Pliska Stud. Math. Bulgar. **14** (2003), 71–80

PLISKA
STUDIA MATHEMATICA
BULGARICA

LOCAL-MEDIAN METHOD OF FORECASTING FOR REGRESSION TIME SERIES UNDER OUTLIERS

Yu. Kharin, V. Maevskiy

The local-median method of forecasting under the regression model with outliers is analyzed in this paper. The breakdown point is evaluated, the distribution function of the local-median forecast is given.

1. Introduction

Forecasting under regression model of data is an actual problem in economics, engineering, biology and other areas. If the regression model is linear w.r.t. regression parameters, then the standard method of least squares (LS) is used for forecasting [1]. The LS-method works well at the validity of all hypothetical model assumptions [2], [7], however in real situations there are outliers in the raw data [6], and the performance of the LS-forecast decreases significantly [2], [5]. Therefore it is important to develop robust algorithms not so sensitive to distortions [1], [2], [8] and to analyze their robustness. The local-median method proposed in [9] and investigated in this paper has some attractive properties among numerous robust methods: at first, it is not necessary to know the distortion level in the observed data; at second, it is not necessary to construct the robust estimator for the vector of regression parameters, only the robust forecast is constructed.

2002 *Mathematics Subject Classification:* 62M20, 62-07, 62J05, 62P20
Key words: Robustness, Forecasting, Regression, Outliers

2. Local-median method of forecasting under distorted regression model

Let the observations $\{x_t\}$ on the investigated stochastic dynamic system be described by the regression equation:

$$x_t = \theta^{0\prime}\psi(z_t) + u_t + \xi_t \nu_t, \tag{1}$$

where $t \in Z$ be a discrete time moment, $z_t \in R^M$ be a nonrandom observable input vector of factors (predictors) at the moment t, $\psi(z) = (\psi_i(z)) : R^M \to R^m$ be the vector of m linearly independent functions, $\theta^0 = (\theta_i^0) \in R^m$ be the vector of m unknown true values of parameters of the model (coefficients of the regression), $u_t \in R$ be a random error at the moment t, $\nu_t \in R$ be the outlier at the moment t, $\{\xi_t\}$ be i.i.d. Bernoulli random variables, that describe presence of the outliers, $P\{\xi_t = 1\} = \varepsilon$, $P\{\xi_t = 0\} = 1 - \varepsilon$, $\varepsilon \in [0; 0.5)$ be the probability of an outlier appearance (proportion of outliers in the observed sample). Random errors $\{u_t\}$ are i.i.d. random variables, $E\{u_t\} = 0$, $D\{u_t\} = \sigma^2 < +\infty$; $\{\nu_t\}$ are i.i.d. random variables, $E\{\nu_t\} = a_t$, $D\{\nu_t\} = K\sigma^2 < +\infty$, $K \geq 0$. The random variables $\{u_t\}$, $\{v_t\}$, $\{\xi_t\}$ are independent in total.

Define: T is the observation time; $\{t_1^{(l)}, ..., t_n^{(l)}\} \subset \{1, 2, ..., T\}$ is a subset of n $(m \leq n \leq T)$ observed time moments $(l = 1, ..., L)$, where L is the number of different subsets of time moments $(m \leq L \leq L_+ = \binom{T}{n})$; $\Psi = (\psi_j(z_t))$, $j = 1, ..., m$, $t = 1, ..., T$; $\Psi_n^{(l)} = (\psi_j(z_{t_i^{(l)}}))$, $i = 1, ..., n$, $j = 1, ..., m$ is the $(n \times m)$ - submatrix of the $(T \times m)$ - matrix Ψ, $\left|\Psi_n^{(l)\prime}\Psi_n^{(l)}\right| \neq 0$; $X = (x_1, x_2, ..., x_T)' \in R^T$ is the observed sample; $X_n^{(l)} = (x_{t_1^{(l)}}, x_{t_2^{(l)}}, \ldots, x_{t_n^{(l)}})' \in R^n$ is the subsample of the sample X; $a^{(l)} = (a_{t_1^{(l)}}, a_{t_2^{(l)}}, \ldots, a_{t_n^{(l)}})' \in R^n$.

