]. \tag{26}$$

Proof. We can expand the sum of the squares of first finite differences as follows:

$$\sum_{t=1}^{n-1} (\Delta x_t)^2 = \sum_{t=1}^{n-1} (x_{t+1}-x_t)^2 = \sum_{t=1}^{n-1} [(x_{t+1}-M_{t+1}) + (M_t - x_t) + (M_{t+1}-M_t)]^2,$$

$$\sum_{t=1}^{n-1} (\Delta x_t)^2 = \sum_{t=1}^{n-1} (x_{t+1}-M_{t+1})^2 + \sum_{t=1}^{n-1} (x_t - M_t)^2 + (n-1)(M_{t+1}-M_t)^2 - $$
$$-2\sum_{t=1}^{n-1} (x_{t+1}-M_{t+1})(x_t - M_t),$$

$$\sum_{t=1}^{n-1} (\Delta x_t)^2 = (n-1)[S^2_{t+1} S^2_t - 2r_1 S_t S_{t+1} + (M_{t+1}-M_t)^2],$$

which was to be proved.

When the sample size n is large the parameters S^2_t and S^2_{t+1} will differ from each other only insignificantly, just as M_{t+1} will not differ much from M_t. It is therefore legitimate to look for a new, simpler formula for large n in which only the general sample variance of X_t would appear instead of S^2_t and S^2_{t+1}. To do this, however, we need to know the order of the error committed by using such an approximation. The derivation of formulas for those errors is easy but requires some tedious algebra, which will be mostly omitted here. First we shall try to express S^2_{t+1} and S^2_t in terms of the parameter S^2 defined as follows:

$$S^2 = \frac{1}{n}\sum_{t=1}^{n} (x_t - M)^2 \quad \text{where} \quad M = \frac{1}{n}\sum_{t=1}^{n} x_t.$$

We can write the following identity:

$$S^2 - S_{t+1}^2 = \frac{1}{n}\sum_{t=1}^{n}(x_t - M)^2 - \frac{1}{n-1}\sum_{t=1}^{n}(x_{t+1} - M_{t+1})^2$$

$$= \frac{(n-1)\sum_{t=1}^{n}(x_t - M)^2 - n\sum_{t=1}^{n-1}(x_{t+1} - M_{t+1})^2}{n(n-1)}$$

$$= \frac{(n-1)\sum_{t=1}^{n}x_t^2 - n(n-1)M^2 - n\sum_{t=1}^{n-1}x_{t+1}^2 + n(n-1)M_{t+1}^2}{n(n-1)}$$

$$= \frac{nx_1^2 - \sum_{t=1}^{n}x_t^2 - n(n-1)M^2 + n(n-1)M_{t+1}^2}{n(n-1)}$$

$$= \frac{x_1^2 - M^2}{n-1} - \frac{S^2}{n-1} - (M^2 - M_{t+1}^2)$$

and finally

$$S_{t+1}^2 = \frac{M^2 - x_1^2}{n-1} + \frac{nS^2}{n-1} + (M^2 - M_{t+1}^2). \tag{27}$$

By using similar arguments it can be shown that

$$S_t^2 = \frac{M^2 - x_n^2}{n-1} + \frac{nS^2}{n-1} + (M^2 - M_t^2), \tag{28}$$

so that

$$S_t^2 + S_{t+1}^2 = \frac{2nS^2}{n-1} + \frac{M^2 - x_1^2}{n-1} + \frac{M^2 - x_n^2}{n-1} + (M^2 - M_t^2) + (M^2 - M_{t+1}^2). \tag{29}$$

Now the probem is reduced to finding the order of magnitude of the expressions standing on the right side of equation (29). These calculations are quite simple. We have for example

$$M_t - M = \frac{1}{n-1}\sum_{t=1}^{n-1}x_t - \frac{1}{n}\sum_{t=1}^{n}x_t = \frac{n\sum_{t=1}^{n-1}x_t - (n-1)\sum_{t=1}^{n}x_t}{n(n-1)} \tag{30}$$

$$= \frac{-nx_n + \sum_{t=1}^{n}x_t}{n(n-1)} = \frac{M - x_n}{n-1} = O\left(\frac{S}{n}\right)$$

because the difference $M-x_n$ is of the order of S. In a similar way it can be shown that the following relations hold true:

$$M_{t+1}-M = o\left(\frac{S}{n}\right), \tag{31}$$

$$M_t^2-M^2 = O\left(\frac{MS}{n}\right), \tag{32}$$

$$(M_{t+1}-M_t)^2 = o\left(\frac{S}{n}\right), \tag{33}$$

$$\frac{M^2-x_n^2}{n-1}-\frac{M^2-x_1^2}{n-1} = O\left(\frac{MS}{n}\right). \tag{34}$$

