1.3 Uniform convergence in probability

In many contexts it is desirable to obtain uniform convergence of $\hat{f}$ over a specified subset of the range of X_i. The next theorem gives conditions under which $\hat{f}$ converges to f uniformly in probability over a compact subset of $\mathbb{R}$.

Theorem 13. *Assume that: a) $|K(u)| \leq C$ for all $u \in \mathbb{R}$; b) $\int |K(u)|du < \infty$; c) $|u||K(u)| \to 0$ as $|u| \to \infty$; d) for any $u, u' \in \mathbb{R}$ with $u \neq u'$ we have $|K(u') - K(u)| < C|u' - u|$ for some arbitrary constant $C < \infty$ (Lipschitz condition of order 1 in $\mathbb{R}$); e) $h_n \to 0$, $nh_n^3 \to \infty$ and $\frac{nh_n}{\log n} \to \infty$ as $n \to \infty$. Then, if f is continuous in G, a compact subset of $\mathbb{R}$*

$$\sup_{x\in G} |\hat{f}(x) - E(\hat{f}(x))| = O_p\left(\left(\frac{\log n}{nh_n}\right)^{1/2}\right).$$

Proof. Since G is a compact subset of $\mathbb{R}$, it is closed and bounded. Therefore, there exists $x_0 \in \mathbb{R}$ and some $r > 0$ such that $G \subseteq B(x_0, r) = \{x \in \mathbb{R} : |x - x_0| < r\}$. In addition, for any two $x, x' \in G$ we have $|x - x'| < 2r$. For every $x \in G$ consider $B(x, r_x)$ and note that the collection $\{B(x, r_x) :$ for $x \in G$ and $r_x > 0\}$ is an open covering for G. By the Heine-Borel Theorem every open cover of G contains a finite subcover, which we denote by $\{B(x_k, r_{x_k}) :$ for $x_k \in G$, $r_{x_k} > 0$ and $k = 1, \cdots, m$, $m \in \mathbb{N}.\}$. Now, by the triangle inequality

$$|\hat{f}(x) - E(\hat{f}(x))| \leq |\hat{f}(x_k) - E(\hat{f}(x_k))| + |\hat{f}(x) - \hat{f}(x_k)| + |E(\hat{f}(x)) - E(\hat{f}(x_k))|.$$

Given d), for $x \in B(x_k, r_{x_k})$ we have

$$|\hat{f}(x) - \hat{f}(x_k)| \leq \frac{C}{h_n^2}|x_k - x| < C\frac{r_{x_k}}{h_n^2} \text{ and } \left|E(\hat{f}(x)) - E(\hat{f}(x_k))\right| \leq \frac{C}{h_n^2}|x_k - x| < C\frac{r_{x_k}}{h_n^2}.$$

Hence,

$$|\hat{f}(x) - E(\hat{f}(x))| \leq |\hat{f}(x_k) - E(\hat{f}(x_k))| + 2C\frac{r_{x_k}}{h_n^2}.$$

Since for every $x \in G$ there exists $B(x_k, r_{x_k})$ that contains x,

$$\sup_{x\in G} |\hat{f}(x) - E(\hat{f}(x))| \leq \max_{1\leq k\leq m} |\hat{f}(x_k) - E(\hat{f}(x_k))| + 2C\frac{r_{x_k}}{h_n^2}$$

Now, for every n, let $m := l_n$ and $r_{x_k} := r_n$ with $G \subset \cup_{k=1}^{l_n} B(x_k, r_n)$ with $r > l_n r_n$ or $l_n < r/r_n$. Letting $r_n = \frac{h_n^{3/2}}{n^{1/2}}$ we have

$$\frac{(nh_n)^{1/2}}{log(n)^{1/2}} \sup_{x\in G} |\hat{f}(x) - E(\hat{f}(x))| \leq \frac{(nh_n)^{1/2}}{log(n)^{1/2}} \max_{1\leq k\leq l_n} |\hat{f}(x_k) - E(\hat{f}(x_k))| + \frac{2C}{log(n)^{1/2}}$$

and to prove the theorem it suffices to show that for all $\epsilon > 0$ there exists $\Delta_\epsilon > 0$ such that

$$P\left(\frac{(nh_n)^{1/2}}{log(n)^{1/2}} \max_{1\le k\le l_n} |\hat{f}(x_k) - E(\hat{f}(x_k))| > \Delta_\epsilon\right) < \epsilon \text{ for all } n.$$