Define now the l-th local LS-estimator for θ^0, based on the l-th subsample $X_n^{(l)}$:

$$\hat{\theta}^{(l)} = (\Psi_n^{(l)\prime}\Psi_n^{(l)})^{-1}\Psi_n^{(l)\prime}X_n^{(l)}, \; l = 1, ..., L, \tag{2}$$

and the family of L local forecasts of the future state $x_{T+\tau}$ for $\tau \geq 1$, based on local LS-estimators (2):

$$\hat{x}_{T+\tau}^{(l)} = \hat{\theta}^{(l)\prime}\psi(z_{T+\tau}), \; l = 1, ..., L. \tag{3}$$

The LM-forecast introduced in [9] is the sample median of L local forecasts (3):

$$\hat{x}_{T+\tau} = S(X) = \text{med}\{\hat{x}_{T+\tau}^{(1)}, \ldots, \hat{x}_{T+\tau}^{(L)}\}. \tag{4}$$

Note, that the subsample size n and the number L of subsamples are parameters of the LM-method. If $n = T$, $L = 1$, then the LM-forecast is equivalent to the traditional LS-forecast. If $L = \binom{T}{n}$, then all subsamples of size n from the initial sample of size T are used in (2) - (4) to construct the LM-forecast.

3. Breakdown point

Let us evaluate the breakdown point for the LM-forecast (4) in the Hampel sense [2]. Define the breakdown point as the maximal portion ε^* of "arbitrary large" outliers in the sample X, when the forecast statistic $S(X)$ can not be made "arbitrary large" by varying of outliers values:

(5) $$\varepsilon^* = \max\{\varepsilon \in [0,1] \,|\, \forall X_{(\varepsilon)}\ |S(X_{(\varepsilon)})| \le C < +\infty\},$$

where $X_{(\varepsilon)} = \{x_t : 1 \le t \le T, \xi_t = 1\}$ is a subsample of observations (1) distorted by outliers.

Theorem 1. If $L = L_+ = \binom{T}{n}$, then the breakdown point (5) of the LM-forecast (4) under the distorted model (1) is the unique root of the n-th order algebraic equation w.r.t. $\varepsilon \in [0, 1 - nT^{-1}]$

(6) $$\prod_{t=0}^{n-1}(1 - \varepsilon - \frac{t}{T}) = (1-\alpha)\prod_{t=0}^{n-1}(1 - \frac{t}{T}),$$

where

(7) $$\alpha = \lfloor (L-1)/2 \rfloor / L = 1/2 + O(1/\binom{T}{n}).$$

Proof. At first, let us show, that the unique root ε_r of the equation (6) in the segment $[0, 1 - nT^{-1}]$ exists, so it is sufficient to find the breakdown point in the segment $[0, 1 - nT^{-1}]$. Let us consider the function $f(\varepsilon) = \prod_{t=0}^{n-1}(1 - \varepsilon - t/T) - (1-\alpha)\prod_{t=0}^{n-1}(1 - t/T)$ defining the equation (6) in the equivalent form $f(\varepsilon) = 0$. As its derivative $f'(\varepsilon) = -\sum_{p=0}^{n-1}\prod_{t=0, t\neq p}^{n-1}(1 - \varepsilon - t/T) < 0$ is negative for $\varepsilon \le 1 - n/T$, then the function $f(\varepsilon)$ decreases strictly monotonically w.r.t. ε in the segment $[0, 1 - nT^{-1}]$.

Let us consider the following cases: a) $n = T$. In this case, $\varepsilon_r = 0$ and it belongs to the segment $[0, 1 - nT^{-1}]$, which is the singleton; b) $n = 1$, $T > 1$. In this case, $\varepsilon_r = \alpha \in [0, 1 - T^{-1}]$ as $\alpha \le 1/2$; c) $2 \le n \le T$-1. In this case, $f(0) = \alpha\prod_{t=0}^{n-1}(1 - t/T) > 0$, $f(1 - nT^{-1}) = n!T^{-n}(1 - (1\text{-}\alpha)\binom{T}{n}) < 0$, as $T > 2$, $n \le T - 1$, $\alpha \le 1/2$.

From these facts and the strict monotonicity of the function $f(\varepsilon)$ we get the existence and uniqueness of the root ε_r.

