

PAPER • OPEN ACCESS

Mathematical modelling in cyber-physical construction systems

To cite this article: Boris Titarenko and Roman Titarenko 2019 *J. Phys.: Conf. Ser.* **1425** 012158

View the [article online](#) for updates and enhancements.

You may also like

- [Eclectic approach as idea of e-educounseling preliminary system model](#)
Jumail, M F Noordin, N M Ibrahim et al.
- [Special issue on applied neurodynamics: from neural dynamics to neural engineering](#)
Hillel J Chiel and Peter J Thomas
- [Information and Statistical Structure in Spike Trains](#)
Jonathan D Victor and Emery N Brown



ECS
The Electrochemical Society
Advancing solid state & electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Mathematical modelling in cyber-physical construction systems

Boris Titarenko¹ and Roman Titarenko²

¹Moscow State University of Civil Engineering, Yaroslavskoe shosse, 26, Moscow, 129337, Russia

²Synergy University, Russia

Abstract. This paper deals with the problem of mathematical foundation for information block of cyber-physical system (CPS). CPS requires the involvement of a variety of information technologies. Their effectiveness is related to methods of data processing of contaminated data sets. The robust estimation means minimizing asymptotic bias of the estimates in the presence of contaminating observations. In the contrast to the approach of P. Huber and J. Tukey, where robustness is minimizing of the asymptotic variance of estimates, minimax asymptotic bias method is based on the maximum likelihood approach assuming an arbitrary contaminating distribution. Examples of this approach for the most common probability distributions and the ways of applying this approach are also given.

1. Introduction

The construction of a cyber-physical system requires the involvement of a variety of information technologies. Their effectiveness is directly related to the quality and methods of processing used information about the state of the system and methods for predicting its behavior. The mathematical apparatus here is the methods of mathematical statistics. The specificity of the building system is that it functions in a situation of great uncertainty of its internal and external environment. Classical information processing methods oriented to its "Gaussian" nature may fail, leading to a distortion of estimates of system parameters and a decrease in their efficiency and a decrease in the accuracy of forecasts. This defect allows eliminating robust methods of statistical information processing, which are widely used at present.

The work considers a robust approach that minimizes the bias of estimates in a situation of arbitrary "contamination", conducts a theoretical study of the properties of the proposed modification of the classical maximum likelihood method. The development of a computational algorithm based on it and its application for processing information in the information block of a cyber-physical building system will improve the reliability of the estimates obtained.

The direction developing the robust procedures of data analysis that has important applied significance is one of the actively developing areas in modern mathematical statistics.

The emergence of a robust approach to estimate parameters of probability distributions is associated with the names of the well-known mathematical statisticians J. Tukey and P. Huber. The problem is to estimate parameters of probability distribution in the presence of foreign (contaminating) observations with different unknown probability distribution in the sample.

Within the framework of the model for the distribution function of the contaminated aggregate there was suggested a model, where random variable has a distribution $P_\varepsilon(y) = (1-\varepsilon)P(y) + \varepsilon H(y)$, where theoretical and contaminating distributions of $P(y)$ and $H(y)$ are symmetric. All distribution



functions P_ε are combined into a single class K_ε . If $P(y)$ is a function of the normal distribution, this model of "gross errors" shows the situation when the approximately $(1-\varepsilon)N$ observations are the subject to the normal law (as in the classical case), and some εN of them are gross errors. Value ε (intensity of contamination or the probability of gross errors occurrence) is considered as a known number.

The concept of robust estimation was firstly introduced in the works devoted to the evaluation of the location parameter symmetric distribution [1], [2] and [3]. Robust estimation is an estimation of having the smallest asymptotic variance in K_ε class.

Over the last 50 years the scope of application of robust methods has been expanded. There was given a detailed overview of robust methods, which set out the main directions of their development and application [4] and [5].

At present, a robust approach is widely used in various branches of mathematical statistics general problems, see, e.g., [6] and [7].

Most of the modern works on this subject are devoted to the problems of multivariate statistical analysis and estimation of their parameters in case of the presence of gross errors in data or their contamination by extraneous data, see, e.g., [8] and [9].