By Lemma 6 in the Appendix, we have

$$P\left(\frac{(nh_n)^{1/2}}{log(n)^{1/2}} \max_{1\le k\le l_n} |\hat{f}(x_k) - E(\hat{f}(x_k))| > \Delta_\epsilon\right) \le \sum_{k=1}^{l_n} P\left(\frac{(nh_n)^{1/2}}{log(n)^{1/2}} |\hat{f}(x_k) - E(\hat{f}(x_k))| > \Delta_\epsilon\right).$$

Now, using Lemma 7 in the Appendix (Bernstein's Inequality) with $\Delta = \left(\frac{\log n}{nh_n}\right)^{1/2} \frac{\Delta_\epsilon}{\sigma_n}$, $W_{in} = \frac{1}{nh_n}K\left(\frac{X_i - x_k}{h_n}\right) - E\left(\frac{1}{nh_n}K\left(\frac{X_i - x_k}{h_n}\right)\right)$, we have that if $|K(x)| < C$, $|W_{in}| \le \frac{2C}{nh_n}$ and

$$\begin{aligned}
P\left(\frac{(nh_n)^{1/2}}{(log(n))^{1/2}} |\hat{f}(x_k) - E(\hat{f}(x_k))| > \Delta_\epsilon\right) &< 2\exp\left(-\frac{\frac{\log n}{nh_n}\frac{\Delta_\epsilon^2}{\sigma_n^2}}{2 + \frac{2}{3}\frac{2C}{nh_n}\frac{1}{\sigma_n}\frac{\Delta_\epsilon}{\sigma_n}\left(\frac{\log n}{nh_n}\right)^{1/2}}\right) \\
&= 2\exp\left(-\frac{\frac{\log n}{nh_n}\Delta_\epsilon^2}{2\sigma_n^2 + \frac{2}{3}\frac{2C}{nh_n}\Delta_\epsilon\left(\frac{\log n}{nh_n}\right)^{1/2}}\right) \\
&= 2\exp\left(-\frac{\log n\ \Delta_\epsilon^2}{2nh_n\sigma_n^2 + \frac{4}{3}C\Delta_\epsilon\left(\frac{\log n}{nh_n}\right)^{1/2}}\right)
\end{aligned}$$

where $\sigma_n^2 = V(\hat{f}(x_k)) = \frac{1}{nh_n}\int K^2(v)f(x_k + h_n v)dv - \frac{1}{n}\left(\int K(v)f(x_k + h_n v)dv\right)^2$. Since f is uniformly continuous on G (continuity on a compact set implies uniform continuity), $nh_n\sigma_n^2 \to f(x_k)\int K^2(v)dv$ for all x_k and, provided that $\frac{nh_n}{\log n} \to \infty$, we have $c_n := 2nh_n\sigma_n^2 + \frac{4}{3}C\Delta_\epsilon\left(\frac{\log n}{nh_n}\right)^{1/2} \to 2f(x_k)\int K^2(v)dv$. Now,

$$\begin{aligned}
P\left(\frac{(nh_n)^{1/2}}{(log(n))^{1/2}} \max_{1\le k\le l_n} |\hat{f}(x_k) - E(\hat{f}(x_k))| > \Delta_\epsilon\right) &< l_n 2\exp\left(-\log n\ \frac{\Delta_\epsilon^2}{c_n}\right) \\
&< 2\frac{n^{1/2}}{h_n^{3/2}} r n^{-\Delta_\epsilon^2/c_n} \text{ since } l_n < \frac{r}{r_n} = \frac{n^{1/2}}{h_n^{3/2}} \\
&= 2\left(\frac{1}{nh_n^3}\right)^{1/2} n^{1-\frac{\Delta_\epsilon^2}{c_n}} r \to 0,
\end{aligned}$$

provided $\Delta_\epsilon^2 > 2f(x_k)\int K^2(v)dv$ and $nh_n^3 \to \infty$. □

Remark 10. *1. From Theorem 13, we can conclude that*

$$\frac{1}{h_n}\sup_{x\in G}|\hat{f}(x)-E(\hat{f}(x))| = \left(\frac{\log n}{nh_n^3}\right)^{1/2} O_p(1).$$