At second, let us construct the equation (6). It is known [2], that the breakdown point α of the sample median of the local forecasts (3) is

$$\alpha = \lfloor (L-1)/2 \rfloor / L = \begin{cases} 1/2 - L^{-1}, \text{ if } L \text{ is even} \\ 1/2 - (2L)^{-1}, \text{ if } L \text{ is odd} \end{cases}.$$

Denote $\beta = \varepsilon T \in N$ the number of distorted observations. If $\varepsilon > 1 - nT^{-1}$, that is if $\beta > T - n$, then all local forecasts based on subsamples of size n will be distorted and therefore the local-median forecast (4) will be distorted, so in (5) we can consider $\varepsilon \in [0, 1 - nT^{-1}]$ only.

The total number of subsamples that do not contain distorted observations equals to $\binom{T-\beta}{n}$. The value α of the breakdown point of the sample median means that the number of local forecasts based on the distorted subsamples must not be greater than αL. So we get the condition on the maximal number of nondistorted observations in the sample X: $\binom{T-\beta}{n} \geq (1-\alpha)\binom{T}{n}$. Using (5) and equivalent transformations we come to the equation (6).

Let us evaluate the breakdown point ε^* in some special cases.

Corollary 1. If $n \leq \alpha T$, then $\varepsilon^* \geq T^{-1} > 0$.

Proof. The function $f(\varepsilon)$ is nonnegative at the point $\varepsilon = nT^{-1}$ for $n \leq \alpha T$. Using the inequality $\alpha \leq 1/2$ we get $\varepsilon^* \in [T^{-1}, 1 - nT^{-1}]$ in the same way as in the proof of Theorem 1.

Corollary 2. If $n = m = 1$, then $\varepsilon^* = \alpha$; if $n = 2$, then

$$\varepsilon^* = 1 - (2T)^{-1} - ((1 - (2T)^{-1})^2 - \alpha(1 - T^{-1}))^{1/2}.$$

Corollary 3. If n is fixed and $T \to \infty$, then

$$\varepsilon^* \to 1 - 2^{-1/n}, \tag{8}$$

and the optimal subsample size maximizing the breakdown point (5) is $n^* = m$.

Proof. Putting $T \to \infty$ in (6) we get (8). The limit value of the breakdown point in (8) is the monotonically decreasing function of n, so the optimal subsample size n^* is the minimal admissible value, i.e. m.

4. Probability distribution of the LM-forecast

Introduce the notation:

$$g^{(l)} = (g_i^{(l)}) = \Psi_n^{(l)} (\Psi_n^{(l)\prime} \Psi_n^{(l)})^{-1} \psi(z_{T+\tau}) \in R^n, \ l = 1, ..., L; \ x_{T+\tau}^0 = \theta^{0\prime} \psi(z_{T+\tau});$$

$$\phi(z|\mu; \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp(-(z-\mu)^2/(2\sigma^2)), \ z \in R, \tag{9}$$

is the Gaussian probability density function (p.d.f.); $\Phi(z|\mu;\sigma^2)$ is the corresponding Gaussian distribution function; Π_r is the set of $\binom{n}{r}$ combinations of r elements from n elements $A_n = \{1, ..., n\}$.

Theorem 2. If the distorted regression model (1) takes place, and $\{u_t\}$, $\{\nu_t\}$ have Gaussian distributions, then the p.d.f. of the l-th local forecast (3) is the mixture of 2^n Gaussian distributions:

$$(10) \qquad p_{\hat{x}^{(l)}_{T+\tau}}(z) = \sum_{r=0}^{n} (1-\varepsilon)^r \varepsilon^{n-r} \sum_{(k_1,\dots,k_r)\in\Pi_r} p_{k_1,\dots,k_r}(z),\ z \in R,\ l = \overline{1, L},$$

$$p_{k_1,\dots,k_r}(z) = \phi(z|x^0_{T+\tau} + \sum_{i=r+1}^{n} g^{(l)}_{k_i} a^{(l)}_{k_i}; \sigma^2(\sum_{i=1}^{r}(g^{(l)}_{k_i})^2 + (K+1)\sum_{i=r+1}^{n}(g^{(l)}_{k_i})^2)),$$

where $\{k_{r+1}, ..., k_n\} = A_n \backslash \{k_1, ..., k_r\}$ is the subset of $n-r$ elements from A_n not included into the combination $(k_1, ..., k_r)$.

