A note on a Parzen–Rosenblatt type density estimator for circular data

Carlos Tenreiro*
CMUC, Department of Mathematics, University of Coimbra, Coimbra, Portugal

27 May 2024

Abstract

Using the close connection between the Parzen–Rosenblatt estimator for linear data and the recently proposed Parzen–Rosenblatt type estimator for circular data, we establish some asymptotic properties of this last estimator, such as asymptotic unbiasedness, weak and strong pointwise consistency, and weak and strong uniform consistency.

KEYWORDS: circular data; density estimation; Parzen–Rosenblatt type estimator.

AMS 2020 SUBJECT CLASSIFICATIONS: 62H11, 62G07

*ORCID ID: 0000-0002-5495-6644. E-mail: tenreiro@mat.uc.pt URL: http://www.mat.uc.pt/∼tenreiro/ Postal address: Department of Mathematics, University of Coimbra, Largo D. Dinis, 3000-143 Coimbra, Portugal.

1 Introduction

Given an independent and identically distributed sample of angles $X_1, \ldots, X_n \in [0, 2\pi[$ from some absolutely continuous random variable X with unknown probability density function f, the standard kernel estimator of f is defined, for $\theta \in [0, 2\pi[$, by

$$\tilde{f}_n(\theta; L, g) = \frac{c_g(L)}{n} \sum_{i=1}^{n} L\bigg(\frac{1 - \cos(\theta - X_i)}{g^2}\bigg), \tag{1}$$

where $L : [0, \infty[\to \mathbb{R}$ is a bounded function satisfying some additional conditions, $g = g_n$ is a sequence of strictly positive numbers such that $g \to 0$, as $n \to \infty$, and $c_g(L)$, depending on the kernel L and the bandwidth g, is chosen so that $\tilde{f}_n(\cdot; g)$ integrates to unity. Kernel estimators of this form for estimating densities of q-dimensional unit spheres, for $q \geq 1$, were initially studied in Beran (1979), Hall, Watson, and Cabrera (1987), Bai, Rao, and Zhao (1988) and Klemela (2000), the last paper being restricted to $q \geq 2$, and more recently by García-Portugués (2013) and García-Portugués, Crujeiras, and González-Manteiga (2013). Quite recently, an alternative kernel density estimator for circular data is proposed and studied in Tenreiro (2022, 2023). Such an estimator, which is close in spirit to the Parzen–Rosenblatt (PR) estimator for linear data (see Rosenblatt, 1956, and Parzen, 1962), is defined, for $\theta \in [0, 2\pi[$, by

$$\hat{f}_n(\theta; K, h) = \frac{d_h(K)}{n} \sum_{i=1}^{n} K_h(\theta - X_i), \tag{2}$$

where $h = h_n$ is a sequence of strictly positive real numbers converging to zero as n tends to infinity, K_h is a real-valued periodic function on $\mathbb{R}$, with period 2π, such that

$$K_h(\theta) = K(\theta/h)/h, \text{ for } \theta \in [-\pi, \pi[,$$

with K a kernel on $\mathbb{R}$, that is, a bounded and integrable real-valued function on $\mathbb{R}$ with $\int_{\mathbb{R}} K(u)du > 0$, and $d_h(K)$ is a normalizing constant depending on the kernel K and the bandwidth h which is chosen so that $\hat{f}_n(\cdot; h)$ integrates to unity. The periodicity imposed on K_h makes estimator (2) well adapted to deal with circular data by automatically correcting the potential boundary problems that may occur at the endpoints of the distribution support when the PR-estimator for linear data is used to estimate f. This can be observed in Figure 1 that shows the contribution to the estimator of an observation X for three different locations of such observation.

From a estimation point of view, the previous estimators are closely linked as established in Tenreiro (2022) (Section 3). In fact, the estimator (1) with kernel L and bandwidth g and the estimator (2) with kernel $K(u) = L(u^2)$ and bandwidth $h = \sqrt{2}g$ share the same first-order asymptotic terms for the corresponding mean integrated squared errors. Some simulation experiments, as those performed by taking as test densities the 20 circular models introduced and detailed described in Oliveira, Crujeiras, and Rodríguez-Casal (2012) (not shown here), have also revealed that the two estimators present similar finite sample behaviours for a large variety

