Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090233

A Convenient Ready-to-Use Algorithm for a Conditional Quantile Estimator

Piotr Kulczycki[1,2,*], *Małgorzata Charytanowicz*[1,3] *and Antoni Leon Dawidowicz*[4]

[1] Polish Academy of Sciences, Systems Research Institute, Center of Information Technology for Data Analysis Methods, Newelska 6, 01-447 Warsaw, Poland
[2] AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Division of Information Technology and Biometrics, Mickiewicza 30, 30-059 Cracow, Poland
[3] John Paul II Catholic University of Lublin, Institute of Mathematics and Computer Science, Konstantynów 1H, 20-708 Lublin, Poland
[4] Jagiellonian University, Institute of Mathematics, Łojasiewicza 6, 30-348 Cracow, Poland

Received: 4 Jun. 2014, Revised: 2 Sep. 2014, Accepted: 4 Sep. 2014
Published online: 1 Mar. 2015

Abstract: This paper contains a complete procedure for calculating the value of a conditional quantile estimator. The concept is based on the nonparametric kernel estimator method, which frees the algorithm from the random variables' distributions. The procedure was worked out in a ready-to-use form – specific formulas for functions and the parameter used were given. The practical implementation of this method is very simple, and its computational complexity is linear with respect to random sample size as well as the dimension of conditioning variable. Thanks to a clear, near intuitive interpretation it can easily be modified or generalized depending on the individual needs of atypical applications. In particular, conditioning variables can be taken into account, not only continuous (real), but also binary, discrete and categorized, or any of their combinations.

Keywords: conditional quantile; nonparametric estimation; ready-to-use numerical algorithm; conditioning variables of continuous, real, binary, discrete and categorized types

1 Introduction

Consider for initial illustration – without mathematical formality for the moment – a random variable with real values. The quantile of the assumed order $r \in (0,1)$ is the number $q_r \in \mathbb{R}$, such that the probability of the random variable taking values less than or equal to q_r is r, while for those greater than or equal is $1-r$. In a particular case when $r = 0.5$ the quantile $q_{0.5}$ is called a median and plays the analogical role to the expectation value. Similarly, the difference of the quantile of the orders of 0.75 and 0.25, i.e. $q_{0.75} - q_{0.25}$, referred to as quarter deviation, fulfills the role of a standard deviation.

In comparisons of the median and quarter deviation properties with the classic expectation value and standard deviation, the former often prevails in practical applications. First, they are robust with respect to outliers often arising from so-called gross errors. If for example a decimal point is mistakenly omitted in giving a person's height, then instead of 1.84 we get 184, which significantly changes the estimator of the expected value calculated on this basis, and especially – due to squaring – of the standard deviation. This type of error can occur quite often in the ubiquitous case of automating data collection, unchecked by humans. The above mistake will have a much smaller effect on the estimation of the quantile, if the erroneous value of 184 will be treated here solely as one of the large values, with no consequences resulting from the fact that it is bigger by a hundred times.

The median has one other valuable property, which is seeing ever more application as robust systems develop; namely, the estimation errors of parameters in complex systems can have an asymmetrical influence on their quality. For example for the case of control of a mechanical object, underestimation of its mass leads to adverse limit cycles, while overestimation results in a more effective sliding trajectory [1]. In this situation, instead of estimating the mass by expected value or median, it is worth using a quantile of order greater than 0.5, to increase the probability of a more advantageous overestimation in place of underestimation. The expected

* Corresponding author e-mail: Piotr.Kulczycki@ibspan.waw.pl

value does not possess any parameter allowing overestimation to be preferred to underestimation or vice-versa.

In specific practical applications, the quality of estimation of all possible characteristics of probability distribution can effectively increase through selection of quantities bearing significant influence on the examined phenomena and their appropriate introduction to the algorithm. Later consideration – in specific calculations – of definite values (for example in heating, the current outside temperature, or the current exchange rate in marketing and management) makes the obtained result significantly more precise. With respect to the quantile, the above leads to the notion of a conditional quantile. This will be exactly formulated and interpreted below.

Let the one-dimensional random variable Y be given, with a distribution characterized by the density f_Y. A quantile of the order $r \in (0,1)$ is every number $q_r \in \mathbb{R}$, such that

$$\int_{-\infty}^{q_r} f_Y(y)dy = r \; . \tag{1}$$

If the support of the function f_Y is connected (e.g. when f_Y is positive), then the quantile is unique.

The above definition will now be generalized for the conditional case. Here, besides the basic (sometimes termed the describing) one-dimensional random variable Y, let also be given the n_W-dimensional random variable W, called hereinafter the conditioning random variable. Their composition $Z = \begin{bmatrix} Y \\ W \end{bmatrix}$ is a random variable of the dimension $n_W + 1$. Assume that distributions of the variables Z and, in consequence, W have densities, denoted below as $f_Z : \mathbb{R}^{n_W+1} \to [0,\infty)$ and $f_W : \mathbb{R}^{n_W} \to [0,\infty)$, respectively. Let also be given the so-called conditioning value, that is the fixed value of a conditioning random variable $w^* \in \mathbb{R}^{n_W}$, such that

$$f_W(w^*) > 0 \; . \tag{2}$$

Then the function $f_{Y|W=w^*} : \mathbb{R} \to [0,\infty)$ given by

$$f_{Y|W=w^*}(y) = \frac{f_X(y,w^*)}{f_W(w^*)} \quad \text{for every } y \in \mathbb{R} \tag{3}$$

constitutes a conditional density of probability distribution of the random variable Y for the conditioning value w^*. The conditional density $f_{Y|W=w^*}$ can therefore be treated as a "classic" density f_Y, whose form has been made more accurate in practical applications with w^* – a concrete value taken by the conditioning variable W in a given situation.

Next, a quantile of the order $r \in (0,1)$ with the condition $w^* \in \mathbb{R}^{n_W}$ is every number $q_{r|w^*} \in \mathbb{R}$, such that

$$\int_{-\infty}^{q_{r|w^*}} f_{Y|W=w^*}(y)dy = r \; . \tag{4}$$

Similarly to the above unconditional case, if the support of the function $f_{Y|W=w^*}$ is connected, then the quantile is unique. The conditional quantile $q_{r|w^*}$ constitutes therefore the refinement of the "classic" quantile q_r by using the information that the conditioning random variable, in a specific situation, took the value w^*.

