A note on inference for kernel estimators of density and regression functions

Ahnaf Rafi

July 14, 2025

1 Introduction

In this note, we explore inference techniques for kernel estimators of density and conditional mean regression functions. We consider pointwise inference, uniform inference results are left for the future, either in a different document or future versions of the present document. The key guiding principle of the results presented in the present note will be as follows. There are standard Gaussian inference results for kernel estimators of density and regression functions. In these, the standard error estimators require the use of the estimated density and regression functions. However, we can formulate more “intuitive” standard error estimators that utilize the fact that kernel estimators are sample averages. These also work in that they are consistent in the appropriate sense and ensure that resulting t-statistics are asymptotically standard normal.

2 Density Estimation

Assumption 2.1. $X, X_1, \ldots, X_n$ are iid random $\mathbb{R}^d$-vectors all with Lebesgue density f.

Assumption 2.2. $K : \mathbb{R}^d \to \mathbb{R}$ is a measurable function such that

(i) $\int |K(u)|^p \, \mathrm{d}u < \infty$ for each $p \in \{1, 2, 4\}$.

(ii) $\int K(u) \, \mathrm{d}u = 1$

(iii) $\lim_{\|u\| \to \infty} \|u\|^d K(u) = 0$.

Assumption 2.3. $\{h_n\}$ is a real sequence such that $h_n > 0$ for every $n \in \mathbb{N}$, $\lim_{n\to\infty} h_n = 0$, and $\lim_{n\to\infty} 1/\left(nh_n^d\right) = 0$.

Remark 2.1. It can be shown that if K is continuous, then Assumption 2.2 (i) for $p = 1$ and (iii) imply that K is bounded. Then Assumption 2.2 (i) for $p \in (1, \infty)$ are immediately implied since $|K(u)|^p \leq (\sup_{v \in \mathbb{R}^d} |K(v)|)^{p-1} |K(u)|$ for every $u \in \mathbb{R}^d$. For our purposes, it will be sufficient to assume Assumption 2.2 (i) for $p \in \{1, 2, 4\}$.

For K satisfying Assumption 2.2, define

$$K_h(u) := \frac{1}{h^d} K\left(\frac{u}{h}\right). \tag{2.1}$$

The kernel density estimator and its mean are

$$\widehat{f}_{n,h}(x) = \frac{1}{n} \sum_{i=1}^{n} K_h\left(X_i - x\right) \text{ and } f_h(x) = \mathrm{E}\left[K_h(X - x)\right] = \int f(x+u) K_h(u) \, \mathrm{d}u. \tag{2.2}$$

Note that the final equality in (2.2) follows from the usual change of variables formula. The following is a usual statement of consistency and asymptotic normality for $\widehat{f}_{n,h_n}(x)$.

Theorem 2.1. *Let the probability density f in Assumption 2.1 be continuous at $x \in \mathbb{R}^d$ and satisfy $f(x) > 0$, let K be a function satisfying Assumption 2.2, and let the sequence $\{h_n\}$ satisfy Assumption 2.3. Then as $n \to \infty$,*

$$\widehat{f}_{n,h_n}(x) \xrightarrow{\mathrm{p}} f(x) \quad \textit{and} \quad \sqrt{nh_n^d}\left(\widehat{f}_{n,h_n}(x) - f_{h_n}(x)\right) \rightsquigarrow \mathrm{N}\left(0, f(x) \int K(u)^2 \, \mathrm{d}u\right). \tag{2.3}$$

Proof of Theorem 2.1. See Section 2.1.1. □

In this note, we will avoid talking about asymptotic normality of $\widehat{f}_{n,h_n} - f(x)$ since the bias term $f_{h_n}(x) - f(x)$ is typically non-zero and requires slightly delicate handling after scaling by $\sqrt{nh_n^d}$. Instead, we focus on the centered quantity $\widehat{f}_{n,h_n} - f_{h_n}(x)$ to focus only on asymptotic normality and standard error estimation. To that end, the typical prescription for inference is based on an application of Slutsky's Theorem and direct use of (2.3): under the conditions of Theorem 2.1,

$$\begin{gathered}\widetilde{t}_n = \frac{\sqrt{nh_n^d}\left(\widehat{f}_{n,h_n}(x) - f_{h_n}(x)\right)}{\sqrt{\widetilde{V}_{n,h_n}(x)}} \rightsquigarrow \mathrm{N}(0,1), \\ \text{where} \quad \widetilde{V}_{n,h_n}(x) = \widehat{f}_{n,h_n}(x) \int K(u)^2 \, \mathrm{d}u.\end{gathered} \tag{2.4}$$

In this note, we follow an alternative route. In particular, $\widehat{f}_{n,h}$ is a sample average and so, we might suspect that usual t-statistic is asymptotically standard normal. A key object in our derivations will be the centered kernel sum

$$S_{n,h}(x) = \sum_{i=1}^{n} \left\{ K\left(\frac{X_i - x}{h}\right) - \mathrm{E}\left[K\left(\frac{X - x}{h}\right)\right]\right\}. \tag{2.5}$$

