KERNEL ESTIMATION OF CUMULATIVE DISTRIBUTION FUNCTION OF A RANDOM VARIABLE WITH BOUNDED SUPPORT

Aleksandra Baszczyńska[1]

ABSTRACT

In the paper methods of reducing the so-called boundary effects, which appear in the estimation of certain functional characteristics of a random variable with bounded support, are discussed. The methods of the cumulative distribution function estimation, in particular the kernel method, as well as the phenomenon of increased bias estimation in boundary region are presented. Using simulation methods, the properties of the modified kernel estimator of the distribution function are investigated and an attempt to compare the classical and the modified estimators is made.

Key words: boundary effects, cumulative distribution function, kernel method, bounded support.

1. Introduction

Nonparametric methods are becoming increasingly popular in statistical analysis of economic problems. In most cases, this is caused by the lack of information, especially historical data, about the economic variable being analysed. Smoothing methods concerning functions, such as density or cumulative distribution, play a special role in a nonparametric analysis of economic phenomena. Knowledge of density function or cumulative distribution function, or their estimates, allows one to characterize the random variable more completely.

Estimation of functional characteristics of random variables can be carried out using kernel methods. The properties of the classical kernel methods are satisfactory, but when the support of the variable is bounded, kernel estimates may suffer from boundary effects. Therefore, the so-called boundary correction is needed in kernel estimation.

---

[1] Department of Statistical Methods, University of Łódź. E-mail: albasz@uni.lodz.pl.

Kernel estimator of cumulative distribution function has to be modified when the support of the variable is defined as $[a,\infty)$, $(-\infty,b]$ or $[a,b]$. Such a situation is frequently observed in an economic analysis, for example, when data are considered only on the positive real line (e.g.: arable land, energy use, CO_2 emission, external debts stocks, current account balance, total reserves, etc.). Near zero, the classical kernel distribution function estimator is poor because of its considerable bias. The bias comes from the behaviour of kernel estimator, which has no knowledge of the boundary and assigns probability on the negative real line. A range of boundary correction methods for kernel distribution function estimator is present in the literature. They are addressed mainly to boundary kernels (Tenreiro, 2013; Tenreiro, 2015) and reflection method (Koláček, Karunamuni, 2009; Koláček, Karunamuni, 2012).

In Section 2 we introduce the kernel method, which for the first time was implemented in density estimation in the late 1950s. The properties of the kernel density estimator, as well as the modifications, are presented taking into account the boundary effects reduction of classical kernel density estimator. In Section 3 some selected methods of distribution function estimation are presented, including the kernel method. Some methods of choosing the smoothing parameter of kernel method and properties of estimator are shown, and methods of boundary correction are used in the case of cumulative distribution function estimation. In Section 4 the results of a simulation study are given and an attempt to compare the considered estimators is made. In addition, the comparison of the values of smoothing parameters is presented. The simulations and the plots were carried out using MATLAB software.

The aim of the paper is to give a detailed presentation of methods of the modified kernel distribution function estimation and to compare the considered methods. The simulation shows that when boundary correction is used in kernel estimation of distribution function, the estimator has better properties.

2. Kernel method

The kernel method originated from the idea of Rosenblatt and Parzen dedicated to density estimation. The Rosenblatt-Parzen kernel density estimator is as follows (cf.: Härdle, 1994; Wand, Jones, 1995; Silverman 1996; Domański et al., 2014):

$$\hat{f}_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right), \tag{1}$$

where: $X_1, X_2,..., X_n$ is the random sample from the population with unknown density function $f(x)$; n is the sample size; h_n is the smoothing parameter, which controls the smoothness of the estimator ($\lim_{n\to\infty} h = 0$, $\lim_{n\to\infty} nh = \infty$).

Throughout this paper the notation $h = h_n$ will be used. $K(u)$ is the weighting function called the kernel function. When $K(u)$ is symmetric and unimodal function and the following conditions are fulfilled:

$$
\begin{cases}
\int\limits_{-\infty}^{\infty} K(u)du = 1, \\[2mm]
\int\limits_{-\infty}^{\infty} uK(u)du = 0, \\[2mm]
\int\limits_{-\infty}^{\infty} u^2 K(u)du = \kappa_2 > 0,
\end{cases}
\tag{2}
$$

the kernel function is called the second order kernel function (or classical kernel function).

