The approximation of grain composition curves by non-parametric statistical methods

T. Niedoba, T. Tumidajski
University of Science and Technology, Krakow, Poland

ABSTRACT: Approximation by classical methods is not always efficient, particularly when the considered random variables distribution functions are complex ones or mixtures of distribution functions. We face this situation in case of approximation of grained materials features distribution functions, mainly grain composition curves. To estimate these curves the non-parametric statistical methods were proposed, among which orthogonal Fourier series method and kernel methods based on Epanechnikov and Gauss kernels were taken into consideration. The effects of resulted curves fitting to empirical distributions are better in comparison with classical methods.

1 INTRODUCTION

Approximations of grained materials features distribution functions are being used in purpose of their application to determine certain numerical characteristics (mean, variance, asymmetry factor) and also to further their application to models of process course and its results simulation. The basic law of processing is the law of conservation of mass, known in processing as balance equation of the following form:

$$g(x) = \int_{w_{\min}}^{w_{\max}} p(w,x) f(w) dw \qquad (1)$$

where $g(x)$ is density function of feature X in product; $p(w,x)$ - function determining probability of transfer feed grains of feature values w to elemental range of feature value; $f(w)$ – density function of feature W in feed. Assuming that the function $p(w,x)$ is known and conducting the precise approximation of function $f(w)$ the optional forecast (simulation) of value $g(x)$ may be given for any value of x.

The classical approximation is often not efficient enough to determine the correct form of distribution function, what is important to estimate the values of statistical parameters. It is visible very clearly in the example of grained materials. It is usually hard to determine the exact statistical distribution function of grain composition. For example, the feed, which goes to processing industrial plant, is composed of many fractions of different characteristics. In effect it gives the mixture of distribution functions, which formulae may be difficult to determine precisely, even if the distribution functions of components are known. Furthermore it may be a completely different kind of distribution function than the original ones. In this case, the approximations by the classical distribution functions, like Weibull, Gaudin-Schuhmann-Andreyev, log-norm give usually a large value of residual deviation, what proves that they do not sufficiently well describe the reality. This may cause errors in estimating certain factors of researched material.

Because of this situation, the methodology of non-parametric statistical researches is presented in the paper. This kind of statistics is a new way of approximation, which allow creating the real distribution function for the researched material and minimize the value of residual deviation. Among non-parametric methods, the method of approximation by Fourier orthogonal series and kernel approximation were taken into consideration. From the kernel methodology, the authors applied the approximation by Epanechnikov and Gauss kernels. In the base of results of materials comminution it occurred that the resulting non-parametric distribution functions are far much better fitted to the data than the classical ones. This methodology may be applied not only to determine the grain composition, but also element composition, which may be even more important from the financial point of view. The summarized comparison of the non-parametric and classical methods is presented at the end of the paper.

2 ESTIMATION OF RANDOM VARIABLES DENSITY FUNCTION BY KERNEL METHODS

One of the main methods of modern non-parametric theory of distribution functions estimation is the method called kernel method. This method was introduced by Rosenblatt [Rosenblatt, 1956], who conducted the naive estimator analysis (derived by Fix and Hodges in 1951) and introduced the term of kernel estimation. Researches over these method for one-dimensional random variable were conducted [Parzen, 1962; Stone, 1984; Watson, 1963; Whitle, 1958].

If we assume as the density function estimator so-called naive estimator given by the equation:

$$\hat{f}(x)=\frac{1}{nh}\sum_{i=1}^{n} w\left(\frac{x-x_i}{h}\right) \tag{2}$$

where: $x_1,...,x_n$ is realization of the sample $X_1,...,X_n$ and $w(x)=\begin{cases}\frac{1}{2} if\ |x|<1\\ 0\ in\ other\ cases\end{cases}$

we obtain discrete function (similarly as in case of histogram). However, it is easy to avoid this disadvantage by replacing the weighted function $w(x)$ by smooth function $K(x)$, called kernel function or simply kernel, which have to fulfill the condition [Efromovich, 1999]

$$\int_{-\infty}^{\infty} K(x)dx=1 \tag{3}$$

Then, for the certain simple random sample $X_1,..., X_n$, kernel estimator of density function is defined as

$$\hat{f}(x)=\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \tag{4}$$

where $x_1,..., x_n$ is realization of random sample $X_1,..., X_n$ and $h>0$ is the parameter called band width or smoothing parameter.

