ESTIMATING PARAMETERS OF LINEAR REGRESSION WITH AN EXPONENTIAL POWER DISTRIBUTION OF ERRORS BY USING A POLYNOMIAL MAXIMIZATION METHOD

This paper considers the application of a method for maximizing polynomials in order to find estimates of the parameters of a multifactorial linear regression provided the random errors of the regression model follow an exponential power distribution. The method used is conceptually close to a maximum likelihood method because it is based on the maximization of selective statistics in the neighborhood of the true values of the evaluated parameters. However, in contrast to the classical parametric approach, it employs a partial probabilistic description in the form of a limited number of statistics of higher orders.

The adaptive algorithm of statistical estimation has been synthesized, which takes into consideration the properties of regression residues and makes it possible to find refined values for the estimates of the parameters of a linear multifactorial regression using the numerical Newton-Rafson iterative procedure. Based on the apparatus of the quantity of extracted information, the analytical expressions have been derived that make it possible to analyze the theoretical accuracy (asymptotic variances) of estimates for the method of maximizing polynomials depending on the magnitude of the exponential power distribution parameters.

Statistical modeling was employed to perform a comparative analysis of the variance of estimates obtained using the method of maximizing polynomials with the accuracy of classical methods: the least squares and maximum likelihood. Regions of the greatest efficiency for each studied method have been constructed, depending on the magnitude of the parameter of the form of exponential power distribution and sample size. It has been shown that estimates from the polynomial maximization method may demonstrate a much lower variance compared to the estimates from a least-square method. And, in some cases (for flat-topped distributions and in the absence of a priori information), may exceed the estimates from the maximum likelihood method in terms of accuracy

Keywords: regression, exponential power distribution, parameter estimation, moments, polynomial maximization method

S. Zabolotnii
Doctor of Technical Sciences,
Associate Professor*
E-mail: zabolotniua@gmail.com
V. Khotunov
PhD*
E-mail: khotunov@gmail.com
A. Chepynoha
PhD, Associate Professor*
E-mail: toxachep@ukr.net
O. Tkachenko
Postgraduate Student*
E-mail: tkachenko.ck@gmail.com
*Department of Computer Engineering
and Information Technology
Cherkasy State Business College
V. Chornovola str., 243, Cherkasy, Ukraine, 18028

1. Introduction

At the beginning of the xix century, K. Gauss published several fundamental works, which introduced, among other things, a series of important statistical concepts such as "a least-square method", "maximum likelihood", as well as substantiated a "normal" distribution. This distribution law, often referred to as Gaussian, is by far the most common technique of probabilistic description. However, based on a series of theoretical limitations, it does not describe all possible variety of statistics. The axiomatic statement underlying Gauss's approach was slightly altered at the beginning of the 20th century in work [1], where a more general law for the distribution of errors was justified, termed as an exponential power distribution (EPD). This probabilistic model is better suited to describe a larger number of phenomena that can be observed in reality. Its versatility is evidenced by the fact that the Laplace distribution, Gaussian distribution, and uniform distributions are separate cases of EPD at certain values of its form parameter. Therefore, various literary sources [2–5] use alternative designations such as the "generalized normal distribution", "generalized Gaussian distribution", "generalized distribution of errors", "Subbotin distribution", etc. In addition, the family of exponential power distributions includes more complex multimodal [6] or asymmetrical [7, 8] modifications.

The properties and aspects of EPD practical application are addressed in many studies. In particular, works [2, 3, 9, 10] consider the methods of finding point and, in [11], interval estimates of EPD parameters. Paper [12] reported the EPD characteristic function in the form of a Taylor series, and [5] investigated the properties of statistics of higher orders of this distribution.

The relevance of the problem under study is due to the fact that EPD is used as a probabilistic model in solving a large number of various practical tasks. In particular, to describe the error of measuring instruments and the results of measurement [13, 14], random noise and interference during the processing of radio [15], video [16], and audio [17] signals; medical [18], biological [19], economic [20] statistics; in risk management [21], etc. Special features in the approximation of real data and verification of statistical hypotheses of correctness from an EPD model are discussed in [22, 23]. Methods of generating random sequences for the implementation of statistical modeling are tackled in works [24–26]. Such a widespread EPD application gave rise to the development of a series of specialized software focused on computer statistical simulation. Among them, we note the library "normalp" [27] in the high-level programming language R, one of the most common among statisticians, which contains a set of functions for the statistical treatment of data acquired from EPD.

2. Literature review and problem statement

It is known that a least-square (LS) method is one of the most popular approaches to solving problems related to regression data analysis. This is due to the fact that a ordinary least-square (OLS) method makes it possible to derive a solution in a closed form. It does not require additional *a priori* information about the probabilistic properties of the error model and, when meeting a series of conditions imposed by the Gauss-Markov theorem, is the best linear unbiased appraiser. One of these conditions is the normalization of the distribution of regression errors, which minimizes the variance of parameter estimates [28]. However, in real conditions, there may be a significant deviation of the data distribution from the Gaussian law, which significantly impairs the accuracy of OLS assessments, in particular, in the presence of individual remote observations (emissions) [29].

One technique to improve the accuracy of regression parameter estimates for non-Gaussian data is to replace the quadratic loss function with another. This idea is not new, since back in 1793 Laplace proposed a method of evaluation based on the least absolute deviations (LAD) method, known as the regression appraiser in space L_1 [30]. Such an appraiser, when compared to OLS, is more effective for leptokurtic error distributions demonstrating tails that are heavier than those in the Gaussian law. A generalization of this approach is the construction of regression appraisers in the L_p space, advanced in works [20, 31–33]. Another approach that ensures robustness is the use of nonparametric estimates of the sign [34] or quantile regression [29], which could be successfully applied to both heavy-tailed and asymmetric distributions. It is known that each above loss function has appropriate distribution models under which the estimates of the maximum likelihood estimation (MLE) method are equivalent to the estimates of methods that minimize corresponding losses. In particular, the OLS appraiser corresponds to MLE when the error distribution is normal; the LAD appraiser is equivalent to MLE for the Laplace distribution, and the L_p appraiser corresponds to MLE when regression errors are described by EPD [35].

It is obvious that the key to the use of MLE is the *a priori* determination of the type of probability distribution. Note that in addition to the EPD family, many different types of symmetric distributions are used to solve the problems of regression analysis: elliptical laws (logistic, Cauchy, Student *t*-distribution) [36–38], various Gaussian distributions [39, 40]. In addition, there are specially developed models of symmetric distributions that make it possible to change the value of the excess coefficient and the severity of the tails of regression errors [41, 42]. The task is generally complicated by the fact that, in addition to selecting an adequate model for the distribution of errors, it is necessary to carry out a joint assessment of their parameters, the uncertainty of which significantly affects the accuracy of assessments of informative regression parameters. Under an assumption that model errors do not depend on predictors, an adaptive approach [43] is often used, which involves a series of sequential steps. At the first stage, the parameters of the deterministic component of the model are evaluated by a simple method that does not require additional information about the specificity of the probabilistic distribution of statistical data. After removing the deterministic component, the type of random component of the model is identified, and estimates of its parameters are found. At the third stage, they find refined (adaptive) assessments of the parameters of the deterministic component of the model, taking into consideration *a posteriori* properties of errors.