The most frequently used Gaussian kernel function $K(u) = \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}u^2\right)$ is a function belonging to this group, although its support is unbounded. It stands in contrast to other kernel functions fulfilling conditions (2), like functions presented in Table 1, for which support is bounded. The indicator function $I(|u| \leq 1)$ is defined as follows: $I(|u| \leq 1) = 1$ for $|u| \leq 1$, $I(|u| \leq 1) = 0$ for $|u| > 1$.

Table 1. Kernel functions

Kernel function	$K(u)$
Uniform	$\dfrac{1}{2} I(|u| \leq 1)$
Triangle	$(1 - |u|) I(|u| \leq 1)$
Epanechnikov	$\dfrac{3}{4}(1 - u^2) I(|u| \leq 1)$
Quartic	$\dfrac{15}{16}(1 - u^2)^2 I(|u| \leq 1)$
Triweight	$\dfrac{35}{32}(1 - u^2)^3 I(|u| \leq 1)$
Cosine	$\dfrac{\pi}{4} \cos\left(\dfrac{\pi}{2}u\right) I(|u| \leq 1)$

Higher order kernel functions (the order of the kernel is the order of the first nonzero moment) can be used, especially in reducing the mean squared error of the estimator. But the higher order kernels properties may sometimes be unacceptable because they may result in taking negative values for the density function estimators.

When the support of random variable is, for example, left-bounded (support of random variable is $[0,\infty)$), the properties of the estimator (1) may differ in boundary region $[0,h]$ and in inner region (h,∞) (cf.: Jones, 1993; Jones, Foster, 1996; Li, Racine, 2007). The estimator (1) is not consistent in boundary region. As a result, the support of the kernel density estimator may differ from the support of the random variable and the estimator may be non-zero for negative values of random variable. Moreover, this situation may appear when the kernel function has unbounded as well as bounded support. Removing boundary effects can be done in various ways. The best known and most often used method is the reflection method, which is characterized by both simplicity and best properties.

Assuming that the support of random variable is $[0,\infty)$, the reflection kernel density estimator, based on reflecting data about zero, has the following form (cf. Kulczycki, 2005):

$$\hat{f}_{nR}(x) = \frac{1}{nh} \sum_{i=1}^{n} \left[K\left(\frac{x - X_i}{h} \right) + K\left(\frac{x + X_i}{h} \right) \right]. \qquad (3)$$

The estimator (3) is a consistent estimator of unknown density function f. Moreover, it integrates to unity and for x close to zero the bias is of order $O(h)$. The analysis of the properties of this estimator is presented in Baszczyńska (2015), among others.

3. Distribution function estimation

Let $X_1, X_2, ..., X_n$ denote independent random variables with a density function f and a cumulative distribution function F. One can estimate the cumulative distribution function (CDF) by:

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty, x]}(X_i), \qquad (4)$$

where I_A is the indicator function of the set A: $I_A(x) = 1$ for $x \in A$, $I_A(x) = 0$ for $x \notin A$.

The empirical distribution function defined by (4) is not smooth, at each point $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ it jumps by $\frac{1}{n}$.

The smoothed version of the empirical distribution estimator is the Nadaraya kernel estimator of CDF:

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{\frac{x-X_i}{h}} K(y)dy = \frac{1}{n}\sum_{i=1}^{n}W\left(\frac{x-X_i}{h}\right), \qquad (5)$$

where h is a smoothing parameter such as $\lim_{n\to\infty} h = 0$, $\lim_{n\to\infty} nh = \infty$ and $W(x) = \int_{-1}^{x} K(t)dt$. Assuming that function $K(x) \geq 0$ is a unimodal, symmetric kernel function of the second order with support $[-1,1]$ (examples of these kernels are presented in Table 1), the properties of function $W(x)$ are the following:

$$W(x) = 0 \text{ for } x \in (-\infty, -1],$$

$$W(x) = 1 \text{ for } x \in [1, \infty),$$

$$\int_{-1}^{1} W^2(x)dx \leq \int_{-1}^{1} W(x)dx = 1, \qquad (6)$$

$$\int_{-1}^{1} W(x)K(x)dx = \frac{1}{2},$$

$$\int_{-1}^{1} xW(x)K(x)dx = \frac{1}{2}\left[1 - \int_{-1}^{1} W^2(x)dx\right].$$