The estimation of the quantile value is not a simple task, nor has it been elaborated sufficiently. A survey of classic methods for the unconditional case can be found in the publications [2,3]. For typical simple applications, positional estimators [4] can be recommended, as can kernel [5,6]. In the case of the conditional quantile the most commonly applied is the quantile regression method [7]. This consists of treating conditioning variables as arguments of regressive function and calculating its value which best approximates that of the quantile of a given order. Here, the linear [8] and kernel-based [9] models are most often used. This offers satisfactory results for the majority of practical cases and is rightly considered as the leading method at this time. However, as usual, there are also disadvantages. For example for two different orders of quantile, the regression functions can intersect, which implies that for a fixed condition, and for two orders of the quantile $r_1 < r_2$ it may occur that the estimator of quantile q_{r_1} is greater that of q_{r_2}, which contradicts monotonicity – a basic feature of quantiles with respect to their order. This leads to many difficulties in application and interpretational misunderstandings. Moreover, to write a program alone which realizes this procedure is not easy, nor its analysis of the method allowing a study of errors and the creation of mutations adapted to individual needs. The above implies the research into many different concepts, some of which lie beyond statistical methods, e.g. [10,11].

This paper presents a procedure for calculating the estimator of a conditional quantile, based on the statistical kernel estimator methodology. Its nonparametric nature implies the worked out procedure is independent of types of random variable distributions. The key advantage is, however, its simplicity and ease of interpretation. The former allows the presentation in Section 3 of a basic version of the algorithm for continuous random variables, in ready-to-use form, so that a user wishing to employ it does not have to go into the mathematical foundations discussed below in Section 2, merely apply the formulas from the third section consecutively. Section 4 describes the results of numerical simulation, while the final Section 5 summarizes the subject material and – thanks to its simplicity of interpretation – offers the possibilities of creating individual modifications suitable in practice for specific particular applications, especially the potential increase in estimation accuracy and generalization of conditioning variables, not only the continuous (real) but, it should be clearly underlined, also binary, discrete and categorical (ordered and unordered), as well as their compositions.

2 Conditional quantile estimator

This section presents the mathematic foundations of the proposed method. Readers not very interested in the formal aspects, can find a description of the investigated procedure in its-ready-to-use form within Section 3.

Let the n-dimensional random variable X be given, with a distribution characterized by the density f. Its kernel estimator $\hat{f} : \mathbb{R}^n \to [0,\infty)$, calculated using experimentally obtained values for the m-element random sample

$$x_1, x_2, \ldots, x_m \,, \tag{5}$$

in its basic form is defined as

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^{m} K\left(\frac{x - x_i}{h}\right), \tag{6}$$

where $m \in \mathbb{N}\backslash\{0\}$, the coefficient $h > 0$ is called a smoothing parameter, while the measurable function $K : \mathbb{R}^n \to [0,\infty)$ of unit integral $\int_{\mathbb{R}^n} K(x)\mathrm{d}x = 1$, symmetrical with respect to zero and having a weak global maximum in this place, takes the name of a kernel. The choice of form of the kernel K and the calculation of the smoothing parameter h is made most often with the criterion of the mean integrated square error.

Thus, the choice of the kernel form has – from a statistical point of view – no practical meaning and thanks to this, it becomes possible to take into account primarily properties of the estimator obtained, or calculational aspects, both advantageous from the point of view of the applicational problem under investigation; for broader discussion see [12] – Section 3.1.3; [13] – Sections 2.7 and 4.5. In practice, for the one-dimensional case (i.e. when $n = 1$), the function K is assumed most often to be the density of a common probability distribution. In the multidimensional case, two natural generalizations of the above concept are used: radial and product kernels. However, the former is somewhat more effective, although from an applicational point of view, the difference is immaterial and the product kernel – significantly more convenient in analysis – is often favored in practical problems. The n-dimensional product kernel K can be expressed as

$$K(x) = K\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \mathcal{K}_1(x_1)\mathcal{K}_2(x_2)\ldots\mathcal{K}_n(x_n) \,, \tag{7}$$

where $\mathcal{K}_i$ for $i = 1,2,\ldots,n$ denotes the previously-mentioned one-dimensional kernels, while the expression h^n appearing in the basic formula (6) should be replaced by $h_1 \cdot h_2 \cdot \ldots \cdot h_n$, the product of the smoothing parameters for particular coordinates.

The fixing of the smoothing parameter h has significant meaning for quality of estimation. Fortunately – from the applicational point of view – many suitable procedures for calculating the value of the parameter h on the basis of random sample (5) have been worked out, although most of them are time-consuming, especially with large sample sizes m. For broader discussion of the above tasks see [12, 13, 14].

The kernel estimators technique will now be used below for the task of conditional quantile estimation, formulated in the Introduction. Let – as defined there – a one-dimensional describing random variable Y as well as the n_W-dimensional conditioning random variable W be given. Suppose also the random sample

$$\begin{bmatrix} y_1 \\ w_1 \end{bmatrix}, \begin{bmatrix} y_2 \\ w_2 \end{bmatrix}, \ldots, \begin{bmatrix} y_m \\ w_m \end{bmatrix}, \tag{8}$$

obtained from the variable $Z = \begin{bmatrix} Y \\ W \end{bmatrix}$. The particular elements of this sample are interpreted as the values y_i taken in measurements from the random variable Y, when the conditioning variable W assumes the respective values w_i. Using the methodology presented in the first part of the section below, on the basis of sample (8) one can calculate $\hat{f}_Z$, i.e. the kernel estimator of density of the random variable Z probability distribution, while the sample

$$w_1, w_2, \ldots, w_m \tag{9}$$

gives $\hat{f}_W$ – the kernel density estimator for the conditioning variable W. The kernel estimator of conditional density of the random variable Y probability distribution for the conditioning value w^*, is defined then – as a natural consequence of formula (3) – as the function $\hat{f}_{Y|W=w^*} : \mathbb{R} \to [0,\infty)$ given by

$$\hat{f}_{Y|W=w^*}(y) = \frac{\hat{f}_X(y,w^*)}{\hat{f}_W(w^*)} \,. \tag{10}$$

If for the estimator $\hat{f}_W$ one uses a kernel with positive values, then the inequality $\hat{f}_W(w^*) > 0$ implied by condition (2) is fulfilled for any $w^* \in \mathbb{R}^{n_W}$.

