Hazard function estimation by local smoothing techniques

Ming-Yen Cheng\textsuperscript{1}, Liang Peng\textsuperscript{2} and Shan Sun\textsuperscript{3}

September 4, 2003

Abstract

In this paper we propose to estimate the hazard function based on local smoothing techniques for both i.i.d and censoring data. Such estimators are known as to have no boundary effects while the estimators based on kernel function have the boundary effect, as pointed out by Müller and Wang (1990). We derive the asymptotic normalities of the local smooth estimators. Also, we study the comparisons between the local smooth estimators and the kernel smooth estimators for most commonly used kernel functions. It turns out that our local smooth estimators with optimal bandwidths produce smaller biases than that of the kernel smooth estimators. However, such estimators have large variances than that of the kernel smooth estimators. To overcome this problem, we apply the variance reduction technique in Cheng, Wu and Yen (2003) to our estimators. The resulted estimators have the same asymptotic biases as the local smooth estimators and smaller asymptotic variances than the kernel estimators.

Keywords. Hazard function, kernel smooth estimation, local polynomial estimation, variance reduction.

\textsuperscript{1}Department of Mathematics, National Taiwan University, Taipei 106, Taiwan. Email: cheng@math.ntu.edu.tw
\textsuperscript{2}School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332-0160, USA. Email: peng@math.gatech.edu
\textsuperscript{3}Department of Mathematics and Statistics, Texas Tech University, Lubbock, Texas 79409 - 1042, USA. Email: ssun@math.ttu.edu
1 Introduction

Hazard function based on i.i.d. or censored data is important. It provides useful information in reliability theory and survival analysis, as well as in the fields as diverse as engineering, medical statistics and geophysics. A variety of inferential procedures have been proposed to estimate the hazard function nonparametrically. Estimators of hazard function based on kernel smooth estimation have been studied extensively in the literature. For related investigations in this direction we refer to Watson and Leadbetter (1964), Murphy (1965), Rice and Rosenblatt (1976), Singpurwalla and Wong (1983) and Patil (1997), under the i.i.d. case. Under the censoring case, see the discussions in Tanner and Wong (1983), Tanner (1983), Schäfer (1985), Liu and Ryzin (1985), Diehl and Stute (1988), Lo, Mack and Wang (1989), Müller and Wang (1990), Patil (1993), Wang (1999). It was pointed out that the drawback of using the kernel smooth estimators are known to have the boundary effect (see by Müller and Wang (1990)).

In this paper we propose another type of smooth estimation for the hazard function based on local smoothing techniques, which may be argued to have no boundary effect. We show that our local smooth estimators have smaller asymptotic biases, but larger asymptotic variances, than the kernel smooth estimators under the case with or without censoring. Proofs of these results are nontrivial. To reduce the variances of our local smooth estimators, we apply the variance reduction technique introduced by Cheng, Wu and Yen (2003). Hence, our variance reduced local smooth estimators are better than the kernel smooth estimators in terms of optimal asymptotic mean squared error or asymptotic bias and variance with the same bandwidth. A numerical study demonstrates that these advantageous asymptotic properties are also apparent in finite sample sizes.

We organize this paper as follows. In Section 2, we establish the weak convergence of the local smooth estimators for both i.i.d and censoring cases. In Section 3, we provide some comparisons between our local smooth estimators and the kernel smooth estimators. In Section 4, we propose variance reduced local smooth estimators and compare them with the kernel smooth estimators. A simulation study is given in Section 5. All proofs are deferred till Section 6.
2 Local smooth estimation

Throughout this paper we assume that

A1) $k(x)$ is a symmetric density function with support $[-1, 1]$;

A2) $f''(x)$ exists and is continuous, where $f$ is defined below;

A3) $h = h(n) > 0$, $h \to 0$ and $\sqrt{nh}h^2 \to b \in [0, \infty)$ as $n \to \infty$.