Function $W(x)$ is a cumulative distribution function because $K(x)$ is a probability density function. For example, when the kernel function is Epanechnikov kernel, the function $W(x)$ has the form:

$$W(x) = \begin{cases} 0 & \text{for} \quad x \leq -1, \\ -\frac{1}{4}x^3 + \frac{3}{4}x + \frac{1}{2} & \text{for} \quad |x| \leq 1, \\ 1 & \text{for} \quad x \geq 1. \end{cases}$$

Assuming additionally that $F(x)$ is twice continuously differentiable, the mean integrated squared error (MISE) of kernel distribution estimator (5) is as follows:

$$MISE[\hat{F}(x)] = E \int_{-\infty}^{+\infty} [\hat{F}(x) - F(x)]^2 dx =$$

$$= \frac{1}{n} \int_{-\infty}^{+\infty} F(x)(1 - F(x)) dx - c_1 \frac{h}{n} + c_2 h^4 + o\left(\frac{h}{n} + h^4\right), \tag{7}$$

where:

$$c_1 = 1 - \int_{-1}^{1} W^2(t) dt, \tag{8}$$

$$c_2 = \frac{\kappa_2^2}{4} \int_{-\infty}^{+\infty} [F^{(2)}(t)]^2 dt, \tag{9}$$

$F^{(s)}$ denotes the sth derivative of the cumulative distribution function.

Kernel distribution estimator (5) is a consistent estimator of the distribution function. The expectation value, bias and variance are, respectively:

$$E[\hat{F}(x)] = F(x) + \frac{1}{2} F^{(2)}(x) h^2 \kappa_2 + o(h^2),$$

$$B[\hat{F}(x)] = \frac{1}{2} F^{(2)}(x) h^2 \kappa_2 + o(h^2),$$

$$D^2[\hat{F}(x)] = \frac{1}{n} F(x)(1 - F(x)) - \frac{1}{n} h f(x) \left[1 - \int_{-1}^{1} W^2(t) dt\right] + o\left(\frac{h}{n}\right).$$

The method of choosing the value of the smoothing parameter in kernel estimation of the cumulative distribution function is of crucial interest, as it is in kernel estimation of the density function. Some procedures used frequently in CDF estimation are presented in Table 2.

Table 2. Methods of choosing the smoothing parameter in kernel estimation of the cumulative distribution function

Method	Smoothing parameter
Cross-validation, CV	$\hat{h}_{CV} = \arg\min_{h \in H_n} CV(h)$, $$CV(h) = \frac{1}{n}\sum_{i=1}^{n} \int_{-\infty}^{+\infty} \left[I_{(-\infty,x]}(X_i) - \hat{F}_{-i}(x,h)\right]^2 dx \,,$$ $\hat{F}_{-i}(x,h)$ is a kernel estimator based on the sample with X_i deleted
Maximal smoothing principle, MSP	$$h_{MS} = \left(\frac{7c_1}{15\kappa_2^2}\right)^{1/3} n^{-1/3} \sqrt{7}\hat{\sigma}^2$$
Plug-in, PI	$$h_{PI} = \left(\frac{c_1}{-\hat{\psi}_1 \kappa_2^2}\right)^{1/3} n^{-1/3}, \; \hat{\psi}_k(g) = \frac{1}{n^2 g}\sum_{i,j=1}^{n} L^{(2k)}\left(\frac{X_i - X_j}{g}\right),$$ g is an initial smoothing parameter, $L^{(2k)}$ is the $2k$th derivative of the initial kernel function L
Iteration, IM	$$h_{IT,j+1} = \frac{4h_{IT,j}}{c_1 n}\sum_{\substack{i,j=1 \\ i\neq j}}^{n}\Omega\left(\frac{X_i - X_j}{h_{IT,j}}\right), \; j = 0,1,\ldots,$$ $$\Omega(u) = (K*K*W*W - 2K*W*W + W*W)(u),$$ $f*g$ denotes convolution

When the random variable has bounded support (without loss of generality one can take $[0,\infty)$), as in the case of the kernel density estimation, the properties of the kernel distribution function get poorer, in comparison with the situation when the support is unbounded.