Note that

$$\widehat{f}_{n,h}(x) - f_h(x) = \frac{S_{n,h}(x)}{nh^d}. \tag{2.6}$$

Clearly since $X_i \stackrel{\text{iid}}{\sim} X$ (Assumption 2.1),

$$\mathrm{E}\left[S_{n,h}(x)\right] = 0 \quad \text{and} \quad \mathrm{Var}\left[S_{n,h}(x)\right] = n\Sigma_h(x), \tag{2.7}$$

where by (A.7) in Theorem A.4,

$$\begin{aligned}\Sigma_h(x) := \mathrm{Var}\left[K\left(\frac{X - x}{h}\right)\right] &= \mathrm{E}\left[K\left(\frac{X - x}{h}\right)^2\right] - \mathrm{E}\left[K\left(\frac{X - x}{h}\right)\right]^2 \\ &= h^d \int f(x + uh) K(u)^2 \, \mathrm{d}u - \left(h^d \int f(x + uh) K(u) \, \mathrm{d}u\right)^2.\end{aligned} \tag{2.8}$$

The following result shows that $S_{n,h_n}(x)/\sqrt{n\Sigma_{h_n}(x)}$ is asymptotically standard normal. The key technical tool is Liapunov's Central Limit Theorem (see for example Billingsley (1995, Theorem 27.3, p. 362) or Pollard (1984, Theorem 18 in Section 4 of Chapter III, p. 51)).

Theorem 2.2. *Let the probability density f in Assumption 2.1 be continuous at $x \in \mathbb{R}^d$ and satisfy $f(x) > 0$, let K be a function satisfying Assumption 2.2, and let the sequence $\{h_n\}$ satisfy Assumption 2.3. Then as $n \to \infty$,*

$$\frac{S_{n,h_n}(x)}{\sqrt{n\Sigma_{h_n}(x)}} \rightsquigarrow \mathrm{N}(0,1). \tag{2.9}$$

Proof of Theorem 2.2. See Section 2.1.2. □

Define

$$\begin{aligned} V_h(x) = \mathrm{Var}\left[K_h(X-x)\right] &= \mathrm{E}\left[K_h(X-x)^2\right] - \mathrm{E}\left[K_h(X-x)\right]^2 \\ &= \mathrm{E}\left[K_h(X-x)^2\right] - f_h(x)^2. \end{aligned}$$

Note that

$$V_h(x) = \frac{1}{h^{2d}}\Sigma_h(x).$$

Since $\widehat{f}_{n,h_n}$ is a sample average of the $K_h\left(X_i - x\right)$'s and $X_i \overset{\text{iid}}{\sim} X$, it follows that

$$\mathrm{Var}\left[\widehat{f}_{nh}(x)\right] = V_h(x)/n.$$

Furthermore,

$$\frac{\widehat{f}_{n,h_n}(x) - f_{h_n}(x)}{\sqrt{V_{h_n}(x)/n}} = \frac{S_{n,h_n}(x)}{\sqrt{n\Sigma_{h_n}(x)}} \rightsquigarrow \mathrm{N}(0,1), \tag{2.10}$$

by (2.9) in Theorem 2.2.

For a feasible implementation of (2.10), we require a consistent estimator (in a sense to be defined later on) of $V_h(x)$. To that end, a reasonable estimator of $V_h(x)$ is the sample variance of the $K_h(X_i - x)$'s:

$$\begin{aligned}\widehat{V}_{n,h}(x) &= \frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)^2 - \left(\frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} K_h(X_i - x)^2 - \widehat{f}_{n,h}(x)^2.\end{aligned} \tag{2.11}$$

Theorem 2.3 shows that using $\widehat{V}_{n,h}(x)$ in standardization achieves asymptotic standard normality for the resulting t-statistic.

Theorem 2.3. *Let the probability density f in Assumption 2.1 be continuous at $x \in \mathbb{R}^d$ and satisfy $f(x) > 0$, let K be a function satisfying Assumption 2.2, and let the sequence $\{h_n\}$ satisfy Assumption 2.3. Then as $n \to \infty$,*

$$\frac{\widehat{V}_{n,h_n}(x)}{V_{h_n}(x)} \xrightarrow{\mathrm{p}} 1 \quad \textit{and so} \quad \frac{\widehat{f}_{n,h_n}(x) - f_{h_n}(x)}{\sqrt{\widehat{V}_{n,h_n}(x)/n}} \rightsquigarrow \mathrm{N}(0,1). \tag{2.12}$$

Proof of Theorem 2.3. See Section 2.1.3. □

2.1 Proofs

2.1.1 Proof of Theorem 2.1

The consistency part of (2.3) follows from the asymptotic normality part of (2.3) combined with Theorem A.4. To see this, note that

$$\widehat{f}_{n,h_n}(x) - f(x) = \widehat{f}_{n,h_n}(x) - f_{h_n}(x) + f_{h_n}(x) - f(x),$$

and by Equation (A.7) in Theorem A.4,

$$\lim_{n\to\infty} f_{h_n}(x) = \lim_{n\to\infty} \int f(x + uh_n) K(u)\, \mathrm{d}u = f(x) \int K(u)\, \mathrm{d}u = f(x),$$

where the last equality follows from Assumption 2.2 (ii). Hence, we are done if we can show $\widehat{f}_{n,h_n}(x) - f_{h_n}(x) \xrightarrow{\text{p}} 0$. But the asymptotic normality part of (2.3) implies that $\widehat{f}_{n,h_n}(x) - f_{h_n}(x) = O_{\text{p}}\left(1/\sqrt{nh_n^d}\right) = o_{\text{p}}(1)$ since by Assumption 2.3, $1/\left(nh_n^d\right) \to 0$ as $n \to \infty$.