Proof. Let us use the method of characteristic functions. According to the model assumptions, characteristic functions of u_t, ν_t, $\xi_t\nu_t$ and $\eta_t = u_t + \xi_t\nu_t$ are

$$(11) \qquad \begin{aligned} &f_{u_t}(\lambda) = \exp(-\tfrac{1}{2}\sigma^2\lambda^2),\ f_{\nu_t}(\lambda) = \exp(ia_t\lambda - \tfrac{K}{2}\sigma^2\lambda^2),\ \lambda \in R, \\ &f_{\xi_t\nu_t}(\lambda) = (1-\varepsilon) + \varepsilon\exp(ia_t\lambda - \tfrac{K}{2}\sigma^2\lambda^2),\ f_{\eta_t}(\lambda) = f_{u_t}(\lambda) f_{\xi_t\nu_t}(\lambda) = \\ &= (1-\varepsilon)\exp(-\tfrac{1}{2}\sigma^2\lambda^2) + \varepsilon\exp(ia_t\lambda - \tfrac{K+1}{2}\sigma^2\lambda^2),\ t = 1, ..., T. \end{aligned}$$

From (1), (3), (9) we have

$$(12) \qquad \hat{x}^{(l)}_{T+\tau} = x^0_{T+\tau} + g^{(l)\prime}\eta^{(l)},$$

where $\eta^{(l)} = (\eta_{t^{(l)}_1}, ..., \eta_{t^{(l)}_n})'$.

Using (11), (12), independence of $\{\eta_t\}$ and properties of characteristic functions we get:

$$(13) \quad f_{\hat{x}^{(l)}_{T+\tau}}(\lambda) = e^{ix^0_{T+\tau}\lambda} \prod_{k=1}^{n}((1-\varepsilon)e^{-\frac{1}{2}(g^{(l)}_k)^2\sigma^2\lambda^2} + \varepsilon e^{ia^{(l)}_k g^{(l)}_k \lambda - \frac{K+1}{2}\sigma^2(g^{(l)}_k)^2\lambda^2}).$$

By the inversion formula we get the p.d.f. of $\hat{x}^{(l)}_{T+\tau}$:

$$p_{\hat{x}^{(l)}_{T+\tau}}(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iz\lambda} f_{\hat{x}^{(l)}_{T+\tau}}(\lambda)d\lambda,\ z \in R.$$

Putting (13) here and making identity transformations we come to (10).

Corollary. The distribution function of the l-th LM-forecast ($l \in \{1, ..., L\}$) satisfies the asymptotic expansion at $\varepsilon \to 0$:

$$F_{\hat{x}^{(l)}_{T+\tau}}(z) = \Phi(z|x^0_{T+\tau}; \sigma^2 \textstyle\sum_{k=1}^n (g_k^{(l)})^2) + \varepsilon(\sum_{i=1}^n \Phi(z|x^0_{T+\tau} + g_i^{(l)} a_i^{(l)};$$
$$\sigma^2 \textstyle\sum_{k=1}^n (g_k^{(l)})^2 + \sigma^2 K (g_i^{(l)})^2) - n\Phi(z|x^0_{T+\tau}; \sigma^2 \sum_{k=1}^n (g_k^{(l)})^2)) +$$
$$(14) \qquad +\varepsilon^2(\textstyle\sum_{i,j=1, i<j}^n \Phi(z|x^0_{T+\tau} + g_i^{(l)} a_i^{(l)} + g_j^{(l)} a_j^{(l)}; \sigma^2 \sum_{k=1}^n (g_k^{(l)})^2 +$$
$$+\sigma^2 K((g_i^{(l)})^2 + (g_j^{(l)})^2)) - (n-1) \textstyle\sum_{i=1}^n \Phi(z|x^0_{T+\tau} + g_i^{(l)} a_i^{(l)}; \sigma^2 \sum_{k=1}^n (g_k^{(l)})^2 +$$
$$+\sigma^2 K (g_i^{(l)})^2) + \tfrac{1}{2} n(n-1)\Phi(z|x^0_{T+\tau}; \sigma^2 \textstyle\sum_{k=1}^n (g_k^{(l)})^2)) + o(\varepsilon^2), \ l = 1, ..., L.$$

Proof. Selecting the main terms $O(1)$, $O(\varepsilon)$ and $O(\varepsilon^2)$ in (10) we get the asymptotic expansion for the p.d.f. This expansion implies the expansion (14) for the distribution function.