Consequently, if $\frac{nh_n^3}{log(n)} \to \infty$ *then* $\frac{1}{h_n}\sup_{x\in G}|\hat{f}(x) - E(\hat{f}(x))| = o_p(1)$ *or* $\sup_{x\in G}|\hat{f}(x) - E(\hat{f}(x))| = o_p(h_n)$. *Note that if* $h_n \propto n^{\delta}$, *then* $nh_n^3 \to \infty$ *requires that* $\delta > -\frac{1}{3}$ *and in this case* $\frac{nh_n^3}{log\, n} \to \infty$.

2. If $\frac{nh_n^3}{log(n)} \to \infty$, *then by the triangle inequality*

$$\begin{aligned}\frac{1}{h_n}\sup_{x\in G}|\hat{f}(x)-f(x)| &\le \frac{1}{h_n}\sup_{x\in G}|\hat{f}(x)-E(\hat{f}(x))| + \frac{1}{h_n}\sup_{x\in G}|E(\hat{f}(x))-f(x))| \\ &= o_p(1) + \frac{1}{h_n}\sup_{x\in G}|E(\hat{f}(x))-f(x))|.\end{aligned}$$

If f *satisfies a Lipschitz condition of order 1 in* G, *that is* $|f(x) - f(x')| < C|x - x'|$ *for all* $x, x' \in G$ *and* $\int K(u)du = 1$, *then* $|E(\hat{f}(x)) - f(x)| \le \int |f(x + h_n u) - f(x)||K(u)|du \le h_n C \int |u||K(u)|du$. *Hence, if* $\int |u||K(u)|du < \infty$, $\frac{1}{h_n}\sup_{x\in G}|E(\hat{f}(x)) - f(x))| = O(1)$ *and* $\sup_{x\in G}|\hat{f}(x) - f(x)| = h_n O_p(1)$.

3. If f *satisfies the Lipschitz condition in remark 2 and* $\int uK(u)du < \infty$, *then for any given* n *the mean squared error in Theorem 9 is bounded above as* $E\left((\hat{f}(x) - f(x))^2\right) \le h_n^2 C_1 + \frac{1}{nh_n}C_2$. *The first term is a bound on the square of the bias and the second is a bound on the variance. The bias bound increases as* h_n *increases whereas the variance bound diminishes as* h_n *increases. This is the bias-variance tradeoff associated with the choice of* h_n *(for fixed* n*). The bound on the right-hand side of the inequality is minimized at* $h_n^* = \left(\frac{C_2}{2C_1}\right)^{1/3} n^{-1/3}$. *Hence, substituting* h_n^* *in the expression for the bound we have* $E\left((\hat{f}(x) - f(x))^2\right) = O(n^{-2/3})$.

4. If $h_n = Cn^{\delta}$ *with* $-1 < \delta < 0$ *then* $\frac{nh_n}{log(n)} = \frac{n^{1+\delta}C}{log(n)}$. *By L'Hôpital's rule and the fact that* $-1 < \delta < 0$ *we have* $\frac{nh_n}{log(n)} \to \infty$ *as* $n \to \infty$.

5. If $h_n = Cn^{\delta}$ *then* $nh_n^3 = n^{1+3\delta}C^3 \to \infty$ *if* $-1/3 < \delta$. *As in remark 3, by L'Hôpital's rule and the fact that* $-1/3 < \delta < 0$ *we have* $\frac{nh_n^3}{log(n)} \to \infty$ *as* $n \to \infty$.

Remarks 10.1 and 10.2 together with Theorem 13 provide the following corollary.