Hence, it remains to show the asymptotic normality part of (2.3). We shall derive the latter as a corollary of Theorem 2.2. By (2.9) in Theorem 2.2

$$Z_{n,h_n}(x) := \frac{S_{n,h_n}(x)}{\sqrt{n\Sigma_{h_n}(x)}} \rightsquigarrow \text{N}(0,1). \tag{2.13}$$

The definition of $\Sigma_h(x)$ is given in (2.8). By (2.6) and (2.13)

$$\widehat{f}_{n,h}(x) - f_h(x) = \frac{S_{n,h}(x)}{nh^d} = \frac{\sqrt{n\Sigma_h(x)}}{nh^d} \cdot Z_{n,h}(x) = \frac{\sqrt{h^{-d}\Sigma_h(x)}}{\sqrt{nh^d}} \cdot Z_{n,h}(x),$$

and so,

$$\sqrt{nh^d}\left(\widehat{f}_{n,h}(x) - f_h(x)\right) = \sqrt{h^{-d}\Sigma_h(x)} \cdot Z_{n,h}(x). \tag{2.14}$$

Now by (2.8),

$$\begin{aligned} h^{-d}\Sigma_h(x) &= h^{-d}\left(h^d \int f(x+uh)K(u)^2 \, \text{d}u - \left(h^d \int f(x+uh)K(u) \, \text{d}u\right)^2\right) \\ &= \int f(x+uh)K(u)^2 \, \text{d}u - h^d \left(\int f(x+uh)K(u) \, \text{d}u\right)^2. \end{aligned}$$

By Theorem A.4,

$$\lim_{n\to\infty} h_n^{-d}\Sigma_{h_n}(x) = f(x) \int K(u)^2 \, \text{d}u.$$

Combine the above display with (2.13), (2.14), and Slutsky's Theorem to get the asymptotic normality part of (2.3). $\square$

2.1.2 Proof of Theorem 2.2

We use Liapunov's Central Limit Theorem to prove (2.9). To that end, define

$$\begin{aligned} b_{n,h}(x)^4 &= \sum_{i=1}^{n} \mathrm{E}\left[\left|K\left(\frac{X_i - x}{h}\right) - \mathrm{E}\left[K\left(\frac{X - x}{h}\right)\right]\right|^4\right] \\ &= n \cdot \mathrm{E}\left[\left|K\left(\frac{X - x}{h}\right) - \mathrm{E}\left[K\left(\frac{X - x}{h}\right)\right]\right|^4\right]. \end{aligned}$$

By Liapunov's Central Limit Theorem,

$$\text{If} \quad \lim_{n\to\infty} \frac{b_{n,h_n}(x)^4}{\left(n\Sigma_{h_n}(x)\right)^2} = 0, \quad \text{then} \quad \frac{S_{n,h_n}(x)}{\sqrt{n\Sigma_{h_n}(x)}} \rightsquigarrow \mathrm{N}(0,1). \tag{2.15}$$

By the C_r inequality (in particular $(|a| + |b|)^4 \leq 8\left(|a|^4 + |b|^4\right)$),

$$b_{n,h}(x)^4 \leq 8n \cdot \left(\mathrm{E}\left[K\left(\frac{X - x}{h}\right)^4\right] + \left|\mathrm{E}\left[K\left(\frac{X - x}{h}\right)\right]\right|^4\right).$$

And so,

$$\frac{b_{n,h}(x)^4}{\left(n\Sigma_h(x)\right)^2} \leq \frac{8\left(\mathrm{E}\left[\left|K\left(\frac{X-x}{h}\right)\right|^4\right] + \left|\mathrm{E}\left[K\left(\frac{X-x}{h}\right)\right]\right|^4\right)}{n\left(\mathrm{E}\left[K\left(\frac{X-x}{h}\right)^2\right] - \mathrm{E}\left[K\left(\frac{X-x}{h}\right)\right]^2\right)}. \tag{2.16}$$

For any $p \in \{1, 2, 4\}$,

$$\mathrm{E}\left[K\left(\frac{X - x}{h}\right)^p\right] = \int f(\xi)\left|K\left(\frac{\xi - x}{h}\right)\right|^p \,\mathrm{d}\xi = h^d \int f(x + uh)K(u)^p \,\mathrm{d}u. \tag{2.17}$$

Plug (2.16) into (2.17) to get

$$\frac{b_{n,h}(x)^4}{\left(n\Sigma_h(x)\right)^2} \leq \frac{8\left(h^d \int f(x + uh)K(u)^4 \,\mathrm{d}u + \left(h^d \int f(x + uh)K(u) \,\mathrm{d}u\right)^4\right)}{n\left(h^d \int f(x + uh)K(u)^2 \,\mathrm{d}u - \left(h^d \int f(x + uh)K(u) \,\mathrm{d}u\right)^2\right)^2}.$$