According to (4) the LM-forecast is an order statistic [4]. To find its probability distribution, let us assume that the local forecasts (3) are constructed in the way that they are independent in total.

Theorem 3. If L local forecasts (3) are independent in total, then the distribution function of the j-th order statistic ($j \in \{1, ..., L\}$) is

$$(15) \quad F_{x_{(j)}}(z) = \prod_{i=1}^{L} (1 - F_{\hat{x}^{(i)}_{T+\tau}}(z)) \sum_{k=j}^{L} (\sum_{\substack{i_1, ..., i_k = 1 \\ i_1 > i_2 > ... > i_k}}^{L} \frac{F_{\hat{x}^{(i_1)}_{T+\tau}}(z) \times ... \times F_{\hat{x}^{(i_k)}_{T+\tau}}(z)}{(1 - F_{\hat{x}^{(i_1)}_{T+\tau}}(z)) \times ... \times (1 - F_{\hat{x}^{(i_k)}_{T+\tau}}(z))}),$$

where $F_{\hat{x}^{(i)}_{T+\tau}}(z)$ is the distribution function of the i-th local forecast (3).

Proof. Denote $\mathbb{I}(z) = \{1, z \geq 0; 0, z < 0\}$ the Heaviside function. Consider the j-th order statistic $x_{(j)}$ based on L local forecasts (3). As the events $\{z > \hat{x}^{(j)}_{T+\tau}\}$ and $\{M(z) \geq j\}$ (where $M(z) = \sum_{i=1}^{L} \mathbb{I}(z - \hat{x}^{(i)}_{T+\tau})$ is the number of the local forecasts that are smaller than z) are equivalent, then the distribution function of the j-th order statistic is $F_{x_{(j)}}(z) = P\{z > x_{(j)}\} = \sum_{k=j}^{L} P\{M(z) = k\}$. Taking into consideration (14), $\chi_i = \mathbb{I}(z - \hat{x}^{(i)}_{T+\tau})$ is the Bernoulli random variable with the parameter $p_i = P\{\chi_i = 1\} = F_{\hat{x}^{(i)}_{T+\tau}}(z)$. The value $M(z)$ is the sum of nonhomogeneous Bernoulli random variables. As the local forecasts (3) are independent in total, we get the characteristic function of $M(z)$:

$$f_M(\lambda) = \textstyle\prod_{i=1}^{L} f_{\chi_i}(\lambda) = \prod_{i=1}^{L} (1 + p_i(e^{i\lambda} - 1)) = (1-p_1)(1-p_2)\dots(1-p_L) +$$
$$+ \textstyle\sum_{k=1}^{L} (1-p_1)(1-p_2)\dots(1-p_L) e^{i\lambda} p_k/(1-p_k) +$$
$$+ \sum_{\substack{k,l=1 \\ k<l}}^{L} (1-p_1)(1-p_2)\dots(1-p_L) e^{2i\lambda} p_k p_l/((1-p_k)(1-p_l)) + \dots + p_1 p_2 \dots p_L e^{Li\lambda}.$$

We have $P\{M(z)=k\} = (\sum_{\substack{i_1,\ldots,i_k=1\\ i_1>i_2>\ldots>i_k}}^{L} \frac{p_{i_1}\ldots p_{i_k}}{(1-p_{i_1})\ldots(1-p_{i_k})}) \prod_{i=1}^{L}(1-p_i)$, and the distribution function is $F_{x_{(j)}}(z) = \prod_{i=1}^{L}(1-p_i)\sum_{k=j}^{L}(\sum_{\substack{i_1,\ldots,i_k=1\\ i_1>i_2>\ldots>i_k}}^{L} \frac{p_{i_1}\ldots p_{i_k}}{(1-p_{i_1})\ldots(1-p_{i_k})})$, so we get (15).

