Smooth Estimates of Distribution Functions with Application
in Environmental Studies

IVANA HOROVÁ
Masaryk University
Dept. of Math. and Statist.
Janáčkovo nám. 2a
602 00 Brno
Czech Republic
horova@math.muni.cz

JAN KOLÁČEK
Masaryk University
Dept. of Math. and Statist.
Janáčkovo nám. 2a
602 00 Brno
Czech Republic
kolacek@math.muni.cz

JIŘÍ ZELINKA
Masaryk University
Dept. of Math. and Statist.
Janáčkovo nám. 2a
602 00 Brno
Czech Republic
zelinka@math.muni.cz

ABDEL EL-SHAARAWI
National Water Research Inst.
867 Lakeshore Road
P.O. Box 5050
Burlington, ON L7R 4A6
Canada
Abdel.El-Shaarawi@ec.gc.ca

Abstract: The most commonly used nonparametric estimate of a cumulative distribution function $F$ is an empirical distribution function $F_n$. But $F_n$ is a step function even in case that $F$ is continuous. The present paper aims to provide a smooth estimate of $F$. Kernel methods seem to be adequate for this purpose. There exist several methods on how to choose a bandwidth, e.g. [1], [2], [3]. We propose a method of bandwidth selection based on a suitable estimate of Mean Integrated Square Error. We also focus on an estimate of a cumulative distribution function in case that random variables $X_1, \ldots, X_n$ are nonnegative. The aforementioned methods are not reliable near the point $x = 0$. In order to avoid this problem we propose a reflection method [5]. A simulation study is conducted to compare methods with and without suppressing boundary effects. The theoretical results are applied to study the distributional characteristics in bioaccumulation of a toxic substance in fish population from Lake Ontario.

Key–Words: kernel distribution, boundary effects, iterative method

Acknowledgements: The research was supported by The Jaroslav Hájek center for theoretical and applied statistics (grant No. LC 06024).

References:


