Int. J. Appl. Math. Comput. Sci., 2005, Vol. 15, No. 3, 393–404

BAYES SHARPENING OF IMPRECISE INFORMATION

PIOTR KULCZYCKI*, MAŁGORZATA CHARYTANOWICZ**

* Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01–447 Warsaw, Poland
e-mail: kulczycki@ibspan.waw.pl

** Institute of Mathematics, Catholic University of Lublin
Al. Racławickie 14, 20–950 Lublin, Poland
e-mail: m.charytanowicz@op.pl

A complete algorithm is presented for the sharpening of imprecise information, based on the methodology of kernel estimators and the Bayes decision rule, including conditioning factors. The use of the Bayes rule with a nonsymmetrical loss function enables the inclusion of different results of an under- and overestimation of a sharp value (real number), as well as minimizing potential losses. A conditional approach allows to obtain a more precise result thanks to using information entered as the assumed (e.g. current) values of conditioning factors of continuous and/or binary types. The nonparametric methodology of statistical kernel estimators freed the investigated procedure from arbitrary assumptions concerning the forms of distributions characterizing both imprecise information and conditioning random variables. The concept presented here is universal and can be applied in a wide range of tasks in contemporary engineering, economics, and medicine.

Keywords: imprecise information, sharpening, conditioning factors, kernel estimators, Bayes decision rule, nonsymmetrical loss function, numerical calculations

1. Introduction

Recently there has been a notable increase in research in the area of imprecise information processing, and its applications in various branches of science and technology. This is often connected with the need for mathematical modeling of complicated systems or poorly defined intuitive notions, difficult to describe using conventional methods. Imprecise information, common in natural language—such as "big changes" or "high quality"—has been employed by people for a long time, and is often the only available kind of information. It was not, however, applied in methods based on classical mathematics, which lead to a loss, or a confusion, of the knowledge held. Using models permitting a much more general class of information created new, hitherto unknown possibilities, although this required the development of original mathematical tools for the representation of data.

One of the fundamental tasks in this type of approach is obtaining a concrete sharp value (real number), based on imprecise opinions, which best possibly characterizes the aspect under consideration. The algorithms of sharpening which are available in the literature do not take into account different (in sign as well as size) effects of an over- and underestimation of results. The method investigated in this paper, based on the Bayes decision rule, allows for the consideration of this aspect, while the result obtained minimizes potential losses. Furthermore, the proposed algorithm enables the influence of various conditioning factors—of continuous and binary types—to be taken into account. It is suited for all problems where defining the characteristics of imprecision can be carried out based on a series of opinions regarding concrete values. Mathematical tools originate from the theory of statistical kernel estimators, which frees the method from the types of distributions of imprecise information as well as conditioning factors.

The goal of this paper is to provide a complete algorithm allowing to calculate the Bayes sharp value of imprecise information for fixed values of conditioning factors, in a form which does not require the user to have a deeper knowledge of theoretical aspects.

2. Mathematical Preliminaries

2.1. Elements of Decision Theory

The main aim of decision theory is to select one concrete element from a set of all decisions possible to make, based solely on the characteristics of the distribution of an imprecision measure. Let then the following be given:

- S – a nonempty set of possible states of nature;

- D – a nonempty set of allowable decisions;
- the mapping $l : S \times D \to \mathbb{R}$ being a loss function whose values $l(s,d)$ can be interpreted as losses incurred where the decision d has been made, while the state s exists in reality.

Moreover, let a measure of imprecise states of nature $s \in S$, characterized by a distribution density denoted by μ, be defined. If for every $d \in D$ the integral $\int_S l(s,d)\mu(s)\mathrm{d}s$ exists, then the mapping $l_b : D \to \mathbb{R}$ described by the formula

$$l_b(d) = \int_S l(s,d)\mu(s)\mathrm{d}s \tag{1}$$

is called a Bayes loss function, and every element $d_b \in D$ fulfilling the condition

$$l_b(d_b) = \inf_{d \in D} l_b(d) \tag{2}$$

becomes a Bayes decision, and the above procedure—a Bayes rule. It requires the distribution of the measure of the imprecision of states of nature to be identified, which in the case of a probabilistic approach can be the density of a probability distribution (Billingsley, 1989), as well as a membership function for fuzzy logic (Kacprzyk, 1986).

2.2. Statistical Kernel Estimators

The concept of statistical kernel estimators of a distribution density of the n-dimensional random variable Z whose n_c coordinates are continuous, while the remaining n_b are binary, is presented below. Variables of both these types will be considered first separately, and then together in one approach.

Let therefore (Ω, Σ, P) be a probability space. Let also the n_c-dimensional continuous random variable $X : \Omega \to \mathbb{R}^{n_c}$ be given, with a distribution characterized by the density f_X. The corresponding kernel estimator $\hat{f}_X : \mathbb{R}^{n_c} \to [0,\infty)$, calculated using experimentally obtained values for the m-element random sample $x_1, x_2, \ldots, x_m$, is in its basic form defined as

$$\hat{f}_X(x) = \frac{1}{mh^{n_c}} \sum_{i=1}^{m} K_c\left(\frac{x - x_i}{h}\right), \tag{3}$$

where $m \in \mathbb{N}\backslash\{0\}$, the coefficient $h > 0$ is called a smoothing parameter, while the measurable function $K_c : \mathbb{R}^{n_c} \to [0,\infty)$ of unit inegral $\int_{\mathbb{R}^{n_c}} \hat{f}_X(x)\mathrm{d}x = 1$, symmetrical with respect to zero and having a weak global maximum at this point, takes the name of a kernel.

Let then the n_b-dimensional binary random variable $Y : \Omega \to \mathbb{B}^{n_b}$, where $\mathbb{B} = \{0,1\}$, be given. Its distribution density kernel estimator $\hat{f}_Y : \mathbb{B}^{n_b} \to [0,1]$, calculated on the basis of experimentally obtained values of the random sample $y_1, y_2, \ldots, y_m$, takes the form

$$\hat{f}_Y(y) = \frac{1}{m} \sum_{i=1}^{m} K_b(y, y_i), \tag{4}$$

where $m \in \mathbb{N}\backslash\{0\}$, and the kernel $K_b : \mathbb{B}^{n_b} \to [0,1]$ is defined by

$$K_b(y, y_i) = \lambda^{n_b - d(y,y_i)}(1-\lambda)^{d(y,y_i)}, \tag{5}$$

while $\lambda \in [0.5, 1]$ fulfils the role of a smoothing parameter for the binary component, whereas the function $d : \mathbb{B}^{n_b} \times \mathbb{B}^{n_b} \to \mathbb{N}$, expressed as $d(y_1, y_2) = (y_1 - y_2)^{\mathrm{T}}(y_1 - y_2)$, refers to a number of coordinates of the vectors y_1 and y_2 which are different.