Corollary. If L is an odd number, then the distribution function of the LM-forecast $\hat{x}_{T+\tau}$ is:

$$F_{\hat{x}_{T+\tau}}(z) = \prod_{i=1}^{L}(1-F_{\hat{x}^{(i)}_{T+\tau}}(z))\times \times\sum_{k=(L+1)/2}^{L}(\sum_{\substack{i_1,\ldots,i_k=1\\ i_1>i_2>\ldots>i_k}}^{L} \frac{F_{\hat{x}^{(i_1)}_{T+\tau}}(z)\ldots F_{\hat{x}^{(i_k)}_{T+\tau}}(z)}{(1-F_{\hat{x}^{(i_1)}_{T+\tau}}(z))\ldots(1-F_{\hat{x}^{(i_k)}_{T+\tau}}(z))}). \tag{16}$$

As it is seen from (14), (16), the expression of the distribution function for the LM-forecast is too complicated to analyze it. To simplify it and to analyze the distribution function, let us consider the special case where the local forecasts are identically distributed.

Denote: $\varphi_0(z) = \varphi(z|x^0_{T+\tau};K_0\sigma^2)$, $\Phi_0(z) = \Phi(z|x^0_{T+\tau};K_0\sigma^2)$, $K_0>0$, $\varphi_i(z) = \varphi(z|x^0_{T+\tau}+m_i;(K_i+K_0)\sigma^2)$, $\Phi_i(z) = \Phi(z|x^0_{T+\tau}+m_i;(K_i+K_0)\sigma^2)$, $K_i\geq 0$, $m_i\in R$, $i=1,\ldots,n$.

Theorem 4. If $\zeta_1,\ldots,\zeta_L$ are $L=2l+1$ i.i.d. random variables (local forecasts) with the p.d.f. $p_\zeta(z) = (1-n\varepsilon)\varphi_0(z)+\varepsilon\sum_{i=1}^{n}\varphi_i(z)$ (and the distribution function $F_\zeta(z) = (1-n\varepsilon)\Phi_0(z)+\varepsilon\sum_{i=1}^{n}\Phi_i(z)$), which is the mixture of $n+1$ Gaussian distributions, then the following asymptotic expansion for the p.d.f. of the sample median $\zeta_{\text{med}} = \text{med}(\zeta_1,\ldots,\zeta_L)$ holds at $\varepsilon\to 0$:

$$\begin{aligned} p_{\text{med}}(z) &= \tfrac{L!}{(l!)^2}\Phi_0(z)^l(1-\Phi_0(z))^l\varphi_0(z)+\varepsilon\tfrac{L!}{(l!)^2}\Phi_0(z)^l(1-\\ &-\Phi_0(z))^l\varphi_0(z)(l(\Phi_0(z)^{-1}-(1-\Phi_0(z))^{-1})(\textstyle\sum_{i=1}^{n}\Phi_i(z))+\\ &+\varphi_0(z)^{-1}(\textstyle\sum_{i=1}^{n}\varphi_i(z))+nl\Phi_0(z)(1-\Phi_0(z))^{-1}-n(l+1))+o(\varepsilon). \end{aligned}\tag{17}$$

Proof. It is known [3] that the p.d.f. of the sample median of L i.i.d. random variables with the p.d.f. $f(x)$ and the distribution function $F(x)$ is

$$p(x) = \tfrac{L!}{l!(L-l-1)!}F(x)^l(1-F(x))^{L-l-1}f(x).$$

Then

$$\begin{aligned} p_{\text{med}}(z) &= \tfrac{L!}{l!(L-l-1)!}((1-n\varepsilon)\Phi_0(z)+\varepsilon\textstyle\sum_{i=1}^{n}\Phi_i(z))^l(1-(1-n\varepsilon)\Phi_0(z)-\\ &-\varepsilon\textstyle\sum_{i=1}^{n}\Phi_i(z))^{L-l-1}((1-n\varepsilon)\varphi_0(z)+\varepsilon\sum_{i=1}^{n}\varphi_i(z)). \end{aligned}$$

By equivalent transformations we come to (17).