Corollary 1. *Assume a)* $|K(u)| \leq C$ *for all* $u \in \mathbb{R}$*; b)* $\int |K(u)|du < \infty$*; c)* $|u||K(u)| \to 0$ *as* $|u| \to \infty$*; d) for any* $u, u' \in \mathbb{R}$ *with* $u \neq u'$ *we have* $|K(u') - K(u)| < C|u' - u|$ *for some arbitrary constant* $C < \infty$ *(Lipschitz condition of order 1 in* $\mathbb{R}$*); e)* $\int |u||K(u)|du < \infty$*; f)* $h_n \to 0$ *and* $\frac{nh_n^3}{log(n)} \to \infty$ *as* $n \to \infty$*. Then, if* f *is continuous and if it satisfies a Lipschitz condition of order 1 in* G*, a compact subset of* $\mathbb{R}$*,*

$$\frac{1}{h_n} \sup_{x \in G} |\hat{f}(x) - f(x)| = O_p(1).$$

2 Univariate density estimation over $[0, \infty)$

In practice, it is common to encounter non-negative random variables where the support of the density f is the half-line $[0, \infty)$. In this case, the Rosenblatt kernel density estimator $\hat{f}(x)$ generally has bad properties. To see the source of the problem, suppose $f \in \Sigma(\beta, C)$, K is even and of order $\lfloor \beta \rfloor$ with compact support $[-1, 1]$. Then, for $x \in [h_n, \infty)$ the bias of $\hat{f}(x)$ is given by

$$\begin{aligned} B(\hat{f}(x)) &= \int_{-1}^{x/h_n} K(u)(f(x - h_n u) - f(x))du - f(x) \int_{x/h_n}^{\infty} K(u)du \\ &= \int_{-1}^{1} K(u)(f(x - h_n u) - f(x))du - f(x) \times 0 \leq h_n^{\beta} C. \end{aligned}$$

But if $x \in [0, h_n)$ (or $x \in [0, n^{-1/(2\beta+1)})$) and for $\lambda \in (0, 1)$,

$$\begin{aligned} B(\hat{f}(x)) &= -f(x) \int_{x/h_n}^{1} K(u)du - h_n f^{(1)}(x) \int_{-1}^{x/h_n} uK(u)du + \cdots \\ &\quad + \frac{(-h_n)^{\lfloor \beta \rfloor}}{\lfloor \beta \rfloor !} \int_{-1}^{x/h_n} f^{(\lfloor \beta \rfloor)}(x - \lambda h_n u) u^{\lfloor \beta \rfloor} K(u)du \neq O(h_n^{\beta}) \\ V(\hat{f}(x)) &\leq \frac{C}{nh_n} \int_{-1}^{x/h_n} K^2(u)du. \end{aligned}$$

Unless $f(x) = 0$, $\hat{f}(x)$ is inconsistent. In the typical case where $\beta = 2$, even if $f(x) = 0$, the rate of decay of the bias in the vicinity of the boundary is h_n, rather than h_n^2, away from the boundary. Hence, the rate of decay of the mean squared error and the optimal bandwidth differ in the vicinity of the boundary. We will now define a new kernel density estimator that does not suffer from this boundary problem.

2.1 Hestenes extension and a new estimator for f

Let $w_1, ..., w_{s+1}$ be pairwise different positive numbers for $s = 0, 1, \cdots$. Of special interest are the decreasing sequence $w_i = 1/i$, $i = 1, ..., s+1$ (used by Hestenes (1941)) and the increasing sequence $w_i = i$. Let the numbers $k_1, ..., k_{s+1}$ be defined from the following system

$$\sum_{i=1}^{s+1} (-w_i)^j k_i = 1, \; j = 0, ..., s. \tag{8}$$

Since this system has the Van-der-Monde determinant

$$\begin{vmatrix} 1 & 1 & ... & 1 \\ -w_1 & -w_2 & ... & -w_{s+1} \\ ... & ... & ... & \\ (-w_1)^s & (-w_2)^s & ... & (-w_{s+1})^s \end{vmatrix} \neq 0,$$