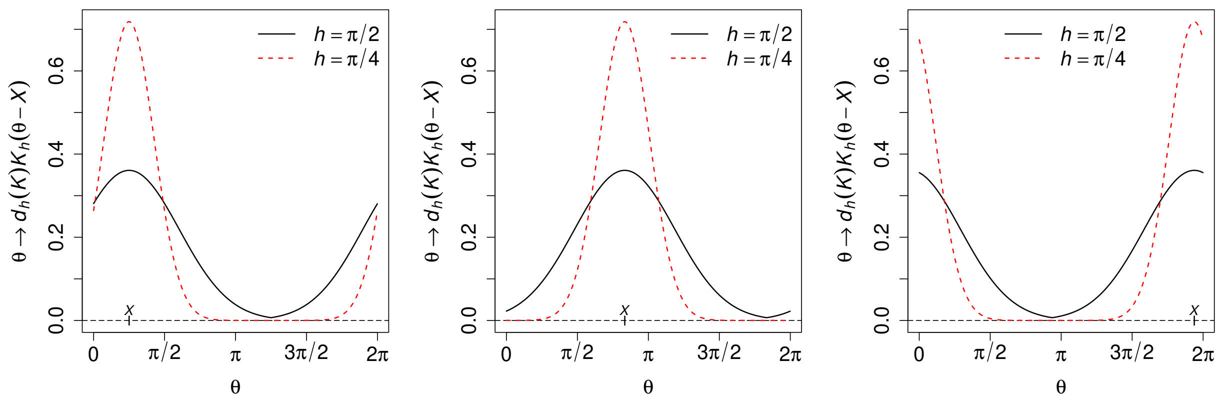

Figure 1: Kernel density estimates $\hat{f}_n(\cdot; K, h)$ based on observation X for $h = \pi/2$ and $h = \pi/4$ when K is the gaussian kernel $K(u) = \exp(-u^2)$.

of underlying distributions even when the sample size is small. The close connection between the two estimators is shown in Figure 2 where estimates for the dragonflies orientation density based on the orientation of 214 dragonflies with respect to the azimuth of the sun (see Batschelet, 1981, pp. 23–24) are computed for three different bandwidths. As most dragonflies have chosen a direction of approximately 90° either to the right or to the left of the sun's rays, the underlying circular density should be bimodal.

Despite the close connection between the standard kernel estimator (1) and the PR-type estimator (2), we think that it is worth drawing the attention of the reader to the simple structure of the PR-type estimator which allows us to derive, with a little extra effort, some of its main asymptotic properties by making use of techniques that are similar to those used, long ago, for the PR-estimator for linear data. This is the main goal of this paper. The rest of the article is as follows. In Section 2 we show that the PR-type estimator is asymptotically unbiased and we derive its asymptotic variance at every point at which f is continuous. For that we follow closely the approach of Parzen (1962) after noticing that the bias and variance of the PR-type estimator (2) at a point θ can be expressed in terms of the convolution between K_h and the circular density f. In Sections 3 and 4 we establish a set of sufficient conditions for the weak and strong pointwise consistency, and for the strong uniform consistency of the PR-type estimator. For that we follow the approaches used by Bhattacharya (1967) and Nadaraya (1965) for the PR-estimator for linear data. Finally, in Section 5, we draw some conclusions. The plots shown in this article were carried out using the R software R Development Core Team (2021).

2 Bias and variance

Taking into account Theorem 2.1 in Tenreiro (2022) and the fact that the class of delta sequence estimators studied in the previous work comprises the class of PR-type estimators, we know that under certain conditions on the kernel K the PR-type estimator $\hat{f}_n(\cdot; K, h)$ defined by (2)