Then, since $\int f(x+uh)K(u)\,\mathrm{d}u = f_h(x)$,

$$\frac{b_{n,h}(x)^4}{\left(n\Sigma_h(x)\right)^2} \le \frac{8\left(\int f(x+uh)K(u)^4\,\mathrm{d}u + h^{3d} f_h(x)^4\right)}{nh^d\left(\int f(x+uh)K(u)^2\,\mathrm{d}u - h^d f_h(x)^2\right)^2}. \tag{2.18}$$

By Theorem A.4, since $\lim_{n\to\infty} h_n = 0$,

$$\lim_{n\to\infty} \int f\left(x+uh_n\right) K(u)^p\,\mathrm{d}u = f(x)\int K(u)\,\mathrm{d}u \quad .$$

Furthermore, Assumption 2.2 (ii) implies that $\int K(u)^p\,\mathrm{d}u > 0$ for $p \in \{2,4\}$. Combine this with the above displayed expression and (2.18) to get

$$\frac{b_{n,h_n}(x)^4}{\left(n\Sigma_{h_n}(x)\right)^2} \le \frac{8}{nh_n} \cdot \frac{f(x)\int K(u)^4\,\mathrm{d}u + o(1) + h_n^{3d}(f(x)+o(1))^4}{\left(f(x)\int K(u)^2\,\mathrm{d}u + o(1) - h_n^d(f(x)+o(1))^2\right)^2} \to 0,$$

since $\lim_{n\to\infty} \max\left\{h_n, 1/\left(nh_n^d\right)\right\} = 0$. The claim in (2.9) now follows by (2.15). □

2.1.3 Proof of Theorem 2.3

The asymptotic standard normality result in (2.12) follows from (2.10), the variance consistency result in (2.12) and Slutsky's Theorem:

$$\frac{\widehat{f}_{n,h_n}(x) - f_{h_n}(x)}{\sqrt{\widehat{V}_{n,h_n}(x)/n}} = \sqrt{\frac{V_{h_n}(x)}{\widehat{V}_{n,h_n}(x)}} \cdot \frac{\widehat{f}_{n,h_n}(x) - f_{h_n}(x)}{\sqrt{V_{h_n}(x)/n}} \rightsquigarrow \mathrm{N}(0,1).$$

Hence, we focus on showing the variance consistency result in (2.12), which is equivalent to showing

$$\frac{\widehat{V}_{n,h_n}(x) - V_{h_n}(x)}{V_{h_n}(x)} \xrightarrow{\mathrm{P}} 0 \iff \widehat{V}_{n,h_n}(x) - V_{h_n}(x) = o_{\mathrm{P}}\left(V_{h_n}(x)\right). \tag{2.19}$$

To that end,

$$\widehat{V}_{n,h}(x) - V_h(x) = f_h(x)^2 - \widehat{f}_{n,h}(x)^2 + \frac{1}{n}\sum_{i=1}^{n}\left\{K_h\left(X_i - x\right)^2 - \mathrm{E}\left[K_h(X-x)^2\right]\right\}$$

$$
\begin{aligned}
&= \left(f_h(x) + \widehat{f}_{n,h}(x)\right)\left(f_h(x) - \widehat{f}_{n,h}(x)\right) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\left\{K_h\left(X_i - x\right)^2 - \mathrm{E}\left[K_h(X-x)^2\right]\right\} \\
&= 2 f_h(x)\left(f_h(x) - \widehat{f}_{n,h}(x)\right) - \left(f_h(x) - \widehat{f}_{n,h}(x)\right)^2 \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\left\{K_h\left(X_i - x\right)^2 - \mathrm{E}\left[K_h(X-x)^2\right]\right\}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\widehat{V}_{n,h}(x) - V_h(x) = R_{n,h}(x) + T_{n,h}(x) \\
&\text{where} \quad R_{n,h}(x) = 2 f_h(x)\left(f_h(x) - \widehat{f}_{n,h}(x)\right) - \left(f_h(x) - \widehat{f}_{n,h}(x)\right)^2, \\
&\text{and} \quad T_{n,h}(x) = \frac{1}{n}\sum_{i=1}^{n}\left\{K_h\left(X_i - x\right)^2 - \mathrm{E}\left[K_h(X-x)^2\right]\right\}.
\end{aligned}
\tag{2.20}
$$

Therefore showing (2.19) is equivalent to showing

$$
R_{n,h_n}(x) = o_{\mathrm{p}}\left(V_{h_n}(x)\right) \quad \text{and} \quad T_{n,h_n}(x) = o_{\mathrm{p}}\left(V_{h_n}(x)\right). \tag{2.21}
$$

By (2.10),

$$
\widehat{f}_{n,h_n}(x) - f_{h_n}(x) = O_p\left(\sqrt{\frac{V_{h_n}(x)}{n}}\right) = O_p\left(\frac{V_{h_n}(x)}{\sqrt{nV_{h_n}(x)}}\right),
$$

and so,

$$
\widehat{f}_{n,h_n}(x) - f_{h_n}(x) = \frac{1}{\sqrt{nV_{h_n}(x)}} \cdot O_p\left(V_{h_n}(x)\right). \tag{2.22}
$$