Taking the above all together, consider the (n_c+n_b)-dimensional random variable $Z \equiv [X,\, Y]^T$, being a composition of the n_c-dimensional random variable X and the n_b-dimensional binary variable Y. The kernel K used for the estimation of the distribution density of the random variable Z takes the form

$$K(z, z_i) = K_c\left(\frac{x - x_i}{h}\right) K_b\left(y, y_i\right), \tag{6}$$

where $z \equiv [x,\, y]^T$ and $z_i \equiv [x_i,\, y_i]^T$ for $i = 1, 2, \ldots, m$. Finally, the kernel estimator $\hat{f}_Z$ of the density of the distribution of the random variable Z, calculated on the basis of the values of the m-element random sample $z_1, z_2, \ldots, z_m$, can be defined as

$$\hat{f}_Z(z) = \frac{1}{mh^{n_c}} \sum_{i=1}^{m} K(z, z_i), \tag{7}$$

where the kernel K is given by the formula (6).

The dependences (3), (4) and (7) constitute a fundamental form of the kernel estimator for the random variables: continuous X, binary Y, and their composition Z, respectively. The tasks concerning the choice of the form of the kernel K_c, the smoothing parameters h and λ, as well as additional procedures which improve the quality of the estimator still need to be investigated.

The choice of the form of the kernel K_c and the fixing of the smoothing parameter h are most often made with the criterion of the mean integrated square error.

Thus, the choice of the kernel form has no practical meaning and hence it is possible to take into account primarily the properties of the obtained estimator (e.g. its class of regularity, the boundary of a support) or aspects of calculations advantageous from the point of view of the applicational problem under study. This feature is particularly worth mentioning, as it seems to be especially convenient during practical investigations concerning kernel estimators.

The fixing of the smoothing parameters h and λ, however, greatly influences the estimation quality. The value of the former, applied to continuous coordinates, can be calculated as a value realizing the minimum of the function $g_c : (0, \infty) \to \mathbb{R}$ of the form

$$g_c(h) = \frac{1}{m^2 h^{n_c}} \sum_{i=1}^{m} \sum_{j=1}^{m} \tilde{K}_c \left(\frac{x_j - x_i}{h} \right) + \frac{2}{m h^{n_c}} K_c(0), \tag{8}$$

where

$$\tilde{K}_c(x) = K_c^{*2}(x) - 2K_c(x), \tag{9}$$

K_c^{*2} being the convolution square of the function K_c, i.e.

$$K_c^{*2}(x) = \int_{\mathbb{R}^{n_c}} K_c(u) K_c(x-u)\, \mathrm{d}u. \tag{10}$$

Similarly, the smoothing parameter λ, used for binary coordinates, is calculated in practice based on the maximum likelihood criterion and given as a value realizing a minimum of the function $g_b : [0.5, 1] \to \mathbb{R}$ defined as

$$g_b(\lambda) = -\sum_{i=1}^{m} \log \hat{f}_Y^{-i}(y_i), \tag{11}$$

where

$$\hat{f}_Y^{-i} = \frac{1}{m-1} \sum_{\substack{j=1 \\ j \neq i}}^{m} K_b(y, y_j). \tag{12}$$

It is worth noting that even for multidimensional random variables, the functions g_c and g_b are real (i.e. one-dimensional).

In the case of the basic definition of the kernel estimator for the continuous random variable (3), the influence of the smoothing parameter on particular kernels is the same. Particularly advantageous—in practical applications—results are obtained thanks to the individualization of this effect, which can be achieved through the so-called modification of the smoothing parameter. This relies on introducing the positive modifying parameters $s_1, s_2, \ldots, s_m$ mapped on particular kernels, described by the formula

$$s_i = \left(\frac{\hat{f}_X(x_i)}{\tilde{s}} \right)^{-\alpha}, \tag{13}$$

where $\alpha \in [0, 1]$, while $\tilde{s}$ denotes the geometrical mean of the numbers $\hat{f}_X(x_1)$, $\hat{f}_X(x_2), \ldots, \hat{f}_X(x_m)$ given in the form of the logarithmic equation

$$\log(\tilde{s}) = \frac{1}{m} \sum_{i=1}^{m} \log \left(\hat{f}_X(x_i) \right) \tag{14}$$

and, finally, defining the kernel estimator with the modification of the smoothing parameter in the following form:

$$\hat{f}_X(x) = \frac{1}{m h^{n_c}} \sum_{i=1}^{m} \frac{1}{s_i^{n_c}} K_c \left(\frac{x - x_i}{h s_i} \right). \tag{15}$$

Thanks to the above procedure, the areas in which the kernel estimator assumes small values (e.g. in the range of "tails") are additionally flattened, while the areas connected with large values are peaked. The result is an improvement in the characterization of the distribution under investigation. Moreover, an estimator with the modification of the smoothing parameter proves to be less sensitive to the precision of the fixing of this parameter value, and the difference in efficiency between particular types of kernel is also lessened. Based on indications for the criterion of the mean integrated square error, one can primarily assume $\alpha = 0.5$.

As regards the basic form of the kernel estimator of the continuous random variable (3), the smoothing parameter has the same influence on particular coordinates of this variable. Taking into account the possibility of sizable differences in scales of the above coordinates, for some of these the value of the parameter may turn out to be too small, whereas for others—too big. Because of this, a linear transformation is applied:

$$X \equiv RY, \tag{16}$$

where the matrix R is nonsingular. In practice it is most often assumed that

$$R = \sqrt{\mathrm{Cov}(X)}, \tag{17}$$

where the matrix $\mathrm{Cov}(X)$ denotes a covariance matrix of the variable X. Following the transformation (16), the kernel estimator takes the form

$$\hat{f}_X(x) = \frac{1}{m h^{n_c} \det(R)} \sum_{i=1}^{m} K_c \left(R^{-1} \frac{x - x_i}{h} \right). \tag{18}$$

As a result, the scales of particular coordinates become equal, and the shapes of kernels stretch out in a direction defined by proper correlation coefficients.

The concepts of kernel estimators with the modification of the smoothing parameter (15) and with the linear transformation (18) can be joined in a natural manner.