$k_1, ..., k_{s+1}$ are uniquely defined. Let $s \in \mathbb{N}$ and $\Omega \subseteq \mathbb{R}$. The class of functions $f : \Omega \to \mathbb{R}$ which are s-times differentiable with $\left| f^{(s)}(x) \right| \leq C$ for some $0 < C < \infty$ and continuous is denoted by $\mathcal{C}_b^s(\Omega)$. If $f \in \mathcal{C}_b^s$ on its domain $D_f = [0, \infty)$, its Hestenes extension to $(-\infty, 0)$ is given by

$$\phi_s(x) = \sum_{j=1}^{s+1} k_j f(-w_j x), \; x < 0. \tag{9}$$

Note that if f is a density function, ϕ_s is not a density, but a linear combination of densities $w_j f(-w_j x)$ with coefficients k_j / w_j. Assuming that f has s right-hand derivatives $f(0+), ..., f^{(s)}(0+)$ at zero ($s = 0$ means continuity), we see that the following sewing conditions at zero are satisfied due to (8):

$$\phi_s^{(m)}(0-) = \sum_{j=1}^{s+1} (-w_j)^m k_j f^{(m)}(0+) = f^{(m)}(0+), \; m = 0, 1, \cdots, s.$$

Now, define g_s on $\mathbb{R}$ by

$$g_s(x) = \begin{cases} f(x), & x \geq 0 \\ \phi_s(x), & x < 0 \end{cases}, \tag{10}$$

with g_s being s times differentiable.

Definition 7. *Let $f \in \mathcal{C}_b^s$, K be an even kernel and $\{X_i\}_{i=1}^n$ be an IID sequence of random variables with density f. We define the estimator of $\hat{f}_s(x)$, for $x \geq 0$, by*

$$\hat{f}_s(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \left[K\left(\frac{x - X_i}{h_n} \right) + \sum_{j=1}^{s+1} \frac{k_j}{w_j} K\left(\frac{x + X_i / w_j}{h_n} \right) \right]. \tag{11}$$

Remark 11. *1. When $s = 0$ in (11), $\hat{f}_0(x)$ is the "reflection estimator" from Schuster (1985), i.e.,*

$$\hat{f}_0(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \left[K\left(\frac{x - X_i}{h_n}\right) + K\left(\frac{x + X_i}{h_n}\right) \right].$$

2. We note that Schuster's estimator does not depend on s, the index on $\mathcal{C}_b^s$. Thus, knowledge that $s > 0$ is not used in constructing his estimator, whereas it is central in the definition of $\hat{f}_s(x)$.

Theorem 14. *Suppose $f \in \mathcal{C}_b^s$ on $D_f = [0, \infty)$ and K is a kernel with $\int K(u)du = 1$. Then, the bias of $\hat{f}_s(x)$ has the representation*

$$E(\hat{f}_s(x)) - f(x) = \int_{\mathbb{R}} K(t) \left[g_s(x - h_n t) - g_s(x)\right] dt, \; x \in D_f. \tag{12}$$

Proof. By the IID assumption

$$\begin{aligned} E\left(\hat{f}_s(x)\right) &= \frac{1}{h_n} E\left[K\left(\frac{x - X_1}{h_n}\right) + \sum_{j=1}^{s+1} \frac{k_j}{w_j} K\left(\frac{x + X_1/w_j}{h_n}\right) \right] \\ &= \frac{1}{h_n} \left[\int_0^\infty K\left(\frac{x - t}{h_n}\right) f(t)dt + \sum_{j=1}^{s+1} \frac{k_j}{w_j} \int_0^\infty K\left(\frac{x + t/w_j}{h_n}\right) f(t)dt \right]. \end{aligned} \tag{13}$$

In the first integral let $u = \frac{x-t}{h_n}$, in the others $u = \frac{x+t/w_j}{h_n}$. Then

$$\begin{aligned} E\left(\hat{f}_s(x)\right) &= \frac{1}{h_n} \left[-h_n \int_{x/h_n}^{-\infty} K(u) f(x - h_n u)du + h_n \sum_{j=1}^{s+1} k_j \int_{x/h_n}^{\infty} K(u) f\left(-w_j(x - h_n u)\right) du \right] \\ &= \left[\int_{-\infty}^{x/h_n} K(u) f(x - h_n u)du + \int_{x/h_n}^{\infty} K(u) \sum_{j=1}^{s+1} k_j f\left(-w_j(x - h_n u)\right) du \right]. \end{aligned}$$