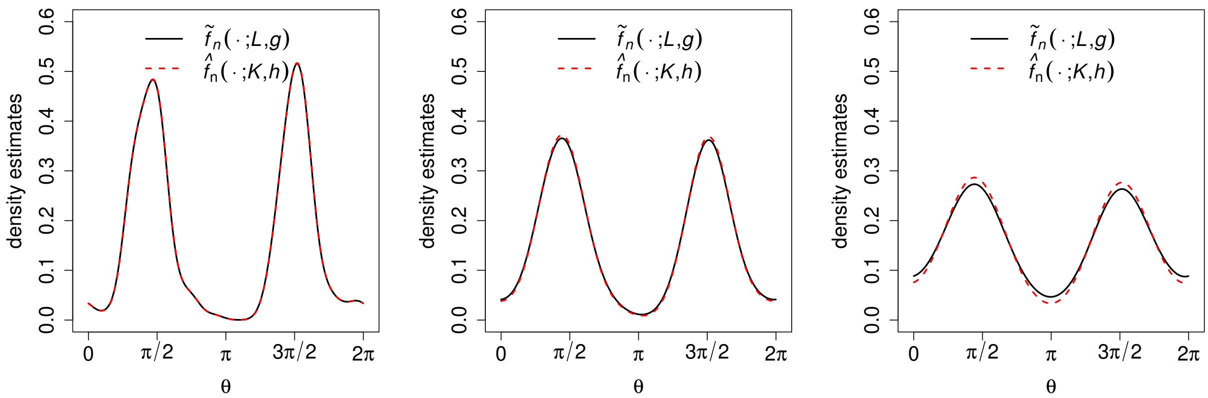

Figure 2: Estimates for the dragonflies orientation density by using estimators $\tilde{f}_n(\cdot; L, g)$ and $\hat{f}_n(\cdot; K, h)$ where $L(t) = \exp(-t)$, $K(u) = \exp(-u^2)$, $h = \sqrt{2}g$ and $g = 0.2$ (left), $g = 0.4$ (centre) and $g = 0.6$ (right).

is asymptotically unbiased and an expression for its asymptotic variance can be given when the probability density function f is continuous on $[0, 2\pi[$. In this section we establish the same asymptotic properties under less restrictive conditions on K and f. With this purpose in mind we employ an alternative and simpler approach that follows closely the one used by Parzen (1962) for the PR-estimator for linear data. For that, we begin by noticing that, similarly to the PR-estimator for linear data, the bias and variance of $\hat{f}_n(\cdot; K, h)$ at a point θ can be expressed in terms of the convolutions between K_h or $(K^2)_h$ and the circular density f. In fact, for all $\theta \in [0, 2\pi[$, we have

$$\mathrm{E}\hat{f}_n(\theta; K, h) = d_h(K) \int_0^{2\pi} K_h(\theta - u) f(u) du = d_h(K)(K_h * f)(\theta) \tag{3}$$

and

$$\mathrm{Var}\hat{f}_n(\theta; K, h) = (nh)^{-1} d_h(K)^2 \big(((K^2)_h * f)(\theta) - h(K_h * f)(\theta)^2 \big), \tag{4}$$

where for the sake of simplicity we are also denoting by f the periodic extension of f to $\mathbb{R}$ given by $f(\theta) = f(\theta - 2k\pi)$, whenever $\theta \in [2k\pi, 2(k+1)\pi[$, for some $k \in \mathbb{Z}$ (see Tenreiro, 2023, Equations (10) and (11)). Recall that if α and β are real-valued functions with period 2π defined on $\mathbb{R}$, the convolution of α and β is defined, for $x \in \mathbb{R}$, by

$$(\alpha * \beta)(x) = \int_0^{2\pi} \alpha(x - y)\beta(y) dy,$$

whenever this integral exists. As the integrand is periodic with period 2π, the previous definition does not depend on the considered interval of integration with length 2π. The convolution $(\alpha * \beta)(x)$ exists for almost every $x \in \mathbb{R}$ whenever α and β are integrable on $[0, 2\pi[$, and it exists for every $x \in \mathbb{R}$ if one of the functions α or β is bounded. Moreover, it exists and is continuous for every $x \in \mathbb{R}$, whenever α and β are square integrable on $[0, 2\pi[$. Obviously, the convolution is a periodic function if it exists (see Butzer and Nessel, 1971, §0.4).

The following result is a version of the Bochner's lemma for real-valued periodic functions with period 2π (Bochner, 1955, p. 2; see also Parzen, 1962, p. 1067, and Bosq and Lecoutre, 1987, p. 61). It is of crucial importance to deal with the convolutions that appear in the equations (3) and (4).