Since

$$
V_h(x) = \frac{1}{h^d}\left(\int f(x+uh)K(u)^2 \,\mathrm{d}u - h^d\left(\int f(x+uh)K(u)\,\mathrm{d}u\right)^2\right), \tag{2.23}
$$

by Theorem A.4 it follows that

$$\frac{1}{nV_{h_n}(x)} = \frac{h_n^d}{n} \cdot \frac{1}{f(x)\int K(u)^2\,\mathrm{d}u + o(1)}.$$

Combine with this (2.22) to conclude that

$$\widehat{f}_{n,h_n} - f_{h_n}(x) = o_{\mathrm{p}}\left(V_{h_n}(x)\right) \quad \text{and so} \quad R_{n,h_n}(x) = o_{\mathrm{p}}\left(V_{h_n}(x)\right). \tag{2.24}$$

For the second part of (2.21), we will use Chebyshev's inequality. Note that (2.23)

$$\begin{aligned} T_{n,h_n}(x) =& O_{\mathrm{p}}\left(\sqrt{\mathrm{E}\left[T_{n,h_n}(x)^2\right]}\right) = O_{\mathrm{p}}\left(V_h(x)\right)\frac{\sqrt{\mathrm{E}\left[T_{n,h_n}(x)^2\right]}}{V_h(x)} \\ =& O_{\mathrm{p}}\left(V_{h_n}(x)\right)\frac{h_n^d\sqrt{\mathrm{E}\left[T_{n,h_n}(x)^2\right]}}{\left(\int f\left(x+uh_n\right)K(u)^2\,\mathrm{d}u - h_n^d\left(\int f\left(x+uh_n\right)K(u)\,\mathrm{d}u\right)^2\right)}, \end{aligned}$$

and so, by Theorem A.4

$$T_{n,h_n}(x) = O_{\mathrm{p}}\left(V_{h_n}(x)\right)\frac{\sqrt{h_n^{2d}\cdot\mathrm{E}\left[T_{n,h_n}(x)^2\right]}}{f(x)\int K(u)^2\,\mathrm{d}u + o(1)} \tag{2.25}$$

Now, since $K_h\left(X_i - x\right)$ are iid,

$$\begin{aligned} \mathrm{E}\left[T_{n,h}(x)^2\right] =& \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^n\left\{K_h\left(X_i-x\right)^2 - \mathrm{E}\left[K_h(X-x)^2\right]\right\}\right)^2\right] \\ =& \frac{\mathrm{E}\left[\left(K_h\left(X-x\right)^2 - \mathrm{E}\left[K_h(X-x)^2\right]\right)^2\right]}{n} \\ =& \frac{\mathrm{E}\left[K_h\left(X-x\right)^4\right] - \mathrm{E}\left[K_h(X-x)^2\right]^2}{n} \\ =& \frac{\frac{1}{h^{3d}}\int f(x+uh)K(u)^4\,\mathrm{d}u - \left(\frac{1}{h^d}\int f(x+uh)K(u)^2\,\mathrm{d}u\right)^2}{n}. \end{aligned}$$

Therefore,

$$h^{2d}\mathrm{E}\left[T_{n,h}(x)^2\right] = \frac{\int f(x+uh)K(u)^4\,\mathrm{d}u - h^d\left(\int f(x+uh)K(u)^2\,\mathrm{d}u\right)^2}{nh^d},$$

and so by Theorem A.4,

$$h_n^{2d}\mathrm{E}\left[T_{n,h_n}(x)^2\right] = \frac{f(x)\int K(u)^4\,\mathrm{d}u + o(1)}{nh_n^d}. \tag{2.26}$$

Combine (2.26) with (2.25) to see that

$$\begin{aligned} T_{n,h_n}(x) &= O_{\mathrm{p}}\left(V_{h_n}(x)\right)\frac{\sqrt{f(x)\int K(u)^4\,\mathrm{d}u + o(1)}}{f(x)\int K(u)^2\,\mathrm{d}u + o(1)}\cdot\frac{1}{nh_n^d}, \qquad (2.27)\\ \text{and so} \quad T_{n,h_n}(x) &= o_{\mathrm{p}}\left(V_{h_n}(x)\right). \end{aligned}$$

Now, (2.21) follows from (2.24) and (2.27). □

References

Billingsley, Patrick. 1995. *Probability and Measure.* Wiley Series in Probability and Statistics. Wiley.

Bochner, S. 1955. *Harmonic Analysis and the Theory of Probability.* Dover Books on Mathematics. Dover Publications.

Pagan, Adrian, and Aman Ullah. 1999. *Nonparametric Econometrics.* Themes in Modern Econometrics. Cambridge University Press.

Parzen, Emanuel. 1962. "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics* 33 (3): 1065–1076.

Pollard, David. 1984. *Convergence of Stochastic Processes.* Springer Series in Statistics. Springer New York.

A Approximation Results for Kernel Estimators

Here, we consider approximation by convolution. The exposition here is mainly based on Pagan and Ullah (1999, pp. 362–365).[1] I should note that Pagan and Ullah (1999) borrow from Parzen (1962) who in turn borrows from Bochner (1955).

Assumption A.1. $\{\kappa_h : h \in (0, \infty)\}$ is a family of functions mapping $\mathbb{R}^d$ to $\mathbb{R}$ satisfying the following.