In the case when for the estimators $\hat{f}_Z$ and $\hat{f}_W$ the product kernel (7) is used, applying in pairs the same positive kernels to the estimator $\hat{f}_W$ and to the last n_W coordinates of the estimator $\hat{f}_Z$, then the expression for the kernel estimator of conditional density becomes particularly helpful for practical applications. Formula (10) can then be specified to the form

$$\hat{f}_{Y|W=w^*}(y) = \tag{11}$$

$$= \frac{\frac{1}{h_0}\sum_{i=1}^{m} \mathcal{K}_0\left(\frac{y-y_i}{h_0}\right)\mathcal{K}_1\left(\frac{w_1^*-w_{i,1}}{h_1}\right)\ldots\mathcal{K}_{n_W}\left(\frac{w_{n_W}^*-w_{i,n_W}}{h_{n_W}}\right)}{\sum_{i=1}^{m}\mathcal{K}_1\left(\frac{w_1^*-w_{i,1}}{h_1}\right)\ldots\mathcal{K}_{n_W}\left(\frac{w_{n_W}^*-w_{i,n_W}}{h_{n_W}}\right)} \,,$$

where $h_0, h_1, \ldots, h_{n_W}$ represent – respectively – smoothing parameters mapped to particular coordinates of the random variable Z (the first h_0 connotes with the describing variable Y, and the rest $h_1,\ldots,h_{n_W}$ with subsequent coordinates of the conditioning variable W), while the coordinates of the vectors w^* and w_i are denoted as

$$w^* = \begin{bmatrix} w_1^* \\ w_2^* \\ \vdots \\ w_{n_W}^* \end{bmatrix} \quad \text{and} \quad w_i = \begin{bmatrix} w_{i,1} \\ w_{i,2} \\ \vdots \\ w_{i,n_W} \end{bmatrix} \quad \text{for } i=1,2,\ldots,m \ . \tag{12}$$

Define the so-called conditioning parameters d_i for $i=1,2,\ldots,m$ by the following formula:

$$d_i = \mathscr{K}_1\left(\frac{w_1^* - w_{i,1}}{h_1}\right)\ldots\mathscr{K}_{n_W}\left(\frac{w_{n_W}^* - w_{i,n_W}}{h_{n_W}}\right) . \tag{13}$$

If one uses the kernels $\mathscr{K}_1,\mathscr{K}_2,\ldots,\ \mathscr{K}_{n_W}$ with positive values, these parameters are also positive. So the kernel estimator of conditional density (11) can be presented in the form

$$\hat{f}_{Y|W=w^*}(y) = \frac{1}{h_0\sum_{i=1}^m d_i}\sum_{i=1}^m d_i\mathscr{K}_0\left(\frac{y-y_i}{h_0}\right) . \tag{14}$$

The value of the parameter d_i characterizes the “distance” of the given conditioning value w^* from w_i – that of the conditioning variable for which the i-th element of the random sample was obtained. Then estimator (14) can be interpreted as the linear combination of kernels mapped to particular elements of a random sample obtained for the variable Y, when the coefficients of this combination characterize how representative these elements are for the given value w^*. The factor $\sum_{i=1}^m d_i$ norms the value of the estimator with the aim of ensuring a unit integral.

With respect to the definition of a conditional quantile (4), its natural estimator is the solution of the following equation with the argument $\hat{q}_{r|w^*}$:

$$\int_{-\infty}^{\hat{q}_{r|w^*}} \hat{f}_{Y|W=w^*}(y)dy = r \ . \tag{15}$$

For the estimator of conditional density $\hat{f}_{Y|W=w^*}$ appearing above, the kernel estimator given in the form (14) will be used. Moreover, as $\mathscr{K}_0$ contained within, one should choose a continuous kernel of positive values, and also so that the function $\mathscr{I}:\mathbb{R}\to\mathbb{R}$ such that $\mathscr{I}(w) = \int_{-\infty}^{w}\mathscr{K}_0(u)du$ can be expressed by a relatively simple analytical formula. Equation (15) can be expressed then equivalently in the following form:

$$\sum_{i=1}^m d_i\mathscr{I}\left(\frac{\hat{q}_{r|w^*} - y_i}{h_0}\right) - r\sum_{i=1}^m d_i = 0 \ . \tag{16}$$

If the left side of the above equation is denoted by L, i.e.

$$L(\hat{q}_{r|w^*}) = \sum_{i=1}^m d_i\mathscr{I}\left(\frac{\hat{q}_{r|w^*} - y_i}{h_0}\right) - r\sum_{i=1}^m d_i \ , \tag{17}$$

then $\lim_{\hat{y}_{w^*}\to-\infty} L(\hat{q}_{r|w^*}) < 0$, $\lim_{\hat{y}_{w^*}\to\infty} L(\hat{q}_{r|w^*}) > 0$, the function L is (strictly) increasing and its derivative is simply expressed by

$$L'(\hat{q}_{r|w^*}) = \frac{1}{h_0}\sum_{i=1}^m d_i\mathscr{K}_0\left(\frac{\hat{q}_{r|w^*} - y_i}{h_0}\right) . \tag{18}$$

In this situation, the solution of equation (16) can be effectively calculated on the basis of Newton’s algorithm [15] as the limit of the sequence $\{\hat{q}_{r|w^*,j}\}_{j=0}^{\infty}$ defined by

$$\hat{q}_{r|w^*,0} = \frac{\sum_{i=1}^m d_i y_i}{\sum_{i=1}^m d_i} \tag{19}$$

$$\hat{q}_{r|w^*,j+1} = \hat{q}_{r|w^*,j} - \frac{L(\hat{q}_{r|w^*,j})}{L'(\hat{q}_{r|w^*,j})} \quad \text{for } j=0,1,\ldots \ , \tag{20}$$

with the functions L and L' being given by dependencies (17)-(18), whereas a stop criterion takes on the form

$$\left|\hat{q}_{r|w^*,j} - \hat{q}_{r|w^*,j-1}\right| \le 0.01\hat{\sigma}_Y \ , \tag{21}$$

while $\hat{\sigma}_Y$ denotes the estimator of the standard deviation of the random variable Y.

3 Algorithm for conditional quantile estimator – summary

The section below gives an explicit algorithm for calculating the conditional quantile estimator according to the concept presented in the previous point. Its basic form will be given, suitable for immediate application without the need for getting into theoretical aspects or specific research.