The kernel estimator methodology can also be applied to other probabilistic characteristics. Below the concept of kernel estimators for a regression function is presented. In this way, the arbitrary assumptions concerning the form of this function, a necessary part of the classical approach, may be avoided. Let m elements $(x_i, y_i) \in \mathbb{R} \times \mathbb{R}$ with $i = 1, 2, \ldots, m$ be given, where x_i can denote any fixed numbers as well as the realizations

of the one-dimensional random variable X, while y_i are realizations of the one-dimensional random variable Y. It is assumed that the function $g : \mathbb{R} \to \mathbb{R}$ exists and that the equality

$$y_i = g(x_i) + \varepsilon_i \tag{19}$$

is true, where ε_i stand independent random variables with zero expected value and finite variance. The concept of kernel estimators for the regression function is based on the use of the widely-known idea of a classic regression function, while here as the weights of particular elements (x_i, y_i), the proper values of the kernels mapped for the arguments x_i are taken. In the case most commonly used in practice, the regression function kernel estimator $\hat{g} : \mathbb{R} \to \mathbb{R}$ is thus given by

$$\hat{g}(x) = \frac{1}{mh} \sum_{i=1}^{m} \frac{[\hat{s}_2(x) - \hat{s}_1(x)(x_i - x)] y_i K_c\left(\frac{x - x_i}{h}\right)}{\hat{s}_2(x)\hat{s}_0(x) - (\hat{s}_1(x))^2}, \tag{20}$$

where

$$\hat{s}_r(x) = \frac{1}{mh} \sum_{i=1}^{m} (x_i - x)^r K_c\left(\frac{x_i - x}{h}\right) \tag{21}$$

for $r = 0, 1, 2$.

Detailed information on kernel estimators can be found in the monographs (Kulczycki, 2005; Silverman, 1986; Wand and Jones, 1995). Exemplary practical applications are presented in the papers (Kulczycki, 2000; 2002a; 2002b; Kulczycki and Wiśniewski, 2002). When applying particular numerical procedures, it is worth using the textbooks (Brandt, 1999; Dahlquist and Bjorck, 1983; Kiełbasiński and Schwetlick, 1994; Stoer and Bulirsch, 1987).

3. Characteristics of Imprecise Information

This section specifies the form of imprecise information which will be investigated later in this paper. Let therefore the n-dimensional random variable $Z : \Omega \to \mathbb{R}^n$ representing a random factor be given, and $U = \mathbb{R}^{n+1}$ mean a vector space whose first coordinate represents the imprecise quantity V and the remaining n components denote further coordinates of the random variable Z (continuous and binary). Let a measure connected with the imprecise quantity V be given in the subspace of the first coordinate, and a probabilistic measure of the random variable Z in the subspace spanned by the following n coordinates. Therefore, in the whole space U one can define the product measure—it is assumed that its distribution has the density $M : \mathbb{R}^{n+1} \to [0, \infty)$. When this density is calculated on the basis of many concrete values obtained successively for different realizations of the conditioning variable Z, and when this is a summable function, one can estimate it using the kernel estimators methodology described in Section 2.2. The random sample required there constitutes $(n+1)$-dimensional vectors of the form $[v_{i,\{w_i\}}, z_{i,1}, z_{i,2}, \ldots, z_{i,n}]^T$ for $i = 1, 2, \ldots, m$. The first coordinate $v_{i,\{w_i\}}$ represents imprecise information, where the notation used means that for the element v_i the opinion w_i is achieved. The further n coordinates $z_{i,1}, z_{i,2}, \ldots, z_{i,n}$ denote the next components of the realizations of the conditioning variable Z, for which the above values v_i and w_i are obtained.

Such an approach will be illustrated using the task of fixing a correct dosage of medicine with an imprecision of a fuzzy type, which is typical in medical practice. The opinion of particular sizes of the dose of medicine is intuitive and expressed verbally—a type of data occurring often in medical practice—and so it is represented here by a fuzzy number V. In the example discussed now, for any group of patients (indexed by $i = 1, 2, \ldots, m$) it will be required to obtain a verbal opinion of the effectiveness for the dose v_i prescribed to particular persons, e.g. "very good", "good", "acceptable", "bad". To these opinions, the values of the constants w_i can be mapped. A proper dosage depends mainly on the age and body mass (continuous variables), and gender (binary variable), which for particular patients are $z_{i,1}, z_{i,2}, z_{i,3}$. If the entire population is considered, these factors are random, represented by the three-dimensional random variable Z, with the coordinates Z_1, Z_2, Z_3. The product measure is defined on the space of the form $U \equiv [V, Z_1, Z_2, Z_3]^T$ and can be characterized applying statistical kernel estimators. For its calculation, one can make use of deterministic data of the type $[v_{i,\{w_i\}}, z_{i,1}, z_{i,2}, z_{i,3}]^T$ obtained on the basis of m clinical cases.

If one particular patient is taken for whom a correct dose of medicine is to be decided, then the realization $Z(\omega)$ of the conditioning random variable becomes fixed. It is worth noting that in the set of the obtained conditioning values $[z_{i,1}, z_{i,2}, z_{i,3}]^T$, there may not be a value equal to the realization $Z(\omega)$, or their low number may not be enough for responsible inference. On the other hand, the kernel estimator allows proper averaging of the data collected.

In practice, the patient must receive a fixed dose of medicine. Underestimating the dose most often has quite a different influence on the general cost of treatment from overestimation. With regard to this, the applied procedure should take an asymmetry of errors into account. The procedure worked out in further parts of this paper fulfils these demands.

4. Bayes Sharpening of Imprecise Information

The formula for the sharp value will be investigated below based on the Bayes decision rule (Section 2.1). As an integration measure occurring in the definition of the Bayes loss function (1), the product measure with the density M, considered in the previous section, with fixed realization of the conditioning variable Z, is assumed. In practical applications, such a density can be calculated using statistical kernel estimators (Section 2.2).

Let therefore the following be given: the sets of states of nature $S = \mathbb{R}$ and allowable decisions $D = \mathbb{R}$, as well as the loss function

$$l(\hat{v}, v) = \begin{cases} -p(\hat{v} - v) & \text{for} \quad \hat{v} - v \leq 0, \\ q(\hat{v} - v) & \text{for} \quad \hat{v} - v \geq 0, \end{cases} \tag{22}$$

where $v \in S$ and $\hat{v} \in D$, while the parameters p and q are positive and—it is worth underling—not necessarily equal. If the random factor $\omega \in \Omega$ (and therefore also the value of the random variable $Z(\omega)$) is fixed, then for the loss function (22), the Bayes loss function (1) takes the form

$$l_b(\hat{v}) = \int_{-\infty}^{\hat{v}} q(\hat{v} - v) M_{|Z(\omega)}(v)\,\mathrm{d}v - \int_{\hat{v}}^{\infty} p(\hat{v} - v) M_{|Z(\omega)}(v)\,\mathrm{d}v, \tag{23}$$

where $M_{|Z(\omega)}$ denotes the product measure density M with a fixed random factor $Z(\omega)$. It is readily shown that it attains a minimum for the value of the argument $\hat{v}$ fulfilling the following criterion:

$$\int_{-\infty}^{\hat{v}} M_{|Z(\omega)}(v)\mathrm{d}v = \frac{p}{p+q} \int_{-\infty}^{\infty} M_{|Z(\omega)}(v)\,\mathrm{d}v. \tag{24}$$

As the parameters p and q are positive, then $0 < p/(p+q) < 1$. Therefore it is easy to see that, if the function $M_{|Z(\omega)}$ has a connected support, then thanks to the continuity of the integral on the left-hand side of the above formula, the solution of the equation (24) exists and is unique. It is also worth noting that if the function $M_{|Z(\omega)}$ were to be multiplied by a nonzero constant, the above solution would remain the same. Finally, the value of the solution of the criterion (24) constitutes the desired sharp value of imprecise information with the fixed random factor $Z(\omega)$.