In addition, Yuan and Zhong as well as Toman used it for factor analysis, using real and simulated data sets with known contamination schemes to demonstrate the performance of the forward search algorithm [10] and [6]; Todorov used it to solve the problem of robust discriminatory analysis [11]; Nevolaianen – for cluster analysis with the emphasis on nonparametric and robust approaches to the inference on the marginal distribution [12].

The idea to define the quality of the estimate through its maximum asymptotic bias belongs to P. Huber. This approach was implemented in applied to one-dimensional symmetric distributions, for which θ is a center parameter.

This paper shows more general contamination model than the one of J. Tukey and P. Huber, and also establishes a new approach based on the method of minimizing the largest asymptotic bias. Examples of its application for the most common probability distributions are also given.

The suggested approach may be applied in the development of information technologies, creation of automated control systems, quality assessment of cyber-physical systems.

Management decisions in these areas of activity involve a thorough analysis of information and its processing using the robust methods, so the application of the suggested method in situations of uncertainty will allow to obtain more reliable estimates of the parameters and will contribute to make the right decisions.

This paper has the following structure. In Section 2 minimax asymptotic bias method of obtaining robust estimates with minimax asymptotic bias is developed. Section 3 contains the results and examples. Section 4 contains discussion and perspectives of possible applications.

2. Minimax asymptotic bias method

The idea to define the quality of the estimate through its maximum asymptotic bias belongs to P. Huber. This approach was implemented by Huber applied to one-dimensional symmetric distributions, for which θ is a center parameter [13]. Consider an application of Huber's ideas to the more general case of any one-dimensional parameter distributions.

Let $p(x, \theta)$ is the density of the distribution of some random variables that depends on one unknown parameter θ , and $x_1, \dots, x_N$ is the N set of its implementations. The maximum likelihood estimate of the parameter θ in this case is the result of solving the problem

$$\prod_{i=1}^N p(x_i, \hat{\theta}) = \max.$$

If $p(x, \theta)$ is continuously differentiable according to θ , then estimate $\hat{\theta}$ is a root of the equation

$\sum_{i=1}^N \frac{p'_\theta(x_i, \hat{\theta})}{p(x_i, \hat{\theta})} = 0$, where $p'_\theta(x_i, \theta)$ is a derivative $p(x, \theta)$ according to θ .

Let's consider such estimates θ , which are the roots of the equation

$$\sum_{i=1}^N f(x_i, \hat{\theta}) = 0 \quad (1)$$

where $f(x, \theta)$ is the function that is not necessarily equal to $\frac{p'_\theta}{p}$. In case when the considered set is

contaminated, it may be that the choice of that function $f(x, \theta)$ will lead to more stable estimates.

Let $f(x, \theta)$ is continuously differentiable according to θ . Expanding $f(x, \hat{\theta})$ in a Taylor series, we get (1)

$$\sum_i f(x_i, \theta) + (\hat{\theta} - \theta) \sum_i f'_\theta(x_i, \theta) \approx 0,$$

from where

$$\hat{\theta} - \theta \approx - \frac{\sum_{i=1}^N f(x_i, \theta)}{\sum_{i=1}^N f'_\theta(x_i, \theta)} = \frac{\frac{1}{N} \sum_{i=1}^N f(x_i, \theta)}{\frac{1}{N} \sum_{i=1}^N f'_\theta(x_i, \theta)}. \quad (2)$$

Consider the behavior of the right part of (2) in the conditions when the set of x_i is contaminated. Denote $p(x, \theta)$ as a true distribution function, $H(x, \theta)$ as the distribution function of gross errors, ε as the intensity of contamination, $P_\varepsilon(x, \theta) = (1 - \varepsilon)P(x, \theta) + \varepsilon H(x, \theta)$ as the distribution function of the contaminated elements x_i .

If the random variables $f(x_i, \theta)$ and $f'_\theta(x_i, \theta)$ have mathematical expectations, then when $N \rightarrow \infty$ according to the enhanced law of large numbers, the numerator and denominator in (2) converge to their mathematical expectations, so

$$\hat{\theta} - \theta \rightarrow - \frac{E[f(x, \theta)]}{E[f'_\theta(x, \theta)]} = - \frac{\int f(x, \theta) dP_\varepsilon(x, \theta)}{\int f'_\theta(x, \theta) dP_\varepsilon(x, \theta)} = m. \quad (3)$$

It should be noted that here and in what follows it is assumed that if the limits of integration aren't specified, the integral is taken over the entire possible values of x .