2.1. The case without censoring. Let $X_1, \cdots, X_n$ be independent and identically distributed survival times with distribution function $F(x)$ and density function $f(x)$. Our aim is to estimate the hazard function $\lambda(x) = \frac{f(x)}{1 - F(x)}$ using local smooth techniques. We apply local smoothing techniques, see for example Fan and Gijbels (1996), to estimate the derivative of $\Lambda(x) = -\log(1 - F(x))$, i.e., $\lambda(x)$ as follows.

Let $F_n(x) = \frac{1}{n}\sum_{i=1}^{n}I(X_i \leq x)$ be the empirical distribution of the sample $\{X_1, X_2, ..., X_n\}$ and define $\Lambda_n(x) = -\log(1 - F_n(x))$. Observe the following regression model:

$$\Lambda_n(X_i) = \Lambda(X_i) + \text{error}, i = 1, \cdots, n,$$

and let $(\hat{a}, \hat{b}, \hat{c})$ be the value of $(a, b, c)$ that minimizes the following kernel weighted squared errors:

$$\sum_{j=1}^{n}\{\Lambda_n(X_j) - a - b(X_j - x) - c(X_j - x)^2\}^2 k\left(\frac{x - X_j}{h}\right).$$

Then our new local smooth estimator for $\lambda(x)$ is defined as $\hat{\lambda}_n(x) = \hat{b}$ and has the following explicit expression

$$\hat{\lambda}_n(x) = \frac{\sum_{j=1}^{n}\Lambda_n(X_j)k\left(\frac{x - X_j}{h}\right)\left[\Delta_{n,1}(x) + (X_j - x)\Delta_{n,2}(x) + (X_j - x)^2\Delta_{n,3}(x)\right]}{\Delta_n(x)},$$

where

$$\Delta_{n,1}(x) = \sum_{j=1}^{n}\Lambda_n(X_j)k\left(\frac{x - X_j}{h}\right),$$

$$\Delta_{n,2}(x) = \sum_{j=1}^{n}\Lambda_n(X_j)k\left(\frac{x - X_j}{h}\right)\left[1 + \frac{x - X_j}{h}\right],$$

$$\Delta_{n,3}(x) = \sum_{j=1}^{n}\Lambda_n(X_j)k\left(\frac{x - X_j}{h}\right)\left[\frac{x - X_j}{h} + \frac{(x - X_j)^2}{2h^2}\right].$$
where \( s_{n,l}(x) = \sum_{j=1}^{n} (x - X_j)^l k\left(\frac{x - X_j}{h}\right) \), \( l = 0, 1, 2, 3, 4 \), and

\[
\begin{align*}
\Delta_n(x) &= s_{n,1}^3(x)s_{n,4}(x) + s_{n,1}(x)s_{n,2}^3(x) + s_{n,0}(x)s_{n,1}(x)s_{n,3}^2(x) \\
&\quad - 2s_{n,1}(x)s_{n,2}(x)s_{n,3}(x) - s_{n,0}(x)s_{n,1}(x)s_{n,2}(x)s_{n,4}(x) \\
\Delta_{n,1}(x) &= s_{n,1}(x)s_{n,2}(x)s_{n,3}(x) - s_{n,1}^2(x)s_{n,4}(x) \\
\Delta_{n,2}(x) &= s_{n,1}(x)s_{n,2}^2(x) - s_{n,0}(x)s_{n,1}(x)s_{n,4}(x) \\
\Delta_{n,3}(x) &= s_{n,1}^2(x)s_{n,2}(x) - s_{n,0}(x)s_{n,1}(x)s_{n,3}(x)
\end{align*}
\]

Define

\[
c_1 = \int_{-1}^{1} s^2 k(s) \, ds, \quad c_2 = \int_{-1}^{1} s^4 k(s) \, ds, \quad c_3 = 2 \int_{-1}^{1} \left\{ \int_{-1}^{t} k(s)k(t)s^2 \, ds \right\} dt
\]

(2.2)

The following theorem provides us with the weak convergence of the local smooth estimator \( \hat{\lambda}_n(x) \).

**Theorem 1.** Under regularity conditions A1) – A3), we have for \( 1 - F(x) > 0 \)

\[
\sqrt{nh}\{\hat{\lambda}_n(x) - \lambda(x)\} \xrightarrow{d} N\left( \frac{b\lambda''(x)c_2}{6c_1}, \frac{f(x)c_3}{[1 - F(x)]^2c_1^2} \right)
\]

as \( n \to \infty \), where \( b \) is defined in condition A3 and \( c_1, c_2 \) and \( c_3 \) are defined in (2.2).