For x in boundary region $x \in [0,h]$, let $c = x/h$, $0 \le c \le 1$, the expectation value and the variance of estimator (5) are the following:

$$_B E\big[\hat{F}(x)\big] = F(x) + hf(0)\int\limits_{-1}^{-c} W(t)dt +$$

$$+ h^2 f^{(1)}(0)\left\{\frac{c^2}{2} + c\int\limits_{-1}^{-c} W(t)dt - \int\limits_{-1}^{-c} tW(t)dt\right\} + o\big(h^2\big),$$

$$_B D^2\big[\hat{F}(x)\big] = \frac{1}{n}F(x)\big(1 - F(x)\big) - \frac{1}{n}hf(0)\left[\int\limits_{-1}^{c} W^2(t)dt - c\right] + o(h).$$

It is to note that in boundary region the estimator is not consistent, but variance is of the same order.

The reflection kernel distribution estimator has the form (cf. Horovà et al., 2012):

$$\hat{F}_R(x) = \frac{1}{n}\sum_{i=1}^{n}\left[W\left(\frac{x - X_i}{h}\right) - W\left(\frac{x + X_i}{h}\right)\right]. \tag{10}$$

The generalized reflection kernel distribution estimator, improving the bias of the estimator and holding onto low variance, is the following (cf. Karunamuni, Alberts, 2005; Karunamuni, Zhang, 2008):

$$\hat{F}_{GR}(x) = \frac{1}{n}\sum_{i=1}^{n}\left[W\left(\frac{x - g_1(X_i)}{h}\right) - W\left(\frac{x + g_2(X_i)}{h}\right)\right], \tag{11}$$

where g_1 and g_2 are cubic polynomials with such coefficients that the bias of the estimator is of order $O\big(h^2\big)$.

In boundary region the expectation value and variance of the estimator (11) are, respectively:

$$E\big[\hat{F}_{GR}(x)\big] =$$
$$= F(x) +$$
$$+ h^2\left\{f^{(1)}(0)\left[\frac{c^2}{2} + 2c\int\limits_{-1}^{-c} W(t)dt - \int\limits_{-c}^{c} tW(t)dt\right] - f(0)g_1^{(2)}(0)\int\limits_{-1}^{c}(c - t)W(t)dt - f(0)g_2^{(2)}(0)\int\limits_{-1}^{-c}(c + t)W(t)dt\right\}$$
$$+ o\big(h^2\big),$$

$$D^2\big[\hat{F}_{GR}(x)\big] = \frac{1}{n}F(x)\big(1 - F(x)\big) - \frac{1}{n}hf(0)\left[\int\limits_{-1}^{c} W^2(t)dt - 2\int\limits_{-1}^{c} W(t)W(t - 2c)dt + \int\limits_{-1}^{-c} W^2(t)dt\right] + o(h).$$

4. Results of the simulation study

The objective of the simulation study was to compare properties of chosen estimators of the distribution function. The estimators were considered in a special situation when the support of the random variable is bounded. The comparison was made through the graphical representation of the results of the estimation. This form of presenting the estimator is of crucial importance, especially from the user's point of view. The graph provides a fast, comprehensive and readable form of presenting the functional characteristic of the random variable, even for inexperienced users.

In the simulation study the following populations with Weibull distribution $W(0, \delta, \gamma)$ with different scale and shape parameters were examined:

$W1(0,1,0.1)$,

$W2(0,1,0.5)$,

$W3(0,1,1)$,

$W4(0,1,2)$,

$W5(0,1,3.4)$,

$W6(0,1,5)$,

$W7(0,4,1)$,

$W8(0,4,2)$.

The use of a wide range of distribution parameters ensures that varied populations are considered in the study. The difference between populations can be seen, for example, in location, dispersion, asymmetry and kurtosis.