It can also be shown that if $x_1, x_2, \ldots, x_n$ and $v_1, v_2, \ldots, v_n$ represent the observed values of the variables X_t and V_t respectively, and if $\operatorname{cov}(x_t, v_t)$ denotes the sample covariance of these two variables, then

$$\sum_{t=1}^{n-1}(x_{t+1}-x_t)(v_{t+1}-v_t) = 2\operatorname{cov}(x_t, v_t)-\operatorname{cov}(x_{t+1}, v_t)-\operatorname{cov}(x_t, v_{t+1})+O(n^{-1}). \tag{35}$$

Using the relations written above and theorems 4 and 6 we can rewrite the formula for $D^2(a_1)$. We shall write it down assuming that n is sufficiently large to make the terms of the order $O(n^{-1})$ small. We get

$$D^2(a_1) = \frac{\sigma_Z^2(1-\varrho_1)}{(n-1)(r_1-1)S^2}\left(1+\sum_{v=1}^{n-2}\varkappa_v\varrho_v^*\right). \tag{36}$$

This formula is adequate for comparisons of efficiency of the total differential method with the classical method. As is known, the variance of the estimator of the regression coefficient a_1 of the regression function $E(Y_t) = a_1X_t+a_0$ is given by([3])

$$D_c^2(a_1) = \frac{\sigma_Z^2}{nS^2}\left(1+\sum_{v=1}^{n-1}r_v\varrho_v\right). \tag{37}$$

We can compare now the efficiency of the total differential method of estimation and we come to the following

THEOREM 7. *Let the regression function be $Y_t = a_1x_t+a_0+Z_t$ where x_t is a non-random variable while Z_t forms a stochastic process, stationary to the second degree with $E(Z_t) = 0$ and with finite variance σ_Z^2. Let $D^2(a_1)$*

([3]) Formula (37) was obtained first by H. Wold. See [6].

and $D_c^2(a_1)$ denote respectively the variance of the total differential method estimator of a_1 and the variance of the estimator of a_1 obtained by the classic device. If for $n \to \infty$ the ratio

$$\frac{\sum_{v=1}^{n-2} \varkappa_v \varrho_v^*}{\sum_{v=1}^{n-1} r_v \varrho_v}$$

tends to a real number R or if its absolute value is bounded, then the efficiency of the total differential method estimator of a_1 is given for $n \to \infty$ and such that $O(n^{-1}) \approx$ zero, by:

$$\eta = \frac{D^2(a_1)}{D_c^2(a_1)} = \frac{1-\varrho_1}{1-r_1} \cdot \frac{\sum_{v=1}^{n-2} \varkappa_v \varrho_v^*}{\sum_{v=1}^{n-1} r_v \varrho_v}. \tag{38}$$

As is easily seen, the efficiency of the total differential method depends on the values of the four sequences of autoregression coefficients: those of ΔZ_t, Z_t, ΔX_t and X_t. When, however, the efficiency of the total differential method is discussed, another point must be emphasized. As is well known, very often in practical applications of the regression analysis with more than one independent variable a serious difficulty arises — the collinearity in the statistical data. When dealing with time series data the collinearity is much more likely to occur because usually we find a strong correlation between the variables which will play the role of independent in our regression. The use of the total differential method may sometimes permit us to overcome that difficulty. In fact, it is possible that even when there is a very strong positive (or negative) correlation between X_t and V_t, it may not exist for their first differences — as can be seen from formula (35). This possibility is a strong argument in favour of the new method when time series data are to be analysed by statistical methods.