Thus, estimate $\hat{\theta}$ in general case is asymptotically biased. It is reasonable to seek such $f(x, \theta)$ so that this m bias possibly occurs less.

Therefore, let's demand, first of all, that in the absence of the contamination estimate $\hat{\theta}$ would be asymptotically unbiased.

This requires that

$$\int f(x, \theta) dP(x, \theta) = \int f(x, \theta) p(x, \theta) dx = 0. \quad (4)$$

Differentiating the left part according to θ , we get

$$\int f'_\theta(x, \theta) p(x, \theta) dx + \int f(x, \theta) p'_\theta(x, \theta) dx = 0. \quad (5)$$

Therefore, if there is contamination, the value of bias is

$$m = - \frac{\int f(x, \theta) dP_\varepsilon(x, \theta)}{\int f'_\theta(x, \theta) dP_\varepsilon(x, \theta)} = \frac{\varepsilon \int f(x, \theta) dH(x, \theta)}{(1 - \varepsilon) \int f(x, \theta) p'_\theta(x, \theta) dx - \varepsilon \int f'_\theta(x, \theta) dH(x, \theta)}. \quad (6)$$

Thus, the bias of an estimate has ε order. Then, neglecting terms of ε order in the denominator (6), let's find

$$m \approx \varepsilon \frac{\int f(x, \theta) dH(x, \theta)}{\int f(x, \theta) p'_\theta(x, \theta) dx}. \quad (7)$$

We will look for a function $f(x, \theta)$ that with the worst contamination provides the smallest modulus of asymptotic bias of M estimate, i.e. we solve with every θ the problem of $M = \inf_f \sup_H |m|$.

It is easy to notice that the worst H contamination, giving the largest bias will be concentrated at the points of the maximal in modulus value f :

$$\sup_H \left| \int f(x, \theta) dH \right| = \sup_x |f(x, \theta)|.$$

$$\text{Thus, the optimum } f(x, \theta) \text{ is the solution of the problem } M = \varepsilon \frac{\sup_x |f(x, \theta)|}{\left| \int f(x, \theta) p'_\theta(x, \theta) dx \right|} = \inf_f \quad (8)$$

under the delimitation (4). The solution of this problem is not unique if f is one of the solutions, then $k(\theta)f(x, \theta)$ with $k(\theta) \neq 0$ is also a solution. Therefore, without delimitation of generality, we can assume that $\sup_x |f(x, \theta)| = 1$,

and to solve problem

$$\left| \int f(x, \theta) p'_\theta(x, \theta) dx \right| = \sup_f \quad (10)$$

under the delimitations (4) and (9). Moreover, once with f solution of the problem will be $-f$, it's possible to omit the modulus sign in (10). We get the following problem

$$\int f(x, \theta) p'_\theta(x, \theta) dx = \sup_f$$

under conditions $\int f(x, \theta) p(x, \theta) dx = 0, \sup_x |f(x, \theta)| = 1$.

Therefore, the optimal f will be in some $\lambda(\theta)$ solution of the problem

$$\int f(x, \theta) p'_\theta(x, \theta) dx - \lambda(\theta) \int f(x, \theta) p(x, \theta) dx = \int f(x, \theta) [p'_\theta(x, \theta) - \lambda(\theta) p(x, \theta)] dx = \sup$$

under the delimitation (9). The solution of this problem is obvious; f should be equal to 1 or -1 depending on if the expression in square brackets is positive or negative:

$$f(x, \theta) = \text{sign} [p'_\theta(x, \theta) - \lambda(\theta) p(x, \theta)] \quad (11)$$

where

$$\text{sign}(a) = \begin{cases} +1, & \text{if } a > 0; \\ -1, & \text{if } a < 0; \\ 0, & \text{if } a = 0. \end{cases}$$

Unknown $\lambda(\theta)$ will be defined from the condition (4)

$$\int \text{sign} [p'_\theta(x, \theta) - \lambda(\theta) p(x, \theta)] p(x, \theta) dx = 0. \quad (12)$$

