

Density Estimation Using Gumbel Kernel Estimator

Javaria Ahmad Khan, Atif Akbar*

Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan

Email: jakhan0@yahoo.com, *atifakbar@bzu.edu.pk

How to cite this paper: Khan, J.A. and Akbar, A. (2021) Density Estimation Using Gumbel Kernel Estimator. *Open Journal of Statistics*, 11, 319-328.

<https://doi.org/10.4236/ojs.2021.112018>

Received: March 14, 2021

Accepted: April 19, 2021

Published: April 22, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this article, our proposed kernel estimator, named as Gumbel kernel, which broadened the class of non-negative, asymmetric kernel density estimators. Such kernel estimator can be used in nonparametric estimation of the probability density function (*pdf*). When the density functions have limited bounded support on $[0, \infty)$ and they are liberated of boundary bias, always non-negative and obtain the optimal rate of convergence for the mean integrated squared error (MISE). The bias, variance and the optimal bandwidth of the proposed estimators are investigated on theoretical grounds as well as on simulation basis. Further, the applicability of the proposed estimator is compared to Weibull kernel estimator, where performance of newly proposed kernel is outstanding.

Keywords

Asymmetrical Kernels, Boundary Problems, Density Estimation, Flood Data, Gumbel Kernel Estimator

1. Introduction

To investigate the properties and features of data or in anomaly detection, density estimation performs a vital role. For this purpose, the nonparametric kernel density estimation or curve estimation is a famous technique. Nonparametric estimation has certain advantages over parametric estimation, e.g. the problem of priori distribution choice, possibility of using non-homogenous data, no functional form and the most important is allocation of weights, etc. [1]. In nonparametric density estimation, boundary bias is a very serious issue. It affects the performance of the estimator at boundary points due to boundary effects, then from the interior points. Such problem is happened; when smoothing is carried out near the boundary and fixed symmetric kernel allocate weights outside the

density support. That's why in some cases, parametric method for curve estimation performs better than nonparametric estimation [2]. Such problem is happened when variables represent some sort of physical measure such as time or length. These variables thus have a natural lower boundary, e.g. time of birth, etc. So, when smoothing is carried out near the boundary and then fixed, symmetric kernels allocate weights outside the density support and due to this, boundary bias arise [3].

There is a vast literature on removing boundary effects in nonparametric method. As yet there appears to be no single dominating solution that corrects the boundary problem for all shapes of densities. Some common techniques are reflection of data, introduced by Schuster [4]. Similarly, Silverman [5] proposed negative-reflection. Eubank and Speckman [6] suggested semi-parametric model. Chen [7] suggested the solution of this problem by replacing the symmetric kernels by the asymmetric Beta kernel which never assigns weight outside the support. Many others used Chen's idea and proposed other kernels, *i.e.* Gamma [8], lognormal [9], Inverse Gaussian [10], Weibull [11], etc.

By following Chen [7], we are going to propose a new class of density estimator named as a Gumbel kernel estimator along with its bias, variance and optimal bandwidth, which will be the keen addition in category of asymmetrical kernel(s) that solve the problem of boundary bias. The Gumbel distribution is a particular case of the Generalized Extreme Value (GEV) distribution, and also known as the log Weibull or Fisher-Tippett distribution. GEV is a family of continuous probability distributions which combines the Gumbel, Fréchet and Weibull families and also known as type I, II and III extreme value distributions. The common functional form for all 3 distributions was discovered by McFadden [12].

This paper is organized as follows. In the first Section, we present some information about kernel smoothing and in Section 2 we presented the proposed kernel. In third Section, we investigated the bias, variance and optimal bandwidth of the Gumbel kernel estimator. The performance of the proposed estimator will be tested via real and simulated data sets in Section 4, while Section 5 concludes.

2. Gumbel Kernel Estimator

Let X_1, \dots, X_n be a random sample from a distribution with an unknown probability density function f which has bounded support on $[0, \infty)$. Representation of pdf of Gumbel (μ, β) is

$$f(j) = \frac{1}{\beta} e^{-(z+e^{-z})}, \quad j > 0, \quad (1)$$

where $z = \frac{j - \mu}{\beta}$ and $\beta > 0$. The mean and variance of J are equal to $\mu + \beta\gamma$ and $\frac{\beta^2 \pi^2}{6}$, where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

As $\mu = x$ and $s = h^{\frac{1}{2}}$, the class of Gumbel kernel considered is:

$$K_{\text{Gumbel}\left(x, h^{\frac{1}{2}}\right)}(j) = \frac{1}{\sqrt{h}} e^{-\left(\frac{j-x}{\sqrt{h}} + e^{-\left(\frac{j-x}{\sqrt{h}}\right)}\right)} \tag{2}$$

where h is bandwidth satisfying the condition that $h \rightarrow 0$ and $nh \rightarrow \infty$ as $n \rightarrow \infty$. If a random variable X has a pdf $K_{\text{Gumbel}\left(x, h^{\frac{1}{2}}\right)}(x)$, then $E(X) = x + \sqrt{h}\gamma$

and the variance is $\text{var}(X) = \frac{h\pi^2}{6}$.

The corresponding estimator of pdf is

$$\hat{f}_{\text{Gumbel}}(X) = n^{-1} \sum_{i=1}^n K_{\text{Gumbel}\left(x, h^{\frac{1}{2}}\right)}(X_i) \tag{3}$$

This estimator is easy to use and similar to following kernels for comparison: Gamma 1 and Gamma 2 kernels by Chen [7] are;

$$K_{\text{Gam1}(x/b+1, b)}(y) = \frac{y^{x/b} \exp\{-y/b\}}{\Gamma\{x/b+1\} b^{x/b+1}}, \tag{4}$$

and

$$K_{\text{Gam2}(\rho_b, b)}(y) = \frac{y^{\rho_b(x)-1} \exp\{-y/b\}}{\Gamma\{\rho_b(x)\} b^{\rho_b(x)}} \tag{5}$$

where

$$\rho_b(x) = \begin{cases} x/b & \text{if } x \in [2b, \infty); \\ \frac{1}{4}(x/b)^2 + 1 & \text{if } x \in [0, 2b). \end{cases} \tag{6}$$

Beta kernel by Chen [7] is;

$$K_{B(x/b+1, 1-x/b+1)}(y) = \frac{y^{x/b} (1-y)^{(1-x)/b}}{B\{x/b+1, (1-x)/b+1\}}, \tag{7}$$

where, B is Beta function.

Birnbaum-Saunders and Log-Normal kernels by Jin and Kawczak [9] are;

$$K_{BS(b^{1/2}, x)}(y) = \frac{1}{\sqrt{2\pi b}} \left(\sqrt{\frac{1}{xy}} + \sqrt{\frac{x}{y^3}} \right) \exp\left[-\frac{1}{2b} \left(\frac{y}{x} - 2 + \frac{x}{y} \right) \right], \tag{8}$$

and

$$K_{LN(\ln x, 4\ln(1+b))}(y) = \frac{1}{\sqrt{8\pi \ln(1+b)} y} \exp\left[-\frac{(\ln y - \ln x)^2}{8 \ln(1+b)} \right] \tag{9}$$

Inverse Gaussian and Reciprocal Inverse Gaussian kernels by Scaillet [10] are;

$$K_{IG(x, 1/b)}(y) = \frac{1}{\sqrt{2\pi b y^3}} \exp\left[-\frac{1}{2bx} \left(\frac{y}{x} - 2 + \frac{x}{y} \right) \right], \tag{10}$$

and

$$K_{RIG(\ln x, 4\ln(1/b))}(y) = \frac{1}{\sqrt{2\pi b y}} \exp\left[-\frac{x-b}{2b}\left(\frac{y}{x-b} - 2 + \frac{x-b}{y}\right)\right] \quad (11)$$