Consider – as has been the case up to now – the describing random variable Y of values in $\mathbb{R}$ and the conditioning random variable W with values in $\mathbb{R}^{n_W}$. Denote then the required m-element random sample in the form

$$\begin{bmatrix} y_1 \\ w_{1,1} \\ w_{1,2} \\ \vdots \\ w_{1,n_W} \end{bmatrix}, \begin{bmatrix} y_2 \\ w_{2,1} \\ w_{2,2} \\ \vdots \\ w_{2,n_W} \end{bmatrix}, \ldots, \begin{bmatrix} y_m \\ w_{m,1} \\ w_{m,2} \\ \vdots \\ w_{m,n_W} \end{bmatrix} . \tag{22}$$

Their elements represent the values y_i measured from the describing variable Y when the conditioning variable W assumed the values $w_i = \begin{bmatrix} w_{i,1} \\ w_{i,2} \\ \vdots \\ w_{i,n_W} \end{bmatrix}$, respectively for $i = 1,2,\ldots,m$.

The kernel will be assumed here in the Cauchy form:

$$\mathscr{K}_0(u) = \mathscr{K}_1(u) = \ldots = \mathscr{K}_{n_W}(u) = \frac{2}{\pi}\frac{1}{(1+u^2)^2} . \tag{23}$$

In the quantile estimation task investigated here its “heavy tails” behave favorably. The primitive of kernel (23) takes the form:

$$\mathscr{I}(u) = \frac{1}{\pi}\arctan(u) + \frac{u}{\pi(1+u^2)} + \frac{1}{2}\,. \tag{24}$$

The smoothing parameters will now be established. The simple but effective method based on the Gauss distribution [12] – Section 3.1.5; [13] – Section 3.2.1 will be used. First, the estimators of standard deviations are calculated for the variable Y and the particular coordinates of the variable W:

$$\hat{\sigma}_Y = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m} y_i^2 - \frac{1}{m(m-1)}\left(\sum_{i=1}^{m} y_i\right)^2} \tag{25}$$

and

$$\hat{\sigma}_{W_j} = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m} w_{i,j}^2 - \frac{1}{m(m-1)}\left(\sum_{i=1}^{m} w_{i,j}\right)^2} \tag{26}$$

for particular coordinates $j = 1,2,\ldots,n_W$. Then, the smoothing parameters can be obtained from formulas:

$$h_0 = \left(\frac{8\sqrt{\pi}}{3} C_K \frac{1}{m}\right)^{1/5} \hat{\sigma}_Y \tag{27}$$

and

$$h_j = \left(\frac{8\sqrt{\pi}}{3} C_K \frac{1}{m}\right)^{1/5} \hat{\sigma}_{W_j} \tag{28}$$

for particular coordinates $j = 1,2,\ldots,n_W$, while for the Cauchy kernel (23) the constant C_K equals

$$C_K = \frac{5}{4\pi}\,. \tag{29}$$

Now – after fixing the smoothing parameters (27)-(28) – for the specific conditioning value

$$w^* = \begin{bmatrix} w_1^* \\ w_2^* \\ \vdots \\ w_{n_W}^* \end{bmatrix}\,, \tag{30}$$

random sample (22) and kernel form (23), one can calculate from formula (13) the conditioning parameters $d_1, d_2, \ldots, d_m$.

At present all quantities necessary for the use of iterative algorithm (19)-(21), from which the value of the conditional quantile estimator is obtained, have been defined. Thus, in formula (22) a value is given for $\hat{\sigma}_Y$, completing stop condition (25). The functions L and L', contained in formulas (19)-(20), are given as (17) and (18), and the functions $\mathscr{K}_0$ and $\mathscr{I}$ by (23) and (24). The smoothing $h_0, h_1, \ldots, h_{n_W}$ and conditioning $d_1, d_2, \ldots, d_m$ parameters are obtained above, whereas the numbers $w_{i,j}$ and y_i can be found in sample (22), and the quantile order r is arbitrarily assumed.

It is worth underlining in particular that the algorithm above is of linear computational complexity both with respect to the sample size m and the conditioning vector dimension n_W, i.e. $O(m\,n_W)$. The convergence of Newton's algorithm is quadratic; its application in the task investigated here most often requires 5-8 iterations. Both of these features result in relatively short computational times, for large sizes of sample (5) too.

4 Numerical verification

The correct functioning and positive properties of the algorithm presented in this paper were confirmed with detailed numerical verification.

Assume for transparency of the results' interpretation that $n_W = 1$, and let the tested random variable $X = \begin{bmatrix} Y \\ W \end{bmatrix}$ have a distribution being the sum of two Gauss factors with expected values, covariance matrixes and shares, respectively,

$$E_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad Cov_1 = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix} \qquad 60\% \tag{31}$$

$$E_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad Cov_2 = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix} \qquad 40\%\,. \tag{32}$$

As can be seen, the describing Y and conditioning W variables are positively correlated. The former is of asymmetrical bimodal distribution and the latter – standard Gauss. The expected values of the variables Y and W as well as their standard deviations are

$$E_Y = -0.2,\ \sigma_Y = \sqrt{2} \tag{33}$$

$$E_W = 0,\ \sigma_W = 1\,. \tag{34}$$

The results acquired with the algorithm from Section 3 for $w^* = 0$, $w^* = 1$, $w^* = 2$, so for the modal value of a conditioning random variable as well as at the first and second standard deviation, are presented in Tab. 1, 2 and 3, respectively. Each of their cells shows the obtained values of the estimator, calculated on the basis of 100 tests and recorded in the classic formula: “mean value $\pm$ standard deviation”. The symbol ∞ denotes there the analytically achieved theoretical value.

In Tab. 1-3 the results have been shaded where the mean estimation error is greater than 10% of the standard deviation of describing value σ_Y, i.e. 0.141 (see formula (33)) or where the standard deviation of the estimation error is greater than 20% of σ_Y, i.e. 0.282. One can note that for modal value $w^* = 0$ satisfactory results are obtained for samples of sizes 20-50, and for extreme orders $r = 0.05$ or $r = 0.95$ from 200. For the first standard deviation $w^* = 1$ this occurs respectively for sizes 100-200 and 200-1000. And finally for the second standard deviation $w^* = 2$, generally a sample of 200-1000 is required, while extreme orders can necessitate even 1000-2000. Taking into account the conditional character of the considered problem and the assumed quality criteria, the need for such size seems reasonable in practice. It is worth underlining that the estimation of quantile of the extreme orders $r = 0.05$ or $r = 0.95$, or also for the second standard deviation $w^* = 2$, is a task of no small difficulty.