Prove that the solution of this equation exists and it is unique. To that end, note that

$$\varphi(\lambda) = \int \text{sign} [p'_\theta(x, \theta) - \lambda p(x, \theta)] p(x, \theta) dx = 1 - P\left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} \leq \lambda\right) - P\left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} < \lambda\right).$$

$$\text{Let } R(\lambda, \theta) = P\left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} \leq \lambda\right). \quad (13)$$

Suppose that $\frac{p'_\theta(x, \theta)}{p(x, \theta)}$ is a continuous function from x that takes each of its possible values only with

$$\text{zero probability, i.e. for any } \lambda \quad P\left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} = \lambda\right) = 0.$$

Then $R(\lambda, \theta)$ is a continuous monotonic function from λ , i.e. $R(\lambda_2, \theta) > R(\lambda_1, \theta)$ with $\lambda_2 > \lambda_1$ and $\varphi(\lambda) = 1 - 2R(\lambda, \theta)$.

$$\text{If we additionally require that } \lim_{\lambda \rightarrow \infty} P\left(\left|\frac{p'_\theta(x, \theta)}{p(x, \theta)}\right| > \lambda\right) = 0,$$

$$\text{then } \lim_{\lambda \rightarrow \infty} R(\lambda, \theta) = 1; \quad \lim_{\lambda \rightarrow -\infty} R(\lambda, \theta) = 0; \quad (14)$$

$$\text{so } \lim_{\lambda \rightarrow \infty} \varphi(\lambda) = -1; \quad \lim_{\lambda \rightarrow -\infty} \varphi(\lambda) = 1.$$

Since $R(\lambda)$, and hence $\varphi(\lambda)$ is continuous and monotone, it follows that the solution $\lambda(\theta)$ of the equation $\varphi(\lambda) = 0$ exists and it is unique.

Other properties of the obtained solution can be obtained if we assume that $\frac{p'_\theta(x, \theta)}{p(x, \theta)}$ is a monotone

non-increasing function according to θ . In this case, function $R(\lambda, \theta)$ is also non-increasing and $\lambda(\theta)$ as a root of the equation $R(\lambda, \theta) = \frac{1}{2}$ will be a non-decreasing function. In this case, function

$$f(x, \theta) = \text{sign}\left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} - \lambda(\theta)\right) \text{ is also non-increasing according to } \theta.$$

There is no need to consider the case when $\frac{p'_\theta(x, \theta)}{p(x, \theta)}$ increases with θ because it is impossible. Indeed,

$$\text{in this case it would be } \frac{\partial}{\partial \theta} \left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} \right) \geq 0, \quad \text{from where } \frac{p''_\theta(x, \theta)}{p(x, \theta)} - \left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} \right)^2 \geq 0$$

$$\text{and, therefore } \int p''_\theta(x, \theta) dx \geq \int \left(\frac{p'_\theta(x, \theta)}{p(x, \theta)} \right)^2 p(x, \theta) dx > 0,$$

that is impossible because the left part is the second derivative according to θ from the integral θ and therefore it is equal to zero.

At the optimal choice $f(x, \theta)$ it's possible in accordance with (8) to define the maximum asymptotic bias of an estimate $\hat{\theta}$. We've got taking into account (4)

$$M = \varepsilon \left(\left| \int p'_\theta(x, \theta) - \lambda(\theta) p(x, \theta) dx \right| \right)^{-1}, \quad (15)$$

moreover, with the made assumptions there is a continuous function from θ on the right.

Our result inadequately justified. The fact that the optimal $f(x, \theta)$ examined as a function from θ , discontinuous, and in such cases, the neglected term in the denominator of (6) could lead to a significant worsening of the estimate. However, in this case, it is not like that. First of all, with some

$H(x, \theta)$ the received estimate with accuracy to the terms of ε^2 order is achievable. It is, for example, when $H(x, \theta)$ is concentrated at the points where $f(x, \theta) = 1$ and which are not the points of discontinuity $f(x, \theta)$. In this case, the second term in bias happens is equal to $\frac{M}{1 - \varepsilon}$.

Prove that with accuracy to terms of ε^2 order bias larger than M , in general case is impossible. With this proof we will not use the approximate expression (2) for $\hat{\theta} - \theta$, and will directly estimate this difference. Such proof can be carried out only under certain restrictions on the true distribution density of $p(x, \theta)$.