Erlang kernel by Salha, *et al.* [13] is;

$$K_{E\left(x, \frac{1}{b}\right)}(y) = \frac{1}{\Gamma\left(1 + \frac{1}{b}\right)} \left[\frac{1}{x}\left(1 + \frac{1}{b}\right)\right]^{-\frac{b+1}{b}} y^{\frac{1}{b}} \exp\left(-\frac{y}{x}\left(1 + \frac{1}{b}\right)\right) \quad (12)$$

Weibull kernel by Salha, *et al.* [11] is;

$$K_{w\left(x, \frac{1}{b}\right)}(y) = \frac{\Gamma(1+b)}{bx} \left[\frac{y\Gamma(1+b)}{x}\right]^{\frac{1}{b}-1} \exp\left(-\left(\frac{y\Gamma(1+b)}{x}\right)^{\frac{1}{b}}\right) \quad (13)$$

3. Bias, Variance and Optimal Bandwidth

Theorem 1 (Bias)

The bias of proposed estimator is given by;

$$\text{Bias}\left\{\hat{f}_{\text{Gumbel}}(x)\right\} = h\left(\frac{\gamma f'(x)}{\sqrt{h}} + \frac{1}{2}f''(x)\frac{\pi^2}{6}\right) + o(1) \quad (14)$$

Proof:

$$E\left(\hat{f}_{\text{Gumbel}}(x)\right) = \int_0^\infty K_{\text{Gumbel}\left(x, h^{\frac{1}{2}}\right)}(X_i) f(x) dx = E(f(\xi_x)),$$

where ξ_x follows a Gumbel distribution with scale parameter $h^{\frac{1}{2}}$ and shape parameter x .

The Taylor expansion about μ_x for $f(\xi_x)$ is:

$$f(\xi_x) = f(\mu_x) + f'(\mu_x)(\xi_x - \mu_x) + \frac{1}{2}f''(\mu_x)(\xi_x - \mu_x)^2 + o(1).$$

So, $E(f(\xi_x)) = f(x) + \frac{1}{2}f''(x)\text{var}(\xi_x) + o(1)$.

$$E(f(\xi_x)) = f(x) + \sqrt{h}\left(\gamma f'(x) + \frac{\sqrt{h}\pi^2}{12}f''(x)\right) + o(1).$$

Hence,

$$\text{Bias}\left(\hat{f}_{\text{Gumbel}}(x)\right) = \sqrt{h}\left(\gamma f'(x) + \frac{\sqrt{h}\pi^2}{12}f''(x)\right) + o(1)$$

Theorem 2 (Variance)

The variance of the proposed estimator is given by:

$$\text{var}\left(\hat{f}_{\text{Gumbel}}(x)\right) = \frac{1}{2h^{\frac{1}{2}}}\left(x + \frac{\sqrt{h}\gamma}{2}f\left(x + \frac{\sqrt{h}\gamma}{2}\right)\right) \quad (15)$$

Proof:

$$\begin{aligned} \text{var}(\hat{f}_{\text{Gumbel}}(x)) &= \frac{1}{n} \text{var} \left[K_{\text{Gumbel}\left(x, \frac{1}{h^2}\right)}(X_i) \right] \\ &= n^{-1} \left[E \left[K_{\text{Gumbel}\left(x, \frac{1}{h^2}\right)}^2(X_i) \right] \right] + o(n^{-1}) \end{aligned}$$

Let η_x be a Gumbel $\left(x, \frac{1}{h^2}\right)$ distributed random variable. Hence

$$\mu_x = E[\eta_x] = x + \frac{\sqrt{h\gamma}}{2} \text{ and } V_x = \text{var}[\eta_x] = \frac{h\pi^2}{24}. \text{ We have}$$

$$E \left[K_{\text{Gumbel}\left(x, \frac{1}{h^2}\right)}^2(X_i) \right] = J_h E[\eta_x f(\eta_x)],$$

where, $J_h = \frac{1}{2h^2}$. By Taylor expansion of $\eta_x f(\eta_x)$ we get:

$$\begin{aligned} \eta_x f(\eta_x) &= \eta_x f(\eta_x) + [(\eta_x f'(\eta_x) + f(\eta_x))](\eta_x - \mu_x) \\ &\quad + \frac{1}{2}[\eta_x f''(\eta_x) + 2f'(\eta_x)](\eta_x - \mu_x)^2 + o(h^{-2}) \end{aligned}$$