It is worth noting that, with respect to the equality

$$\frac{p}{p+q} = \frac{p/q}{p/q+1}, \tag{25}$$

there is no need to define the values of the parameters p and q separately, rather only the ratio p/q. In practical applications one can make use of this property in a variety of ways, e.g. assuming $p = 1$ and appropriately adapting the value of the parameter q, or, conversely, assuming $q = 1$ and changing p, in relation to a more suitable interpretation of a practical problem under investigation. Thus, it is necessary to identify only one parameter: the ratio p/q.

In the algorithm under investigation, the product measure density M will be calculated using statistical kernel estimators. There will be applied the main kernel estimators formula (7) with the modification of the smoothing parameter (15) as well as the linear transformation (18) taken into account, assuming therefore the following form:

$$\hat{f}_Z\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \frac{1}{mh^{n_c+1}\det(R)} \sum_{i=1}^{m} \frac{1}{s_i^{n_c+1}} \times K_c\left(R^{-1}\frac{x - x_i}{hs_i}\right) K_b(y, y_i). \tag{26}$$

What is more, for the needs of further investigations, this dependence is generalized to

$$\hat{f}_Z\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \frac{1}{h^{n_c+1}\det(R)\sum\limits_{i=1}^{m} w_i} \sum_{i=1}^{m} \frac{w_i}{s_i^{n_c+1}} \times K_c\left(R^{-1}\frac{x - x_i}{hs_i}\right) K_b(y, y_i), \tag{27}$$

while the coefficients w_i for $i = 1, 2, \ldots, m$ are nonnegative and not all equal to zero. It is worth noting that, if $w_i \equiv 1$, then the formula (27) reduces to the form (26).

Let therefore the $(n+1)$-dimensional space $U = [V, Z]^T$ be given, whose first coordinate represents the imprecise information V, while the remaining n coordinates constitute further components of the conditioning random variable $Z = [X, Y]^T$, where X is an n_c-dimensional continuous variable and Y — an n_b-dimensional binary variable. Introducing the natural notation $z = [x, y]^T$, for any $u \in U$ one can write

$$u = \begin{bmatrix} v \\ z \end{bmatrix} = \begin{bmatrix} v \\ x \\ y \end{bmatrix}. \tag{28}$$

The product measure density M is characterized using the kernel estimator calculated on the basis of m elements:

$$u_i = \begin{bmatrix} v_{i,\{w_i\}} \\ z_i \end{bmatrix} = \begin{bmatrix} v_{i,\{w_i\}} \\ x_i \\ y_i \end{bmatrix} \tag{29}$$

$$\hat{v}_0 = \frac{1}{\sum\limits_{i=1}^{m} w_i} \sum_{i=1}^{m} w_i v_i, \tag{33}$$

$$\hat{v}_{k+1} = \hat{v}_k + \frac{\frac{p}{p+q}\frac{1}{h^{n_c+1}\det(R)} \sum\limits_{i=1}^{m} \frac{w_i}{s_i^{n_c+1}} K_b\big(Y(\omega), y_i\big) \int\limits_{-\infty}^{\infty} K_c\left(R^{-1}\frac{\left[\begin{smallmatrix} v \\ X(\omega) \end{smallmatrix}\right] - \left[\begin{smallmatrix} v_i \\ x_i \end{smallmatrix}\right]}{hs_i}\right) \mathrm{d}v - \sum\limits_{i=1}^{m} w_i K_b\big(Y(\omega), y_i\big) I_c\left(R^{-1}\frac{\left[\begin{smallmatrix} \hat{v}_k \\ X(\omega) \end{smallmatrix}\right] - \left[\begin{smallmatrix} v_i \\ x_i \end{smallmatrix}\right]}{hs_i}\right)}{\frac{1}{h^{n_c+1}\det(R)} \sum\limits_{i=1}^{m} \frac{w_i}{s_i^{n_c+1}} K_b\big(Y(\omega), y_i\big) K_c\left(R^{-1}\frac{\left[\begin{smallmatrix} \hat{v}_k \\ X(\omega) \end{smallmatrix}\right] - \left[\begin{smallmatrix} v_i \\ x_i \end{smallmatrix}\right]}{hs_i}\right)} \tag{34}$$

for $k = 0, 1, \ldots$

for $i = 1, 2, \ldots, m$, of which the particular values v_i (together with their opinions w_i) are composed, obtained for different values of the conditioning variable $z_i = [x_i, y_i]^T$. According to the definition (27), the function $\hat{M} : \mathbb{R}^{n_c+1} \times \mathbb{B}^{n_b} \to [0, \infty)$ given as

$$\hat{M}\left(\left[\begin{array}{c} v \\ z \end{array}\right]\right) = \hat{M}\left(\left[\begin{array}{c} v \\ x \\ y \end{array}\right]\right) = \frac{1}{h^{n_c+1}\det(R)\sum\limits_{i=1}^{m} w_i} \sum_{i=1}^{m} \frac{w_i}{s_i^{n_c+1}} \times K_c\left(R^{-1}\frac{\left[\begin{array}{c} v \\ x \end{array}\right] - \left[\begin{array}{c} v_i \\ x_i \end{array}\right]}{hs_i}\right) K_b(y, y_i) \tag{30}$$

is the kernel estimator of the density M. The measurable function $K_c : \mathbb{R}^{n_c+1} \to [0, \infty)$ fulfils the condition $\int_{\mathbb{R}^{n_c+1}} K_c([u, x]^T)\mathrm{d}u\, \mathrm{d}x = 1$, which for any fixed $x \in \mathbb{R}^{n_c}$ guarantees the existence of its primitive $I_c : \mathbb{R} \to [0, \infty)$ with respect to the first coordinate, i.e.

$$I_c\left(\left[\begin{array}{c} v \\ x \end{array}\right]\right) = \int_{-\infty}^{v} K_c\left(\left[\begin{array}{c} u \\ x \end{array}\right]\right) \mathrm{d}u. \tag{31}$$