3. Results and examples

Theorem. Let $p(x, \theta)$ is a distribution density of a random variable with the following properties:

1) distribution function of a random variable $\frac{p'_\theta(x, \theta)}{p(x, \theta)}$, i.e. function $R(\lambda, \theta)$ is specified by the equation (13) that is continuous and satisfies (14);

2) $p(x, \theta)$ is twice differentiating according to θ and $\int |p''_\theta(x, \theta)| dx < A(\theta)$ (16)

where $A(\theta)$ is a constrained continuous function.

Let $f(x, \theta)$ is defined by formula (11) and (4) and

$$\mu(\theta) = \int f(x, \theta) p'_\theta(x, \theta) dx = \int |p'_\theta(x, \theta) - \lambda(\theta) p(x, \theta)| dx. \quad (17)$$

Then for any $0 < \delta < \frac{1}{2}$ there is such ε_0 that with $0 \leq \varepsilon \leq \varepsilon_0$ and $N \rightarrow \infty$ the solution of $\hat{\theta}$ equation

(1) falls into the interval $(\theta - \gamma, \theta + \gamma)$ where $\gamma = \frac{\varepsilon}{\mu(\theta)(1 - \delta)}$ with probability 1. If $\frac{p'_\theta(x, \theta)}{p(x, \theta)}$ is a monotone non-increasing function θ , then this solution is unique.

Proof. There is a solution of the equation (1) in the interval $(\theta - \gamma, \theta + \gamma)$, if

$$\frac{1}{N} \sum_{i=1}^N f(x_i, \theta - \gamma) > 0 > \frac{1}{N} \sum_{i=1}^N f(x_i, \theta + \gamma).$$

When $N \rightarrow \infty$, the right and left parts of the inequalities with probability 1 tend to their mathematical expectations. The theorem will be proved if we show that

$$E[f(x, \theta - \gamma)] > 0 > E[f(x, \theta + \gamma)].$$

Let's prove the left part of this inequality, since the right one is proved in the similar way.

Due to $\mu(\theta)$ continuity and positivity it is possible to choose such $\varepsilon_1 > 0$ that with $|t - \theta| < \varepsilon_1$

$$|\mu(t) - \mu(\theta)| < \frac{\delta}{3} \mu(\theta). \quad (18)$$

Then, due to $A(\theta)$ continuity and positivity it is possible to choose such $\varepsilon_2 > 0$ that with $|t - \theta| < \varepsilon_2$

$$A(t) < \frac{\delta \mu^2(\theta)}{3 \varepsilon_2}. \quad (19)$$

Now let $\varepsilon_0 = \min \left(\frac{\delta}{3}, \frac{2\varepsilon_1}{\mu(\theta)}, \frac{2\varepsilon_2}{\mu(\theta)} \right)$. Then with $\delta \leq \frac{1}{2}$ it will be defined that γ will be $\gamma < \frac{2\varepsilon}{\mu(\theta)}$

(20)

and therefore $\gamma < \varepsilon_1, \gamma < \varepsilon_2$.

(21)

Since any $\theta | f(x, \theta) | \leq 1$, then

$$E[f(x, \theta - \gamma)] = (1 - \varepsilon) \int f(x, \theta - \gamma) p(x, \theta) dx + \varepsilon \int f(x, \theta - \gamma) dH(x, \theta) \geq (1 - \varepsilon) \int f(x, \theta - \gamma) p(x, \theta) dx - \varepsilon$$

Expanding $p(x, \theta)$ in a Taylor series with the residual term in integral form, we obtain

$$E[f(x, \theta - \gamma)] \geq (1 - \varepsilon) \gamma \mu(\theta - \gamma) + \int_{\theta - \gamma}^{\theta} \int p''_{\theta}(x, t) f(x, \theta - \gamma) (t - \theta + \gamma) dx dt - \varepsilon = (1 - \varepsilon) \gamma \mu(\theta) \left(1 - \frac{\delta}{3} \right) -$$

$$- \frac{\delta \mu^2(\theta) \gamma^2}{3\varepsilon_2} - \varepsilon > \gamma \mu(\theta) (1 - \delta) - \varepsilon = 0. \quad \square$$

If $\frac{p'_{\theta}(x, \theta)}{p(x, \theta)}$ is a monotone non-increasing function θ , then, as it was shown above, the same will be

$f(x, \theta)$ and the left part (1). In this case, equation (1) has a unique solution $\hat{\theta}$ that proves the theorem. Consequently, with accuracy to terms of order ε^2 bias of an estimate is larger than $M = \varepsilon \left(\int |p'(x, \theta) - \lambda(\theta) p(x, \theta)| dx \right)^{-1}$ and in general case it is impossible.