In the first integral we have $x - h_n u > 0$ and $f(x - h_n u) = g_s(x - h_n u)$; in the second one $x - h_n u < 0$, so $\sum_{j=1}^{s+1} k_j f\left(-w_j(x - h_n u)\right) = g_s(x - h_n u)$. Hence,

$$E\left(\hat{f}_s(x)\right) = \int_{\mathbb{R}} K(u) g_s(x - h_n u)du. \tag{14}$$

Since $\int_{\mathbb{R}} K(t)dt = 1$, this implies (12). $\square$

Remark 12. *The integral representation for bias obtained in Theorem 14 depends on the extension g_s, not the density f. Consequently, existing results for smooth functions (not densities) on $\mathbb{R}$ allow us to obtain bias estimates.*

Theorem 15. *Assume that* $\int K(t)dt = 1$, $\int_{\mathbb{R}} K(t)t^j dt = 0$, *for* $j = 1, ..., s-1$, $\int_{\mathbb{R}} |K(t)t^s| \, dt < \infty$ *and* $\left|f^{(s)}(x)\right| < C$ *for all* $x \geq 0$, *then*

$$E\left(\hat{f}_s(x)\right) - f(x) = O(h^s) \text{ for all } x \in D_f. \tag{15}$$

Proof. Since $\int K(t)dt = 1$ we have from Theorem 14

$$E(\hat{f}_s(x)) - f(x) = \int K(t)(g_s(x - h_n t) - g_s(x))dt = \int K(t) \left(g_s^{(1)}(x)(-h_n t) + \frac{1}{2!} g_s^{(2)}(x)(-h_n t)^2 + \cdots + \frac{1}{s!} g_s^{(s)}(x - th\tau)(-h_n t)^s \right) dt$$

for some $\tau \in (0, 1)$. If $\int_{\mathbb{R}} K(t)t^j dt = 0$, for $j = 1, ..., s-1$ then

$$|E(\hat{f}_s(x)) - f(x)| \leq \frac{h_n^s}{s!} \int |t|^s |K(t)||g^{(s)}(x - th\tau)|dt \leq C h_n^s,$$

where the last inequality follows from the assumptions that $\int_{\mathbb{R}} |K(t)t^s| \, dt < \infty$, $\left|f^{(s)}(x)\right| < C$ for all $x \geq 0$ and the structure of $g^{(s)}$. $\square$

Remark 13. *1. In the density estimation literature, it is usually assumed that* $f \in \mathcal{C}_b^2$. *In this case, under the conditions in Theorem 15 we have* $E\hat{f}_2(x) - f(x) = \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} t^2 K(t)dt + o(h_n^2)$ *for all* $x \in D_f$. *This expression, similar to what is obtained for the classical Rosenblatt-Parzen estimator when* $D_f = \mathbb{R}$, *contrasts with what is obtained for Schuster's reflection estimator. In particular, for boundary points, viz.,* $x = ch$ *for* $0 \leq c < \infty$, *we have (see Marron and Ruppert (1994))*

$$\begin{aligned} E\hat{f}_0(x) - f(x) &= \int K(t)(g_0(x - h_n t) - g_0(x))dt = \int_{-\infty}^{x/h_n} K(t)f(x - h_n t)dt \\ &+ \int_{x/h_n}^{\infty} K(t)f(-(x - h_n t))dt - \int K(t)g_0(x)dt \\ &= \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} t^2 K(t)dt - 2hf^{(1)}(x)\left(c\mu_{0,-c} + \mu_{1,-c}\right) \\ &+ 2h^2 f^{(2)}(x)(c^2\mu_{0,-c} + c\mu_{1,-c}) + o(h_n^2) \end{aligned}$$

where $\mu_{\ell,-c} = \int_{-\infty}^{-c} u^\ell K(u)du$ *for* $\ell = 0, 1, 2$. *The additional bias terms result from the fact that Schuster's estimator does not use the additional smoothness (beyond continuity) of* f *in its construction.*