Note, that the p.d.f. of the sample mean of the random variables $\zeta_1, \ldots, \zeta_L$ under Theorem 4 conditions is: $p_{\text{mean}}(z) = \varphi(z|x^0_{T+\tau}; \sigma^2(K_0+\varepsilon\sum_{i=1}^n K_i)/L)$. Let us define the ratio of the p.d.f. for the sample median with respect to the p.d.f. for the sample mean: $\varkappa(z) = p_{\text{med}}(z)/p_{\text{mean}}(z)$. This function shows relative concentration of these two probability densities.

To evaluate the function $\varkappa(z)$ two series of numerical experiments were developed.

Example 1. The dependence of $\varkappa(z)$ on the distortion level ε and z is evaluated for $x^0_{T+\tau} = 0$, $l = 10$, $L = 21$, $\sigma = 1$, $K_0 = 1$, $n = 5$, $K_i = 10$, $m_i = 0$, $i = 1, ..., n$, $z \in [-1, 1]$, $\varepsilon \in [0, 0.1]$. The dependence of $\varkappa(z)$ on ε and z is plotted on Figure 1.

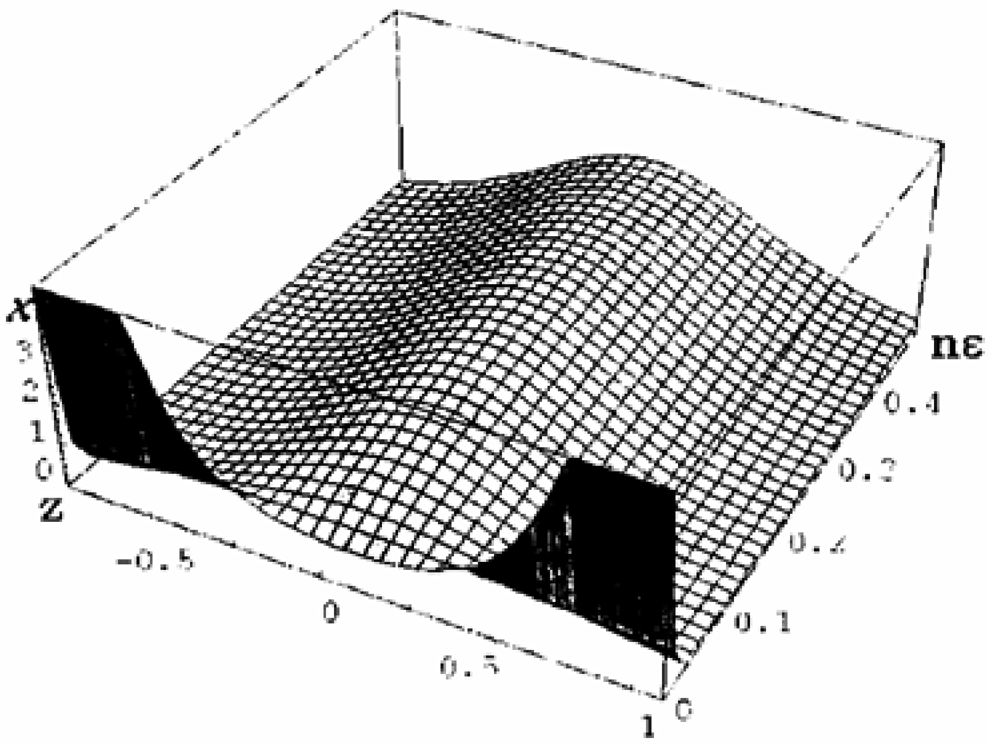

Fig. 1. The dependence of $\varkappa$ on z and ε

We can see, that $p_{\text{med}}(z)$ has more "heavy tails" with respect to $p_{\text{mean}}(z)$ at small values of the distortion level ε and it has more "light tails" when the level ε of the distortions is increasing.

Example 2. The dependence $\varkappa(z)$ on the absolute values of distortions is evaluated for $x^0_{T+\tau} = 0$, $l = 10$, $L = 21$, $\sigma = 1$, $K_0 = 1$, $n = 5$, $\varepsilon = 0.05$, $z \in [-1, 1]$, $K_i = K^*$, $m_i = 0$, $i = 1, ..., n$, $K^* \in [1, 100]$. The dependence of $\varkappa(z)$ on K^* and z is plotted on Figure 2.