Lemma 1. *For $h > 0$, let φ_h be a real-valued periodic function on $\mathbb{R}$ with period 2π such that,*

$$\varphi_h(\theta) = \varphi(\theta/h)/h, \text{ for } \theta \in [-\pi, \pi[,$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is bounded and integrable on $\mathbb{R}$, and let $g : \mathbb{R} \to \mathbb{R}$ a periodic function with period 2π.

a) If g is bounded on $[0, 2\pi[$, or, in alternative, if g is integrable on $[0, 2\pi[$ and φ is such that $\lim_{|u|\to\infty} u\varphi(u) = 0$, then

$$(\varphi_h * g)(\theta) \to g(\theta) \int_{\mathbb{R}} \varphi(u)du, \; h \to 0,$$

at every point $\theta \in [0, 2\pi[$ at which g is continuous.

b) If g is continuous on $[0, 2\pi[$, then the previous convergence is uniform on $[0, 2\pi[$.

Proof: By using standard arguments, for every $h > 0$ and $0 < \delta < \pi$ we have

$$\left|(\varphi_h * g)(\theta) - g(\theta) \int_{\mathbb{R}} \varphi(u)du\right| \leq \sup_{|z|\leq\delta} |g(\theta - z) - g(\theta)| \int_{\mathbb{R}} |\varphi(u)|du + 3||g||_\infty \int_{|u|>\delta/h} |\varphi(u)|du,$$

whenever g is bounded on $[0, 2\pi[$, where we denote $||g||_\infty = \sup_{x\in\mathbb{R}} |g(x)|$, and

$$\left|(\varphi_h * g)(\theta) - g(\theta) \int_{\mathbb{R}} \varphi(u)du\right| \leq \sup_{|z|\leq\delta} |g(\theta - z) - g(\theta)| \int_{\mathbb{R}} |\varphi(u)|du + \delta^{-1} \sup_{|u|>\delta/h} |u\varphi(u)| \int_0^{2\pi} |g(u)|du + 2|g(\theta)| \int_{|u|>\delta/h} |\varphi(u)|du,$$

whenever g is integrable on $[0, 2\pi[$. The stated result follows from the previous inequalities. ■

Taking into account that

$$d_h(K)^{-1} = \int_{-\pi/h}^{\pi/h} K(y)dy \to \int_{\mathbb{R}} K(y)dy, \text{ as } h \to 0, \tag{5}$$

the following result is a consequence of Lemma 1 and equations (3) and (4).

Theorem 1. *Let f be bounded on $[0, 2\pi[$, or, in alternative, let K be such that $\lim_{|u|\to\infty} uK(u) = 0$. If $h \to 0$ then the* PR*-type estimator is asymptotically unbiased, i.e.,*

$$\lim_{n\to\infty} \mathrm{E}\hat{f}_n(\theta; K, h) = f(\theta),$$

and has a variance satisfying

$$\lim_{n\to\infty} nh\mathrm{Var}\hat{f}_n(\theta;K,h) = f(\theta)\Big(\int_{\mathbb{R}} K(u)du\Big)^{-2}\int_{\mathbb{R}} K(u)^2du,$$

at all points $\theta \in [0, 2\pi[$ *at which* f *is continuous. Moreover, if* f *is continuous on* $[0, 2\pi[$*, then the previous limits are uniform on* $[0, 2\pi[$.

As we can conclude from Parzen (1962, pp. 1067, 1069), the previous asymptotic variance coincides with that of the PR-estimator for linear data when K is such that $\int_{\mathbb{R}} K(u)du = 1$.

3 Pointwise consistency

Taking into account that the mean squared error of the PR-type estimator may be written as

$$\mathrm{E}\big(\hat{f}_n(\theta;K,h) - f(\theta)\big)^2 = \mathrm{Var}\hat{f}_n(\theta;K,h) + \big(\mathrm{E}\hat{f}_n(\theta;K,h) - f(\theta)\big)^2, \tag{6}$$

under the conditions of Theorem 1 we conclude that $\hat{f}_n(\cdot;K,h)$ is consistent in quadratic mean at all points $\theta \in [0, 2\pi[$ at which f is continuous if in addition to satisfying $h \to 0$, the bandwidth also satisfies the classical condition $nh \to \infty$. By making use of an exponential inequality, we prove in this section that under stronger conditions on the bandwidth h the PR-type estimator is strongly consistent at every point of continuity of f. Apparently, the use of a similar exponential inequality in the context of the almost sure convergence of the PR-estimator for linear data goes back to Bhattacharya (1967). The proof we give is similar to that of Theorem 1 in Bai et al. (1988) where the strong pointwise consistency of the standard kernel estimator (1) is established.