Let us notice that from the equations (3) and (4) results the condition that

$$\int_{-\infty}^{\infty} \hat{f}(x)dx=1 \tag{5}$$

Furthermore $\hat{f}$ fulfill the same conditions of differentiability as were overlaid on kernel.

The measure of fitting of estimator $\hat{f}(x)$ to the real distribution density function $f(x)$ is so called mean squared error (MSE) [Silverman 1986]

$$MSE\left(\hat{f}\right)=E_f\left(\hat{f}(x)-f(x)\right)^2 \tag{6}$$

The question occurs, if there is a possibility of choosing such optimal kernel and optimal band width h for which the integrated mean squared error is the lowest for any estimated distribution density function.

The answer is positive [Silverman 1986] if the estimated density function fulfills some regularity conditions (is twice differentiable and second derivative fulfill the Lipschitz condition of any grade). It was proven that the only one, asymptotically optimal kernel is Epanechnikov kernel given by the following formulae:

$$K(x)=\begin{cases}\frac{3}{4\sqrt{5}b}\left(1-\frac{1}{5b^2}x^2\right) & for\ |x|\le\sqrt{5}\\ 0 & for\ |x|>\sqrt{5}\end{cases} \tag{7}$$

where b – scale parameter $\left(\int_{-\infty}^{\infty} t^2K(t)dt=b^2\right)$

Other kernel is so-called Gauss kernel, given by the formulae:

$$K(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right), x\in R \tag{8}$$

The selection of optimal band width will be done as following [Gajek and Kałuszka 2000]:
If we apply Epanechnikov kernel:

$$h_{opt}=1{,}056\sigma n^{-\frac{1}{5}} \tag{9}$$

In case of Gauss kernel:

$$h_{opt}=1{,}066\sigma n^{-\frac{1}{5}} \tag{10}$$

where n – sample quantity; σ - standard deviation.
Practically σ is being changed by s calculated in base of random sample

$$s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{11}$$

3 ESTIMATION OF RANDOM VARIABLES DISTRIBUTION FUNCTIONS BY ORTHOGONAL FOURIER SERIES METHOD

One of other known methods of modern non-parametric theory of one-dimensional variable distribution function estimation is approximation by using Fourier series [Chentsov, 1962; Efron and Tibshirani, 1996; Schwartz, 1967; Wahba, 1981; Watson, 1969]. This method was introduced by Chentsov in 1962 to approximation of density function $f(x)$ in base of random sample.

Later [Chentsov, 1980; Efron and Tibshirani, 1996], the estimation of function $g(x)=\log(f(x))$ by this method was recommended. Majority of results of these (and other) works may be found in books [Devroye and Györfi, 1985; Efromovich, 1999; Hart, 1997; Tompson and Tapia, 1990; Silverman, 1986].

Let assume the sequence $\{\varphi_n(x)\}$ to be functional orthonormal sequence in range [0,1], what means that

$$\int_0^1 \varphi_n{}^2(x)dx=1 \text{ i } \int_0^1 \varphi_j(x)\varphi_k(x)dx=0\ for\ j\neq k \tag{12}$$