2. *Theorems 14 and 15, require the existence of densities and their derivatives up to order s at $x = 0$. As such, in the case of densities that either diverge to infinity, or have derivatives that diverge to infinity near the boundary, the bias order we have derived do not hold. In particular, when $f(x) \to \infty$ as $x \to 0$ (a pole at $x = 0$) the bias diverges to infinity at $x = 0$. If $f(0)$ is finite, but $f^{(1)}(x) \to \infty$ as $x \to 0$, the bias decays at a speed that is slower than h^s.*

The following theorem provides the order of the variance for $\hat{f}_s(x)$. It is necessary to consider the boundary $x = 0$ and interior points ($x > 0$) separately. Let $k_0 = w_0 = -1$. The quantities

$$\Gamma = \int_{\mathbb{R}} K^2(t)dt \text{ and } \Gamma_l = \int_0^\infty \left[\sum_{j=0}^{s+1} \frac{k_j}{w_j} K\left(\frac{t}{w_j}\right) \right]^2 dt,$$

will appear in variance expressions inside the domain and at the left boundary, respectively.

Theorem 16. *Suppose the conditions in Theorem 15 hold, $\sup_{x \geq 0} f(x) < \infty$, $\sup_x |K(x)| < \infty$ and $\int_{\mathbb{R}} |K(t)|\, dt < \infty$. Then, I) for fixed $x > 0$*

$$V\left(\hat{f}_s(x)\right) = \frac{1}{nh}\left\{f(x)\Gamma + o(1)\right\}, \tag{16}$$

II) at the left boundary

$$V\left(\hat{f}_s(0)\right) = \frac{1}{nh}\left\{f(0)\Gamma_l + o(1)\right\}. \tag{17}$$

Proof. By the IID assumption, the variance of $\hat{f}_s(x)$ is given by $V(\hat{f}_s(x)) = \frac{1}{n}\left(E(u_1^2) - E(u_1)^2\right)$ where

$$u_1 = I_{\{X_1 \geq 0\}} \frac{1}{h_n}\left[K\left(\frac{x - X_1}{h_n}\right) + \sum_{j=1}^{s+1} \frac{k_j}{w_j} K\left(\frac{x + X_1/w_j}{h_n}\right)\right].$$

Since $E(u_1) = E(\hat{f}_s(x)) \to f(x)$, we have that $\frac{1}{n}E(u_1)^2 = O(n^{-1})$. Now,

$$\begin{aligned}
\frac{1}{n}E(u_1^2) &= \frac{1}{nh_n}\left(\frac{1}{h_n}E\left(I_{\{X_1 \geq 0\}}K^2\left(\frac{x - X_1}{h_n}\right)\right)\right. \\
&\quad + \frac{2}{h_n}\sum_{j=1}^{s+1} E\left(I_{\{X_1 \geq 0\}}\frac{k_j}{w_j}K\left(\frac{x - X_1}{h_n}\right)K\left(\frac{x + X_1/w_j}{h_n}\right)\right) \\
&\quad \left. + \frac{1}{h_n}E\left(\sum_{j=1}^{s+1}\frac{k_j}{w_j}I_{\{X_1 \geq 0\}}K\left(\frac{x + X_1/w_j}{h_n}\right)\right)^2\right) \\
&= \frac{1}{nh_n}\left(I_{1n} + I_{2n} + I_{3n}\right)
\end{aligned}$$

By a change of variable in integration

$$I_{1n} = \int_{-\infty}^{x/h_n} K^2(u) f(x - h_n u) du \to \begin{cases} f(0) \int_{-\infty}^{0} K^2(u) du, & \text{if } x = 0 \\ f(x) \int_{-\infty}^{\infty} K^2(u) du & \text{if } x > 0 \end{cases}$$

as $n \to \infty$.

$$\begin{aligned} I_{3n} &= \sum_{j=1}^{s+1} \sum_{l=1}^{s+1} \frac{k_j}{w_j} \frac{k_l}{w_l} \frac{1}{h_n} E\left(I_{\{X_1 \geq 0\}} K\left(\frac{x + X_1/w_j}{h_n}\right) K\left(\frac{x + X_1/w_l}{h_n}\right)\right) \\ &= \sum_{j=1}^{s+1} \sum_{l=1}^{s+1} \frac{k_j}{w_j} \frac{k_l}{w_l} I_{jln}. \end{aligned}$$