Theorem 2. *Under the general conditions of Theorem 1, if* $h \to 0$ *and*

$$\sum_{n=1}^{\infty} \exp(-\gamma nh) < \infty, \text{ for all } \gamma > 0,$$

we have

$$\hat{f}_n(\theta;K,h) \xrightarrow{a.s.} f(\theta),$$

at every point $\theta \in [0, 2\pi[$ *of continuity of* f.

Proof: From Theorem 1, it is enough to prove that $\hat{f}_n(\theta;K,h) - \mathrm{E}\hat{f}_n(\theta;K,h) \xrightarrow{a.s.} 0$, at every point $\theta \in [0, 2\pi[$ of continuity of f. We have

$$\hat{f}_n(\theta;K,h) - \mathrm{E}\hat{f}_n(\theta;K,h) = \frac{1}{n}\sum_{i=1}^{n} Z_{i,n},$$

where $Z_{i,n} = d_h(K)\big(K_h(\theta - X_i) - \mathrm{E}K_h(\theta - X_i)\big)$ is such that

$$|Z_{i,n}| \le 2h^{-1}|d_h(K)|\,||K||_\infty$$

and

$$E(Z_{i,n}^2) \leq h^{-1} d_h(K)^2 ((K^2)_h * f)(\theta).$$

Therefore, using Lemma 1 and the equation (5) we conclude that there exist constants $C, D > 0$ such that $|Z_{i,n}| \leq Ch^{-1}$ and $E(Z_{i,n}^2) \leq Dh^{-1}$ for n large enough. Using now the Theorem 3 of Hoeffding (1963), we deduce that

$$P\big(|\hat{f}_n(\theta; K, h) - \mathrm{E}\hat{f}_n(\theta; K, h)| \geq \epsilon\big) \leq 2\exp(-c\epsilon^2 nh/D),$$

for all $\epsilon \in]0, D/C[$ with $c = 2\log 2 - 1$. The stated result follows now from the Borel–Cantelli lemma. ■

4 Global consistency

The overall quality of a nonparametric density estimator can be assessed through different performance measures. Making use of the equalities (3), (4) and (6), Tenreiro (2023) (Theorem 2) established that the PR-type estimator is a mean integrated square error consistent estimator of f, whenever the bandwidth h is such that $h \to 0$ and $nh \to \infty$, that is,

$$\mathrm{E} \int_0^{2\pi} \big\{\hat{f}_n(\theta; K, h) - f(\theta)\big\}^2 d\theta \to 0,$$

for all square integrable density f on $[0, 2\pi[$. In the following result, which is inspired by the corresponding result obtained by Nadaraya (1965) for the PR-estimator for linear data, we establish sufficient conditions for the almost sure uniform convergence of the PR-type estimator. The kernel K is assumed to be of bounded variation on $\mathbb{R}$ (see Cohn, 1980, §4.4), a condition that holds for the normal kernel $K(u) = \exp(-u^2)$, and for many other kernels such as the polynomial kernels $K(u) = (1 - u^2)^p \mathbb{1}(|u| \leq 1)$, with $p \in \mathbb{N}$.

Theorem 3. *Let K be of bounded variation on $\mathbb{R}$ and f continuous on $[0, 2\pi[$. If $h \to 0$ and*

$$\sum_{n=1}^{\infty} \exp(-\gamma n h^2) < \infty, \text{ for all } \gamma > 0, \tag{7}$$

we have

$$\sup_{\theta \in [0, 2\pi[} \big|\hat{f}_n(\theta; K, h) - f(\theta)\big| \xrightarrow{a.s.} 0.$$

Proof: From Theorem 1, it is enough to prove that $\sup_{\theta \in [0, 2\pi[} \big|\hat{f}_n(\theta; K, h) - \mathrm{E}\hat{f}_n(\theta; K, h)\big| \xrightarrow{a.s.} 0$. For $\theta \in [0, 2\pi[$, we have

$$\hat{f}_n(\theta; K, h) = \frac{d_h(K)}{h} \int_0^{2\pi} L_{\theta,h}(y) dF_n(y)$$

and

$$\mathrm{E}\hat{f}_n(\theta; K, h) = \frac{d_h(K)}{h} \int_0^{2\pi} L_{\theta,h}(y) dF(y),$$

where F_n is the empirical distribution function associated to the observations $X_1, \ldots, X_n$, F is the distribution function of X, and, for $y \in \mathbb{R}$,

$$\begin{aligned} L_{\theta,h}(y) &= K\left(\frac{\theta - y - 2\pi}{h}\right) \mathbb{1}_{]-\infty,\theta-\pi]}(y) \\ &\quad + K\left(\frac{\theta - y}{h}\right) \mathbb{1}_{]\theta-\pi,\theta+\pi]}(y) \\ &\quad + K\left(\frac{\theta - y + 2\pi}{h}\right) \mathbb{1}_{]\theta+\pi,+\infty[}(y). \end{aligned}$$