The results obtained in the three enclosed tables, related to the subsequent conditioning values $w^* = 0$, $w^* = 1$, $w^* = 2$, will now be compared. The first of them corresponds to the modal value of the conditioning variable W (see formulas (31)-(32)). Because the variables Y and W are correlated positively, for $w^* = 1$ the estimators values become bigger than for $w^* = 0$. When $w^* = 2$ these values increase even more. In the cases $w^* = -1$, $w^* = -2$ the estimators values are respectively smaller than for $w^* = 0$, which because of symmetry have been omitted from Tab. 1-3. It should be underlined that the dependence of the estimator value on the given conditioning value, considered in this paragraph, constitutes the essence of the conditional approach investigated in this paper. Note that in applicational tasks indicating the concrete conditioning value w^* can significantly improve the accuracy of the model of reality under investigation.

Table 1: Values of conditional quantile estimator for the conditioning value $w^* = 0$ (for 1000 samples with notion "mean value $\pm$ standard deviation").

m	$r=0.05$	$r=0.1$	$r=0.3$	$r=0.5$	$r=0.7$	$r=0.9$	$r=0.95$
20	-2.211 ± 0.391	-1.856 ± 0.393	-1.000 ± 0.376	-0.288 ± 0.434	0.553 ± 0.478	1.557 ± 0.442	1.939 ± 0.434
50	-2.187 ± 0.291	-1.834 ± 0.264	-1.000 ± 0.254	-0.297 ± 0.301	0.554 ± 0.332	1.566 ± 0.306	1.958 ± 0.330
100	-2.168 ± 0.226	-1.810 ± 0.201	-1.006 ± 0.191	-0.309 ± 0.234	0.552 ± 0.251	1.565 ± 0.235	1.964 ± 0.252
200	-2.127 ± 0.168	-1.782 ± 0.150	-1.006 ± 0.143	-0.322 ± 0.182	0.553 ± 0.198	1.552 ± 0.183	1.938 ± 0.198
500	-2.093 ± 0.114	-1.760 ± 0.101	-1.004 ± 0.101	-0.329 ± 0.131	0.556 ± 0.146	1.532 ± 0.129	1.907 ± 0.139
1000	-2.070 ± 0.086	-1.747 ± 0.077	-1.006 ± 0.075	-0.337 ± 0.099	0.555 ± 0.112	1.524 ± 0.094	1.892 ± 0.102
2000	-2.053 ± 0.069	-1.734 ± 0.060	-1.004 ± 0.058	-0.339 ± 0.076	0.558 ± 0.084	1.515 ± 0.071	1.879 ± 0.080
5000	-2.036 ± 0.046	-1.723 ± 0.041	-1.004 ± 0.039	-0.345 ± 0.053	0.562 ± 0.059	1.508 ± 0.050	1.864 ± 0.055
10000	-2.024 ± 0.035	-1.715 ± 0.031	-1.004 ± 0.029	-0.349 ± 0.042	0.562 ± 0.046	1.502 ± 0.036	1.856 ± 0.040
∞	-1.988 ± 0.000	$-1.691 \pm .0000$	-1.003 ± 0.000	-0.361 ± 0.000	0.565 ± 0.000	1.483 ± 0.000	1.822 ± 0.000

Table 2: Values of conditional quantile estimator for the conditioning value $w^* = 1$ (for 1000 samples with notion "mean value $\pm$ standard deviation").

m	$r=0.05$	$r=0.1$	$r=0.3$	$r=0.5$	$r=0.7$	$r=0.9$	$r=0.95$
20	-1.648 ± 0.404	-1.270 ± 0.425	-1.438 ± 0.474	0.254 ± 0.547	1.046 ± 0.584	1.997 ± 0.571	2.365 ± 0.546
50	-1.587 ± 0.315	-1.228 ± 0.304	-0.404 ± 0.325	0.297 ± 0.402	1.119 ± 0.430	2.113 ± 0.416	2.482 ± 0.420
100	-1.554 ± 0.266	-1.180 ± 0.238	-0.396 ± 0.244	0.303 ± 0.312	1.146 ± 0.333	2.164 ± 0.312	2.560 ± 0.328
200	-1.512 ± 0.202	-1.131 ± 0.178	-0.376 ± 0.184	0.319 ± 0.246	1.174 ± 0.266	2.169 ± 0.236	2.555 ± 0.250
500	-1.455 ± 0.147	-1.120 ± 0.129	-0.360 ± 0.128	0.317 ± 0.169	1.192 ± 0.183	2.174 ± 0.162	2.545 ± 0.177
1000	-1.419 ± 0.109	-1.094 ± 0.096	-0.350 ± 0.096	0.320 ± 0.129	1.207 ± 0.138	2.172 ± 0.117	2.537 ± 0.130
2000	-1.386 ± 0.084	-1.068 ± 0.073	-0.334 ± 0.072	0.329 ± 0.099	1.223 ± 0.107	2.180 ± 0.090	2.540 ± 0.100
5000	-1.358 ± 0.060	-1.046 ± 0.053	-0.326 ± 0.050	0.334 ± 0.069	1.237 ± 0.078	2.184 ± 0.062	2.537 ± 0.070
10000	-1.343 ± 0.046	-1.034 ± 0.041	-0.320 ± 0.039	0.334 ± 0.055	1.243 ± 0.060	2.182 ± 0.048	2.533 ± 0.052
∞	-1.288 ± 0.000	-0.991 ± 0.000	-0.303 ± 0.000	0.339 ± 0.000	1.265 ± 0.000	2.183 ± 0.000	2.522 ± 0.000

In practice it may occur that the amount of data acquired for the purposefully used conditioning value w^*, or even in its neighborhood, is completely insufficient for reliable statistical inference. Table 4 shows the results corresponding to Tab. 1, where the random sample contained no elements for the conditioning variable from the interval $[-0.25, 0.25]$, so in the range $w^* \pm 0.25\sigma_W$, therefore with the width of half the standard deviation. Comparing Tab. 1 and 4 one can conclude that – apart from extreme orders of quantile – the required sample size practically does not increase. The above is worth underlining from a practical point of view, and arises from the general averaging features of kernel estimators.