If random variable has values from the range [a, b] then, by linear transformation, we change this range into the range [0, 1], by using the following equation:

$$X'=\frac{1}{b-a}X-\frac{a}{b-a} \tag{13}$$

Let $f(x)$ be the density function of researched distribution. Then, by using Fourier series theory, we may present $f(x)$ by infinite sum in the form

$$f(x)=\sum_{j=0}^{\infty} Q_j \varphi_j(x), \tag{14}$$

where Q_j are the Fourier coefficients determined by the formulae:

$$Q_j=\int_0^1 f(x)\varphi_j(x)dx,\ j=1,2,\mathrm{K} \tag{15}$$

whereas

$$\varphi_j(x)=\begin{cases} 1\ dla\ j=0 \\ \sqrt{2}\cos \pi jx\ dla\ j=1,2,... \end{cases} \tag{16}$$

The estimator of researched distribution density function is partial sum of Fourier series

$$f_J(x)=\sum_{j=0}^{J} Q_j \varphi_j(x) \tag{17}$$

where J is so-called cut parameter [Tarter and Lock, 1990].

The statistical problem of density function f estimation rely on determination the estimators for:

a) Fourier coefficients;
b) cut parameter J;
c) weight coefficients ω_j (if it is requested).

Let X to be researched random variable and $x_1,...x_n$ n-element realization of random sample of this variable. We assume that the random variable X accepts the values from the range [0, 1].

As the estimator of Fourier coefficients Q_j the following formulae is assumed

$$\hat{Q}_j=\frac{1}{n}\sum_{i=1}^{n}\varphi_j(x_i) \tag{18}$$

This estimator is unbiased and heavy consistent one [Efromovich, 1999].

The next step is to choose the cut parameter J. As the criterion of selection, the MSE error should be used (6), which measures the global fitting of estimator. For cut parameter, the following estimator is accepted:

$$\hat{J}=\arg\min_{0\le J\le J_n}\sum_{j=0}^{J}\left(\frac{2\hat{Q}_0}{n}-\hat{Q}_j{}^2\right) \tag{19}$$

for which $\arg\min_{0\le J\le J_n}(a_s)$ accepts the values equal to s coefficient for which $(a_0,...a_s)$ is the lowest element and J_n is the integral part of number $(4+0{,}5\ln n)$ [Efromovich, 1999].

4 ESTIMATION OF GRAIN COMPOSITION CURVES BY NON-PARAMETRIC STATISTICAL METHODS

The grain composition of 10- stage porphyry crushing products were selected to the estimation. The results are presented in table 1.

4.1 *Kernel methods*

One of the most important issues in mineral processing is estimation of crushing products grain

Table 1. The juxtaposition of ten various time porphyry crushing results .