The samples $X_1, X_2,..., X_n$ of size $n = 10,20,...,100$ were drawn from each population and the following estimators of the distribution function were calculated: empirical distribution function (4), kernel distribution function (5) and reflection kernel distribution function (11). For kernel estimators, Gaussian, Epanechnikov and quartic kernels were used, with Silverman's practical rule (RR), maximal smoothing principle (MSP), plug-in method (PI) and iteration method (IM) used for choosing the smoothing parameter.

Some results for medium size sample $n = 50$ drawn from selected populations, where Epanechnikov kernel and Silverman's rule were used in kernel estimators, are presented in Figures 1-8.
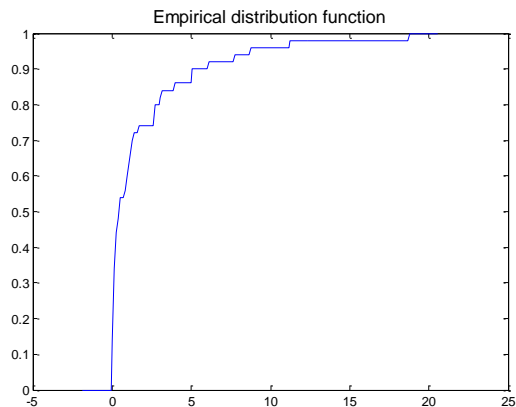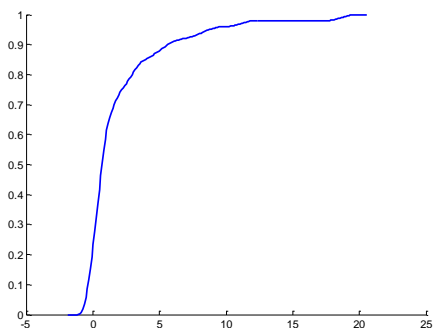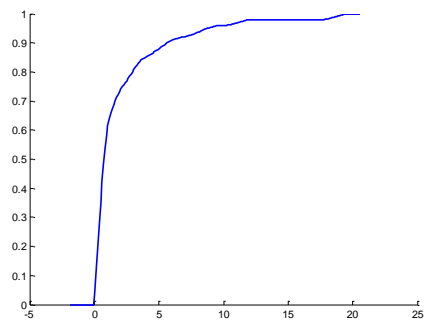
Figure 1. Empirical distribution function, sample size $n = 50$ from $W2(0,1,0.5)$ population

Kernel estimator

Reflection kernel estimator

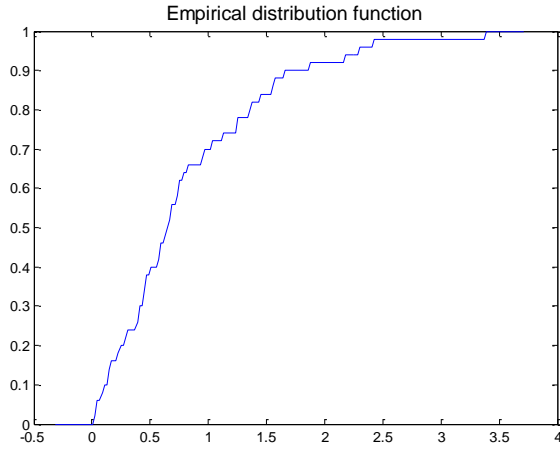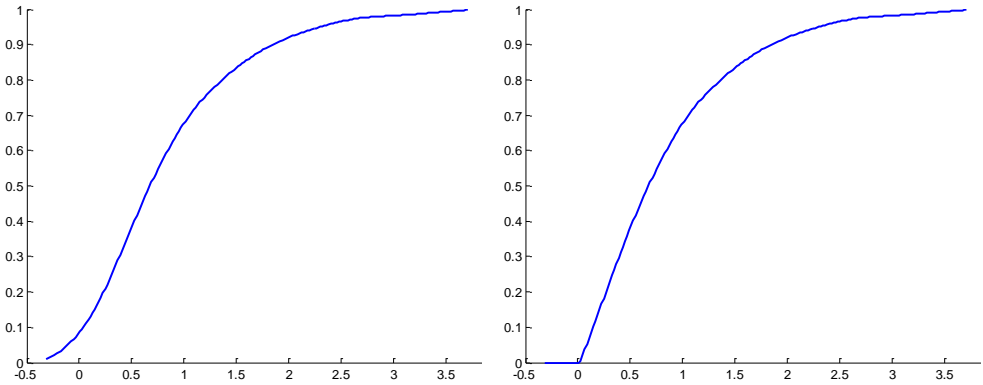**Figure 2**. Kernel distribution function estimators, sample size $n = 50$ from $W2(0,1,0.5)$ population