Now, for $j \neq l$ and $x > 0$, a change of variable $u = \frac{x + X_1/w_j}{h_n}$ we have

$$I_{jln} = w_j \int_{x/h_n}^{\infty} K(u) K\left(u \frac{w_j}{w_l} + \frac{x(1 - w_j/w_l)}{h_n}\right) f((h_n u - x) w_j) du.$$

Since $|K(u)| = o(1)$ as $u \to \infty$ and $h_n \to 0$, we have that $I_{jln} \to 0$ as $n \to \infty$. If $x = 0$, then

$$I_{jln} = w_j \int_{0}^{\infty} K(u) K\left(u \frac{w_j}{w_l}\right) f(h_n u w_j) du \to f(0) w_j \int_{0}^{\infty} K(u) K\left(u \frac{w_j}{w_l}\right) du.$$

For $j = l$ and $x > 0$

$$I_{jjn} = w_j \int_{x/h_n}^{\infty} K^2(u) f((h_n u - x) w_j) du \to 0$$

and for $x = 0$, $I_{jjn} \to w_j f(0) \int_0^\infty K^2(u) du$. In all,

$$I_{3n} = \begin{cases} o(1) & \text{if } x > 0 \\ \sum_{j=1}^{s+1} \frac{k_j^2}{w_j} f(0) \int_0^\infty K^2(u) du + f(0) \sum_{j=1}^{s+1} \sum_{l=1, l \neq j}^{s+1} \frac{k_l}{w_l} k_j \int_0^\infty K(u) K\left(u \frac{w_j}{w_l}\right) du & \text{if } x = 0 \end{cases}$$

Similar arguments give,

$$I_{2n} = \begin{cases} o(1) & \text{if } x > 0 \\ 2 f(0) \sum_{j=1}^{s+1} k_j \int_0^\infty K(u) K(-u w_j) du & \text{if } x = 0 \end{cases}.$$

Combining the orders and expressions for I_{1n}, I_{2n}, I_{3n} we have, for $x > 0$

$$V(\hat{f}_s(x)) = \frac{f(x)}{n h_n} \int_{\mathbb{R}} K^2(u) du + o((n h_n)^{-1})$$

and for $x = 0$

$$V(\hat{f}_s(0)) = \frac{f(0)}{nh_n}\left(\int_{-\infty}^{0} K^2(u)du + \sum_{j=1}^{s+1}\frac{k_j^2}{w_j}\int_0^\infty K^2(u)du \right.$$
$$+ f(0)\sum_{j=1}^{s+1}\sum_{l=1,l\neq j}^{s+1}\frac{k_l}{w_l}k_j\int_0^\infty K(u)K\left(u\frac{w_j}{w_l}\right)du$$
$$\left.+2f(0)\sum_{j=1}^{s+1}k_j\int_0^\infty K(u)K(-uw_j)du\right) + o((nh_n)^{-1}).$$

□

Remarks. 1. It is a direct consequence of Theorem 23 that, if $\int |K(u)|^{2+\delta}du < C$ for some $\delta > 0$, by Liapounov's Central Limit Theorem we have for $x > 0$

$$\sqrt{nh}\left(\hat{f}_s(x) - E\left(\hat{f}_s(x)\right)\right) \xrightarrow{d} N\left(0, f(x)\Gamma\right), \tag{18}$$

and for $x = 0$,

$$\sqrt{nh}\left(\hat{f}_s(0) - E\left(\hat{f}_s(0)\right)\right) \xrightarrow{d} N\left(0, f(0)\Gamma_l\right). \tag{19}$$

2. For $f \in \mathcal{C}_b^2$, equation (19) together with the expression for $E(\hat{f}_2(x))$ gives

$$\sqrt{nh}\left(\hat{f}_2(0) - \left(f(0) + \frac{h^2}{2}f^{(2)}(0+)\int_{\mathbb{R}} t^2K(t)dt + o(h^2)\right)\right) \xrightarrow{d} N\left(0, f(0)\Gamma_l\right),$$

which can be used to construct confidence intervals and conduct hypothesis testing.