So,

$$\begin{aligned} E[\eta_x f(\eta_x)] &= x + \frac{\sqrt{h\gamma}}{2} f\left(x + \frac{\sqrt{h\gamma}}{2}\right) \\ &\quad + \frac{1}{2} \left[x + \frac{\sqrt{h\gamma}}{2} f''\left(x + \frac{\sqrt{h\gamma}}{2}\right) + 2f'\left(x + \frac{\sqrt{h\gamma}}{2}\right) \right] \frac{h\pi^2}{24} + o(h^{-2}) \\ &= x + \frac{\sqrt{h\gamma}}{2} f\left(x + \frac{\sqrt{h\gamma}}{2}\right) + h \left[\frac{x + \frac{\sqrt{h\gamma}}{2} f''\left(x + \frac{\sqrt{h\gamma}}{2}\right) \pi^2}{24} + \frac{f'\left(x + \frac{\sqrt{h\gamma}}{2}\right)}{12} \right] \\ &\quad + o(h^{-2}x^{-2}) \\ &= x + \frac{\sqrt{h\gamma}}{2} f\left(x + \frac{\sqrt{h\gamma}}{2}\right) + o(h^{-2}x^{-2}) \end{aligned}$$

Therefore,

$$\text{var}(\hat{f}_{\text{Gumbel}}(x)) = \frac{1}{2h^2} \left[x + \frac{\sqrt{h\gamma}}{2} f\left(x + \frac{\sqrt{h\gamma}}{2}\right) \right] + o(h^{-2}x^{-2})$$

Optimal Bandwidth

To proceed for optimal bandwidth, initially Mean Squared Error (MSE) and Mean Integrated Squared Error (MISE) are derived as;

As we know Mean square errors for Gumbel kernel estimator is

$$\begin{aligned} \text{MSE}[\hat{f}_{\text{Gumbel}}(x)] &= \text{Bias}^2[\hat{f}_{\text{Gumbel}}(x)] + \text{var}[\hat{f}_{\text{Gumbel}}(x)] \\ &= h \left(\left(\gamma f'(x) + \frac{\sqrt{h}\pi^2}{12} f''(x) \right)^2 \right) + \frac{1}{2\sqrt{h}} \left[x + \frac{\sqrt{h}\gamma}{2} f \left(x + \frac{\sqrt{h}\gamma}{2} \right) \right] \end{aligned}$$

We can approximate MISE to be:

$$\text{MISE}[\hat{f}_{\text{Gumbel}}(x)] = hA + \frac{1}{2h}B \tag{16}$$

where, $A = \int \left(\gamma f'(x) + \frac{\sqrt{h}\pi^2}{12} f''(x) \right)^2 dx$ and $B = \int x + \frac{\sqrt{h}\gamma}{2} f \left(x + \frac{\sqrt{h}\gamma}{2} \right) dx$.

To find optimal bandwidth, now we minimize Equation (16) with respect to h , so we have

$$\begin{aligned} \frac{d\text{MISE}[\hat{f}_{\text{Gumbel}}(x)]}{dh} &= A - \frac{h^{-\frac{3}{2}}}{4} B \tag{17} \\ \frac{d^2\text{MISE}[\hat{f}_{\text{Gumbel}}(x)]}{dh^2} &= \frac{3}{8h^{\frac{5}{2}}} B > 0 \end{aligned}$$

Setting (17) equal zero yields an optimal bandwidth h_{opt} for the given *pdf* and kernel:

$$h_{opt} = \left[\frac{4 \int \left(\gamma f'(x) + \frac{\sqrt{h}\pi^2}{12} f''(x) \right)^2 dx}{\int x + \frac{\sqrt{h}\gamma}{2} f \left(x + \frac{\sqrt{h}\gamma}{2} \right) dx} \right]^{-\frac{2}{3}}$$

4. Applications

In this section, the performance of the proposed estimators in estimating the *pdf* is observed through real life data as well as by simulation study.