As K is of bounded variation on $\mathbb{R}$ and integrable, the function $L_{\theta,h}$ is also of bounded variation on $\mathbb{R}$ with $\lim_{y\to-\infty} L_{\theta,h}(y) = 0$. Moreover, we have

$$V_{L_{\theta,h}}(\mathbb{R}) \le 3V_K(\mathbb{R}) + 4||K||_\infty,$$

where $V_g(\mathbb{R})$ denotes the variation on $\mathbb{R}$ of the real-valued function g. On integrating by parts (see Cohn, 1980, pp. 163–164), we find

$$\hat{f}_n(\theta; K, h) = L^+_{\theta,h}(2\pi) - \frac{d_h(K)}{h} \int_0^{2\pi} F_n(y^-) d\mu_{\theta,h}(y) \quad a.s.$$

and

$$\mathrm{E}\hat{f}_n(\theta; K, h) = L^+_{\theta,h}(2\pi) - \frac{d_h(K)}{h} \int_0^{2\pi} F(y^-) d\mu_{\theta,h}(y),$$

where $L^+_{\theta,h}$ is the right-continuous function given by $L^+_{\theta,h}(y) = L_{\theta,h}(y^+)$, and $\mu_{\theta,h}$ is the signed measure induced by $L^+_{\theta,h}$ on $\mathscr{B}(\mathbb{R})$. Therefore, for $\theta \in [0, 2\pi[$, we have

$$\begin{aligned} \left|\hat{f}_n(\theta; K, h) - \mathrm{E}\hat{f}_n(\theta; K, h)\right| &\le \frac{|d_h(K)|}{h} \int_0^{2\pi} |F_n(y^-) - F(y^-)| d|\mu_{\theta,h}|(y) \\ &\le \frac{|d_h(K)|}{h} ||F_n - F||_\infty |\mu_{\theta,h}|(\mathbb{R}), \end{aligned}$$

where $|\mu_{\theta,h}|$ is the variation of $\mu_{\theta,h}$ and

$$|\mu_{h,\theta}|(\mathbb{R}) = V_{L^+_{\theta,h}}(\mathbb{R}) \le V_{L_{\theta,h}}(\mathbb{R}) \le 3V_K(\mathbb{R}) + 4||K||_\infty.$$

Finally we get

$$\sup_{\theta\in[0,2\pi[} \left|\hat{f}_n(\theta; K, h) - \mathrm{E}\hat{f}_n(\theta; K, h)\right| \le C_K \frac{1}{h} ||F_n - F||_\infty, \quad a.s.,$$

where the constant C_K depends only on the kernel K. The stated result follows now from the Borel–Cantelli lemma and the exponential inequality

$$P\big(||F_n - F||_\infty \ge \epsilon\big) \le 2\exp(-2\epsilon^2 n), \tag{8}$$

which is valid for all $\epsilon > 0$ and $n \in \mathbb{N}$ (see Massart, 1990). ∎

From the previous proof, we also conclude that, together with the condition $h \to 0$, the condition

$$nh^2 \to \infty \tag{9}$$

is sufficient for the weak uniform consistency of the PR-type estimator, that is,

$$\sup_{\theta \in [0,2\pi[} \left| \hat{f}_n(\theta; K, h) - f(\theta) \right| \xrightarrow{p} 0,$$

for K and f under the conditions of Theorem 3. Note also that if the law of the iterated logarithm

$$\limsup_{n \to \infty} \sqrt{\frac{2n}{\log \log n}} \, ||F_n - F||_\infty \leq 1, \quad a.s.$$