Finally, it is worth noting that in any case shown in Tab. 1-4, as the sample size increased, the obtained parameter value converged to the theoretical, 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 wished, although this requires the assurance of a sufficient random sample size. In practice, therefore, the necessity of the right compromise between these quantities is called for. This empirically proves the strong consistency of the estimator worked out. The formal proof for the unconditional case was presented in the paper [5] – for the conditional case considered here, the proof is analogous however but technically and formally more complicated. Some basic theoretical investigations can also be found in the article [16].

Table 3: Values of conditional quantile estimator for the conditioning value $w^* = 2$ (for 1000 samples with notion "mean value $\pm$ standard deviation").

m	$r=0.05$	$r=0.1$	$r=0.3$	$r=0.5$	$r=0.7$	$r=0.9$	$r=0.95$
20	-1.171 ± 0.523	-0.733 ± 0.584	0.196 ± 0.758	0.854 ± 0.863	1.502 ± 0.914	2.249 ± 0.825	2.568 ± 0.779
50	-1.010 ± 0.373	-0.601 ± 0.434	0.273 ± 0.590	0.939 ± 0.700	1.653 ± 0.756	2.494 ± 0.698	2.812 ± 0.650
100	-0.919 ± 0.333	-0.546 ± 0.356	0.273 ± 0.435	0.975 ± 0.535	1.746 ± 0.597	2.672 ± 0.554	2.983 ± 0.525
200	-0.863 ± 0.280	-0.518 ± 0.295	0.280 ± 0.279	0.974 ± 0.442	1.775 ± 0.481	2.746 ± 0.465	3.064 ± 0.457
500	-0.808 ± 0.256	-0.475 ± 0.237	0.300 ± 0.247	0.992 ± 0.335	1.831 ± 0.369	2.806 ± 0.337	3.169 ± 0.369
1000	-0.767 ± 0.203	-0.430 ± 0.180	0.318 ± 0.190	0.996 ± 0.263	1.855 ± 0.274	2.826 ± 0.251	3.186 ± 0.272
2000	-0.726 ± 0.159	-0.412 ± 0.143	0.332 ± 0.143	1.003 ± 0.202	1.880 ± 0.232	2.841 ± 0.188	3.201 ± 0.208
5000	-0.701 ± 0.117	-0.383 ± 0.103	0.347 ± 0.106	1.011 ± 0.152	1.906 ± 0.170	2.852 ± 0.130	3.206 ± 0.145
10000	-0.675 ± 0.090	-0.363 ± 0.079	0.360 ± 0.080	1.018 ± 0.117	1.917 ± 0.131	2.861 ± 0.100	3.215 ± 0.111
∞	-0.588 ± 0.000	-0.291 ± 00.00	0.397 ± 0.000	1.039 ± 0.000	1.965 ± 0.000	2.883 ± 0.000	3.222 ± 0.000

The above results were compared to those obtained by other methodologies, in particular quantile regression [8] and neural networks [11]. When using the first of these, the calculations produced better effects than with the procedure investigated here. This was especially true for mean – in respect to numerous random samples – values of examined estimators, while their variances were similar. Tables 1-4 show the negative influence only of the asymptotic kind of unbiased smoothing parameter h value. It should however be noted that the practical implementation, analysis, or possible specialized modifications the quantile regression method are significantly more difficult, in particular when conditioning variables are not only continuous. And finally, the results obtained using the procedure presented here were comparable to those of the neural networks method, as – again – the algorithm from this paper is simpler and more convenient for implementation, analysis and possible inclusion of conditioning variables other than continuous (Section 5).

Table 4: Values of conditional quantile estimator for the conditioning value $w^* = 0$ without data from the interval $[-0.25, 0.25]$ (for 1000 samples with notion "mean value $\pm$ standard deviation").

m	$r=0.05$	$r=0.1$	$r=0.3$	$r=0.5$	$r=0.7$	$r=0.9$	$r=0.95$
20	-2.291 ± 0.399	-1.917 ± 0.409	-1.032 ± 0.387	-0.309 ± 0.437	0.520 ± 0.473	1.549 ± 0.452	1.946 ± 0.444
50	-2.264 ± 0.292	-1.888 ± 0.268	-1.021 ± 0.265	-0.303 ± 0.303	0.542 ± 0.330	1.565 ± 0.309	1.972 ± 0.327
100	-2.260 ± 0.234	-1.886 ± 0.217	-1.033 ± 0.203	-0.331 ± 0.222	0.510 ± 0.247	1.567 ± 0.240	1.978 ± 0.251
200	-2.210 ± 0.177	-1.831 ± 0.163	-1.011 ± 0.158	-0.297 ± 0.181	0.548 ± 0.199	1.585 ± 0.193	1.961 ± 0.204
500	-2.168 ± 0.127	-1.812 ± 0.120	-1.004 ± 0.111	-0.294 ± 0.126	0.560 ± 0.141	1.578 ± 0.133	1.962 ± 0.140
1000	-2.160 ± 0.099	-1.809 ± 0.093	-1.007 ± 0.088	-0.301 ± 0.091	0.556 ± 0.107	1.566 ± 0.106	1.956 ± 0.111
2000	-2.126 ± 0.085	-1.793 ± 0.081	-1.003 ± 0.073	-0.308 ± 0.071	0.554 ± 0.083	1.559 ± 0.084	1.947 ± 0.090
5000	-2.117 ± 0.076	-1.782 ± 0.072	-1.001 ± 0.064	-0.314 ± 0.054	0.551 ± 0.065	1.549 ± 0.075	1.930 ± 0.077
10000	-2.102 ± 0.074	-1.770 ± 0.074	-1.000 ± 0.062	-0.316 ± 0.045	0.554 ± 0.053	1.544 ± 0.073	1.922 ± 0.075
∞	-1.988 ± 0.000	-1.691 ± 0.000	-1.003 ± 0.000	-0.361 ± 0.000	0.565 ± 0.000	1.483 ± 0.000	$1.822\pm.0000$

5 Final remarks and summary

The procedure presented in this paper has been given in its basic form, easier to implement and calculationally more convenient. A clearer interpretation means it is possible to make individual modifications and generalizations, which may be useful in particular atypical tasks [17].