4.1. Suicide Data Example

We take suicide data given in Silverman [5] to inspect the performance of new developed kernel. The data gives the lengths of the treatment spells (in days) of control patients in suicide study.

We used the logarithm of the data to draw **Figure 1** using data driven bandwidth, named as normal scale rule (NSR) by Silverman [5]. The NSR is given by;

$$h_{NSR} = 0.79Rn^{-\frac{1}{5}}, \tag{18}$$

where R is inter-quartile range, which results in 0.4894. It can be observed that Gumbel kernel performed very well, especially near end points and free of boundary bias.

4.2. Flood Data Example

Further, we used the flood data given in Gumbel [14], to exhibit the practical performance of the Gumbel estimator. The data give the discharge per second of the Rhone River (Europe).

Here fixed bandwidth which is 1,000,000, is used. **Figure 2** shows that the performance of new proposed kernel estimator, which is acceptable.

4.3. Simulation Study

In this section we wish to investigate the finite sample properties of the two asymmetric kernel estimators; Gumbel and Weibull, which belong to family of extreme value distributions. The experiments are based on 1000 random samples of length $n = 3^5 = 243$, $n = 486$ and $n = 972$. For each simulated sample and each estimator considered, mean squared errors (MSE) are reported in **Table 1**, for extreme value distributions, namely Frechet, Weibull and Gumbel distributions and various parameter values by using bandwidth given as [8].

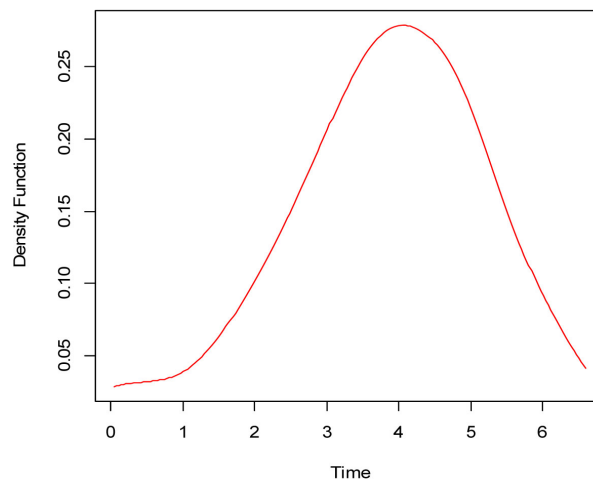


Figure 1. The Gumbel kernel estimator for the suicide data.

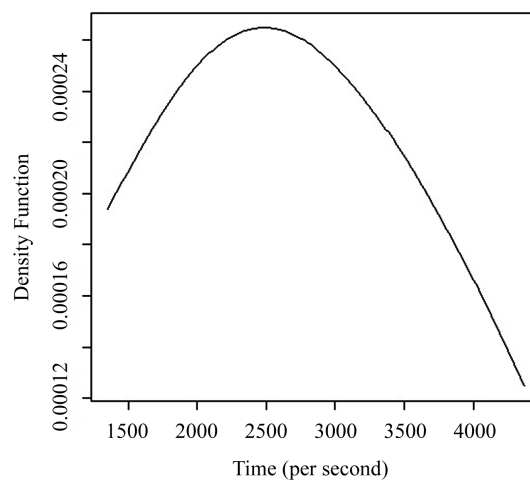


Figure 2. The Gumbel kernel estimator for the flood data.

Table 1. Mean square errors.