Example 1. Let it be necessary to estimate the mathematical expectation θ of a random variable distributed according to the normal law with the known dispersion S . In this case

$$p(x, \theta) = \frac{1}{\sqrt{2\pi S}} e^{-\frac{(x-\theta)^2}{2S}};$$

$$p'_{\theta}(x, \theta) = p(x, \theta) \frac{x - \theta}{S}; \quad p''_{\theta}(x, \theta) = p(x, \theta) \left(\left(\frac{x - \theta}{S} \right)^2 - \frac{1}{S} \right)$$

.It is easy to notice that $\frac{p'_{\theta}(x, \theta)}{p(x, \theta)}$ is a decreasing function θ and the variable $\frac{p'_{\theta}(x, \theta)}{p(x, \theta)}$ has a normal

distribution with an average equal to 0 and dispersion $\frac{1}{S}$. Hypotheses of the theorem are fulfilled

here, and therefore, the optimal $f(x, \theta)$ is given by the formulas (11) and (4). In this case

$$f(x, \theta) = \text{sign} \left(\frac{p'_{\theta}(x, \theta)}{p(x, \theta)} - \lambda(\theta) \right) = \text{sign} \left(\frac{x - \theta}{S} - \lambda(\theta) \right).$$

Unknown $\lambda(x, \theta)$ will be found as a median of a random variable $\frac{x - \theta}{S}$ and will be equal to 0. Thus,

$$f(x, \theta) = \text{sign}(x - \theta)$$

Therefore, the optimal estimate $\hat{\theta}$ will be defined from the solution of the equation

$$\sum_{i=1}^N \text{sign}(x_i - \hat{\theta}) = 0, \text{ and hence, will be a sample median of } \text{Med}(x_1, \dots, x_N), \text{ i.e.}$$

$$\hat{\theta} = \begin{cases} x_{(k+1)}, & \text{if } N = 2k + 1; \\ \frac{x_{(k+1)} + x_{(k+1)}}{2}, & \text{if } N = 2k. \end{cases}$$

Example 2. Let it be necessary to estimate a mean-square deviation σ of a random variable, distributed according to the normal law with mathematical expectation m . In this case

$$p(x, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}};$$

$$p'_\sigma(x, \sigma) = p(x, \sigma) \left[\frac{(x-m)^2}{\sigma^3} - \frac{1}{\sigma} \right]; \quad p''_\sigma(x, \sigma) = p(x, \sigma) \left[\frac{(x-m)^4}{\sigma^6} - \frac{5(x-m)^2}{\sigma^4} + \frac{2}{\sigma^2} \right].$$

Here $\frac{p'_\theta(x, \sigma)}{p(x, \sigma)}$ isn't already a monotone function σ and the non-uniqueness of the optimal estimate is possible. We have

$$f(x, \sigma) = \text{sign} \left(\frac{(x-m)^2}{\sigma^3} - \frac{1}{\sigma} - \lambda(\sigma) \right),$$

and, as before, $\lambda(x, \sigma)$ is a median of a random variable $\frac{p'_\theta(x, \sigma)}{p(x, \sigma)} = \frac{(x-m)^2}{\sigma^3} - \frac{1}{\sigma}$,

i.e. fulfill the condition

$$\int_{\frac{(x-m)^2}{\sigma^3} - \frac{1}{\sigma} \leq \lambda} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{2}.$$

Perform the change of variables in the integral $x = m + y\sigma$, we get

$$\frac{1}{2} = \int_{\frac{y^2-1}{\sigma} \leq \lambda} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{1+\lambda\sigma}} e^{-\frac{y^2}{2}} dy.$$

It follows that $\sqrt{1+\lambda\sigma}$ is the 75% quantile of the normal distribution and, therefore $\sqrt{1+\lambda\sigma} = a \approx 0,67449$.