VI. An attempt to generalize the total differential method to the non-linear case. In this section we shall make some comments about the possibility of a generalization of the total differential method procedure of estimation also to the case when the regression function is not linear but has a well-defined non-linear functional form and when all independent variables entering into it are known. In particular let us assume that

$$E(Y_t) = F(x_{1t}, x_{2t}, \ldots, x_{ht}; a_1, a_2, \ldots, a_k) + a_0, \tag{39}$$

where $a_0, a_1, \ldots, a_k$ are unknown parameters which must be estimated from our statistical data. We assume furthermore that the function F has derivatives with respect to each of the variables $x_{1t}, x_{2t}, \ldots, x_{ht}$. For these derivatives we shall use the following notation:

$$\frac{\partial F}{\partial x_{it}} = f_i(x_{1t}, x_{2t}, \ldots, x_{ht}; a_1, a_2, \ldots, a_k) = f_{it}. \tag{40}$$

Because of random fluctuations we shall have for different time points

$$y_t = F(x_{1t}, x_{2t}, \ldots, x_{ht}; a_1, a_2, \ldots, a_k) + a_0 + z_t, \tag{41}$$

where z_t is a random element. We assume that its mean is zero and variance σ_Z^2 for all time points t.

The expression

$$\Delta y_t = \sum_{i=1}^{h} f_{it} \Delta x_{it} + \Delta z_t \tag{42}$$

can be interpreted as the total differential of Y_t plus a disturbance term Δz_t. Obviously (42) depends on regression parameters $a_1, a_2, \ldots, a_k$. If the number $n-1$ of the observed first finite differences of observed variables is larger than the number of the parameters entering the right hand of the equation (42), then the estimation of $a_1, a_2, \ldots, a_k$ can be done by applying the least squares method to the expression

$$S_3 = \sum_{t=1}^{n-1} \Big(\Delta y_t - \sum_{i=1}^{h} f_{it} \Delta x_{it}\Big)^2 \tag{43}$$

because this procedure is equivalent to the minimization of the disturbance term Δz_t.

The normal equations obtained as the result of this minimization may not be linear with respect to unknown parameters $a_1, a_2, \ldots, a_k$. Solving a system of several non-linear equations might be a hopeless problem, but for a large class of regression functions this difficulty can be overcome by using the following simple device. Let us suppose that the partial derivatives f_{it} appearing in (42) can be expressed as functions of y_t and of x_{it}'s. Let these functions be:

$$\begin{aligned} f_{1t} &= g_{1t}(y_t, x_{1t}, x_{2t}, \ldots, x_{ht}), \\ f_{2t} &= g_{2t}(y_t, x_{1t}, x_{2t}, \ldots, x_{ht}), \\ &\cdots\cdots\cdots\cdots\cdots \\ f_{ht} &= g_{ht}(y_t, x_{1t}, x_{2t}, \ldots, x_{ht}). \end{aligned} \tag{44}$$

Now, if on substituting (44) into (43) we get normal equations which are linear with respect to the unknown parameters $a_1, a_2, \ldots, a_k$, our problem is solved and estimators of these parameters can be obtained. This possibility of using substitution (44) makes the total differential method useful in non-linear cases because owing to it one can deal with regression functions whose parameters would be impossible to estimate by the classical least squares method.

EXAMPLE 2. Let us suppose that $y_t = \sin ax_t + z_t$. Obviously the parameter a cannot be estimated by the classical method, but it can be by the total differential method. We have namely:

$$\Delta y_t = a\cos ax_t \cdot \Delta x_t + \Delta z_t. \tag{45}$$

But $\cos ax_t = \sqrt{1-\sin^2 ax_t} = \sqrt{1-y_t^2}$. Substituting therefore $\cos ax_t$ in (45) by the last expression we get

$$\Delta y_t = a\sqrt{1-y_t^2}\,\Delta x_t + \Delta z_t.$$

As y_t, Δy_t and Δx_t are known, the parameter a can easily be estimated. Minimizing the expression below with respect to a

$$S = \sum_{t=1}^{n-1} (\Delta y_t - a\sqrt{1-y_t^2}\,\Delta x_t)^2$$

we get for the estimator of a

$$a = \frac{\sum\limits_{t=1}^{n-1} \sqrt{1-y_t^2}\,\Delta y_t \Delta x_t}{\sum\limits_{t=1}^{n-1} (1-y_t^2)(\Delta x_t)^2}.$$

EXAMPLE 3. Let us suppose that $y_t = \sqrt{ax_t+1} + z_t$. The estimation of the parameter a can be done by the total differential method because

$$\Delta y_t = \frac{a}{2\sqrt{ax_t+1}}\Delta x_t + \Delta z_t \qquad \text{but} \qquad f_t = \frac{a}{2\sqrt{ax_t+1}} = \frac{a}{2y_t}$$

and therefore

$$\Delta y_t = \frac{a}{2y_t}\Delta x_t + \Delta z_t,$$

whence it is easily seen that the resulting equation obtained after the minimization of (43) will be linear with respect to the parameter a.