By applying kernel estimators to find the density $M_{|Z(\omega)}$, the criterion (24) then takes the form

$$\sum_{i=1}^{m} w_i K_b\left(Y(\omega), y_i\right) I_c\left(R^{-1}\frac{\left[\begin{array}{c} \hat{v} \\ X(\omega) \end{array}\right] - \left[\begin{array}{c} v_i \\ x_i \end{array}\right]}{hs_i}\right) = \frac{p}{p+q}\frac{1}{h^{n_c+1}\det(R)} \sum_{i=1}^{m} \frac{w_i}{s_i^{n_c+1}} K_b\big(Y(\omega), y_i\big) \times \int_{-\infty}^{\infty} K_c\left(R^{-1}\frac{\left[\begin{array}{c} v \\ X(\omega) \end{array}\right] - \left[\begin{array}{c} v_i \\ x_i \end{array}\right]}{hs_i}\right) \mathrm{d}v. \tag{32}$$

The solution of the above equation constitutes at last the desired sharp value of imprecise information with the fixed random factor $Z(\omega) = [X(\omega), Y(\omega)]^T$. If a kernel assuming positive values is used to construct the estimator, then the left-hand side is a function of the argument $\hat{v}$, continuous and strictly increasing from zero to a positive value, and, thanks to the dependence $0 < p/(p+q) < 1$, greater than the value of the right-hand side. The solution of (32) thus exists and is unique.

In practice, this solution can be calculated effectively based on Newton's numerical method. The desired sharp value is then given as a limit of the sequence $\{\hat{v}_k\}_{k=0}^{\infty}$ defined by the formulas (33) and (34), with the stopping condition $|\hat{v}_k - \hat{v}_{k-1}| \leq 0,01\hat{\sigma}_V$, where $\hat{\sigma}_V$ denotes an estimator of the standard deviation obtained for the sample $v_1, v_2, \ldots, v_m$. The dependences presented above constitute a fundamental version of Newton's algorithm. Depending on the individual character of a concrete applica-

tion, one can make use of many modifications available in the literature (Dahlquist and Bjorck, 1983; Stoer and Bulirsch, 1987).

In the following section, a comprehensive set of formulas necessary for a direct application of the above algorithm when using the Cauchy kernel is given.

5. Formulas for the Cauchy Kernel

5.1. Cauchy Kernel

As has been mentioned in Section 2.2, the assumed form of the kernel K_c has no practical meaning from the statistical point of view and, thanks to this, in applications, a suitable choice for the needs of a concrete task is possible. In the procedure investigated above, it is demanded that the kernel K_c assumes positive values, and its primitive with respect to the first coordinate is expressed by a convenient analytical formula. These properties ensure the Cauchy kernel defined as

$$K_c\left(\begin{bmatrix} v \\ x \end{bmatrix}\right) = K_c\left(\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}\right) = \frac{1}{A}\frac{1}{(1+v^2+x_1^2+x_2^2+\cdots+x_{n_c}^2)^a}, \tag{35}$$

while $a = [(n_c+3)/2]$ and $A = S_{n_c+1}\int_0^\infty r^{n_c}/(1+r^2)^a \mathrm{d}r$, where $[b]$ denotes the integer part of the number $b \in \mathbb{R}$, with S_{n_c+1} being the surface of the (n_c+1)-dimensional unit sphere. Concrete formulas for the Cauchy kernel (35) for particular dimensions n_c are included in Tab. 1.

The use of the algorithm investigated in this paper also requires the calculation of the primitive I_c, for the kernel K_c, with respect to the first coordinate v. The relevant formulas are given in Tab. 2. Furthermore, Tab. 3 contains the values of the definite integral with respect to this coordinate, within the limits from $-\infty$ to $+\infty$.

The above dependences are presented in Tabs. 1–3 for the basic form of the kernel estimator (3). In the general case for the version (27), following the modification of the smoothing parameter and the linear transformation, it is necessary to make a proper change of variables. The relevant formulas are presented in Sections 5.2 and 5.3, for the cases of $n_c = 1$ and $n_c \geq 2$, respectively.

Table 1. Forms of the Cauchy kernel for various dimensions n_c.

n_c	a	A	$K_c\left(\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}\right)$
1	2	π	$\frac{1}{\pi}\frac{1}{(1+v^2+x_1^2)^2}$
2	2	π^2	$\frac{1}{\pi^2}\frac{1}{(1+v^2+x_1^2+x_2^2)^2}$
3	3	$\frac{\pi^2}{2}$	$\frac{2}{\pi^2}\frac{1}{(1+v^2+x_1^2+x_2^2+x_3^2)^3}$
4	3	$\frac{\pi^3}{2}$	$\frac{2}{\pi^3}\frac{1}{(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2)^3}$
5	4	$\frac{\pi^3}{6}$	$\frac{6}{\pi^3}\frac{1}{(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)^4}$
6	4	$\frac{\pi^4}{6}$	$\frac{6}{\pi^4}\frac{1}{(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^4}$

Table 3. Definite integral of the Cauchy kernel with respect to the first coordinate from $-\infty$ to $+\infty$.

n_c	$\int_{-\infty}^{\infty} K_c\left(\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}\right)\mathrm{d}v$
1	$\frac{1}{2(1+x_1^2)^{3/2}}$
2	$\frac{1}{2\pi(1+x_1^2+x_2^2)^{3/2}}$
3	$\frac{3}{4\pi(1+x_1^2+x_2^2+x_3^2)^{5/2}}$
4	$\frac{3}{4\pi^2(1+x_1^2+x_2^2+x_3^2+x_4^2)^{5/2}}$
5	$\frac{15}{8\pi^2(1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)^{7/2}}$
6	$\frac{15}{8\pi^3(1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^{7/2}}$

5.2. Case of n_c=1

Let the elements of the matrix inverse of the transformation matrix R be denoted thus:

$$R^{-1} = P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \tag{36}$$

Table 2. Primitive for the Cauchy kernel with respect to the first coordinate.