Thus, the quality of kernel estimator (6) can be improved by applying the procedure for modifying the smoothing parameter [12] – Section 3.1.6; [14] – Section 5.3.1. This consists of the appropriate individualization of the above parameter for particular kernels, by suitably thinning those in the areas of dense sample (5) elements and additionally smoothing those where elements are sparse, especially "tails" of distribution. The intensity of the above procedure can be decided by appropriately fixing the value of a given parameter. In the conditional quantile estimation method presented here, this procedure can prove particularly useful for extreme values of the quantile order (i.e. close to 0 and 1), for less numerous random samples, although it has the computational complexity $O(m^2 n_W)$.

In the case of complicated, irregular, multimodal distributions of the describing and conditioning variables, one can apply a more complex procedure for calculating the smoothing parameter value than those defined by formulas (27)-(28). Here, particularly worth recommending is the plug-in method [12] – Section 3.1.5; [13] – Section 3.6.1 used separately for the variable Y and particular coordinates of the variable W. Its computational complexity, however, is $O(m^2 n_W)$. For the Cauchy kernel (23) the constants required there are $\int_{-\infty}^{\infty} u^2 \mathscr{K}_i(u)du = 1$ and $\int_{-\infty}^{\infty} \mathscr{K}_i(u)^2 du = 5/4\pi$ for $i = 0, 1, \ldots, n_W$.

Newton's method (19)-(21) can also be applied in numerous mutations available in literature, in particular those lessening the number of iterations as well as extending the convergence area. A broad review of concepts available on this subject can be found in the monographs [18,19]. It must, however, be underlined that in the research undertaken, the problem of no convergence for sizes of sample (5) guaranteeing a satisfactory estimation quality (see previous section) did not arise, largely due to the choice of the starting point in form (19).

As mentioned in Section 2, the kernel K can assume many form, which in practice has a slight influence on the estimator's quality. A survey of the most commonly used kernels can be found in the publications [12] – Section 3.1.3; [13] – Sections 2.7 and 4.5. It is worth remembering the exponential kernel:

$$\mathscr{K}_0(u) = \mathscr{K}_1(u) = \ldots = \mathscr{K}_{n_W}(u) = \frac{e^{-u}}{(1+e^{-u})^2} , \qquad (35)$$

due to its particularly convenient primitive function

$$\mathscr{I} = \frac{1}{1+e^{-u}} . \qquad (36)$$

In this case, the constant C_K existing in formulas (27)-(28) amounts to $C_K = 3/(2\pi^4)$.

The kernel estimator's definition (6) has been presented in Section 2 for continuous (real) random variables. Similarly, one can construct kernel estimators for binary, discrete and categorized (including ordered) variables, as well as any of their compositions. As an example the case for binary and basic type of continuous coordinates combined, will be described below.

Let then V denote an n_V-dimensional binary random variable with the density distribution $f_V : \{0,1\}^{n_V} \to [0,\infty)$, mapping for each n_V-dimensional vector with binary values, a probability of its occurrence. Its kernel estimator can be calculated based on the random sample values

$$v_1, v_2, \ldots, v_m \qquad (37)$$

and is defined by the formula

$$\hat{f}_V(v) = \frac{1}{m} \sum_{i=1}^{m} B(v, v_i) , \qquad (38)$$

while the function $B : \{0,1\}^{n_V} \times \{0,1\}^{n_V} \to [0,1]$ is given by

$$B(v, v_i) = \delta^{k-d(v,v_i)} (1-\delta)^{d(v,v_i)} , \qquad (39)$$

with $0.5 \leq \delta \leq 1$, whereas the function $d : \{0,1\}^{n_V} \times \{0,1\}^{n_V} \to \mathbb{N}$ is

$$d(v, v_i) = (v - v_i)^T (v - v_i) . \qquad (40)$$

The value of the function d equals therefore the number of these coordinates of the vectors v and v_i, for which those vectors differ, so the quantity $n_v - d(v, v_i)$, denotes the number of coordinates whose vectors are identical and represent "similarity" of binary vectors. The function B takes the role played in definition (6) by kernel K and is called a binary kernel, while the constant δ has analogous meaning as the previous parameter h and is therefore referred to as a binary smoothing parameter. In practice its value is fixed using optimizing criterions.

Finally, consider the $(n_X + n_V)$-dimensional random variable $Z = \begin{bmatrix} X \\ V \end{bmatrix}$ as a composition of the n_X-dimensional (continuous) variable X, considered previously in this paper, and the above described n_V-dimensional binary variable V, maintaining the same assignations as before. If the kernel estimator $\hat{f}_Z : \mathbb{R}^{n_X} \times \{0,1\}^{n_V} \to [0,\infty)$ of the random variable Z is calculated on the basis of the m-element random sample

$$\begin{bmatrix} x_1 \\ v_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ v_2 \end{bmatrix}, \ldots, \begin{bmatrix} x_m \\ v_m \end{bmatrix} , \qquad (41)$$

then it is defined as

$$\hat{f}_Z\left(\begin{bmatrix} x \\ v \end{bmatrix}\right) = \frac{1}{mh^n} \sum_{i=1}^{m} K\left(\frac{x - x_i}{h}\right) B(v, v_i) . \qquad (42)$$

So if in the procedure for estimating the conditional quantile value investigated here, the conditioning vector consists of the continuous and binary coordinates, then the above concept can be easily taken into account through multiplying by the binary kernel B the factors of the sums of the nominator and denominator in formula (11). This results in its entry to the definition of conditioning parameters (13) with no other changes in the described algorithm, which should be clearly noted.

The literature concerning the construction of kernel estimators for binary, discrete and categorical (ordered as well) coordinates as well as their compositions (especially with continuous variables) is quite broad and varied. For the first case, it is worth quoting the classic monographs [12] – Section 3.1.8; [14] – Section 6.1.4 as well as the classic paper [20], and for the second [21,22]. Issues connected with categorical variables can be found in the publications [23,24,25]. After introducing a binary, discrete and/or categorized variable to the algorithm worked out here, it undergoes practically no changes – apart from technical ones resulting from calculational differences – as it is briefly described above for binary variables. This property particularly should be underlined considering the modern data analysis tasks, which more and more often take advantage of the many different configurations for particular types of attributes.