In general, as in the linear case, the estimation procedure will have two steps. In the first, estimates $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ of $a_1, a_2, \ldots, a_k$ are obtained, while the second step is devoted to the estimation of the constant term a_0. We put $a_1 = \boldsymbol{a}_1$, $a_2 = \boldsymbol{a}_2, \ldots, a_k = \boldsymbol{a}_k$ and then minimize the expression

$$S_4 = \sum_{t=1}^{n} \big(y_t - F(x_{1t}, \ldots, x_{ht}; \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k) - a_0\big)^2$$

with respect to α_0. The estimator a_0 of α_0 will be that value of α_0 for which S_4 has the smallest value.

The total differential method estimators, got in the non-linear case in the way mentioned above, may happen to be biased. This, however, is not a serious difficulty and it can easily be remedied because in general it would not be too difficult to find the expected values of estimators.

Let us assume in particular that we find

$$E(a_j) = \alpha_j b_j, \quad j = 0, 1, \ldots, k,$$

where $b_j \neq 1$, so that a_j is biased. However, if instead of a_j we take as the estimator for α_i the expression a_j^* defined as

$$a_j^* = \frac{a_j}{b_j},$$

we get an unbiased estimator of α_j.

References

[1] O. Anderson, *Die Korrelationsrechnung in der Konjunkturforschung*, Bonn 1929.

[2] Student, *The elimination of spurious correlation due to position in time and space*, Biometrika 1914.

[3] G. Tintner, *The variate-difference method*, Cowles Commission Monograph No. 6, New York 1940.

[4] H. Wold, *Demand analysis*, New York 1952.

[5] — *A study in the analysis of stationary time series*, Uppsala 1954, 2 ed.

[6] — *On least square regression with autocorrelated variables and residuals*, Bull. Inst. Intern. Stat. 32 (1952), part two.

Praca wpłynęła 18. 12. 1958

Z. PAWŁOWSKI (Warszawa)

O METODZIE ESTYMACJI ZA POMOCĄ RÓŻNICZKI ZUPEŁNEJ I JEJ EFEKTYWNOŚCI W PRZYPADKU REGRESJI LINIOWEJ

STRESZCZENIE

Przedmiotem artykułu jest omówienie pewnej szczególnej metody estymacji parametrów funkcji regresji stosowanej w przypadku, gdy materiałem statystycznym są szeregi czasowe. Autor nazywa tę metodę *metodą estymacji za pomocą różniczki*

zupełnej, gdyż w procesie estymacji podstawową rolę odgrywa wyrażenie, które może być traktowane jako różniczka zupełna zmiennej zależnej względem zmiennych niezależnych funkcji regresji $X_{1t}, X_{2t}, \ldots, X_{ht}$. Omawiana metoda estymacji stosowana była już niekiedy w przeszłości przez różnych autorów, jednak nie badano nigdy jej efektywności i innych własności otrzymywanych za jej pomocą estymatorów.

Rozdział II omawia sposób szacowania parametrów regresji, w przypadku gdy funkcja regresji ma postać (1). Dla poszczególnych, obserwowanych momentów czasu zachodzić więc będzie związek (2), gdzie Z_t jest składnikiem przypadkowym o średniej zero i stałej dla wszystkich t wariancji σ_z^2. O zmiennych $X_{1t}, X_{2t}, \ldots, X_{ht}$ zakładamy, iż nie są one losowe. Parametry $\alpha_0, \alpha_1, \ldots, \alpha_h$ są nieznane i należy je oszacować na podstawie zebranych materiałów statystycznych dotyczących wartości $Y, X_1, X_2, \ldots, X_h$ w różnych okresach (momentach) czasu.

Wprowadzając pierwsze skończone różnice obserwowanych zmiennych, określone wzorami (4) i (5), szacuje się parametry $\alpha_1, \alpha_2, \ldots, \alpha_h$ za pomocą metody najmniejszych kwadratów. Uzyskane w ten sposób estymatory $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_h$ są zgodne i nieobciążone, jeżeli tylko rzeczywiście funkcja regresji Y_t względem X_{it} ma postać (1). Stały składnik α_0 jest szacowany w ten sposób, że przyjmuje się $\alpha_1 = \mathbf{a}_1$, $\alpha_2 = \mathbf{a}_2, \ldots, \alpha_h = \mathbf{a}_h$, a następnie minimizuje się wartość wyrażenia (9) względem α_0.