(i) For each $h \in (0, \infty)$, κ_h is integrable with $\int |\kappa_h(u)| \, \mathrm{d}u =: \overline{\kappa}_h < \infty$.

(ii) There is a $h_* \in (0, \infty)$ and $\overline{\kappa} \in (0, \infty)$ such that for every $h \in (0, h_*]$, $\overline{\kappa}_h \leq \overline{\kappa}$.

(iii) For each $\delta \in (0, \infty)$, $\lim_{h \to 0} \int_{\|u\| \geq \delta} |\kappa_h(u)| \, \mathrm{d}u = 0$.

(iv) For each $h \in (0, \infty)$, $\int \kappa_h(u) \, \mathrm{d}u = \underline{\kappa}_h$ and $\lim_{h \to 0} \underline{\kappa}_h = \underline{\kappa} \in \mathbb{R}$.

Remark A.1. If the family $\{\kappa_h\}$ satisfies $\kappa_h \geq 0$ almost everywhere for each $h \in (0, \infty)$, then Assumption A.1 (i) and (iv) imply (ii), since in this case $\underline{\kappa}_h = \overline{\kappa}_h$ for every $h \in (0, \infty)$ and we can choose h_* to satisfy $|\underline{\kappa}_h - \underline{\kappa}| \leq 1$ for every $h \in (0, h_*]$. Then set $\overline{\kappa} = |\underline{\kappa}| + 1$.

Lemma A.1. *Let $\kappa : \mathbb{R}^d \to \mathbb{R}$ be an integrable function. Define $\kappa_h(u) := h^{-d} \kappa(u/h)$. Then $\{\kappa_h\}$ satisfies Assumption A.1.*

Proof of Lemma A.1. By the change of variables $v = u/h$,

$$
\begin{aligned}
\underline{\kappa}_h &= \int \kappa_h(u) \, \mathrm{d}u = \int h^{-d} \kappa(u/h) \, \mathrm{d}u = \int \kappa(v) \, \mathrm{d}v \quad \forall h \in (0, \infty), \\
\overline{\kappa}_h &= \int |\kappa_h(u)| \, \mathrm{d}u = \int h^{-d} |\kappa(u/h)| \, \mathrm{d}u = \int |\kappa(v)| \, \mathrm{d}v =: \overline{\kappa},
\end{aligned}
$$

1. In my PDF copy of Pagan and Ullah (1999), this corresponds to pp. 381–384 (in PDF).

$$\int_{\|u\|\geq\delta} |\kappa_h(u)| \, \mathrm{d}u = \int_{\|u\|\geq\delta} h^{-d} |\kappa(u/h)| \, \mathrm{d}u = \int_{\|v\|\geq\delta/h} |\kappa(v)| \, \mathrm{d}v \to 0$$

$$\text{since} \quad \{v : \|v\| \geq \delta/h\} \searrow \emptyset, \quad \text{as} \quad h \to 0.$$

$\square$

For a measurable function $g : \mathbb{R}^d \to \mathbb{R}$, consider the convolution based approximant

$$g_h(x) := T_h[g](x) := \int g(x+u)\kappa_h(u) \, \mathrm{d}u \tag{A.1}$$

for a family $\{\kappa_h\}$ satisfying Assumption A.1.

Assume henceforth that for the point x in question, the integral in (A.1) exists and is finite. The natural question of primitive assumptions for this existence and finiteness. We first provide a pointwise error bound for $g_h - g$.

Theorem A.1. *Let $\{\kappa_h\}$ satisfy Assumption A.1 (i)-(iii). Let $h, \delta \in (0, \infty)$, $x \in \mathbb{R}^d$, and let $g : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Assume the integral in (A.1) exists and is finite. Then*

$$\begin{aligned} |g_h(x) - g(x) \cdot \underline{\kappa}_h| \leq \left\{ \sup_{\|u\|<\delta} |g(x+u) - g(x)| \right\} \cdot \overline{\kappa} + \int_{\|u\|\geq\delta} |g(x+u)| \cdot |\kappa_h(u)| \, \mathrm{d}u \\ + |g(x)| \int_{\|u\|\geq\delta} |\kappa_h(u)| \, \mathrm{d}u. \end{aligned} \tag{A.2}$$

Proof of Theorem A.1.

$$|g_h(x) - g(x) \cdot \underline{\kappa}_h| = \left| \int (g(x+u) - g(x))\kappa_h(u) \, \mathrm{d}u \right| \leq \int |g(x+u) - g(x)| \, |\kappa_h(u)| \, \mathrm{d}u.$$

Split the integral on the right:

$$\begin{aligned} |g_h(x) - g(x) \cdot \underline{\kappa}_h| \leq & \int_{\|u\|<\delta} |g(x+u) - g(x)| \, |\kappa_h(u)| \, \mathrm{d}u \\ & + \int_{\|u\|\geq\delta} |g(x+u) - g(x)| \, |\kappa_h(u)| \, \mathrm{d}u \end{aligned}$$

$$\le \left[\sup_{\|u\|<\delta} |g(x+u) - g(x)|\right] \int |\kappa_h(u)| \ \mathrm{d}u$$
$$+ \int_{\|u\|\ge\delta} |g(x+u)| \, |\kappa_h(u)| \ \mathrm{d}u + |g(x)| \int_{\|u\|\ge\delta} |\kappa_h(u)| \ \mathrm{d}u.$$