The adaptive approach can be based not only on likelihood maximization. An alternative is to use higher-order statistics, such as moments or cumulants, for the probabilistic description of regression models [44–47]. The application of this notation in combination with the adaptive approach to evaluation is the basis of a quadratic modification of the least-square method [48]. Different variations of this appraiser demonstrate their effectiveness only in the asymmetry of distributions, using as additional information the values of *a posteriori* estimates of asymmetry coefficients and the excess of regression residues [49–51].

The principle of adaptive evaluation by taking into consideration the properties of statistics of higher orders underlying the second-order least squares estimator (SLS) is similar to the idea of using a polynomial maximization method (PMM) to evaluate the regression parameters [52]. Paper [53] shows that when using, as the basis functions of power transformations, the PMM-estimation for the power $S=2$ demonstrates similar accuracy with the estimates of the SLS. The common property is also that in the symmetry of the distribution of statistical data, SLS (like PMM at $S=2$) degenerates into OLS. Therefore, it cannot be used to evaluate regression parameters with symmetric errors of the EPD type [53, 54].

We also note a certain conceptual similarity of PMM to MLE, which implies that both methods employ the principle of maximizing certain selective statistics in the vicinity of the true values of the parameters that are evaluated. However, in contrast to MLE, PMM uses not the density of distribution, but a simpler probabilistic description in the form of the final number of moments or cumulants to form such statistics. Note the functionality of PMM, which was successfully used to assess the shear of symmetrical [55–57] and

asymmetrical [58] distributions, to determine the moment of changes (disorder) of the properties of random sequences in the *a posteriori* problem statement [59], and others. In particular, works [60, 61] consider the use of cubic (at the power of polynomial *S*=3) modification of PMM to solve the problem of adaptive evaluation of the scalar parameter of experimental data acquired from EPD. Their results suggest that PMM estimates may have significantly lower variances both compared to nonparametric estimates (median, middle, and mid-range) and to the estimates of maximum likelihood. But, since both the linear multifactor regression model and the principles of finding assessments of components of its vector parameter are more complex and differ from those discussed in the cited works, the issue of the effectiveness of applying a polynomial maximization method for such a task remains unresolved.

3. The aim and objectives of the study

The purpose of this study is to synthesize the algorithms of polynomial evaluation of parameters of linear regression and to analyze their accuracy depending on the properties of regression errors having exponential power distribution. This would make it possible to make a reasoned choice (taking into consideration the specificity of real problems) between the proposed solution and approaches based on classical methods.

To accomplish the aim, the following tasks have been set:
– to apply a polynomial maximization method to synthesize algorithms for the adaptive evaluation of linear regression parameters;
– to analyze the theoretical accuracy of PMM estimations depending on the values of regression error parameters;
– to carry out, by statistical modeling, a comparative analysis of the effectiveness (according to a criterion of the amount of variance) of MLE estimates with respect to OLS and MLE.

4. Using a polynomial maximization method to evaluate regression parameters

Suppose there is a linear multifactor regression model of N observations that describes the dependence of the values of the target variable $\mathbf{Y}$ on the regressors $\mathbf{X}$

$$y_v = f_v(\boldsymbol{\theta}, \mathbf{X}) + \xi_v, \ v = \overline{1, N}, \tag{1}$$

where the deterministic component

$$f_v(\boldsymbol{\theta}, \mathbf{X}) = \sum_{p=0}^{Q-1} a_p x_{p,v}, \tag{2}$$

is the linear function with a vector parameter $\theta = \{a_0, a_1, ..., a_{Q-1}\}$, whose components are subject to evaluation.

It is assumed that regression errors ξ_v represent a sequence of independent and equally distributed random variables, which are adequately described by the exponential power distribution in the following form

$$w(\xi) = \frac{1}{2\sigma_\beta p^{1/\beta} \Gamma(1 + 1/\beta)} \exp\left(-\frac{|\xi|^\beta}{\beta \sigma_\beta^\beta}\right), \tag{3}$$

where the values of the parameters are of the scale σ_β and form β, which may be *a priori* unknown.

To solve the set problem of regression analysis, one can use a PMM modification for the statistical evaluation of vector parameters with unevenly distributed data [52]. It is based on the property of maximizing the functionality in the form of a stochastic polynomial of the S order in a general form

$$L_{SN} = \sum_{v=1}^{N} \sum_{i=1}^{S} \phi_i(y_v) \int\limits_{}^{a} k_{iv}(a) \mathrm{d}z - \sum_{i=1}^{S} \sum_{v=1}^{N} \int\limits_{}^{a} \Psi_{iv} k_{iv}(a) \mathrm{d}z, \tag{4}$$

in the vicinity of the true value of some parameter a to be evaluated. As the basis functions of polynomial (4), one can use power transformations $\phi_i(y_v) = y_v^i$. It is assumed that for random values y_v there is a sequence of initial moments α_i of the corresponding order

$$\Psi_{iv} = E\{y_v^i\} = \alpha_{iv}, \ i = \overline{1, S}, \ v = \overline{1, N}, \tag{5}$$

and there are twice differentiated for the parameter a.

By analogy with MLE, the estimate of the parameter a can be derived by solving the following equation

$$\frac{d}{da} L_{SN}\bigg|_{a=\hat{a}} = \sum_{i=1}^{S} \sum_{v=0}^{N} k_{iv} \left[y_v^i - \alpha_{iv}\right]\bigg|_{a=\hat{a}} = 0. \tag{6}$$

The optimal k_{iv} coefficients that maximize functionality (4) and minimize the variance of estimates (for the corresponding order of the polynomial S) of the parameters are derived by solving the following system of linear algebraic equations

$$\sum_{j=1}^{S} k_{jv} F_{(i,j)v} = \frac{d}{da} \alpha_{iv}, \ i = \overline{1, S}, \ v = \overline{1, N}, \tag{7}$$

where $F_{(i,j)v} = \alpha_{(i+j)v} - \alpha_{iv} \alpha_{jv}, \ i, j = \overline{1, S}$.

To find estimates of the vector parameter $\theta = \{a_0, a_1, ..., a_{Q-1}\}$, it is necessary to use Q polynomials $L_{SN}^{(p)}, \ p = \overline{0, Q-1}$ for each component of the vector parameter a_p. In this case, every p-th stochastic polynomial $L_{sn}^{(p)}$ as a function of the parameter a_p at known values of other components of the vector also has a maximum in the vicinity of the true value of this parameter at $N \rightarrow \infty$. Thus, the desired estimates can be found by solving the system of Q equations in the following form

$$\sum_{i=1}^{S} \sum_{v=1}^{N} k_{iv}^{(p)} \left[y_v^i - \alpha_{iv}\right]\bigg|_{a_p = \hat{a}_p} = 0, \ p = \overline{0, Q-1}. \tag{8}$$