n_c	$I_c\left(\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}\right)$
1	$\frac{1}{\pi(1+x_1^2)^{3/2}}\left[\frac{1}{2}\arctan\left(\frac{v}{\sqrt{1+x_1^2}}\right)+\frac{v\sqrt{1+x_1^2}}{2(1+v^2+x_1^2)}+\frac{\pi}{4}\right]$
2	$\frac{1}{\pi^2(1+x_1^2+x_2^2)^{3/2}}\left[\frac{1}{2}\arctan\left(\frac{v}{\sqrt{1+x_1^2+x_2^2}}\right)+\frac{v\sqrt{1+x_2^2+x_3^2}}{2(1+v^2+x_1^2+x_2^2)}+\frac{\pi}{4}\right]$
3	$\frac{1}{\pi^2(1+x_1^2+x_2^2+x_3^2)^{5/2}}\left[\frac{3}{4}\arctan\left(\frac{v}{\sqrt{1+x_1^2+x_2^2+x_3^2}}\right)+\frac{3v\sqrt{1+x_1^2+x_2^2+x_3^2}}{4(1+v^2+x_1^2+x_2^2+x_3^2)}+\frac{v}{2(1+v^2+x_1^2+x_2^2+x_3^2)^2}+\frac{3\pi}{8}\right]$
4	$\frac{1}{\pi^3(1+x_1^2+x_2^2+x_3^2+x_4^2)^{5/2}}\left[\frac{3}{4}\arctan\left(\frac{v}{\sqrt{1+x_1^2+x_2^2+x_3^2+x_4^2}}\right)+\frac{3v\sqrt{1+x_1^2+x_2^2+x_3^2+x_4^2}}{4(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2)}+\frac{v}{2(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2)^2}+\frac{3\pi}{8}\right]$
5	$\frac{1}{\pi^3(1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)^{7/2}}\left[\frac{15}{8}\arctan\left(\frac{v}{\sqrt{1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2}}\right)+\frac{15v\sqrt{1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2}}{8(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)}+\frac{5v}{4(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)^2}+\frac{v(1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)^{5/2}}{(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2)^3}+\frac{15\pi}{16}\right]$
6	$\frac{1}{\pi^4(1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^{7/2}}\left[\frac{15}{8}\arctan\left(\frac{v}{\sqrt{1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2}}\right)+\frac{15v\sqrt{1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2}}{8(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)}+\frac{5v}{4(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^2}+\frac{v(1+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^{5/2}}{(1+v^2+x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^3}+\frac{15\pi}{16}\right]$

moreover

$$d_{jk} = \frac{p_{jk}}{hs_i} \tag{37}$$

for $j = 1, 2$ and $k = 1, 2$,

$$a_0 = d_{11}^2 + d_{21}^2, \tag{38}$$

$$b_{01} = 2d_{11}d_{12} + 2d_{21}d_{22}, \tag{39}$$

$$a_1 = d_{12}^2 + d_{22}^2, \tag{40}$$

as well as

$$E = \left(a_1 - \frac{b_{01}^2}{4a_0}\right)(x_1 - x_{i,1})^2. \tag{41}$$

It is readily shown that $a_0 > 0$ and $E \geq 0$. Then the particular quantities occurring in the criterion (32) and the Newton's algorithm formulas (33)–(34) equal respectively

$$K_c\left(R^{-1}\frac{\begin{bmatrix} v \\ x_1 \end{bmatrix} - \begin{bmatrix} v_i \\ x_{i,1} \end{bmatrix}}{hs_i}\right) = K_c\left(\begin{bmatrix} d_{11}(v - v_i) + d_{12}(x_1 - x_{i,1}) \\ d_{21}(v - v_i) + d_{22}(x_1 - x_{i,1}) \end{bmatrix}\right), \tag{42}$$

$$I_c\left(R^{-1}\frac{\begin{bmatrix} v \\ x_1 \end{bmatrix}-\begin{bmatrix} v_i \\ x_{i,1} \end{bmatrix}}{hs_i}\right)$$

$$=\frac{1}{\sqrt{a_0}}I_c\left(\sqrt{a_0}(v-v_i)+\frac{b_{01}}{2\sqrt{a_0}}(x_1-x_{i,1})\right)_{\left|x_1^2=E\right.} \quad (43)$$

$$\int_{-\infty}^{\infty} K_c\left(R^{-1}\frac{\begin{bmatrix} v \\ x_1 \end{bmatrix}-\begin{bmatrix} v_i \\ x_{i,1} \end{bmatrix}}{hs_i}\right)\mathrm{d}v$$

$$=\frac{1}{\sqrt{a_0}}\int_{-\infty}^{\infty} K_c\left(\begin{bmatrix} v \\ \sqrt{E} \end{bmatrix}\right)\mathrm{d}v, \quad (44)$$

while the expressions on the right-hand sides of the above dependences can be sequentially read from Tabs. 1–3, where the symbol $I_c(\,\cdot\,)_{\left|x_1^2=E\right.}$ in the formula (43) means replacing x_1^2, in the proper equality in Tab. 2, with the value of E.

5.3. Case of $n_c \geq 2$

Let the elements of the matrix inverse of R be denoted in the following manner:

$$R^{-1}=P=\left[p_{jk}\right]_{\substack{j=1,2,\ldots,n_c+1\\ k=1,2,\ldots,n_c+1}}, \quad (45)$$

moreover

$$d_{jk}=\frac{p_{jk}}{hs_i} \quad (46)$$

for $j=1,2,\ldots,n_c+1$ and $k=1,2,\ldots,n_c+1$,

$$a_j=\sum_{k=1}^{n_c+1} d_{k\,j+1}^2 \quad (47)$$

for $j=0,1,\ldots,n_c$,

$$b_{pq}=\sum_{k=1}^{n_c+1} d_{k\,p+1}d_{k\,q+1} \quad (48)$$

for $p=0,1,\ldots,n_c-1$ and $q=1,2,\ldots,n_c$,

as well as

$$E=\sum_{j=1}^{n_c} a_j(x_j-x_{i,j})^2$$
$$+2\sum_{p=1}^{n_c-1}\sum_{q=p+1}^{n_c} b_{pq}(x_p-x_{i,p})(x_q-x_{i,q})$$
$$-\frac{1}{a_0}\sum_{j=1}^{n_c} b_{0j}^2(x_j-x_{i,j})^2$$
$$-\frac{2}{a_0}\sum_{p=1}^{n_c-1}\sum_{q=p+1}^{n_c} b_{0p}b_{0q}(x_p-x_{i,p})(x_q-x_{i,q}). \quad (49)$$