Stage										
I	Class [μm]	0-5	5-6,3	6,3-8	8-10	10-12,5	12,5-16	16-18	18-20	20-25
	d_i	2,5	5,65	7,15	9	11,25	14,25	17	19	22,5
	n_i	10,05	1,94	2,52	6,01	13,16	26,92	9,42	14,4	15,58
II	Class [μm]	0-4	4-5	5-6,3	6,3-8	8-10	10-12,5	12,5-16	16-18	18-20
	d_i	2	4,5	5,65	7,15	9	11,25	14,25	17	19
	n_i	14,57	6,65	6,03	7,83	18,34	20,91	18,4	3,11	4,16
III	Class [μm]	0-1,6	1,6-2	1-3,15	3,15-4	4-5	5-6,3	6,3-8	8-10	10-12,5
	d_i	0,8	1,8	2,575	3,575	4,5	5,65	7,15	9	11,25
	n_i	16,31	2,67	8	6,21	16,03	14,11	14,85	16,45	5,37
IV	Class [μm]	0-1	1-1,6	1,6-2	2-3,15	3,15-4	4-5	5-6,3	6,3-8	8-10
	d_i	0,5	1,3	1,8	2,575	3,575	4,5	5,65	7,15	9
	n_i	17,23	6,44	3,76	11,59	9,42	25,33	14,32	8,3	3,61
V	Class [μm]	0-0,8	0,8-1	1-1,6	1,6-2	2-3,15	3,15-4	4-5	5-6,3	6,3-8
	d_i	0,4	0,9	1,3	1,8	2,575	3,575	4,5	5,65	7,15
	n_i	17,66	3,03	7,59	4,57	14,2	12,95	27,77	8,81	3,42
VI	Class [μm]	0-0,315	0,315-0,4	0,4-0,5	0,5-0,63	0,63-0,8	0,8-1	1-1,6	1,6-2	2-3,15
	d_i	0,575	0,3575	0,45	0,565	0,715	0,9	1,3	1,8	2,575
	n_i	19,02	4,08	4,26	5,34	5,19	9,94	31,02	11,62	9,53
VII	Class [μm]	0-0,25	0,25-0,315	0,315-0,4	0,4-0,5	0,5-0,63	0,63-0,8	0,8-1	1-1,6	1,6-2
	d_i	0,125	0,2825	0,3575	0,45	0,565	0,715	0,9	1,3	1,8
	n_i	20,53	2,79	5,43	5,9	8,65	8,77	17,73	27,15	3,05
VIII	Class [μm]	0-0,16	0,16-0,25	0,25-0,315	0,315-0,4	0,4-0,5	0,5-0,63	0,63-0,8	0,8-1	1-1,6
	d_i	0,08	0,205	0,2825	0,3575	0,45	0,565	0,715	0,9	0,3
	n_i	33,25	21,06	8,53	14,48	10,01	7,05	3,01	1,82	0,79
IX	Class [μm]	0-0,1	0,1-0,125	0,125-0,16	0,16-0,25	0,25-0,315	0,315-0,4	0,4-0,5	0,5-0,63	0,63-0,8
	d_i	0,05	0,1125	0,1425	0,205	0,2825	0,3575	0,45	0,565	0,715
	n_i	27,5	8,7	1,87	25,23	9,61	14,39	7,83	3,57	1,3
X	Class [μm]	0-0,08	0,08-0,1	0,1-0,125	0,125-0,16	0,16-0,25	0,25-0,315	0,315-0,4	0,4-0,5	0,5-0,63
	d_i	0,04	0,09	0,1125	0,1425	0,205	0,2825	0,3575	0,45	0,515
	n_i	26,81	3,59	7,02	4,93	25,25	9,96	13,01	6,24	3,19

where: d_i – middle of i-th class, i=1,2,...,9; n_i – percentage part of i-th class.

composition curves. In this paper, the kernel estimation by Epanechnikov and gauss kernels were applied to estimate grain composition density functions. Then we determine the suitable grain composition curves and compare their „fitting" to the curves given by parametric estimation and by Fourier series estimation.

To evaluate the fitting level, the mean residual error were applied, calculated from the formulae:

$$s_r = \sqrt{\frac{\sum_{i=1}^{k}\left[F(d_i)-\hat{F}(d_i)\right]^2}{k-2}} \tag{20}$$

where: $\hat{F}$ - distribution function given by kernel estimation; F – empirical distribution function; k – number of classes; d_i – grain diameter.

In case of kernel methods, the formulae (4) becomes as following:

$$\hat{f}(d) = \frac{1}{nh}\sum_{i=1}^{k} K\left(\frac{d-d_i}{h}\right)n_i \tag{21}$$

where: n – sample quantity; n_i – i-th class quantity; h – optimal band width; other symbols as in the formulae (19).

a) for Epanechnikov kernel:

The kernel estimator of researched distribution density function (of random variable D) is obtained by applying Epanechnikov kernel into equation (21). The form of this estimator is as following:

$$\hat{f}(d) = \frac{1}{nh}\sum_{i\in D_1(d)} \frac{3}{4\sqrt{5}}\left[1-\frac{1}{5h^2}(d-d_i)^2\right]n_i \tag{22}$$

where: $D_1(d) = \left\{i \in N : d_i - h\sqrt{5} \le d \le d_i + h\sqrt{5}\right\}$;

$h = 1{,}056\sigma n^{-\frac{1}{5}}$ (equation (9)); other symbols as in (20).