Rozdział III poświęcony jest dowodom nieobciążoności estymatorów parametrów $\alpha_0, \alpha_1, \ldots, \alpha_h$.

W rozdziale IV znajduje się wariancję estymatora parametru α_1 funkcji regresji $E(Y_t) = \alpha_1 X_t + \alpha_0$ wyznaczonego za pomocą metody różniczki zupełnej. Wariancja ta jest dana wzorem (19), w którym ϱ_v^* oznaczają współczynniki autokorelacji ΔZ_t i ΔZ_{t+v}, a współczynniki $\varkappa_v$ są określone wzorem (18). Ponieważ wariancja ta dąży do zera, gdy $n \to \infty$, a estymator $\mathbf{a}_1$ jest nieobciążony, przeto wynika stąd, że estymator ten jest również i zgodny. W podobny sposób dowodzi się, że zgodny jest również estymator stałego składnika α_0, $\mathbf{a}_0$.

Celem porównania efektywności metody estymacji z klasycznymi metodami estymacji parametrów regresji liniowej za pomocą różniczki zupełnej wzór (19) tak się przekształca, by wariancja estymatora wyrażona była za pomocą parametrów dla Y_t i X_{it} $(i = 1, 2, \ldots, h)$, a nie za pomocą parametrów ich pierwszych różnic. Wykazuje się, że odrzucając składniki rzędu $O(n^{-1})$ i wyższych, wariancję tego estymatora można przedstawić w postaci (36). Wzór (38) daje stosunek wariancji estymatora $\mathbf{a}_1$ do wariancji estymatora parametru α_1 uzyskanego za pomocą metody klasycznej. Jak widać, efektywność nowej metody zależy od współczynników autokorelacji ϱ_v^* (zmiennej ΔZ_t i ΔZ_{t+v}), ϱ_v (zmiennych Z_t i Z_{t+v}), r_v (zaobserwowanych dla X_t i X_{t+v}) oraz od współczynników $\varkappa_v$.

Ostatni rozdział poświęcony jest próbie uogólnienia metody estymacji za pomocą różniczki zupełnej, w przypadku gdy funkcja regresji nie jest liniowa. Autor wskazuje na fakt, iż za pomocą nowej metody można szacować parametry takich funkcji regresji, które nie pozwalają na użycie klasycznej metody najmniejszych kwadratów, jak na przykład funkcji $E(Y_t) = \sin \alpha x_t$.

З. ПАВЛОВСКИЙ (Варшава)

О МЕТОДЕ ПОЛНОГО ДИФФЕРЕНЦИАЛА И ЕГО ЭФФЕКТИВНОСТИ В СЛУЧАЕ ЛИНЕЙНОЙ РЕГРЕССИИ

РЕЗЮМЕ

Работа посвящена специальному методу оценки параметров функции регрессии

$$E(Y_t) = \sum_{i=1}^{h} a_i x_{it} + a_0,$$

а также сравнению эффективности данного метода (называемого автором *методом полного дифференциала*) с классическим методом оценки параметров функции регрессии. Беря первые разности отмеченных значений переменных определённые формулами (4) и (5), оценки $a_1, a_2, \ldots, a_h$ параметров $\alpha_1, \alpha_2, \ldots, \alpha_h$ выводятся путём минимализации выражения S_1 по отношению $a_1, a_2, \ldots, a_h$. Оценка свободного члена a_0 получается при помощи применения метода наименьших квадратов к выражению $\sum_{t=1}^{n}(y_t - \sum_{i=1}^{h} a_i x_{it})^2$.

Автор доказывает, что полученные таким образом оценки являются состоятельными и несмещёнными. В дальнейшей части работы исследуется эффективность метода полного дифференциала для случая $h = 1$. Эффективность эта, выраженная формулой (38), зависит от значения четырех последовательностей коэффициентов автокорреляции: от значения независимой переменной x_t, от первых разностей этой переменной, от значения остаточного члена z_t, а также от первых разностей этого члена.

Последняя часть работы является попыткой обобщения метода полного дифференциала на случай нелинейных функций регрессии.