2.2. The case with censoring. Let \( X_1, \ldots, X_n \) be independent and identically distributed random variables with distribution function \( F(x) \) and density function \( f(x) \), and \( Y_1, \ldots, Y_n \) be independent and identically distributed random variables with distribution function \( G(y) \) and density function \( g(y) \). Suppose \( X_i^f \) and \( Y_i^f \) are independent and our observations are \( Z_i = \min(X_i, Y_i) \) with censoring indicators \( \delta_i = I(X_i \leq Y_i) \) for \( i = 1, \ldots, n \). Thus \( \delta_i = 1 \) indicates the survival time \( X_i \) for the \( i \)th individual is observed while \( \delta_i = 0 \) indicates \( X_i \) is not observed but it is known to be greater than \( Y_i \). Our aim is to estimate the hazard function \( \lambda(x) = \frac{f(x)}{1 - F(x)} \), which is of importance in many lifetime studies.

In estimating the distribution function \( F \), a popular nonparametric estimator \( F_n^* \) based on the right censored data \( \{Z_i, \delta_i\}, i = 1, \ldots, n \), is the well-known Kaplan-Meier (Kaplan and
Meier, 1958) estimator given by
\[ F_n^*(x) = \begin{cases} 1 - \prod_{j=1}^{n} \frac{N(Z_j)}{1 + N(Z_j)} I[Z_j \leq x, \delta_j = 1] & \text{if } x < \max(Z_1, \cdots, Z_n) \\ 1 & \text{elsewhere,} \end{cases} \]
where \( N(u) = \sum_{j=1}^{n} I(Z_j > u) \). The large sample properties of the product-limit estimator \( F_n^*(x) \) have drawn much attention in the literature; see Chen and Lo (1997) and references cited therein.

Define \( \Lambda_n^*(x) = -\log(1 - F_n^*(x)) \) and \( \Lambda(x) = -\log(1 - F(x)) \). Observe the following regression model:
\[ \Lambda_n^*(Z_i) = \Lambda(Z_i) + \text{error}, i = 1, \cdots, n. \]
For those \( i' \)'s such that \( \delta_i = 1 \), i.e., \( Z_i = X_i \), we apply local smoothing techniques to estimate the derivative of \( \Lambda(x) \), i.e., \( \lambda(x) \). That is, let \( (\hat{a}, \hat{b}, \hat{c}) \) be the value of \( (a, b, c) \) that minimizes the following kernel weighted squared errors:
\[ \sum_{j=1}^{n} \delta_j \{ \Lambda_n^*(Z_j) - a - b(Z_j - x) - c(Z_j - x)^2 \}^2 k\left( \frac{x - Z_j}{h} \right). \]
Then our local smooth estimator for \( \lambda(x) \) is defined as \( \hat{\lambda}_n^*(x) = \hat{b} \) and has the following explicit expression
\[ \hat{\lambda}_n^*(x) = \frac{\sum_{j=1}^{n} \Lambda_n^*(Z_j) k\left( \frac{x - Z_j}{h} \right) \left[ \Delta_{n,1}(x) + (Z_j - x) \Delta_{n,2}(x) + (Z_j - x)^2 \Delta_{n,3}(x) \right]}{\Delta_n(x)}, \quad (2.3) \]
where \( s_{n,t}(x) = \sum_{j=1}^{n} \delta_j (x - Z_j)^l k\left( \frac{x - Z_j}{h} \right), l = 0, 1, 2, 3, 4 \), and \( \Delta_n(x), \Delta_{n,1}(x), \Delta_{n,2}(x) \) and \( \Delta_{n,3}(x) \) are defined as in (2.1). The following theorem provides us the weak convergence of the local smooth estimator \( \hat{\lambda}_n^*(x) \).