	Gumbel	Weibull
<i>n</i> = 243		
Gumbel (3, 1)	0.0017	0.0054
Gumbel (10, 1)	0.0026	0.0213
Gumbel (100, 1)	0.0013	0.0279
Frechet (3, 1, 1)	0.0009	0.0086
Frechet (5, 1, 1)	0.0015	0.0140
Frechet (10, 1, 1)	0.0006	0.0098
Weibull (25.713, 1)	7.0497e-05	0.0001
Weibull (76.318, 1)	8.9387e-06	6.8821e-06
Weibull (350.832, 1)	1.0607e-06	2.0332e-06
<i>n</i> = 486		
Gumbel (3, 1)	0.0007	0.0034
Gumbel (10, 1)	0.0015	0.0127
Gumbel (100, 1)	0.0009	0.0342
Frechet (3, 1, 1)	0.0006	0.0071
Frechet (5, 1, 1)	0.0008	0.0074
Frechet (10, 1, 1)	0.0011	0.0126
Weibull (25.713, 1)	9.7541e-05	0.0006
Weibull (76.318, 1)	8.4365e-06	7.2863e-06
Weibull (350.832, 1)	4.7844e-07	5.0882e-07
<i>n</i> = 972		
Gumbel (3, 1)	0.0010	0.0021
Gumbel (10, 1)	0.0014	0.0137
Gumbel (100, 1)	0.0012	0.0241
Frechet (3, 1, 1)	0.0006	0.0056
Frechet (5, 1, 1)	0.0005	0.0052
Frechet (10, 1, 1)	0.0005	0.0074
Weibull (25.713, 1)	8.5902e-05	0.0007
Weibull (76.318, 1)	8.1771e-06	5.4514e-06
Weibull (350.832, 1)	4.15146e-07	3.1866e-08

Here in **Table 1**, variety of randomly selected location parameter (small/medium/large) is examined with constant scale parameter. We may observe that the Gumbel kernel estimator performs better than Weibull kernel estimator un-animously almost for all density estimates with all different parameters and different sample sizes. For both Gumbel and Weibull kernel, MSEs decreased as sample size increased. In graphical representation, we present Gumbel kernel with real density. It can be examined in **Figure 3** that the performance of the

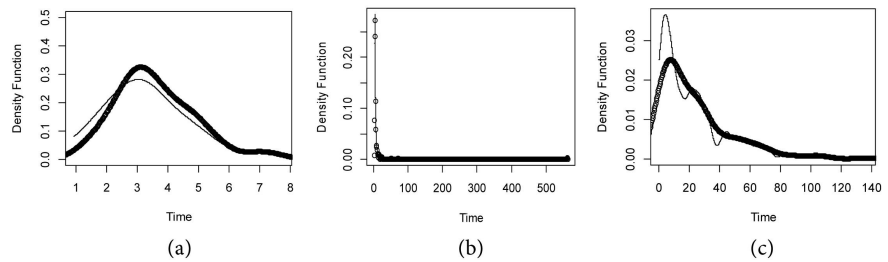


Figure 3. The Gumbel kernel estimator of the density functions with different distributions. (where solid line shows the real density and other line represents the density estimated by Gumbel kernel). (a) Gumbel (3, 1) (b) Frechet (3, 1, 1) (c) Weibull (25.713, 1).

Gumbel estimator is acceptable at the boundary near the zero with different densities. In the interior, the behavior of the *pdf* estimator becomes more similar as we get away from zero in any extreme value distribution case.

5. Conclusion

In this paper, we have proposed a new kernel estimator for probability density functions for (*iid*) data $[0, \infty)$, namely Gumbel kernel. Such densities are encountered in a wide variety of applications to describe extreme wind speeds, sea wave heights, floods, rainfall, age at death, minimum temperature, rainfall during droughts, electrical strength of materials, air pollution problems, geological problems, naval engineering etc. [4]. Gumbel kernel is free of boundary bias, non-negative, with natural varying shape. We showed that the bias depends on the smoothing parameter h and the estimated point x , and it goes to zero as $h \rightarrow 0$, also it gets smaller for the values of x closed to zero. The variance of the new proposed kernel estimator was investigated, and we noticed that it depends also on h and x . On the other hand, it goes to zero as $h \rightarrow 0$, and gets large at the values of x close to zero.