It should be noted that when using polynomials of the $S \geq 2$ power, finding the PMM estimates of a vector parameter in most cases requires the use of numerical methods for solving the systems of nonlinear equations. Our work employs an approach based on a Newton-Rafson iterative numerical procedure. This numerical method is often used in similar situations when using PMM to find estimates for the linear regression parameters [62, 63]. It is based on the principle of linearization by decomposing the left-hand part of each nonlinear equation of system (8) into a Taylor series in the neighborhood of the true value of the vector θ. If limited to the first two terms of the series, it is possible to record the following linear system in a matrix form

$$\hat{\theta}_{(S)}^{(k+1)} = \hat{\theta}_{(S)}^{(k)} - \left[\mathbf{Z}_S\left(\mathbf{Y} / \hat{\theta}^{(k)}\right)\right]^{-1} \mathbf{F}_S\left(\mathbf{Y} / \hat{\theta}^{(k)}\right), \tag{9}$$

which can be used for the iterative search of valuation values. To derive system (9), it is necessary to calculate the ma-

trix-column $\mathbf{F}_s(\mathbf{Y}/\theta^{(k)})$, which is composed of the elements of the left-hand part of each nonlinear equation in system (8), and a square matrix

$$\mathbf{Z}_S\left(\mathbf{Y}/\hat{\theta}^{(k)}\right) = \mathbf{H}_S\left(\mathbf{Y}/\hat{\theta}^{(k)}\right) + \mathbf{J}_S\left(\hat{\theta}^{(k)}\right)$$

with components

$$H_{SN}^{(p,q)} = \sum_{i=1}^{S} \frac{\partial}{\partial a_q} k_{iv}^{(p)} \left[\sum_{v=1}^{N} y_v^i - N\alpha_{iv}\right],$$

where the elements of the matrix of the amount of information extracted

$$J_{SN}^{(p,q)} = \sum_{v=1}^{N} \sum_{i=1}^{S} k_{i,v}^{(p)} \frac{\partial}{\partial a_q} \alpha_{iv}, \ \ p,q = \overline{0,Q-1}. \quad (10)$$

To start the iterative procedure, it is assumed that there is some initial approximation $\theta^{(1)}$, for which "rough" assessments can be selected, to be found by a simpler evaluation method [52].

5. Finding PMM estimates for the linear multifactorial regression parameters

Works [53, 54] show that with the symmetry of the distribution of the model of regression residues, the PMM estimations, derived by using the $S=1$ and $S=2$ order polynomials, are equivalent to the OLS estimates. Therefore, we consider the case of finding the PMM estimates based on the use of the $S=3$ order polynomials. The ratio for the first 6 initial moments, necessary to build a system of equations (6), can be represented in the form:

$$\alpha_{1v} = f_v(\boldsymbol{\theta}, \mathbf{X}), \ \ \alpha_{2v} = \left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 + \mu_2,$$

$$\alpha_{3v} = \left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^3 + 3 f_v(\boldsymbol{\theta}, \mathbf{X})\mu_2,$$

$$\alpha_{4v} = \left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^4 + 6\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 \mu_2 + 3\mu_2^2 + \mu_4,$$

$$\alpha_{5v} = \left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^5 + 10\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^3 \mu_2 + $$
$$+5\left(3\mu_2^2 + \mu_4\right) f_v(\boldsymbol{\theta}, \mathbf{X}),$$

$$\alpha_{6v} = \left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^6 + 15\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^4 \mu_2 + $$
$$+45\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 \mu_2^2 + 15\mu_2^3 + $$
$$+15\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 \mu_4 + 15\mu_2\mu_4 + \mu_6, \quad (11)$$

where μ_i are the central moments of regression errors $\downharpoonright_v$, which depend on the parameters of scale σ_β and form β of distribution (3) and can be calculated from a general formula given in [27]:

$$\mu_i = \Gamma\left[(i+1)\beta^{-1}\right]\Gamma\left[\beta^{-1}\right]^{\frac{i-2}{2}} \Gamma\left[3\beta^{-1}\right]^{\frac{i}{2}} \sigma_\beta^2. \quad (12)$$

To find the optimal coefficients that minimize the variance of PMM estimates at $S=3$, the following expressions are necessary for derivatives from the first 3 moments:

$$\frac{\partial}{\partial \theta}\alpha_1 = 1. \ \ \frac{\partial}{\partial \theta}\alpha_2 = 2 f_v(\boldsymbol{\theta}, \mathbf{X}) + \mu_2,$$

$$\frac{\partial}{\partial \theta}\alpha_3 = 3\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 + 3\mu_2. \quad (13)$$

By solving a system of equations (7) analytically by a Kramer method (taking into consideration expressions (11) and (13)), we obtain:

$$k_{1,v}^{(p)} = x_{p,v} \frac{3\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 \left(\mu_4 - 3\mu_2^2\right) + 3\mu_4\mu_2 - \mu_6}{\mu_2^{-2}\left(\mu_4^2 - \mu_2\mu_6\right)},$$

$$k_{2,v}^{(p)} = x_{p,v} \frac{-3 f_v(\boldsymbol{\theta}, \mathbf{X})\left[\mu_4 - 3\mu_2\right]}{\mu_2^{-2}\left(\mu_4^2 - \mu_2\mu_6\right)},$$

$$k_{3,v}^{(p)} = x_{p,v} \frac{\mu_4 - 3\mu}{\mu_2^{-2}\left(\mu_4^2 - \mu_2\mu_6\right)}, \ \ p = \overline{0,Q-1}, \quad (14)$$

Substituting the coefficients (14) in (8), after certain transformations, the system of equations for finding the estimates can be written as:

$$\sum_{v=1}^{N}\left\{x_v^p \begin{bmatrix} A\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^3 + \\ +B_v\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 + \\ +C_v\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right] + D_v \end{bmatrix}\right\} = 0, \ \ p = \overline{0,Q-1}, \quad (15)$$

where

$$A = 1, \ B_v = -3 y_v,$$

$$C_v = 3 y_v^2 - \frac{\mu_6 - 3\mu_4\mu_2}{\mu_4 - 3\mu_2^2},$$

$$D_v = y_v \frac{\mu_6 - 3\mu_4\mu_2}{\mu_4 - 3\mu_2^2} - y_v^3.$$

It is obvious that the PMM estimates can only be found through a numerical solution to equation systems, in particular, using the Newton-Rafson iterative procedure (9), for which it is necessary to calculate the following elements

$$Z_{3N}^{(p,q)} = \sum_{v=1}^{N} \sum_{v=1}^{N}\left\{x_v^{p+g} \begin{bmatrix} 3A\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right]^2 + \\ +2B_v\left[f_v(\boldsymbol{\theta}, \mathbf{X})\right] + C_v \end{bmatrix}\right\}. \quad (16)$$

As an initial approximation, those OLS estimates can be used whose formation does not require additional information about the properties of regression residues.

6. Accuracy of PMM estimates of the linear multifactorial regression parameters

Since, at $S=1$, the PMM estimates of the components of vector parameter θ of linear regression model (1) are equivalent to OLS estimates, their accuracy, in this case, coincides. The variation matrix containing the variance value of such estimates takes the following form

$$\mathbf{V}_{(OLS)} = \mathbf{V}_{(PMM1)} = \mu_2\left[\mathbf{X}\mathbf{X}^T\right]^{-1}. \quad (18)$$

It is known that such linear estimates are optimal (according to the criterion of the minimum variance) only for

the situation when errors of the regression model follow a Gaussian distribution, which is known to be a separate case of EPD (3) at $\beta=2$.

For the analytical calculation of the amount of variance in the parameter estimates by a maximum likelihood method, we use a matrix of the amount of information according to Fisher, calculated for regression models with EPD [64]. Taking into consideration (18), the corresponding variation matrix can be represented in the following form

$$\mathbf{V}_{(\text{MLE})} = g_{\text{MLE/OLS}}\mathbf{V}_{(\text{OLS})},$$

where

$$g_{\text{MLE/OLS}} = \frac{\Gamma\left[1+\beta^{-1}\right]^2}{\Gamma\left[2-\beta^{-1}\right]\Gamma\left[3\beta^{-1}\right]}. \tag{19}$$

To derive analytical expressions describing the variances in the PMM estimates of components of the vector parameter θ, we use the matrix $\mathbf{J}_s(\theta)$ of the amount of information obtained when applying the S order stochastic polynomials, consisting of elements (10). It has been proven in [52] that with an increase in the order of stochastic polynomial the variance of PMM estimates decreases while the amount of information extracted at $S\to\infty$ tends to the information by Fisher. In this case, the variances in the PMM estimates of the components of vector parameter θ in an asymptotic case (at $N\to\infty$) are elements of the main diagonal of the variation matrix, which is inverse to $\mathbf{J}_s(\theta)$.

Using expressions (12) for EPD moments and (14) for optimal coefficients, we can show that the elements of the variation matrix of PMM estimates at $S=3$ differ from the elements of the matrix of linear OLS estimates (18) by the following coefficient

$$\mathbf{V}_{(\text{PMM3})} = g_{\text{PMM3/OLS}}\mathbf{V}_{(\text{OLS})},$$

that can be represented in the form

$$g_{\text{PMM3/OLS}} =$$
$$= \frac{\Gamma\left[p^{-1}\right]^2\left(\Gamma\left[5\beta^{-1}\right]^2-\Gamma\left[3\beta^{-1}\right]\Gamma\left[7\beta^{-1}\right]\right)}{\Gamma\left[3\beta^{-1}\right]\left(9\Gamma\left[3\beta^{-1}\right]^3-6\Gamma\left[\beta^{-1}\right]\Gamma\left[3\beta^{-1}\right]\Gamma\left[5\beta^{-1}\right]+\Gamma\left[\beta^{-1}\right]^2\Gamma\left[7\beta^{-1}\right]\right)}. \tag{20}$$

It is necessary to note the versatility of expressions (20) and (19) since they are equivalent to similar formulae describing the ratio of variances in the MLE estimates and PMM estimates to the variance of linear estimation in the form of arithmetic mean shift of EPD [61].

The analysis of these dependences, shown in Fig. 1, reveals that for the leptokurtic (sharp-topped) distributions at $\beta<2$ the relative efficiency of MLE can be significantly higher than other methods. In the case of Gaussian distribution at $\beta=2$, the accuracy of all three methods (MLE, PMM, and OLS) is the same. As the value of the form parameter $\beta>2$ increases, for a sufficiently wide range of values, the theoretical efficiency of MLE and PMM does not actually differ but can significantly exceed OLS. For the essentially plateau-kurtic (flat-topped) distributions at $\beta>2$, the relative efficiency of MLE increases slightly relative to PMM.

Note that the estimate of the coefficient of reduction of variance (20) can be represented not only as a function dependent on the parameter of the form of EPD but also

as dependence on the values of the central moments μ_r or dimensionless cumulant coefficients γ_r:

$$g_{\text{PMM3/OLS}} = \frac{\mu_2\mu_6-\mu_4^2}{9\mu_2^3-6\mu_2\mu_4+\mu_6} = 1-\frac{\gamma_4^2}{6+9\gamma_4+\gamma_6}, \tag{21}$$

where

$$\gamma_4 = \frac{\mu_4}{\mu_2^2}-3,$$

$$\gamma_6 = \frac{\mu_6}{\mu_2^3}-15\frac{\mu_4}{\mu_2^2}+30.$$

This representation makes it possible to more simply estimate the value of the efficiency of PMM estimates depending on the degree of non-Gaussian regression errors [54].
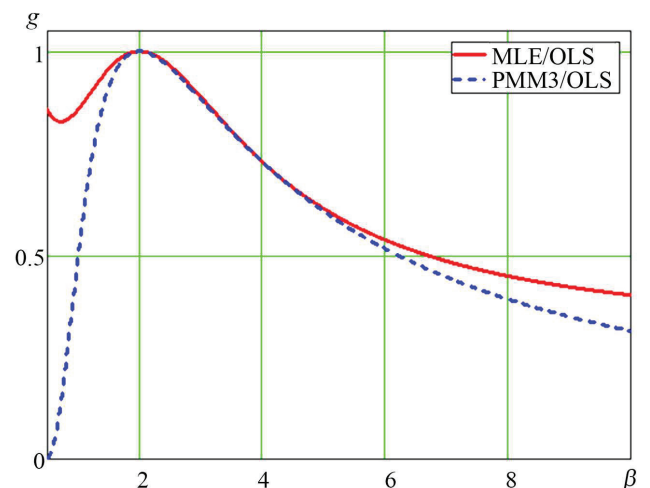

Fig. 1. Dependence of the coefficient of reduction in the variance $g_{\text{MLE/OLS}}$ and $g_{\text{PMM3/OLS}}$ on the EPD β form parameter

7. Statistical modeling by Monte Carlo method

Based on the statistical modeling by the Monte Carlo method, we have experimentally verified the theoretical results of our analysis of the effectiveness of various methods for evaluating the parameters of a linear regression model with EPD errors. Similar to works [60, 61], its implementation employed the R language with the software module "normalp package". This module contains a set of functions for generating random values from EPD, as well as statistical estimation of linear regression parameters with EPD errors by a maximum likelihood method [27].

Since the theoretical values of the coefficients of the ratio of the variance in estimates (20) and (19) are the same for all components of the vector parameter θ, we used, as their quantitative empirical analogs, geometric averaging in the following form:

$$\hat{g}_{\text{MLE/OLS}} = \left(\prod_{p=0}^{Q-1}\frac{\hat{\sigma}^2_{(a_p)\text{MLE}}}{\hat{\sigma}^2_{(a_p)\text{OLS}}}\right)^{\frac{1}{Q}},$$

$$\hat{g}_{\text{PMM3/OLS}} = \left(\prod_{p=0}^{Q-1}\frac{\hat{\sigma}^2_{(a_p)\text{PMM3}}}{\hat{\sigma}^2_{(a_p)\text{OLS}}}\right)^{\frac{1}{Q}}, \tag{22}$$

where

$$\hat{\sigma}^2_{(a_p)\text{OLS}}, \quad \hat{\sigma}^2_{(a_p)\text{MLE}}, \quad \hat{\sigma}^2_{(a_p)\text{PMM3}}$$

are the values of the variance of the parameter estimates obtained using the appropriate method averaged on the basis of M experiments.

It should be noted that statistical modeling was carried out for two fundamentally different situations in terms of the presence of *a priori* information on the characteristics of the random component of the regression model. In the first (idealized, from a practical point of view) case, the value of the parameters of EPD distribution (3) was considered *a priori* known, and their values were used in finding the PMM estimates. This information also made it possible to calculate, based on (12), the values of paired moments up to the 6th order, required to find PMM estimates at power $S=3$. It is obvious that, in real conditions, such *a priori* information is practically absent. Therefore, in the second case, an adaptive approach is used. As mentioned above, it is based on the hypothesis about the similarity of the distribution of a random component of the regression model and regression residues obtained using linear evaluation methods, in particular, OLS. Thus, instead of *a priori* parameter values necessary for both the use of MLE and PMM, their *a posteriori* estimates, found on regression OLS residues, can be used. It is obvious that this approach is asymptotically-approximate and its accuracy improves with an increase in the sample size N.