**Theorem 2.** Under regularity conditions A1) - A3) and \( g(y) \) is continuous, we have for \( 1 - F(x) > 0 \)
\[ \sqrt{nh} \{ \hat{\lambda}_n^*(x) - \lambda(x) \} \overset{d}{\to} N\left( \frac{\hat{b} \lambda'(x)c_2}{6c_1}, \frac{f(x)c_3}{[1 - F(x)]^2[1 - G(x)]^2c_1^2} \right) \]
as \( n \to \infty \), where \( b \) is defined in condition A3 and \( c_1, c_2 \) and \( c_3 \) are given by (2.2).
3 Comparisons between kernel smooth estimators and local smooth estimators

In this section, we study some asymptotic properties of the kernel smooth estimators and local smooth estimators for both i.i.d and censored cases.

3.1. The case without censoring. Under the case of no censoring, the kernel smooth estimator studied by Singpurwala and Wong (1983) is defined as

\[ \hat{\lambda}_n(x) = \frac{1}{h} \sum_{j=1}^{n} k\left(\frac{X(j) - x}{h}\right)/(n - j + 1), \]

where \( X_{(1)} \leq \cdots \leq X_{(n)} \) denote the order statistics of \( X_1, \ldots, X_n \). Under the regularity conditions \( A1 - A3 \), Singpurwala and Wong (1983) proved that for \( 1 - F(x) > 0 \)

\[ \sqrt{nh}(\hat{\lambda}_n(x) - \lambda(x)) \overset{d}{\rightarrow} N\left(\frac{b\lambda''(x)c_1}{2}, \frac{f(x)c_4}{[1 - F(x)]^2}\right) \]

as \( n \rightarrow \infty \), where \( c_1 \) is given by (2.2), \( b \) satisfies condition \( A3 \) and

\[ c_4 = \int_{-1}^{1} k^2(x)dx. \]

Hence, by minimizing the asymptotic mean squared error, we obtain that the local optimal bandwidth for \( \hat{\lambda}_n(x) \) is

\[ \bar{h}_{opt} = n^{-1/5}\left\{ \frac{f(x)c_4}{[1 - F(x)]^2[\lambda''(x)]^2c_4^2} \right\}^{1/5}. \quad (3.1) \]

Thus the optimal asymptotic mean squared error of \( \hat{\lambda}_n(x) \) is given by

\[ amse(\hat{\lambda}_n(x), \bar{h}_{opt}) = n^{-4/5}\left\{ \frac{f(x)}{[1 - F(x)]^2} \right\}^{4/5}\left\{ \lambda''(x) \right\}^{2/5} \frac{5c_1^2c_4^2}{4}. \quad (3.2) \]

On the other hand, it follows from Theorem 1 that, by minimizing the asymptotic mean squared error (amse), the local optimal bandwidth for our local smooth estimator \( \hat{\lambda}_n(x) \) is

\[ \hat{h}_{opt} = n^{-1/5}\left\{ \frac{f(x)9c_3}{[1 - F(x)]^2[\lambda''(x)]^2c_2^2} \right\}^{1/5}. \quad (3.3) \]

Thus the optimal asymptotic mean squared error for \( \hat{\lambda}_n(x) \) is given by

\[ amse(\hat{\lambda}_n(x), \hat{h}_{opt}) = n^{-4/5}\left\{ \frac{f(x)}{[1 - F(x)]^2} \right\}^{4/5}\left\{ \lambda''(x) \right\}^{2/5} \frac{5c_1^2c_4^2}{4 \cdot 3^{2/5}c_2^2}. \quad (3.4) \]
3.2. The case with censoring. For the case of censoring, a kernel smooth estimator for \( \lambda(x) \) was proposed by Tanner and Wong (1983) as

\[
\tilde{\lambda}_n^*(x) = \frac{1}{h} \sum_{j=1}^{n} (n - j - 1)^{-1} \delta(j) k\left( \frac{x - Z(j)}{h} \right),
\]

where \( Z(1) \leq \cdots \leq Z(n) \) denote the order statistics of \( Z_1, \cdots, Z_n \) and \( \delta(1), \cdots, \delta(n) \) denote the corresponding censoring indicators. Under the same regularity conditions as in Theorem 2, Müller and Wang (1990) showed that for \( 1 - F(x) > 0 \)