In summary, this paper presents a convenient ready-to-use procedure for calculating the kernel conditional quantile estimator value. In particular Section 3 gives a complete algorithm for use without getting deeper into the mathematical aspects or laborious specific research. Thanks to the use of a nonparametric kernel estimators method, this procedure is distribution free, and due to its averaging feature it becomes robust to lack of data from the neighborhood of the fixed conditioning value. Its computational complexity is linear both in respect to sample size as well as dimension of a conditioning vector. A clear and near-intuitive interpretation allows the creation of individual generalizations and modifications with the aim of altering them to specific unusual practical applications. In particular it is possible to consider continuous, binary, discrete, categorized, and any their compositions, as conditioning variables.

The algorithm's simplicity and ability to include conditioning variables of different types are the main, noteworthy positives of the presented method. In the basic case for continuous variables as well as the ability to use a ready library or calculation bundle in the concrete application, it is necessary however to pay tribute to the classic method of quantile regression as being more precise.

References

[1] P. Kulczycki and R. Wisniewski, Fuzzy Controller for a System with Uncertain Load, Fuzzy Sets and Systems **131**, 185-195 (2002).

[2] C.E. David and S.M. Steinberg, Quantile estimation; in Encyclopedia of Statistical Sciences **7**, Kotz S., Johnson N.L. (Eds), 408-412, Wiley, New York (1986).

[3] T. Dielman, C. Lowry, and R. Pfaffenberger, A comparison of quantile estimators, Communications in Statistics – Simulation and Computation **23**, 355-371 (1994).

[4] R.S. Parrish, Comparison of Quantile Estimators in Normal Sampling, Biometrics **46**, 247-257 (1990).

[5] P. Kulczycki and A.L. Dawidowicz, Kernel Estimator of Quantile, Universitatis Jagiellonicae Acta Mathematica **37**, 101-112 (1999).

[6] S.J. Sheather and J.S. Marron, Kernel quantile estimators. Journal of the American Statistical Association **85**, 410-416 (1990).

[7] R. Koenker, Quantile Regression, Cambridge University Press, Cambridge (2005).

[8] R. Koenker and G. Bassett, Regression quantiles, Econometrica **46**, 33-50 (1978).

[9] I. Takeuchi, Q.V. Le, T. Sears, and A.J. Smola, Nonparametric quantile estimation, Journal of Machine Learning Research **7**, 1231-1264 (2006).

[10] H. White, Nonparametric Estimation of Conditional Quantiles Using Neural Networks, Computing Science and Statistics **24**, 190-199 (1992).

[11] P. Wawrzynski and A. Pacut, Modeling of distributions with neutral approximation of conditional quantiles, 2nd IASTED International Conference on Artificial Intelligence and Applications, Malaga, Spain, 9-12 September 2002, 539-543 (2002).

[12] P. Kulczycki, Estymatory jądrowe w analizie systemowej, WNT, Warszawa (2005).

[13] M.P. Wand and M.C. Jones, Kernel Smoothing, Chapman and Hall, London (1995).

[14] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, London (1986).

[15] D. Kincaid and W. Cheney, Numerical Analysis, Brooks/Cole, Pacific Grove (2002).

[16] M. Samanta, Non-parametric estimation of conditional quantiles, Statistics and Probability Letters **7**, 407-412 (1989).

[17] P. Kulczycki, O. Hryniewicz, and J. Kacprzyk, Techniki informacyjne w badaniach systemowych, WNT, Warszawa (2007).

[18] P. Deuflhard, Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms, Springer, Berlin (2004).

[19] C.T. Kelley, Solving Nonlinear Equations with Newton's Method, SIAM, Philadelphia (2003).

[20] J. Aitchison and C. Aitken, Multivariate Binary Discrimination by the Kernel Method, Biometrika **63**, 413-420 (1976).

[21] I.A. Ahmad and P.B. Cerrito, Nonparametric estimation of joint discrete-continuous probability densities with applications, Journal of Statistical Planning and Inference **41**, 349-364 (1994).

[22] M. Wang and J. van Ryzin, A class of smooth estimators for discrete distributions, Biometrica **68**, 301-309 (1981).

[23] J. Gaosheng, L. Rui, and L. Zhongwen, Nonparametric Estimation of Multivariate CDF with Categorical and Continuous Data, Advances in Econometrics **25**, 291-318 (2009).

[24] Q. Li and J.S. Racine, Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data, Journal of Business and Economic Statistics **26**, 423-434 (2008).

[25] D. Ouyang, Q. Li, and J. Racine, Cross-validation and the estimation of probability distributions with categorical data, Journal of Nonparametric Statistics **18**, 69-100 (2006).

P. Kulczycki is a professor at the Systems Research Institute of the Polish Academy of Sciences, where he is the head of the Centre of Information Technology for Data Analysis Methods, and at the AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, where he is the founder and head of the Division of Information Technology and Biometrics. He has also held the position of visiting professor at the Aalborg University, as well as a professor at the Cracow University of Technology, Faculty of Electrical and Computer Engineering, where he was the founder and head of the Department of Automatic Control and Information Technology. A graduate of electrical engineering at the AGH University of Science and Technology, and applied mathematics at the Jagiellonian University, his field of scientific activity to date is the applicational aspects of information technology and data mining and analysis, mostly connected with the use of modern statistical methods and fuzzy logic in diverse issues of contemporary systems research and control engineering.

M. Charytanowicz is a research scientist at the Systems Research Institute of the Polish Academy of Sciences, as well as at the John Paul II Catholic University of Lublin in the Institute of Mathematics and Computer Science. She obtained her M.Sc. degree in mathematics from the Maria Curie-Sklodowska University in Lublin, and her Ph.D. from the Polish Academy of Sciences, Systems Research Institute, in the area of computer science. Her research interests are programming methods, data analysis and its applications for medicine.

A. L. Dawidowicz is a professor of the Institute of Mathematics of the Jagiellonian University, where he previously obtained M.Sc., Ph.D., and D.Sc. (habilitation) degrees in mathematics. His scientific specialty is mathematical statistics, probability theory and the applications of mathematics in biology. He has been a Cracow city councilor, the president of the Polish Tatra Mountains Society, a member of the Polish National Group of the International Society of Clinical Biostatistics, and currently he is the president of the Cracow branch of the Polish Mathematical Society.