So $\lambda = \frac{a^2-1}{\sigma}$;

$$f(x, \sigma) = \text{sign} \left(\frac{(x-m)^2}{\sigma^3} - \frac{1}{\sigma} - \frac{a^2-1}{\sigma} \right) = \text{sign} \left(\frac{(x-m)^2}{a^2} - \sigma^2 \right).$$

Hence the equation to define the optimal estimate $\hat{\sigma}$ will be

$$\sum_{i=1}^N \text{sign} \left(\frac{(x_i - m)^2}{a^2} - \hat{\sigma}^2 \right) = 0,$$

i.e. $\hat{\sigma}^2$ is the sample median of $\frac{(x_i - m)^2}{a^2}$. As we see, despite of the non-monotonous of $\frac{p'_\theta(x, \theta)}{p(x, \theta)}$, the solution $\hat{\sigma}$ is unique.

4. Discussion and conclusions

The asymptotic bias as a characteristic of robustness is more stable method than any other known, it's a tool for assessing the quality of parameters estimates of probability models in situations of uncertainty type of contaminating distributions. In Huber and Tukey's approach symmetry both the estimated distribution and the symmetric contaminating distributions are postulated. This significantly narrows the application of this approach.

The paper shows the estimates of the median type. Thus, in case of normally distributed aggregate (the most common in practice), the sample median is a robust estimate of mathematical expectation, and

the estimate of variance is the sample median of $\frac{(x_i - \hat{m})^2}{a^2}$, where a is the 75% quintile of the standard normal distribution.

There is the need to improve forecast accuracy in different areas like economics, management, ecology, meteorology, agriculture. Considering this, as well as the fact that hypothetical suggestions, formulated in the form of probability distributions, are often violated in practice, it determines the relevance of the minimax bias approach application.

The suggested approach may be applied in the development of information technologies, creation of automated control systems, quality assessment of cyber-physical systems.

Management decisions in these areas of activity involve a thorough analysis of information and its processing using the robust methods, so the application of the suggested method in situations of uncertainty will allow to obtain more reliable estimates of the parameters and will contribute to make the right decisions.

References

- [1] Tukey, J., 1960. A survey of sampling from contaminated distributions. In *Contributions to Probability and Statistics*, I. Olkin, Ed., Stanford University Press, Stanford.
- [2] Tukey, J., 1962. The future of data analysis. *Annals of Mathematical Statistics* 33, 1–67.
- [3] Huber, P., 1964. Robust estimation of a location parameter. *Annals of Mathematical Statistics* 35, 73–101.
- [4] Morgenthaler, S., 2007. A survey of robust statistics. *Stat. Math. & Appl.* 15, 271–293.
- [5] Huber, P., Ronchetti, E., 2009. *Robust statistics*, second ed. J. Wiley, New Jersey.
- [6] Toman, A., 2014. Robust confirmatory factor analysis based on the forward search algorithm. *Stat. Papers* 55, 233–252.
- [7] Mbaidheen, M., Alawneh, A., 2017. Assessing the risk of corporate strategic change. *International Journal of Information, Business and Management* 9(1), 45–56.
- [8] Agostinelli, C., Leung, A., Yohair, V., Zomer, R., 2015. Robust estimation of multivariate location and scatter in the presence of cellwise and casewise contamination. *TEST* 24, 441–461.
- [9] Cizak, P., 2011. Efficient robust estimation of regression models. *Comput. Stat. Data Anal* 55, 774–788.
- [10] Yuan, K., Zhong, X., 2008. Outliers, leverage observations and influential cases in factor analysis: using robust procedures to minimize their effect. *Sociological Methodology* 38, 329–368.
- [11] Todorov, V., 2007. Robust selection of variables in linear discriminant analysis. *Stat. Math. & Appl* 15, 395–407.
- [12] Nevola, J., Datta, S., Ojc, H., 2014. Inference on the marginal distribution of clustered data with informative cluster size. *Stat. Papers* 55, 71–92.
- [13] Huber, P. 1972. Robust statistics: a review. *Annals of Mathematical Statistics* 43, 1041–1067.