Whereas, distribution function estimator of this variable is given by the formulae:

$$\hat{F}(d) = \frac{1}{n}\left\{\sum_{i\in D_1(d)} \frac{3}{4\sqrt{5}h}\left[\begin{array}{l} d-d_i+\frac{2}{3}h\sqrt{5}+ \\ -\frac{(d-d_i)^3}{15h^2}\end{array}\right]n_i + \sum_{i\in D_2(d)} n_i\right\} \tag{23}$$

where: $D_2(d) = \left\{i \in N; d > d_i + h\sqrt{5}\right\}$; other symbols as in (21).

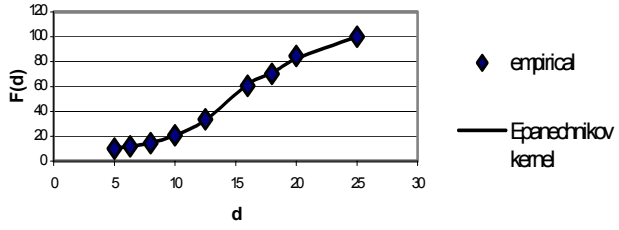

Figure 1. Distribution function plot given by Epanechnikov kernel (stage I) – s_r=1,01.

The example of distribution function plot, given by Epanechnikov kernel, is presented on the figure 1.

b) for Gauss kernel:

In this case we place into equation (21) the following kernel function:

$$K(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \tag{24}$$

The result will be the distribution density function of random variable D, given by the equation:

$$\hat{f}(d) = \frac{1}{nh\sqrt{2\pi}}\sum_{i=1}^{k} n_i e^{-\frac{(d-d_i)^2}{2h^2}} n_i \tag{25}$$

where: n – sample quantity; n_i - i-th class quantity; d – grain diameter; h – optimal band width $h = 1{,}066\sigma n^{-\frac{1}{5}}$ (equation (10)).

And the distribution function estimator is in the form:

$$\hat{F}(d) = \frac{1}{n}\sum_{i=1}^{k}\Phi\left(\frac{d-d_i}{h}\right)n_i \tag{26}$$

where: n, n_i, h, d – as in (25).

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt \tag{27}$$

Similarly as in case of Epanechnikov kernel, the group of distribution functions resulted. The example of this kind of function is presented on figure 2.

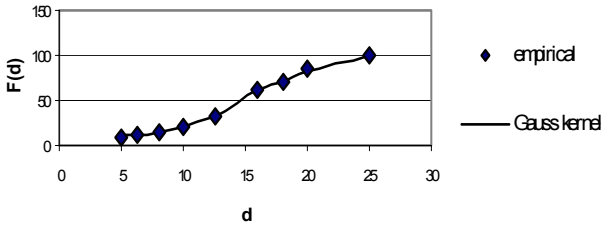

Figure 2. Distribution function plot given by Gauss kernel (stage I) – s_r=0,75.

4.2 *Orthogonal Fourier series method*

In this case, the Fourier estimator of distribution density function of random variable D is in the following form:

$$\hat{f}(d) = \frac{1}{A}\left[1+\sqrt{2}\sum_{j=1}^{\hat{j}}\hat{Q}_j\cos\left(j\frac{\pi d}{A}\right)\right] \tag{28}$$

where: A – right end of the last class; $\hat{Q}_j = \frac{\sqrt{2}}{n}\sum_{i=1}^{k}\cos(\pi j u_i)n_i$; n_i – percentage share of i-th class; $u_i = \frac{d_i}{A}$; k – number of classes.

Table 2. Comparison of mean residual errors for individual approximation methods.