In addition, the performance of the proposed estimators is tested in three applications. In a simulation study, we used different densities of GEV distribution and compared it with Weibull (Extreme value distribution III) kernel estimator on basis of MSE. We observed that the performance of the proposed estimator is excellent, and gives a smaller MSE. Additionally, by using real data examples, we exhibited the practical performance of the new estimator.

From the above discussion, it can be concluded that one of the reason for adaptation of nonparametric method was to control the allocation of weights at boundary points. But boundary bias is still present if symmetrical kernels are used for curve estimation. In this situation, best alternative is to use asymmetrical kernel and Gumbel kernel is finest selection than Weibull kernel, comparatively.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Adamowski, K. and Labatiuk, C. (1987) Estimation of Flood Frequencies by a Non-parametric Density Procedure. In: Singh, V.P., ed., *Hydrologic Frequency Modeling*, Springer, Dordrecht. https://doi.org/10.1007/978-94-009-3953-0_5
- [2] Jou, P.H., Akhoond-Ali, A.M., Behnia, A. and Chinipardaz, R. (2008) Parametric and Nonparametric Frequency Analysis of Monthly Precipitation in Iran. *Journal of Applied Sciences*, **8**, 3242-4348. <https://doi.org/10.3923/jas.2008.3242.3248>
- [3] Cid, J.A and von Davier, A.A. (2015) Examining Potential Boundary Bias Effects in Kernel Smoothing on Equating: An Introduction for the Adaptive and Epanechnikov Kernels. *Applied Psychological Measurement*, **39**, 208-222. <https://doi.org/10.1177/0146621614555901>
- [4] Schuster, E.F. (1985) Incorporating Support Constraints into Nonparametric Estimation of Densities. *Communications in Statistics—Theory and Methods*, **14**, 1123-1136. <https://doi.org/10.1080/03610928508828965>
- [5] Silverman, B.W. (1986) Density Estimation. Chapman and Hall/CRC, London. https://doi.org/10.1007/978-1-4899-3324-9_6
- [6] Eubank, R.L. and Speckman, P. (1990) Curve Fitting by Polynomial-Trigonometric Regression. *Biometrika*, **77**, 1-9. <https://doi.org/10.1093/biomet/77.1.1>
- [7] Chen, S.X. (2000) Beta Kernel Smoothers for Regression Curves. *Statistica Sinica*, **10**, 73-91.
- [8] Chen, S.X. (2000) Probability Density Function Estimation Using Gamma Kernels. *Annals of the Institute of Statistical Mathematics*, **52**, 471-480. <https://doi.org/10.1023/A:1004165218295>
- [9] Jin, X. and Kawczak, J. (2003) Birnbaum-Saunders and Lognormal Kernel Estimators for Modeling Durations in High Frequency Financial Data. *Annals of Economics and Finance*, **4**, 103-124.
- [10] Scaillet, O. (2004) Density Estimation Using Inverse and Reciprocal Inverse Gaussian Kernels. *Nonparametric Statistics*, **16**, 217-226. <https://doi.org/10.1080/10485250310001624819>
- [11] Salha, R.B., Ahmed, E.S. and Alhoubi, I.M. (2014) Hazard Rate Function Estimation Using Weibull Kernel. *Open Journal of Statistics*, **4**, 650-661. <https://doi.org/10.4236/ojs.2014.48061>
- [12] McFadden, D. (1978) Modeling the Choice Residential Location. *Transportation Research Record*, **673**, 72-77.
- [13] Salha, R.B., Ahmed, E.S. and Alhoubi, I.M. (2014) Hazard Rate Function Estimation Using Erlang Kernel. *Pure Mathematical Sciences*, **3**, 141-152. <https://doi.org/10.12988/pms.2014.4616>
- [14] Gumbel, E.J. (1941) The Return Period of Flood Flows. *The Annals of Mathematical Statistics*, **12**, 163-190. <https://doi.org/10.1214/aoms/1177731747>