\[
\sqrt{n} h (\tilde{\lambda}_n^* (x) - \lambda(x)) \overset{d}{\to} N \left( \frac{b \lambda''(x) c_1}{2}, \frac{f(x) c_4}{[1 - F(x)]^2 [1 - G(x)]} \right)
\]

as \( n \to \infty \). Hence, by minimizing the asymptotic mean squared error of \( \tilde{\lambda}_n^* (x) \), the local optimal bandwidth for kernel smooth estimator \( \tilde{\lambda}_n^* (x) \) is given by

\[
\hat{h}_{opt} = n^{-1/5} \left\{ \frac{f(x) c_4}{[1 - F(x)]^2 [1 - G(x)] [\lambda''(x)]^2 c_1^2} \right\}^{1/5}.
\]

Thus, the optimal asymptotic mean squared error for \( \tilde{\lambda}_n^* (x) \) is

\[
amse(\tilde{\lambda}_n^* (x), \hat{h}_{opt}) = n^{-4/5} \left\{ \frac{f(x)}{[1 - F(x)]^2 [1 - G(x)]} \right\}^{4/5} \left\{ \lambda''(x) \right\}^{2/5} \frac{5 c_1^{2/5} c_4^{4/5}}{4}.
\]

On the other hand, it follows from Theorem 2 that, by minimizing the asymptotic mean squared error (amse), the local optimal bandwidth for local smooth estimator \( \hat{\lambda}_n^* (x) \) is

\[
\hat{h}_{opt} = n^{-1/5} \left\{ \frac{f(x) 9 c_3}{[1 - F(x)]^2 [1 - G(x)] [\lambda''(x)]^2 c_2^2} \right\}^{1/5}.
\]

Therefore, the optimal asymptotic mean squared error for \( \hat{\lambda}_n^* (x) \) is given by

\[
amse(\hat{\lambda}_n^* (x), \hat{h}_{opt}) = n^{-4/5} \left\{ \frac{f(x)}{[1 - F(x)]^2 [1 - G(x)]} \right\}^{4/5} \left\{ \lambda''(x) \right\}^{2/5} \frac{5 c_3^{4/5} c_2^{2/5}}{4 \cdot 3^2/5 c_1^2}.
\]

3.3. Comparisons. For the purpose of comparison, we compute the values of \( c_1, \cdots, c_4 \) and the constant factors in the asymptotically optimal bandwidth and amse expressions (3.1)–(3.8) for four commonly used kernels, They are the Epanechnikov, Biweight, Triangular and Uniform kernels. The results are shown in Table 1.
First, let us look at the optimal amse for kernel smooth estimators and our local smooth estimators. From (3.2), (3.4), (3.6) and (3.8), we notice that for both i.i.d case and censored case, the difference in amse between these estimators are mainly the terms $\frac{c_2^{4/5}c_4^{2/5}}{3^{2/5}c_1^{4/5}}$ and $c_1^{2/5}c_4^{4/5}$. Table 1 shows that our local smooth estimators ($\hat{\lambda}_n$ and $\hat{\lambda}_n^*$) have smaller bias terms than that of the kernel estimators ($\tilde{\lambda}_n$ and $\tilde{\lambda}_n^*$), but with large variances in general, for the four commonly used kernel functions. It is also interesting to see that the optimal mean squared error is the same for local smooth estimators with uniform kernel and for the kernel smooth estimators with Epanechnikov kernel. In fact, this optimal mean squared error is the smallest among the four kernels for both local and kernel smooth estimators.