In the last inequality above, bound the integral in the first term by $\overline{\kappa}$ to get (A.2). □

Lemma A.2. *Let $\{\kappa_h\}$ satisfy Assumption A.1 (i)-(iii). Let $x \in \mathbb{R}^d$, and let $g : \mathbb{R}^d \to \mathbb{R}$ be a measurable function that is continuous at x. Assume the integral in (A.1) exists and is finite for $h > 0$ sufficiently small.*

$$\begin{aligned} &\textit{If} \quad \forall \delta > 0, \quad \lim_{h\to 0} \int_{\|u\|\ge\delta} |g(x+u)| \, |\kappa_h(u)| \ \mathrm{d}u = 0, \quad \textit{then} \\ &\qquad\qquad \lim_{h\to\infty} \left(g_h(x) - g(x) \cdot \underline{\kappa}_h\right) = 0. \end{aligned} \tag{A.3}$$

Furthermore if $\{\kappa_h\}$ also satisfies satisfies Assumption A.1 (iv) then the premise in (A.3) also implies

$$\lim_{h\to\infty} g_h(x) = g(x)\underline{\kappa}. \tag{A.4}$$

Proof of Lemma A.2. Under Assumption A.1 (iv), (A.4) is an immediate consequence of (A.3). So we prove (A.3). Since g is continuous at x, the first term in (A.2) can be controlled by choice of δ sufficiently small. For any choice of δ, the second term in (A.2) is controlled by choice of h sufficiently small by (A.3). For any choice of δ, the third term in (A.2) is controlled by choice of h sufficiently small by Assumption A.1 (iii). □

We therefore need some way(s) to show (A.3) to deal with the second term in (A.2). We present two ways to do this.

Theorem A.2. *Let $\{\kappa_h\}$ satisfy Assumption A.1 (i)-(iii). Let $x \in \mathbb{R}^d$, and $g : \mathbb{R}^d \to \mathbb{R}$ be a bounded and measurable function that is continuous at x. Then the integral in (A.1) exists and is finite and furthermore, $\lim_{h\to 0} (g_h(x) - g(x)\underline{\kappa}_h) = 0$. If in addition, $\{\kappa_h\}$ satisfies Assumption A.1 (iv), then $\lim_{h\to 0} g_h(x) = g(x)\underline{\kappa}$*

Proof of Theorem A.2. Existence and finiteness of the integral in (A.1) are immediate consequences of the hypothesis of g being bounded. Furthermore by this hypothesis, given any $\delta > 0$

$$\int_{\|u\|\geq\delta} |g(x+u)|\,|\kappa_h(u)|\;\mathrm{d}u \leq \sup_{y\in\mathbb{R}^d} |g(y)| \int_{\|u\|\geq\delta} |\kappa_h(u)|\;\mathrm{d}u \to 0 \quad \text{as} \quad h \to 0.$$

Both limiting claims in Theorem A.2 now follow from (A.3) in Lemma A.2. □

Assumption A.2. The family $\{\kappa_h\}$ in Assumption A.1 also satisfies the following condition:

$$\forall\delta > 0, \quad \lim_{h\to 0} \sup_{u\in\mathbb{R}^d:\|u\|\geq\delta} \|u\|^d \kappa_h(u) = 0. \tag{A.5}$$

Lemma A.3. *Consider the function κ and the associated family $\{\kappa_h\}$ in Lemma A.1.*

$$\textit{If} \lim_{\|y\|\to\infty} \|y\|^d \kappa(y) = 0, \;\textit{then}\; \{\kappa_h\} \;\textit{satisfies Assumption A.2.} \tag{A.6}$$

Proof of Lemma A.3. Note that $\sup_{\|u\|\geq\delta} \|u\|^d \kappa_h(u) = \sup_{\|u\|\geq\delta} \|u/h\|^d \kappa(u/h)$. Given any $\delta > 0$, the premise in (A.6) ensures that the right hand side can be made arbitrarily small when $h \to 0$. Thus (A.5) is satisfied. □

Theorem A.3. *Suppose the family $\{\kappa_h\}$ satisfies Assumption A.1 (i)-(iii) and Assumption A.2. Let $x \in \mathbb{R}^d$, and let $g : \mathbb{R}^d \to \mathbb{R}$ be an integrable function that is continuous at x. Then for $h > 0$ sufficiently small, the integral in (A.1) exists and is finite. In addition, $\lim_{h\to 0} (g_h(x) - g(x)\underline{\kappa}_h) = 0$. Furthermore, if $\{\kappa_h\}$ satisfies Assumption A.1 (iv) then $\lim_{h\to 0} g_h(x) = g(x)\underline{\kappa}$.*

Proof of Theorem A.3. For existence and finiteness of the integral in (A.1), first note that for any $\delta > 0$,

$$|g(x+u)|\,|\kappa_h(u)| \leq |g(x+u)|\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| < \delta\} + |g(x+u)|\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| \geq \delta\}.$$