One can show that $a_0>0$ and $E\geq 0$. Here the particular quantities occurring in the criterion (32) and the Newton's algoritm formulas (33)–(34) equal respectively

$$K_c\left(R^{-1}\frac{\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}-\begin{bmatrix} v \\ x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,n_c} \end{bmatrix}}{hs_i}\right)$$

$$=K_c\begin{pmatrix} \sqrt{a_0}(v-v_i) \\ +\frac{1}{\sqrt{a_0}}\sum_{j=1}^{n_c} b_{0j}(x_j-x_{i,j}) \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{pmatrix}_{\left|x_1^2+x_2^2+\ldots+n_{n_c}^2=E\right.} \quad (50)$$

$$I_c\left(R^{-1}\frac{\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}-\begin{bmatrix} v \\ x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,n_c} \end{bmatrix}}{hs_i}\right)$$

$$=\frac{1}{\sqrt{a_0}}I_c\Big(\sqrt{a_0}(v-v_i)$$
$$+\frac{1}{\sqrt{a_0}}\sum_{j=1}^{n_c} b_{0j}(x_j-x_{i,j})\Big)_{\left|x_1^2+x_2^2+\ldots+n_{n_c}^2=E\right.} \quad (51)$$

$$\int_{-\infty}^{\infty} K_c\left(\frac{R^{-1}\left(\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix} - \begin{bmatrix} v \\ x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,n_c} \end{bmatrix}\right)}{hs_i}\right) \mathrm{d}v$$

$$= \frac{1}{\sqrt{a_0}} \int_{-\infty}^{\infty} K_c\left(\begin{bmatrix} v \\ x_1 \\ x_2 \\ \vdots \\ x_{n_c} \end{bmatrix}\right) \mathrm{d}v\,|_{x_1^2+x_2^2+\cdots+n_{n_c}^2=E}\,, \quad (52)$$

while the expressions on the right-hand sides of the above dependences can be sequentially read from Tabs. 1–3, where the symbol $(\cdot)_{|x_1^2+x_2^2+\cdots+n_{n_c}^2=E}$ means replacing $x_1^2+x_2^2+\cdots+x_{n_c}^2$, in the proper equality in these tables, with the value E.

6. Numerical Verification

The correct functioning of the algorithm worked out in the previous section has been confirmed by numerical simulations. In the trivial one-dimensional case, when $n_c = 0$, $n_b = 0$ and $w_i \equiv 1$, the investigated concept of the sharpening of imprecise information is reduced to a procedure which calculates the kernel estimator of a quantile. Mean errors in kernel and classic estimators of a quantile were compared in the work (Kulczycki, 2001), where the former seemed to be more precise. In the multidimensional case treated in this paper, taking into account conditioning variables of continuous and binary types, and also coefficients w_i mapped to particular sample elements v_i, the concept of sharpening presented here becomes remarkably general. The numerical verification described below was carried out for the case where imprecision is represented by a probabilistic measure, characterized by an assumed density of a probability distribution.

In each field of Tab. 4, the results obtained for 100 independent samples are listed. These results are presented giving

- a theoretical value for the sharp value, denoted by v;
- the mean of the obtained sharp values $\bar{\hat{v}}$, calculated on the basis of the aforementioned 100 random samples;
- the standard deviation of the obtained sharp values $\tilde{\hat{v}}$ for the same 100 samples.

The last two quantities are expressed in particular fields using the natural notation $\bar{\hat{v}} \pm \tilde{\hat{v}}$. The parameters w_i mapped onto the respective values v_i are obtained using a generator for the symmetrical triangular distribution with the support $[0, 1]$. In this way verbal opinions expressed by people, being vague and nonextreme by nature, are characterized properly. In the description below, the following notation used for the parameter occurring in the criterion (24) and its version (32) is introduced:

$$r = \frac{p}{p+q}. \quad (53)$$

First, the intensity of the smoothing parameter modification procedure, defined by the value of the parameter $\alpha \in [0, 1]$ introduced in the formula (13), was investigated. The universal value $\alpha = 0.5$ seemed to be satisfactory in the majority of cases, particularly in real-life problems, although changes in the value of this parameter contribute to an improvement in the quality in extreme cases. In particular when the conditioning variable value deviates widely from the modal, one can propose the significant decreasing of the parameter α value, even to zero.

Basic research was carried out for $n_c = 1$ and $n_b = 0$, when in this case the two-dimensional random variable $[V,\ X]^T$ has the normal distribution with the expected value $E([V,\ X]^T) = [0,\ 0]^T$ and the covariance matrix $\mathrm{Cov}([\,V,\ X\,]^T) = \left[\begin{smallmatrix} 1 & 0.7 \\ 0.7 & 1 \end{smallmatrix}\right]$. The positive correlation of the variables V and X indicates that an increase in the values of the conditioning variable X should imply the corresponding growth in the sharp values. The obtained results are presented in Tab. 4. They confirm the above expectation as well as the general correctness of the algorithm worked out in this paper. As the sample size increased, the obtained sharp values converged to the theoretical ones, and the standard deviation to zero. The above asymptotical features are of fundamental significance from an applicational point of view, as they prove that it is possible to obtain any precision of sharpening, although this requires a sufficient random sample size. Therefore, in practice, the necessity of a proper compromise between these quantities is called for.

The convergence speed increased as the value of the parameter r given by the formula (53) was closer to 0.5, although even for extreme cases, i.e. when it neaved to 0 or 1, satisfactory results were achieved. The convergence speed also increased, as the value of the conditioning variable came within the range of its modal value. These facts are rather intuitively clear. It is worth noting that the cases where $r = 0.1$ or $r = 0.9$, and also when the value of the conditioning variable appears in the neighborhood of its second standard deviation, generally become very difficult problems, which are naturally associated with greater demands, and for some classical methods they are actually impossible to settle in practice.

Detailed results of numerical simulations, including those for multimodal and nonsymmetrical distributions,

Table 4. Results of a numerical simulation for the normal distribution with the expected value $E([V,\ X]^T) = [0, 0]^T$ and the covariance matrix $\mathrm{Cov}([V,\ X]^T) = \left[\begin{smallmatrix} 1 & 0.7 \\ 0.7 & 1 \end{smallmatrix}\right]$.

$X(\omega) = 0$

r / m	0.1 ($v = -0.92$)	0.3 ($v = -0.37$)	0.5 ($v = 0.00$)	0.7 ($v = 0.37$)	0.9 ($v = 0.92$)
10	-1.151 ± 0.425	-0.460 ± 0.312	-0.025 ± 0.279	0.397 ± 0.285	1.108 ± 0.432
20	-1.028 ± 0.273	-0.387 ± 0.227	0.001 ± 0.239	0.377 ± 0.256	0.989 ± 0.289
50	-0.965 ± 0.188	-0.364 ± 0.182	-0.002 ± 0.185	0.365 ± 0.187	0.957 ± 0.185
100	-0.943 ± 0.145	-0.382 ± 0.144	-0.016 ± 0.144	0.352 ± 0.141	0.918 ± 0.150
200	-0.921 ± 0.112	-0.374 ± 0.119	-0.014 ± 0.122	0.347 ± 0.124	0.892 ± 0.119
500	-0.890 ± 0.081	-0.365 ± 0.089	-0.006 ± 0.095	0.366 ± 0.094	0.885 ± 0.090
1000	-0.887 ± 0.065	-0.365 ± 0.069	0.002 ± 0.073	0.369 ± 0.079	0.886 ± 0.074