Comminution level	RRB	GSA	Log-norm	Cut RRB	Fourier	Epanechnikov kernel	Gauss Kernel
I	5,49	4,3	4,28	3,13	1,26	1,01	0,75
II	3,86	5,3	5,32	2,85	1,90	0,87	0,76
III	4,98	2,8	2,77	2,30	1,82	0,76	0,59
IV	7,08	5,2	5,23	4,46	1,64	0,83	0,69
V	8,04	5,1	5,13	4,00	1,61	0,58	0,52
VI	4,46	2,3	2,28	1,75	1,67	0,79	0,66
VII	4,92	2,0	2,01	1,72	2,13	0,64	0,56
VIII	1,13	9,2	9,24	16,99	1,33	1,22	0,91
IX	3,12	6,7	6,69	8,58	2,47	1,49	1,37
X	3,16	4,0	4,03	4,76	1,92	0,95	0,73

And the Fourier estimator of distribution density function is as following:

$$\hat{F}(d) = \frac{d}{A} + \frac{\sqrt{2}}{\pi} \sum_{j=1}^{\hat{j}} \frac{\hat{Q}_j}{j} \sin\left(\frac{\pi j d}{A}\right) \tag{29}$$

where: A, $\hat{Q}_j$ - as in equation (28).

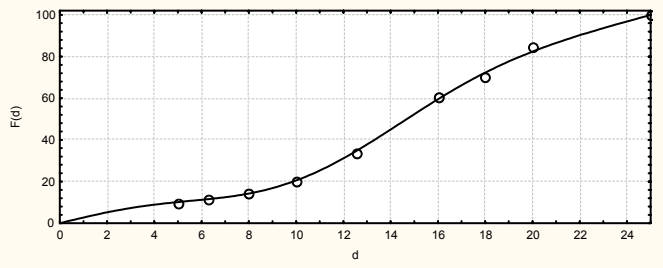

Figure 3. Distribution function plot given by Fourier series approximation (stage I) – s_r=1,26

The results from every methods were compared to the traditional, classical approximations by known distribution functions RRB, GSA, log-norm and cut RRB. The results are shown in the table 2.

5 COAL ENRICHMENT CURVE ESTIMATION BY APPLICATION OF NON PARAMETRIC METHODS

Basing on equation (1) and marking by w density of investigated material and by $p(x,w)$ separation function, which various forms are being applied for various devices and ways of enrichment, we may obtain the theoretical enrichment curve. Let assume that the separation function is given by normal distribution function.

$$p(x,w) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(w-x)}{2\sigma^2}} \tag{30}$$

where σ was calculated from the formulae for probable dissipation $E = \frac{0{,}477}{\sqrt{2\pi}\sigma}$. It was assumed that E=0,14 (device characteristics).

By application of non-parametric methods, we approximate the density function $f(w)$ in base of the formulas (17), (22) or (25) and we obtain that, i.e. for Gauss kernel, the density function is in the following form:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma hn} \sum_{i=1}^{n} \int_{w_{\min}}^{w_{\max}} e^{-\left[\frac{(w-x)^2}{2\sigma^2} + \frac{(w-w_i)^2}{2h^2}\right]} n_i dw \tag{31}$$

If we assume that the coal enrichment is based on difference between specific gravities of researched material grains during jigging and the Gauss kernel was taken to approximate the density function, then the percentage share of fraction (x_j, x_{j+1}) is given by the equation:

$$P(x_i \le x \le x_j) = \int_{x_j}^{x_{j+1}} g(x)dx = \frac{1}{n} \sum_{i=1}^{k} \left[\Phi\left(\frac{x_{j+1} - w_i}{\sqrt{h^2 + \sigma^2}}\right) - \Phi\left(\frac{x_j - w_i}{\sqrt{h^2 + \sigma^2}}\right) \right] \tag{31}$$

where $\Phi(u)$ is normal distribution function N(0,1)

Table 3 contains the example of coal enrichment curve estimation by application of Gauss kernel function. The graphical presentation of results is shown on figure 4.