Next, let us compare the optimal bandwidths for both local ($\hat{\lambda}_n$ and $\hat{\lambda}_n^*$) and kernel smooth estimators ($\tilde{\lambda}_n$ and $\tilde{\lambda}_n^*$). Observe equations (3.1), (3.3), (3.5) and (3.7), we see that, for both i.i.d. case and censoring case, the difference in optimal bandwidths for kernel and local smooth estimators are based on terms $\left(\frac{c_2}{c_1}\right)^{1/5}$ and $\left(\frac{c_2}{c_1}\right)^{1/5}$. From Table 1, we see that the optimal bandwidths for our local smooth estimators are larger than those for the kernel smooth estimators. So, in practice, one may prefer local smooth estimators to kernel smooth estimators since the larger optimal bandwidth will allow more data points in the local model.
4 Variance reduced local smooth estimation

Note that our local smooth estimator has a smaller asymptotic bias, but a larger asymptotic variance, than the kernel smooth estimator under the case with or without censoring. In this section we apply the variance reduction technique in Cheng, Wu and Yen (2003) to our local smooth estimators defined in Section 2.

4.1. The case without censoring. We consider the following variance reduced local smooth estimators

\[ \hat{\lambda}_n(x) = \frac{1 - \sqrt{2}}{4} \hat{\lambda}_n(x - (\sqrt{1/2} + 1)\delta h) + \frac{1}{2} \hat{\lambda}_n(x - \sqrt{1/2}\delta h) + \frac{1 + \sqrt{2}}{4} \hat{\lambda}_n(x - (\sqrt{1/2} - 1)\delta h) \]

where \( \delta > 0 \). This estimator is a linear combination of the three values \( \hat{\lambda}_n(x - (\sqrt{1/2} + 1)\delta h) \), \( \hat{\lambda}_n(x - \sqrt{1/2}\delta h) \) and \( \hat{\lambda}_n(x - (\sqrt{1/2} - 1)\delta h) \), and it is parallel to the form of the variance reduced local linear regression estimator of Cheng, Wu and Yen (2003). The principle of Cheng, Wu and Yen (2003) is to find the maximal relative variance reduction among all points in an interpolation interval of length \( 2\delta h \). In the current hazard estimation context, the covariance structure of the local smooth estimator at different locations is much more complicated than in the regression setting. For simplicity reasons we take \( \hat{\lambda}_n(x) \) the specified form. This may not achieve the most variance reduction. Nevertheless, \( \hat{\lambda}_n(x) \) admits a very simple form and it is shown that \( \hat{\lambda}_n(x) \) enjoys superior performance in both asymptotic and finite sample cases.

To analyze asymptotic properties of the new estimator, define

\[ c_5(a, b) = \int_{-1}^{1} \int_{-1}^{t-a+b} k(s)k(t)(s - b)st dsdt + \int_{-1}^{1} \int_{-1}^{t+a-b} k(s)k(t)(s - a)st dsdt \]  (4.1)

and

\[ c_6(\delta) = \frac{5}{8} c_3 + \frac{1 - \sqrt{2}}{4} c_5((\sqrt{1/2} + 1)\delta, \sqrt{1/2}\delta) - \frac{1}{8} c_5((\sqrt{1/2} + 1)\delta, (\sqrt{1/2} - 1)\delta) + \frac{1 + \sqrt{2}}{4} c_5(\sqrt{1/2}\delta, (\sqrt{1/2} - 1)\delta). \]  (4.2)

First we derive the asymptotic normality for our variance reduced local smooth estimator as follows.

**Theorem 3.** Under regularity conditions A1) - A3), we have for \( 1 - F(x) > 0 \) and \( \delta > 0 \)

\[ \sqrt{n}h \{ \hat{\lambda}_n(x) - \lambda(x) \} \overset{d}{\longrightarrow} N \left( \frac{b\lambda''(x)c_2}{6c_1}, \frac{f(x)c_6(\delta)}{[1 - F(x)]^2c_1^2} \right) \]
as \( n \to \infty \).

Second, we shall compare our variance reduced local smooth estimator \( \tilde{\lambda}_n(x) \) with the kernel smooth estimator \( \lambda_n(x) \) defined in Section 3.1. Since the local smooth estimator \( \lambda_n(x) \) with the uniform kernel and \( \tilde{\lambda}_n(x) \) with the Epanechnikov kernel have the same smallest optimal amse among the four different kernels considered in Section 3.3, we only compare between the variance reduced local smooth estimator with the uniform kernel and the kernel smooth estimator with the Epanechnikov kernel. In this case, we have for \( 0 \leq b - a \leq 2 \)