For statistical modeling of parameter evaluation, the deterministic component of regression model (2) used was a two-factor linear dependence with a parameter vector $\theta=\{1, 3, 1\}$, which were considered unknown and subject to evaluation. Tables 1, 2 give a set of statistical modeling results for different values of the magnitude of the parameter of EPD form $\beta=1, 2, 3, 5, 10$, and sample sizes $N=20, 50, 100, 200$, obtained from $M=10^4$ multiple experiments.

Fig. 2 shows the boundaries that designate the regions of greatest efficiency (based on the criterion of the minimum variance) when applying various evaluation methods: OLS, MLE, and PMM.
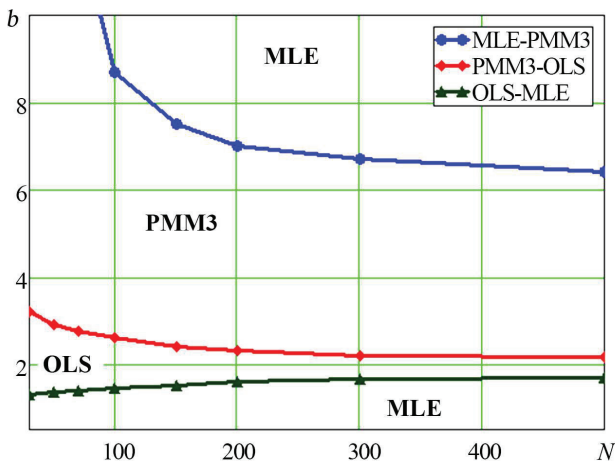

Fig. 2. Efficiency regions of methods for finding estimates of linear regression parameters with EPD errors

These regions are built for the practical situation of the absence of *a priori* information about the true values of the parameters of a random component of the regression model.

They are also derived from statistical modeling by the Monte Carlo method (for $M=10^4$ experiments) at different values of the EPD form parameter β and sample size N.

Table 1

Coefficients of the ratio of variance in estimates in the presence of *a priori* information

β	$\hat{g}_{\text{PMM3/OLS}}$				$g_{\text{PMM3/OLS}}$	$\hat{g}_{\text{MLE/OLS}}$				$g_{\text{MLE/OLS}}$
	N					N				
	20	50	100	200		20	50	100	200	
10	0.98	0.56	0.46	0.44	0.4	0.74	0.49	0.39	0.36	0.31
5	0.88	0.72	0.66	0.64	0.61	0.84	0.71	0.65	0.63	0.61
3	0.95	0.91	0.9	0.89	0.89	0.95	0.91	0.9	0.88	0.88
2	1	1	1	1	1	1	1	1	1	1
1	1.57	1.21	1.19	0.89	0.85	0.91	0.73	0.66	0.61	0.5

Table 2

Coefficients of the ratio of variance in estimates in the absence of *a priori* information

β	$\hat{g}_{\text{PMM3/OLS}}$				$g_{\text{PMM3/OLS}}$	$\hat{g}_{\text{MLE/OLS}}$				$g_{\text{MLE/OLS}}$
	N					N				
	20	50	100	200		20	50	100	200	
10	0.87	0.61	0.49	0.43	0.4	1.06	0.64	0.47	0.39	0.31
5	0.98	0.8	0.69	0.65	0.61	1.15	0.86	0.75	0.7	0.61
3	1.11	0.97	0.94	0.91	0.89	1.26	1.08	0.99	0.93	0.88
2	1.35	1.15	1.07	1.03	1	1.22	1.06	1.04	1.02	1
1	3.2	1.17	0.82	0.82	0.85	0.91	0.72	0.66	0.61	0.5

8. Experiment with real data on regression dependences with EPD errors

To illustrate the feasibility of the proposed approach, we shall use an example involving real data, considered in [27, 64], which also investigates the use of EPD as a model of errors of linear regression. That set, taken from [65], contains data on the box office revenue, in millions of dollars (Gross), and the number of home videos sold in thousands (Video), for a sample of 30 movies. The task is to define the dependence of the number of videos sold on the profit at the box office, which is adequately described by the linear regression (Fig. 3, *a*).

The OLS estimates of the parameters of such a regression model are the following values: $\hat{a}_0^{(\text{OLS})}=76.54$; $\hat{a}_1^{(\text{OLS})}=4.33$. When using, to evaluate the parameters of regression by a maximum likelihood method and visualization of results (Fig. 3, *b*), the functionality of "normalp-package" software [27], we obtain the following estimated values: $\hat{a}_0^{(\text{MLE})}=77.41$; $\hat{a}_1^{(\text{MLE})}=4.37$. By fitting the estimate of the parameter of EPD form of regression OLS residues $\hat{\beta}=2.39$ to (19), we obtain an estimate of the coefficient $\hat{g}_{\text{MLE/OLS}}=0.97$. By estimating the value of cumulant coefficients of the 4th $\hat{\gamma}_4=-0.59$ and 6th $\hat{\gamma}_4=2.14$ orders and using (21), we find the estimated value of the coefficient $\hat{g}_{\text{PMM3/OLS}}=0.88$. At the same time, the refined adaptive PMM3-estimates of the parameters of the regression model are $\hat{a}_0^{(\text{PMM3})}=79.99$; $\hat{a}_1^{(\text{PMM3})}=4.37$.
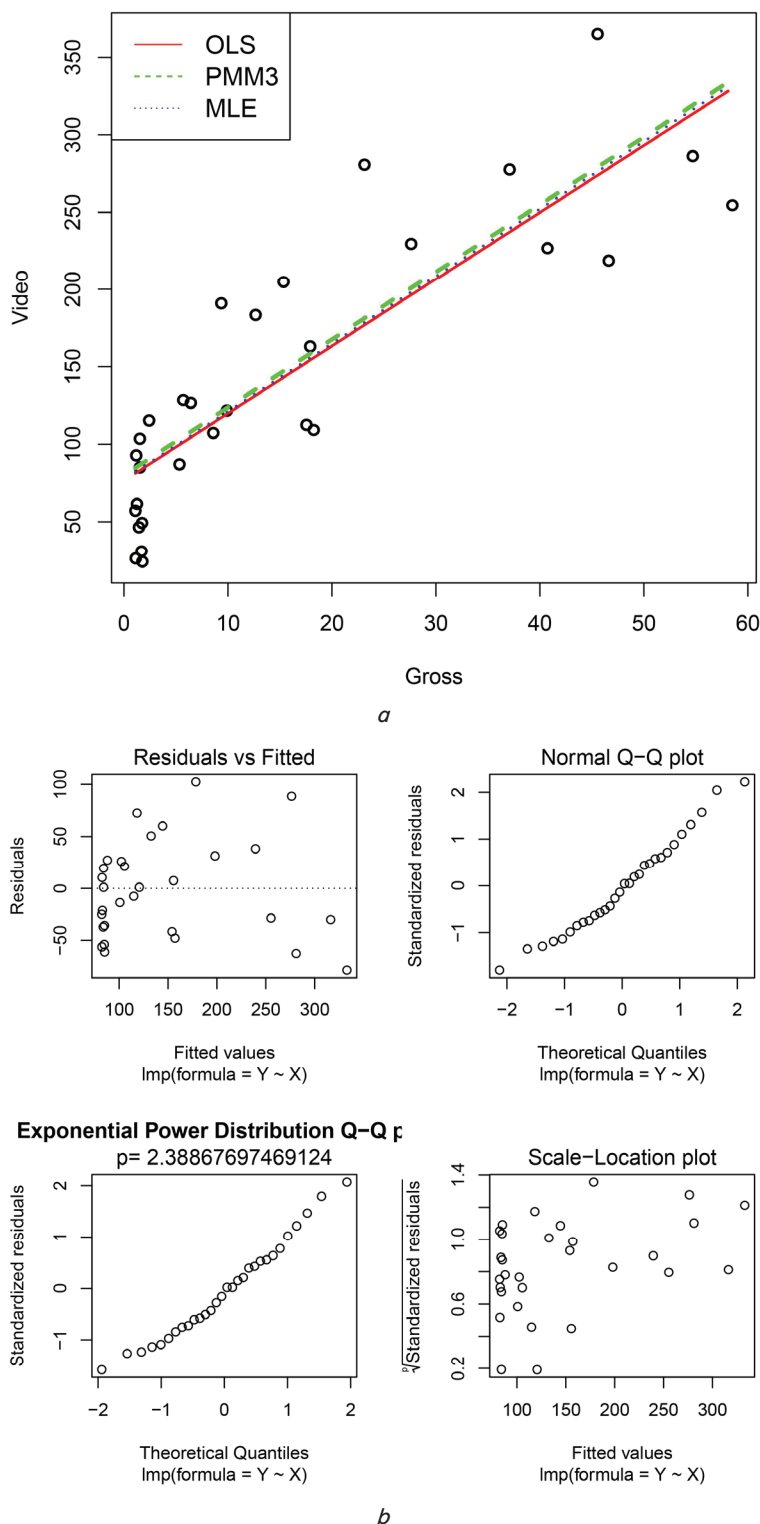
Fig. 3. Linear regression model: *a* — experimental data and dependences built by different methods; *b* — characteristics of OLS residues when using EPD