Table 3. Coal enrichment curve estimation by application of Gauss kernel function

Specific gravity w	Percentage yield in feed	Percentage yield in product γ_i	Separation function
1,30-1,35	49,8	31,58	37,58
1,35-1,40	6,3	3,27	41,05
1,40-1,50	5,5	10,84	51,89
1,50-1,60	5,4	9,61	61,5
1,60-1,75	5,0	9,69	71,19
1,75-1,90	4,2	10,45	81,60
1,90-2,10	23,8	18,56	100

6 CONCLUSIONS

1. In case of lacking the information of which family of distribution functions may the random variable origins (or it is known that it is the mixture

or complex distribution function), the non-parametrical methods may be applied to estimate the

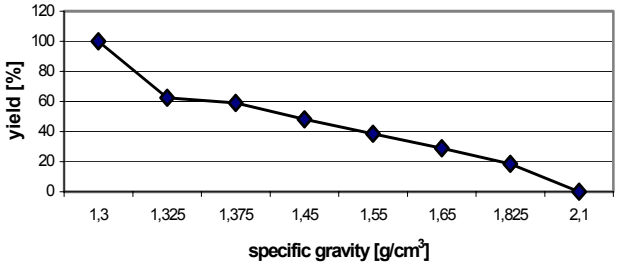

Figure 4. Coal enrichment curve given by Gauss kernel function application

density function. Assuming as the measure of conformity the mean square error it may be said that the curves given by non-parametric estimation methods are usually better fitted to the empirical data than curves given by parametric approximation.

2. The formulas of density functions given by non-parametric methods may be applied to calculate the parameters of researched variable distribution and to calculate the confidence intervals for these parameters. In connection with non-classical Bayesian methods it allows determination the correct formulas of sample size for researched grained material.

3. The non-parametric methods are particularly useful in situation when we are sure that researched random variable is characterized by stable distribution and our goal is to fit the distribution function to empirical one as well as it is possible. We face this kind of situation e.g. in case of balancing the useful component in feed and concentrate.

7 ACKNOWLEDGEMENT

The paper is the effect of the work no. 10.10.100.38

8 REFERENCES

Chentsov N. N.: *Evaluation of an unknown distribution density from observations*, Soviet Math. Dokl. 3, pp. 1159-1562, 1962.

Chentsov N.N.: *Statistical decision rules and optimum inference*, New York, Springer-Verlag, 1980.

Devroye L. and Györfi L.: *Nonparametric density estimation*, The L_1 View, New York, Wiley, 1985.

Efromovich S.: *Nonparametric curve estimation*, New York, Springer-Verlag, 1999.

Efron B. and Tibshirani R.J.: *Using specially designed exponential families for density estimation*, Ann. Statist. 24, pp. 2431-2461, 1996.

Gajek L. and Kałuszka M.: *Wnioskowanie statystyczne*, Warszawa, WNT, 2000.

Hart J.D.: *Nonparametric smoothing and Lack-of fit tests*, New York, Springer, 1997.

Parzen E.: *On estimation of probability density function and mode*, Ann. Math. Statist., nr 33, 1962.

Rosenblatt M.: *Remarks on some nonparametric estimates of a density function*, Ann. Math. Statist. 27, 1956.

Schwartz S.C.: *Estimation of a probability density by an orthogonal series*, Ann. Math. Statist. 38, pp. 1262-1265, 1967.

Silverman B.W.: *Nonparametric density estimation for statistic and data analysis*, London New York, Chapman and Hall, 1986.

Stone C.J.: *An asymptotically optimal window selection rule for kernel density estimates*, Ann. Statist. 12, pp. 1285-1297, 1984.

Tarter M.E. and Lock M.D.: *Model-Free Curve Estimation, Modeling and Simulation,* Philadelphia: SIAM, 1990.

Tompson J.R. and Tapia R.A: *Nonparametric function estimation modeling and simulation*, Philadelphia, SIAM, 1990.

Wahba G.: *Data-based optimal smoothing of orthogonal series density estimates*, Ann. Statist. 9, pp. 146-156, 1981.

Watson G.S.: *On the estimation of the probability density*, I, Ann. Math. Statist. 34, 1963.

Whitle P.: *On the smoothing of probability density functions*, J. Ray Statist. Soc. Ser. B, 1958.