Take any $\varepsilon \in (0, \infty)$ and choose $\delta := \delta_{\varepsilon,x} \in (0, \infty)$ to ensure that $|g(x+u) - g(x)| < \varepsilon$ for $\|u\| < \delta$. Then, $|g(x+u)| < |g(x)| + \varepsilon$ and combining this with the bound in the above display,

$$\begin{aligned}
|g(x+u)|\,|\kappa_h(u)| &\leq (|g(x)| + \varepsilon)\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| < \delta\} + |g(x+u)|\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| \geq \delta\} \\
&\leq (|g(x)| + \varepsilon)\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| < \delta\} \\
&\quad + \frac{|g(x+u)|}{\|u\|^d} \|u\|^d\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| \geq \delta\} \\
&\leq (|g(x)| + \varepsilon)\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| < \delta\} \\
&\quad + \frac{|g(x+u)|}{\delta^d} \|u\|^d\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| \geq \delta\} \\
&\leq (|g(x)| + \varepsilon)\,|\kappa_h(u)| \cdot \mathbf{1}\{\|u\| < \delta\} \\
&\quad + \frac{1}{\delta^d}\left[\sup_{y \in \mathbb{R}^d : \|y\| \geq \delta} \|y\|^d\,|\kappa_h(y)|\right] \cdot |g(x+u)| \mathbf{1}\{\|u\| \geq \delta\}.
\end{aligned}$$

The supremum in the above display exists and is finite for h sufficiently small by (A.5). We know that $|\kappa_h(u)|$ has a finite integral by Assumption A.1 (i). Therefore, the a sufficient condition for the existence and finiteness of the integral in (A.1) is integrability of $|g(x+u)|$ since then $|g(x+u)|\mathbf{1}\{\|u\| \geq \delta\}$ would be integrable. To that end, note that $\int |g(x+u)|\,\mathrm{d}u = \int |g(v)|\,\mathrm{d}v < \infty$ by the change of variables $v = x + u$.

To prove the limit claims, we show (A.3) in Lemma A.2. Given any $\delta > 0$,

$$\begin{aligned}
\int_{\|u\| \geq \delta} |g(x+u)|\,|\kappa_h(u)|\;\mathrm{d}u &= \int_{\|u\| \geq \delta} \frac{|g(x+u)|}{\|u\|^d} \cdot \|u\|^d \cdot |\kappa_h(u)|\;\mathrm{d}u \\
&= \left[\sup_{y \in \mathbb{R}^d : \|y\| \geq \delta} \|y\|^d \cdot |\kappa_h(y)|\right] \cdot \int_{\|u\| \geq \delta} \frac{|g(x+u)|}{\|u\|^d}\,\mathrm{d}u \\
&\leq \frac{1}{\delta^d}\left[\sup_{y \in \mathbb{R}^d : \|y\| \geq \delta} \|y\|^d \cdot |\kappa_h(y)|\right] \cdot \int_{\|u\| \geq \delta} |g(x+u)|\,\mathrm{d}u.
\end{aligned}$$

In addition,

$$\int_{\|u\| \geq \delta} |g(x+u)|\,\mathrm{d}u = \int_{\|x-v\| \geq \delta} |g(v)|\,\mathrm{d}v \leq \int |g(v)|\,\mathrm{d}v.$$

Combining both bounds,

$$\int_{\|u\|\geq\delta} |g(x+u)|\,|\kappa_h(u)|\ \mathrm{d}u \leq \frac{1}{\delta^d}\left[\sup_{y\in\mathbb{R}^d:\|y\|\geq\delta} \|y\|^d \cdot |\kappa_h(y)|\right] \cdot \int |g(v)|\ \mathrm{d}v.$$

Given any $\delta > 0$, the right hand side of the above display tends to 0 as $h \to 0$ by (A.5). Both limiting claims in Theorem A.3 now follow from (A.3) in Lemma A.2. □

Theorem A.4. *Let $\kappa : \mathbb{R}^d \to \mathbb{R}$ be an integrable function with $\lim_{\|y\|\to\infty} \|y\|^d \kappa(y) = 0$. Let $x \in \mathbb{R}^d$, and let $g : \mathbb{R}^d \to \mathbb{R}$ be an integrable function that is continuous at x. Then*

$$\lim_{h\to 0} \int g(x+vh)\kappa(v)\ \mathrm{d}v = \lim_{h\to 0} \frac{1}{h^d} \int g(x+u)\kappa(u/h)\ \mathrm{d}u = g(x) \int \kappa(u)\ \mathrm{d}u. \tag{A.7}$$

Proof of Theorem A.4. Let $\kappa_h(u) := h^{-d}\kappa(u/h)$. The first equality in (A.7) follows from the change of variables $v = u/h$. So we prove the second equality. By Lemma A.1 and Lemma A.3, $\{\kappa_h\}$ satisfies Assumption A.1 and Assumption A.2. By the change of variables $v = u/h$, $\underline{\kappa}_h = h^{-d} \int \kappa(u/h)\ \mathrm{d}u = \int \kappa(v)\ \mathrm{d}v$ for every $h \in (0, \infty)$. Hence, $\lim_{h\to 0} \underline{\kappa}_h = \underline{\kappa} := \int \kappa(v)\ \mathrm{d}v$ is trivially true. Then (A.7) follows from Theorem A.3. □