9. Discussion of results of studying the effectiveness of PMM estimates for regression parameters with EPD errors

The analysis of the theoretical and experimental values of the coefficients of the ratio of the variance in estimates shows that there is a certain correlation of analytical calculations and results obtained through statistical modeling. The difference is due to the fact that the expressions describing the variances in MLE and PMM estimates were obtained for the asymptotic case (at $N{\to}\infty$). Data from Table 1 confirm that as the sample volume increases, the discrepancy between the theoretical and experimental data decreases.

The experimental results also confirm the thesis mentioned above that the accuracy of obtaining the MLE and PMM estimates is significantly influenced by the factor of the presence/absence of *a priori* information on the properties of the error model. For MLE, these are the parameters of EPD; for PMM, these are the paired central moments up to the 6th order. Data from Table 2 reflect an important fact that in the absence of *a priori* information about the properties of EPD for flat-topped distributions (β>2), the efficiency of PMM estimates may be higher than the MLE estimates. This can be explained by the fact that, in these conditions, the impact of uncertainty of the value of the EPD form parameter on MLE is more significant than the impact of uncertainty of central moments on PMM. At the same time, the relative decrease in the variance of PMM estimates (compared to MLE) depends both on the form parameter and on the amount of sample values and is especially significant at small sampling. Such results are especially important from a practical point of view since for most real situations *a priori* information about the true values of the parameters of a random component of the regression model is absent.

Based on the set of our results, the following conclusions can be drawn about the relative effectiveness of PMM (in the absence of *a priori* information about the values of regression error parameters):

1. When the distribution of regression errors is close to Gaussian (β=2) the most effective (both in terms of accuracy and ease of implementation) are estimates of the least-square method.

2. For leptokurtic (sharp-topped) distributions with heavy tails, in general, the maximum likelihood method is more effective. For small sample sizes, the efficiency limit is near β=1.3 and smoothly asymptotically (with growth in N) increases to β=1.7.

3. For plateau-kurtic (flat-topped) distributions for a sufficiently wide range of values of the form parameter β the best efficiency is demonstrated by a polynomial maximization method. For small sample sizes, the lower limit of efficiency (between OLS and PMM) approximately begins at β=3 and parabolically asymptotically (with growth in N) tends to β=2.

4. For large values of the form parameter β>6 (which corresponds to the probabilistic distributions close to a uniform law) at large sample volumes, the maximum likelihood method is more effective, again. The boundary that separates MLE and PMM is even more dependent on the amount of sample values but the difference in their accuracy (in relation to the variance of estimates) is only a few percent.

It should be noted that the current work is an integral part of the study on the feasibility of using a polynomial maximization method to find estimates of the parameters of statistical data with EPD. Its results are correlated to a large degree with the conclusions drawn in [60, 61], which consider a simpler task of assessing the EPD shift parameter when using PMM.

In general, the results reported here provide an opportunity to choose the most effective among the analyzed methods for evaluating the parameters of linear regression depending on the volume of sample data and the properties of regression errors, in case they differ from Gaussian idealization but are symmetrical and adequately described by the EPD model.

The next stage of our research may address a modification of the proposed approach to find estimates of the parameters of regression of the nonlinear type, as well as the evaluation of parameters of autoregressive models.

10. Conclusions

1. The algorithm for finding adaptive estimates of the linear regression parameters, synthesized using a polynomial maximization method is less resource-intensive when compared to the estimation algorithm based on a maximum likelihood method. The simplification is due to the fact that, instead of identifying and evaluating the parameters of the distribution of regression errors, we used the estimates of central moments to the 6th order that are simpler in terms of computation.

2. The theoretical accuracy (of the variance in estimates) of a polynomial maximization method is close enough to the accuracy of a maximum likelihood method and can significantly exceed the accuracy of a least-square method. All three methods are characterized by the same accuracy only if the regression errors abide the Gaussian distribution (as a separate case of EPD).

3. The results of our statistical modeling indicate that for an almost important situation of the absence of *a priori* information about the values of regression error parameters (for plateau-kurtic EPD), the PMM estimates may demonstrate the lowest relative variance. At the same time, the maximum possible increase in accuracy with respect to OLS is up to 60 %, and with respect to MLE – up to 10 %.

Reference

1. Subbotin, M. T. (1923). On the law of frequency of error. Mat. Sb., 31 (2), 296–301.

2. Varanasi, M. K., Aazhang, B. (1989). Parametric generalized Gaussian density estimation. The Journal of the Acoustical Society of America, 86 (4), 1404–1415. doi: https://doi.org/10.1121/1.398700

3. Nadarajah, S. (2005). A generalized normal distribution. Journal of Applied Statistics, 32 (7), 685–694. doi: https://doi.org/10.1080/02664760500079464

4. Giller, G. L. (2005). A Generalized Error Distribution. SSRN Electronic Journal. doi: https://doi.org/10.2139/ssrn.2265027

5. Krasilnikov, A. I. (2019). Family of Subbotin Distributions and its Classification. Electronic modeling, 41 (3), 15–32. doi: https://doi.org/10.15407/emodel.41.03.015

6. Hassan, M. Y., Hijazi, R. H. (2010). A bimodal exponential power distribution. Pakistan Journal of Statistics, 26 (2), 379–396.