$X(\omega) = 1$

r / m	0.1 ($v = -0.22$)	0.3 ($v = 0.33$)	0.5 ($v = 0.70$)	0.7 ($v = 1.07$)	0.9 ($v = 1.62$)
10	-0.582 ± 0.539	0.171 ± 0.400	0.646 ± 0.355	1.130 ± 0.358	1.902 ± 0.504
20	-0.455 ± 0.371	0.233 ± 0.298	0.680 ± 0.262	1.127 ± 0.280	1.829 ± 0.339
50	-0.357 ± 0.258	0.286 ± 0.195	0.696 ± 0.194	1.107 ± 0.206	1.759 ± 0.260
100	-0.313 ± 0.186	0.299 ± 0.163	0.696 ± 0.152	1.091 ± 0.147	1.708 ± 0.175
200	-0.275 ± 0.141	0.307 ± 0.119	0.691 ± 0.112	1.083 ± 0.111	1.676 ± 0.137
500	-0.201 ± 0.099	0.333 ± 0.085	0.687 ± 0.093	1.047 ± 0.098	1.596 ± 0.101
1000	-0.229 ± 0.074	0.319 ± 0.064	0.692 ± 0.067	1.074 ± 0.068	1.623 ± 0.081

$X(\omega) = 2$

r / m	0.1 ($v = 0.48$)	0.3 ($v = 1.03$)	0.5 ($v = 1.40$)	0.7 ($v = 1.77$)	0.9 ($v = 2.32$)
10	0.103 ± 0.846	0.804 ± 0.654	1.364 ± 0.590	1.910 ± 0.576	2.798 ± 0.726
20	0.107 ± 0.554	0.877 ± 0.475	1.379 ± 0.438	1.892 ± 0.401	2.672 ± 0.418
50	0.241 ± 0.417	0.959 ± 0.384	1.443 ± 0.318	1.908 ± 0.342	2.563 ± 0.375
100	0.308 ± 0.328	0.995 ± 0.276	1.460 ± 0.246	1.887 ± 0.263	2.528 ± 0.281
200	0.338 ± 0.301	0.986 ± 0.232	1.421 ± 0.229	1.853 ± 0.220	2.494 ± 0.253
500	0.394 ± 0.219	0.995 ± 0.161	1.396 ± 0.156	1.809 ± 0.157	2.413 ± 0.191
1000	0.378 ± 0.181	0.994 ± 0.118	1.394 ± 0.124	1.805 ± 0.131	2.366 ± 0.176

taking into account binary conditioning factors, inference for the lack of data from the neighborhood of a fixed value of a conditioning variable, as well as an additional procedure for boundaries of a support, can be found in the work (Charytanowicz, 2005).

7. Summary

This paper proposes the concept of the sharpening of imprecise information based on the Bayes decision rule and the methodology of statistical kernel estimators, including conditioning factors of continuous and binary nature. A complete procedure for the algorithm was worked out, consisting in part of defining analytical forms of relevant functions and the fixing of parameter values. Thus, in order to apply the algorithm investigated here, which allows calculating the sharp value, the following steps are required:

1. Obtain m opinions w_i concerning the values v_i, achieved respectively for the values of the conditioning variables z_i with the continuous x_i and binary y_i coordinates.

2. Assume the quotient of losses resulting from an under- and overestimation of a sharp value, i.e. the quantity p/q, and then calculate the value $p/(p+q)$ from the dependency (25).

3. Using Tab. 1, find the appropriate formula for the Cauchy kernel K_c for a given number of continuous coordinates of the conditioning variable n_c.

4. Based on the sample $[v_i,\, x_i]^T$, compute the value of the smoothing parameter h by the minimization of the function (8), while the convolution square K_c^{*2} present there can be defined using the kernel estimator for the regression function (20)–(21).

5. Find the modification parameters s_i based on the dependences (13)–(14).

6. Calculate the covariance matrix, where the linear transformation matrix R is given by the formula (17).

7. Assign the primitive of the kernel K_c with respect to the coordinate v and the value of its definite integral from $-\infty$ to $+\infty$, using Tabs. 2–3 and the procedures investigated in Section 5.

8. As the parameter λ set the value which minimizes the function (11).

9. Define the binary kernel K_b using the dependence (5).

10. Apply the Newton's algorithm (33)–(34), which eventually results in the desired sharp value.

This paper is based on the work (Charytanowicz, 2005), where further information can be found, in particular, derivations of the formulas presented here.

References

Billingsley P. (1989): *Probability and Measure*. — New York: Wiley.

Brandt S. (1999): *Statistical and Computational Methods in Data Analysis*. — New York: Springer.

Charytanowicz M. (2005): *Bayesian sharpening of imprecise information in medical applications*. — Ph.D. thesis, Systems Research Institute, Polish Academy of Sciences, Warsaw, (in Polish).

Dahlquist G. and Bjorck A. (1983): *Numerical Methods*. — Englewood Cliffs: Prentice-Hall.

Kacprzyk J. (1986): *Fuzzy Sets in Systems Analysis*. — Warsaw: PWN, (in Polish).

Kiełbasiński A. and Schwetlick H. (1994): *Numerical Linear Algebra*. — Warsaw: WNT, (in Polish).

Kulczycki P. (2000): *Fuzzy controller for mechanical systems*. — IEEE Trans. Fuzzy Syst., Vol. 8, No. 5, pp. 645–652.

Kulczycki P. (2001): *An algorithm for Bayes parameter identification*. — J. Dynam. Syst. Meas. Contr., Vol. 123, No. 4, pp. 611–614.

Kulczycki P. (2002a): *Statistical inference for fault detection: A complete algorithm based on kernel estimators*. — Kybernetika, Vol. 38, No. 2, pp. 141–168.

Kulczycki P. (2002b): *A test for comparing distribution functions with strongly unbalanced samples*. — Statistica, Vol. LXII, No. 1, pp. 39–49.

Kulczycki P. (2005): *Kernel Estimators in Systems Analysis*. — Warsaw: WNT in press, (in Polish).

Kulczycki P. and Wiśniewski R. (2002): *Fuzzy controller for a system with uncertain load*. — Fuzzy Sets Syst., Vol. 131, No. 2, pp. 185–195.

Silverman B.W. (1986): *Density Estimation for Statistics and Data Analysis*. — London: Chapman and Hall.

Stoer J. and Bulirsch R. (1987): *Introduction to Numerical Analysis*. — New York: Springer.

Wand M.P. and Jones M.C. (1995): *Kernel Smoothing*. — London: Chapman and Hall.

Received: 25 April 2005
Revised: 30 June 2005