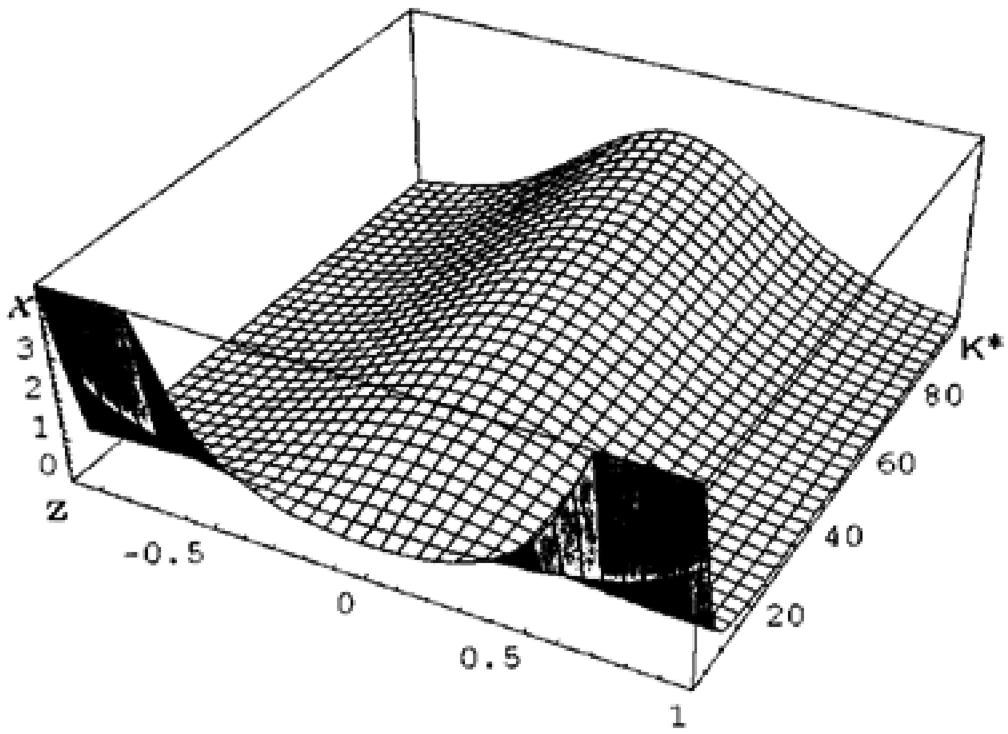

Fig. 2. The dependence of $\varkappa(z)$ on z and K^*

It is seen from Figure 2, that $p_{\text{med}}(z)$ has more "heavy tails" with respect to $p_{\text{mean}}(z)$ at small variance of the outliers $(K^* + K_0)\sigma^2$ and it has more "light tails" when the variance of the outliers $(K^* + K_0)\sigma^2$ is increasing.

This research is supported by the ISTC project B-705.

REFERENCES

[1] E. Chetyrkin. Statistical forecasting. Science, Moscow, 1977 (in Russian).

[2] F. Hampel, E. Ronchetti, P. Rousseeuw, W. Stahel. Robust statistics. The approach based on influence functions. John Wiley and Sons, New York, 1986.

[3] G. I. Ivchenko, Yu. Ya. Medvedev. Mathematical statistics. Higher School, Moscow, 1994 (in Russian).

[4] H. A. David. Order statistics. John Wiley and Sons, New York, 1979, 334p.

[5] J. W. Tukey. A survey of sampling from contaminated distributions. *Contributions to Probability and Statistics, Stanford Univ. Press*, 1960.

[6] M. Romanowski, E. Green. Practical applications of the modified normal distribution. *Bull. Geodesique,* vol. 76, 1965.

[7] P. Huber. Robust statistics. John Wiley and Sons, New York, 1981.

[8] Yu. S. Kharin. Robustness in statistical pattern recognition. Kluwer Academic Publishers, Dordrecht, 1996.

[9] Yu. S. Kharin, R. Fursa. Robustness of time series forecasting for distorted trend model. Computer Data Analysis and Modeling. BSU, Minsk, 1995, pp. 47-58.

Yu. Kharin, V. Maevskiy
Belarussian State University
Minsk, Belarus
e-mail: kharin@bsu.by