Figure 3. Empirical distribution function, sample size $n = 50$ from $W3(0,1,1)$ population

Kernel estimator   Reflection kernel estimator

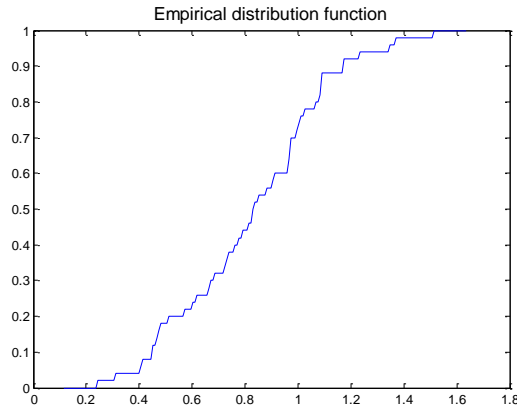**Figure 4**. Kernel distribution function estimators, sample size $n = 50$ from $W3(0,1,1)$ population

Empirical distribution function
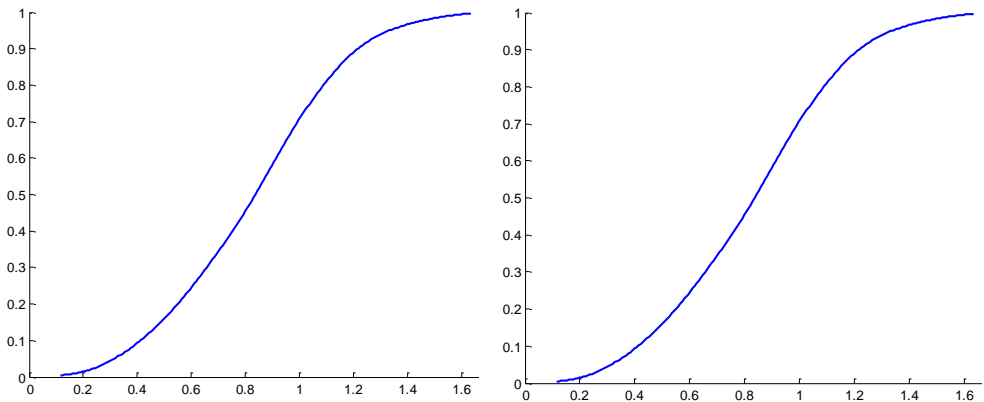
Figure 5. Empirical distribution function, sample size $n = 50$ from $W5(0,1,3.4)$ population

Kernel estimator   Reflection kernel estimator

Figure 6. Kernel distribution function estimators, sample size $n = 50$ from $W5(0,1,3.4)$ population

When the group of samples from populations with the same scale parameter but with shape parameter differences (populations $W1(0,1,0.1) - W6(0,1,5)$) is considered, it can be noticed that the lower the value of shape parameter, the bigger the incompatibility between the support of the random variable and the