7. Komunjer, I. (2007). Asymmetric power distribution: Theory and applications to risk measurement. Journal of Applied Econometrics, 22 (5), 891–921. doi: https://doi.org/10.1002/jae.961

8. Zhu, D., Zinde-Walsh, V. (2009). Properties and estimation of asymmetric exponential power distribution. Journal of Econometrics, 148 (1), 86–99. doi: https://doi.org/10.1016/j.jeconom.2008.09.038

9. Crowder, G. E., Moore, A. H. (1983). Adaptive Robust Estimation Based on a Family of Generalized Exponential Power Distributions. IEEE Transactions on Reliability, R-32 (5), 488–495. doi: https://doi.org/10.1109/tr.1983.5221739

10. Mineo, A. M. (2003). On the estimation of the structure parameter of a normal distribution of order p. Statistica, 63(1), 109–122. doi: https://doi.org/10.6092/issn.1973-2201/342

11. Olosunde, A. A., Soyinka, A. T. (2019). Interval Estimation for Symmetric and Asymmetric Exponential Power Distribution Parameters. Journal of the Iranian Statistical Society, 18 (1), 237–252. doi: https://doi.org/10.29252/jirss.18.1.237

12. Pog ny, T. K., Nadarajah, S. (2010). On the characteristic function of the generalized normal distribution. Comptes Rendus Mathematique, 348 (3-4), 203–206. doi: https://doi.org/10.1016/j.crma.2009.12.010

13. Novitskiy, P. V., Zograf, I. A. (1991). Otsenka pogreshnostey rezul'tatov izmereniy. Leningrad: Izdatel'stvo Energoatomizdat, 304.

14. Lindsey, J. K. (1999). Multivariate Elliptically Contoured Distributions for Repeated Measurements. Biometrics, 55 (4), 1277–1280. doi: https://doi.org/10.1111/j.0006-341x.1999.01277.x

15. Sheluhin, O. I. (1999). Negaussovskie protsessy v radiotehnike. Moscow: Radio i svyaz', 310.

16. Sharifi, K., Leon-Garcia, A. (1995). Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. IEEE Transactions on Circuits and Systems for Video Technology, 5 (1), 52–56. doi: https://doi.org/10.1109/76.350779

17. Dominguez-Molina, J. A., Gonzalez-Farias, G., Rodriguez-Dagnino, R. M. (2001). A practical procedure to estimate the shape parameter in the generalized Gaussian distribution. Available at: http://www.cimat.mx/reportes/enlinea/I-01-18_eng.pdf

18. Saatci, E., Akan, A. (2010). Respiratory parameter estimation in non-invasive ventilation based on generalized Gaussian noise models. Signal Processing, 90 (2), 480–489. doi: https://doi.org/10.1016/j.sigpro.2009.07.015

19. Olosunde, A. A. (2013). On exponential power distribution and poultry feeds data: a case study. Journal of The Iranian Statistical Society, 12 (2), 253–269.

20. Giacalone, M. (2020). A combined method based on kurtosis indexes for estimating p in non-linear Lp-norm regression. Sustainable Futures, 2, 100008. doi: https://doi.org/10.1016/j.sftr.2020.100008

21. Chan, J. S. K., Choy, S. T. B., Walker, S. G. (2021). On The Estimation Of The Shape Parameter Of A Symmetric Distribution. Journal of Data Science, 18 (1), 78–100. doi: https://doi.org/10.6339/jds.202001_18(1).0004

22. Olosunde, A. A., Adegoke, A. M. (2016). Goodness-of-fit-test for Exponential Power Distribution. American Journal of Applied Mathematics and Statistics, 4 (1), 1–8.

23. Giacalone, M., Panarello, D. (2020). Statistical hypothesis testing within the Generalized Error Distribution: Comparing the behavior of some nonparametric techniques. Book of Short Papers SIS 2020, 1338–1343.

24. Johnson, M. E. (1979). Computer Generation of the Exponential Power Distributions. Journal of Statistical Computation and Simulation, 9, 239–240.

25. Nardon, M., Pianca, P. (2009). Simulation techniques for generalized Gaussian densities. Journal of Statistical Computation and Simulation, 79 (11), 1317–1329. doi: https://doi.org/10.1080/00949650802290912

26. Kalke, S., Richter, W.-D. (2013). Simulation of the p-generalized Gaussian distribution. Journal of Statistical Computation and Simulation, 83 (4), 641–667. doi: https://doi.org/10.1080/00949655.2011.631187

27. Mineo, A. M., Ruggieri, M. (2005). A Software Tool for the Exponential Power Distribution: The normalp Package. Journal of Statistical Software, 12 (4). doi: https://doi.org/10.18637/jss.v012.i04

28. Rao, C. R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. Bulletin of Calcutta Mathematical Society, 37, 81–89.

29. Koenker, R., Bassett, G. (1978). Regression Quantiles. Econometrica, 46 (1), 33–50. doi: https://doi.org/10.2307/1913643

30. Barrodale, I., Roberts, F. D. K. (1973). An Improved Algorithm for Discrete l_1 Linear Approximation. SIAM Journal on Numerical Analysis, 10 (5), 839–848. doi: https://doi.org/10.1137/0710069

31. Zeckhauser, R., Thompson, M. (1970). Linear Regression with Non-Normal Error Terms. The Review of Economics and Statistics, 52 (3), 280–286. doi: https://doi.org/10.2307/1926296

32. Mineo, A. (1989). The Norm-P Estimation of Location, Scale and Simple Linear Regression Parameters. Lecture Notes in Statistics, 222–233. doi: https://doi.org/10.1007/978-1-4612-3680-1_26

33. Agrò, G. (1992). Maximum likelihood and Lp-norm estimators. Statistica Applicata, 4 (2), 171–182.

34. Tarassenko, P. F., Tarima, S. S., Zhuravlev, A. V., Singh, S. (2014). On sign-based regression quantiles. Journal of Statistical Computation and Simulation, 85 (7), 1420–1441. doi: https://doi.org/10.1080/00949655.2013.875176

35. Yang, T., Gallagher, C. M., McMahan, C. S. (2018). A robust regression methodology via M-estimation. Communications in Statistics - Theory and Methods, 48 (5), 1092–1107. doi: https://doi.org/10.1080/03610926.2018.1423698

36. Galea, M., Paula, G. A., Bolfarine, H. (1997). Local influence in elliptical linear regression models. Journal of the Royal Statistical Society: Series D (The Statistician), 46 (1), 71–79. doi: https://doi.org/10.1111/1467-9884.00060

37. Liu, S. (2002). Local influence in multivariate elliptical linear regression models. Linear Algebra and Its Applications, 354 (1-3), 159–174. doi: https://doi.org/10.1016/s0024-3795(01)00585-7

38. Ganguly, S. S. (2014). Robust Regression Analysis for Non-Normal Situations under Symmetric Distributions Arising In Medical Research. Journal of Modern Applied Statistical Methods, 13 (1), 446–462. doi: https://doi.org/10.22237/jmasm/1398918480