(see van der Vaart, 1998, p. 268) is used instead of the exponential inequality (8) in the proof of Theorem 3, we easily conclude that it also holds whenever the condition (7) on the bandwidth is replaced by

$$\frac{nh^2}{\log \log n} \to \infty. \tag{10}$$

Nevertheless, any of the conditions (7), (9) and (10) is more restrictive than the condition

$$\frac{nh}{\log n} \to \infty,$$

that, together with $h \to 0$, are, for suitable kernels, sufficient for the strong uniform consistency of the standard kernel estimator (1) as established in Theorem 2 of Bai et al. (1988), and necessary and sufficient for the weak and strong uniform consistency of the PR-estimator for linear data (see Bertrand-Retali, 1978, Silverman, 1978, and Devroye and Wagner, 1980). Using a more refined proof technique as that followed in the last bibliographic reference, we conjecture that it will be possible to show that, for suitable kernels, the conditions $h \to 0$ and $nh/\log n \to \infty$ are sufficient for the strong uniform consistency of the PR-type estimator.

5 Conclusion

Taking advantage of the close connection between the well-known PR-estimator for linear data and the recently proposed PR-type estimator for circular data defined by (2), in this paper we have successfully explored the possibility of establishing asymptotic properties of this last estimator (asymptotic unbiasedness, weak and strong pointwise consistency, and weak and strong uniform consistency) by making use of proof techniques originally proposed for the PR-estimator for linear data.

Funding

Research partially supported by the Centre for Mathematics of the University of Coimbra (funded by the Portuguese Government through FCT/MCTES, doi: 10.54499/UIDB/00324/2020).

References

Bai, Z.D., Rao, C.R., Zhao, L.C. (1988). Kernel estimators of density function of directional data. *J. Multivariate Anal.* 27, 24–39.

Batschelet, E. (1981). *Circular statistics in biology.* London: Academic Press.

Beran, R. (1979). Exponential models for directional data. *Ann. Statist.* 7, 1162–1178.

Bertrand-Retali, M. (1978). Convergence uniforme d'un estimateur de la densité par la méthode du noyau. *Rev. Roumaine Math. Pures Appl.* 13, 361–385.

Bhattacharya, P.K. (1967). Estimation of a probability density function and its derivatives. *Sankhya Ser. A* 29, 373–382.

Bochner, S. (1955). *Harmonic analysis and the theory of probability.* Berkeley: University of California Press.

Bosq, D., Lecoutre, J.-P. (1987). *Théorie de l'estimation fonctionnelle.* Paris: Economica.

Butzer, P.L., Nessel, R.J. (1971). *Fourier analysis and approximation. One-dimensional theory.* Basel: Birkhauser.

Cohn, D.L. (1980). *Measure theory.* Boston: Birkhäuser.

Devroye, L., Wagner, T.J. (1980). The strong uniform consistency of kernel density estimates. In *Multivariate Analysis V*, Krishnainh, P.R. (Ed.), North-Holland, New York, 59–77.

Dvoretzky, A., Kiefer, J., Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* 27, 642–669.

García-Portugués, E. (2013). Exact risk improvement of bandwidth selectors for kernel density estimation with directional data. *Electron. J. Stat.* 7, 1655–1685.

García-Portugués, E., Crujeiras, R.M., González-Manteiga, W. (2013). Kernel density estimation for directional-linear data. *J. Multivariate Anal.* 121, 152–175.

Hall, P., Watson, G.S., Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika* 74, 751–762.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* 58, 13–30.

Klemela, J. (2000). Estimation of densities and derivatives of densities with directional data. *J. Multivariate Anal.* 73, 18–40.

Massart, P. (1990). The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.* 3, 1269–1283.

Nadaraya, E.A. (1965). On non-parametric estimates of density functions and regression curves. *Theory Probab. Appl.* 10, 186–190.

Oliveira, M., Crujeiras, R.M., Rodríguez-Casal, A. (2012). A plug-in rule for bandwidth selection in circular density estimation. *Comp. Stat. Data Anal.* 56, 3898–3908.

Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* 33, 1065–1076.

R Development Core Team (2021). *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.* 27, 832–837.

Silverman, B.W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.* 6, 177–189.

Tenreiro, C. (2022). Kernel density estimation for circular data: a Fourier series-based plug-in approach for bandwidth selection. *J. Nonparametr. Stat.*, 34, 377–406.

Tenreiro, C. (2023). A Parzen-Rosenblatt type density estimator for circular data: exact and asymptotic optimal bandwidths. *Comm. Statist. Theory Methods* doi: 10.1080/03610926.2023.2264996

van der Vaart, A.W. (1998). *Asymptotic statistics*. Cambridge: Cambrige University Press.