support of the kernel distribution function estimator. For high values of shape parameters (for example, in populations $W4(0,1,2) - W6(0,1,5)$) the influence of boundary effects is almost imperceptible. When samples are drawn from populations with high values of shape parameters, kernel functions, used in constructing the distribution function estimator in observations near zero, do not extend beyond the support of the random variable.
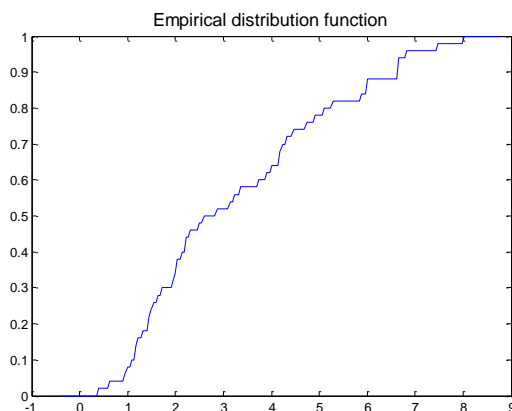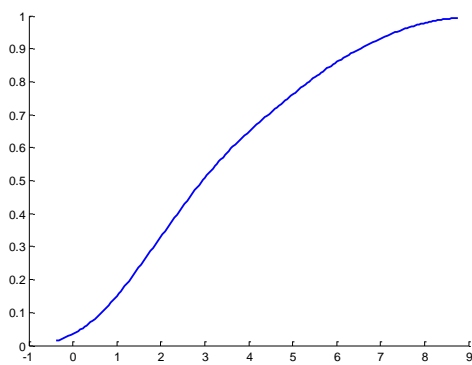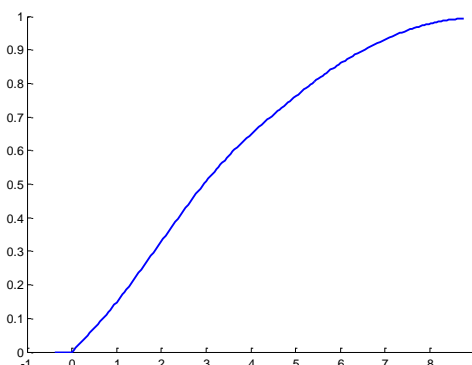

Figure 7. Empirical distribution function, sample size $n = 50$ from $W8(0,4,2)$ population

Kernel estimator   Reflection kernel estimator

Figure 8. Kernel distribution function estimators, sample size $n = 50$ from $W8(0,4,2)$ population

It is worth stressing that when Gaussian kernel function was used in kernel estimator, the boundary effects were bigger, in comparison with other kernel functions. This results directly from the properties of Gaussian kernel which is the only one among the studied kernel functions that has unbounded support.

The kernel distribution function estimators behave in a very similar way, even for very small samples ($n=10$, $n=20$). Hence, the sample size is not the essential factor in the occurrence of boundary effects. Taking into account the values of shape parameters and sample sizes, the same results were observed when for the same populations with bounded random variable, the kernel density function estimators were constructed (cf. Baszczyńska, 2015). However, it must be indicated that boundary effects influence the shape of estimators more strongly in the case of density estimator, in some cases even giving the wrong impression of multimodality.

To extend the study, the dependence between kernel function and smoothing parameter was observed. The results are presented in Table 3.

Table 3. Values of smoothing parameters in kernel distribution function estimation for samples from Weibull distribution populations

Population	Kernel function	Method of smoothing parameter choice			
		RR	MSP	PI	IM
$W1(0,1,0.1)$	Epanechnikov	0.8635	0.9224	0.2655	
	quartic	1.0203	1.0899	0.3137	
$W2(0,1,0.5)$	Epanechnikov	1.0958	1.1706	0.3878	0.1872
	quartic	1.2948	1.3832	0.4582	0.2638
$W3(0,1,1)$	Epanechnikov	0.5852	0.6251	0.4817	0.4974
	quartic	0.6914	0.7386	0.5692	0.5699
$W4(0,1,2)$	Epanechnikov	0.3792	0.4051	0.4105	0.4199
	quartic	0.4481	0.4787	0.4851	0.4885
$W5(0,1,3.4)$	Epanechnikov	0.2771	0.2961	0.3523	0.3267
	quartic	0.3275	0.3498	0.4162	0.3858
$W6(0,1,5)$	Epanechnikov	0.2265	0.2419	0.3178	0.2645
	quartic	0.2676	0.2859	0.3756	0.1986
$W7(0,4,1)$	Epanechnikov	3.0632	3.2722	0.7626	1.8332
	quartic	3.6195	3.8666	0.9011	2.1697
$W8(0,4,2)$	Epanechnikov	1.9802	2.1154	0.7755	1.6573
	quartic	2.3399	2.4996	0.9163	2.0212

In the procedure of kernel estimation of cumulative distribution function, two kernel functions: Gaussian and Epanechnikov functions, influence the kernel estimator in a very similar way. The application of these kernel functions is connected with almost the same values of smoothing parameters. It can indicate that Gaussian and Epanechnikov kernels have similar smoothing properties, although they are characterized by different support. When quartic kernel is used, the smoothing parameter is smaller in comparison with other kernel functions.