39. Bartolucci, F., Scaccia, L. (2005). The use of mixtures for dealing with non-normal regression errors. Computational Statistics & Data Analysis, 48 (4), 821–834. doi: https://doi.org/10.1016/j.csda.2004.04.005

40. Seo, B., Noh, J., Lee, T., Yoon, Y. J. (2017). Adaptive robust regression with continuous Gaussian scale mixture errors. Journal of the Korean Statistical Society, 46 (1), 113–125. doi: https://doi.org/10.1016/j.jkss.2016.08.002

41. Tiku, M. L., Islam, M. Q., Selçuk, A. S. (2001). Nonnormal Regression. II. Symmetric Distributions. Communications in Statistics - Theory and Methods, 30 (6), 1021–1045. doi: https://doi.org/10.1081/sta-100104348

42. Andargie, A. A., Rao, K. S. (2013). Estimation of a linear model with two-parameter symmetric platykurtic distributed errors. Journal of Uncertainty Analysis and Applications, 1 (1). doi: https://doi.org/10.1186/2195-5468-1-13

43. Stone, C. J. (1975). Adaptive Maximum Likelihood Estimators of a Location Parameter. The Annals of Statistics, 3 (2), 267–284. doi: ttps://doi.org/10.1214/aos/1176343056

44. Montfort, K., Mooijaart, A., Leeuw, J. (1987). Regression with errors in variables: estimators based on third order moments. Statistica Neerlandica, 41 (4), 223–238. doi: https://doi.org/10.1111/j.1467-9574.1987.tb01215.x

45. Dagenais, M. G., Dagenais, D. L. (1997). Higher moment estimators for linear regression models with errors in the variables. Journal of Econometrics, 76 (1-2), 193–221. doi: https://doi.org/10.1016/0304-4076(95)01789-5

46. Cragg, J. G. (1997). Using Higher Moments to Estimate the Simple Errors-in-Variables Model. The RAND Journal of Economics, 28, S71. doi: https://doi.org/10.2307/3087456

47. Gillard, J. (2014). Method of Moments Estimation in Linear Regression with Errors in both Variables. Communications in Statistics - Theory and Methods, 43 (15), 3208–3222. doi: https://doi.org/10.1080/03610926.2012.698785

48. Wang, L., Leblanc, A. (2007). Second-order nonlinear least squares estimation. Annals of the Institute of Statistical Mathematics, 60 (4), 883–900. doi: https://doi.org/10.1007/s10463-007-0139-z

49. Chen, X., Tsao, M., Zhou, J. (2010). Robust second-order least-squares estimator for regression models. Statistical Papers, 53 (2), 371–386. doi: https://doi.org/10.1007/s00362-010-0343-4

50. Kim, M., Ma, Y. (2011). The efficiency of the second-order nonlinear least squares estimator and its extension. Annals of the Institute of Statistical Mathematics, 64 (4), 751–764. doi: https://doi.org/10.1007/s10463-011-0332-y

51. Huda, S., Mukerjee, R. (2016). Optimal designs with string property under asymmetric errors and SLS estimation. Statistical Papers, 59 (3), 1255–1268. doi: https://doi.org/10.1007/s00362-016-0819-y

52. Kunchenko, Yu. P., Lega, Yu. G. (1991). Otsenka parametrov sluchaynyh velichin metodom maksimizatsii polinoma. Kyiv: Naukova dumka, 180.

53. Zabolotnii, S., Warsza, Z. L., Tkachenko, O. (2018). Polynomial Estimation of Linear Regression Parameters for the Asymmetric PDF of Errors. Advances in Intelligent Systems and Computing, 758–772. doi: https://doi.org/10.1007/978-3-319-77179-3_75

54. Zabolotnii, S. W., Warsza, Z. L., Tkachenko, O. (2019). Estimation of Linear Regression Parameters of Symmetric Non-Gaussian Errors by Polynomial Maximization Method. Advances in Intelligent Systems and Computing, 636–649. doi: https://doi.org/10.1007/978-3-030-13273-6_59

55. Warsza, Z. (2017). Ocena niepewności pomiarów o rozkładzie trapezowym metodą maksymalizacji wielomianu. Przemysł Chemiczny, 1 (12), 68–71. doi: https://doi.org/10.15199/62.2017.12.6

56. Warsza, Z. L., Zabolotnii, S. W. (2017). A Polynomial Estimation of Measurand Parameters for Samples of Non-Gaussian Symmetrically Distributed Data. Advances in Intelligent Systems and Computing, 468–480. doi: https://doi.org/10.1007/978-3-319-54042-9_45

57. Zabolotnii, S. V., Kucheruk, V. Yu., Warsza, Z. L., Khassenov, A. K. (2018). Polynomial Estimates of Measurand Parameters for Data from Bimodal Mixtures of Exponential Distributions. Vestnik Karagandinskogo universiteta, 2 (90), 71–80.

58. Warsza, Z. L., Zabolotnii, S. (2018). Estimation of Measurand Parameters for Data from Asymmetric Distributions by Polynomial Maximization Method. Advances in Intelligent Systems and Computing, 746–757. doi: https://doi.org/10.1007/978-3-319-77179-3_74

59. Zabolotnii, S. W., Warsza, Z. L. (2016). Semi-parametric Estimation of the Change-Point of Parameters of Non-gaussian Sequences by Polynomial Maximization Method. Advances in Intelligent Systems and Computing, 903–919. doi: https://doi.org/10.1007/978-3-319-29357-8_80

60. Zabolotnii, S. V., Chepynoha, A. V., Bondarenko, Y. Y., Rud, M. P. (2018). Polynomial parameter estimation of exponential power distribution data. Visnyk NTUU KPI Seriia - Radiotekhnika Radioaparatobuduvannia, 75, 40–47, 75, 40–47. doi: https://doi.org/10.20535/radap.2018.75.40-47

61. Zabolotnii, S. V., Chepynoha, A. V., Chorniy, A. M., Honcharov, A. V. (2020). Comparative Analysis of Polynomial Maximization and Maximum Likelihood Estimates for Data with Exponential Power Distribution. Visnyk NTUU KPI Seriia - Radiotekhnika Radioaparatobuduvannia, 82, 44–51. doi: https://doi.org/10.20535/radap.2020.82.44-51

62. Liu, M., Bozdogan, H. (2008). Multivariate regression models with power exponential random errors and subset selection using genetic algorithms with information complexity. European Journal of Pure and Applied Mathematics, 1 (1), 4–37. Available at: https://www.researchgate.net/publication/264947828_Multivariate_regression_models_with_power_exponential_random_errors_and_subset_selection_using_genetic_algorithms_with_information_complexity

63. Atsedeweyn, A. A., Srinivasa Rao, K. (2013). Linear regression model with new symmetric distributed errors. Journal of Applied Statistics, 41 (2), 364–381. doi: https://doi.org/10.1080/02664763.2013.839638

64. Ferreira, M. A., Salazar, E. (2014). Bayesian reference analysis for exponential power regression models. Journal of Statistical Distributions and Applications, 1 (1), 12. doi: https://doi.org/10.1186/2195-5832-1-12

65. Levine, D. M., Krehbiel, T. C., Berenson, M. L. (2000). Business Statistics: A First Course. Prentice-Hall.