For samples from populations with small shape parameter, the kernel distribution estimator with smaller smoothing parameters was used. The bigger the shape parameter, the bigger the smoothing parameter in kernel estimation. In general, using Silverman's reference rule ensures smaller values of smoothing parameter. When the shape parameter of population distribution is small, the iterative method is rather poor, the smoothing parameter is unacceptably big, which is denoted by a grey spot in Table 3.

5. Conclusion

The kernel method is an intuitive, simple and useful procedure, especially in density and distribution function estimation. When the support of the random variable is bounded, this procedure needs modification. The modified kernel distribution function estimator ensures that the estimator is consistent, even in boundary region, and the support of the estimator is the same as the support of the random variable being analysed. In kernel method two parameters should be predetermined: kernel function and smoothing parameter. Quartic kernel function indicates higher values of smoothing parameter. Silverman's reference rule, though based on the assumption that the population distribution is normal, gives smaller values of the smoothing parameter.

REFERENCES

BASZCZYŃSKA, A., (2015). Bias Reduction in Kernel Estimator of Density Function in Boundary Region, Quantitative Methods in Economics, XVI, 1.

DOMAŃSKI, C., PEKASIEWICZ, D., BASZCZYŃSKA, A., WITASZCZYK, A., (2014). Testy statystyczne w procesie podejmowania decyzji [Statistical Tests in the Decision Making Process], Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

HÄRDLE, W., (1994). Applied Nonparametric Regression, Cambridge University Press, Cambridge.

LI, Q., RACINE, J. S., (2007). Nonparametric Econometrics. Theory and Practice, Princeton University Press, Princeton and Oxford.

JONES, M. C., (1993). Simple Boundary Correction for Kernel Density Estimation, Statistics and Computing, 3, pp. 135–146.

JONES, M. C., FOSTER, P. J., (1996). A Simple Nonnegative Boundary Correction Method for Kernel Density Estimation, Statistica Sinica, 6, pp. 1005–1013.

KARUNAMUNI, R. J., ALBERTS, T., (2005). On Boundary Correction in Kernel Density Estimation, Statistical Methodology, 2, pp. 191–212.

KARUNAMUNI, R. J., ZHANG, S., (2008). Some Improvements on a Boundary Corrected Kernel Density Estimator, Statistics and Probability Letters, 78, pp. 497–507.

KOLÁČEK, J., KARUNAMUNI, R. J., (2009). On Boundary Correction in Kernel Estimation of ROC Curves, Australian Journal of Statistics, 38, pp. 17–32.

KOLÁČEK, J., KARUNAMUNI, R. J., (2012). A Generalized Reflection Method for Kernel Distribution and Hazard Function Estimation, Journal of Applied Probability and Statistics, 6, pp. 73–85.

KULCZYCKI, P., (2005). Estymatory jądrowe w analizie systemowej [Kernel Estimators in Systems Analysis], Wydawnictwa Naukowo-Techniczne, Warszawa.

HOROVÀ, I., KOLÁČEK, J., ZELINKA, J., (2012). Kernel Smoothing in MATLAB. Theory and Practice of Kernel Smoothing, World Scientific, New Jersey.

SILVERMAN, B.W., (1996). Density Estimation for Statistics and Data Analysis, Chapman and Hall, London.

TENREIRO, C., (2013). Boundary Kernels for Distribution Function Estimation, REVSTAT Statistical Journal, 11, 2, pp. 169–190.

TENREIRO, C., (2015). A Note on Boundary Kernels for Distribution Function Estimation, http://arxiv.org/abs/1501.04206 [14.11.2015].

WAND, M. P., JONES, M.C., (1995). Kernel Smoothing, Chapman and Hall, London.