\[
    c_5(a, b) = \frac{1}{4} \int_{-1-a+b}^{3} \left\{ \int_{-1}^{t+\frac{a-b}{2}} (s-a)st \, ds \right\} dt + \frac{1}{4} \int_{-1-a+b}^{3} \left\{ \int_{t+\frac{a-b}{2}}^{t} (t-b)st \, ds \right\} dt \\
    + \frac{1}{4} \int_{-1-a+b}^{3} \left\{ \int_{t+\frac{a-b}{2}}^{t} (t-b)st \, ds \right\} dt \\
    = \frac{1}{4} \int_{-1-a+b}^{3} \left\{ \left( \frac{t+\frac{a-b}{2}}{2} \right)^3 + \frac{a(t+\frac{a-b}{2}-1)}{2} \right\} dt + \frac{1}{4} \int_{-1-a+b}^{3} \left( t^3 - \frac{t}{12} - \frac{1}{24} \right) dt \\
    = \frac{1}{15} - \frac{(b-a)^2}{12} + \frac{(b-a)^3}{24} - \frac{(b-a)^5}{480},
\]

and for \( b - a > 2 \), \( c_5(a, b) = 0 \). Therefore,

\[
    c_6(\delta) = \begin{cases} 
        1/15 - \delta^3/48 + 7\delta^5/960 & \text{if } 0 \leq \delta \leq 1 \\
        3/40 - \delta^2/24 + 3\delta^3/48 - \delta^5/960 & \text{if } 1 < \delta \leq 2 \\
        1/24 & \text{if } \delta > 2,
    \end{cases}
\]

i.e.,

\[
    c_6(\delta)/c_1^2 = \begin{cases} 
        3/5 - 3\delta^3/16 + 21\delta^5/320 & \text{if } 0 \leq \delta \leq 1 \\
        27/40 - 3\delta^2/8 + 3\delta^3/16 - 3\delta^5/320 & \text{if } 1 < \delta \leq 2 \\
        3/8 & \text{if } \delta > 2.
    \end{cases}
\]

Notice that \( \tilde{\lambda}_n(x) \) with \( \delta = 0 \) reduces to the original estimator \( \hat{\lambda}_n(x) \). By checking that \( \frac{d}{d\delta}c_6(\delta) \leq 0, \delta > 0 \), we see that \( c_6(\delta)/c_1^2 < \frac{3}{5} \) for any \( \delta > 0 \), i.e., \( \tilde{\lambda}_n(x) \) with the uniform kernel has a smaller asymptotic variance than \( \lambda_n(x) \) with the Epanechnikov kernel. Because both estimators have the same asymptotic bias, we conclude that the variance reduced local smooth estimator \( \tilde{\lambda}_n(x) \) with the uniform kernel is better than the kernel smooth estimator \( \lambda(x) \) with the Epanechnikov kernel in terms of optimal amse or amse with the same bandwidth.

Cheng, Wu and Yen (2003) discussed in detail the choice of the parameter \( \delta \). Larger values of \( \delta \) are preferred so that more variance reductions are achieved if the hazard function
is smooth. Otherwise, if the curve has sharp feature, second order bias may appear and play a role. In that case, smaller values of δ would still provide reasonable amount of variance reductions. Data-driven choice of δ can be done by, for example, cross-validation.

4.2. The case with censoring.

The variance reduced local smooth estimators in this case is defined as
\[ \tilde{\lambda}_n^*(x) = \frac{1}{4} - \sqrt{2} \tilde{\lambda}_n^*(x - (\sqrt{1/2} + 1) \delta h) + \frac{1}{2} \tilde{\lambda}_n^*(x - \sqrt{1/2} \delta h) + \frac{1}{4} \sqrt{2} \tilde{\lambda}_n^*(x - (\sqrt{1/2} - 1) \delta h), \]
where δ > 0. The asymptotic normality of this variance reduced local smooth estimator is given below. The comparison between \( \tilde{\lambda}_n(x) \) and \( \tilde{\lambda}_n(x) \) is similar to the i.i.d case in Section 4.1, hence is omitted here.

**Theorem 4.** Under regularity conditions A1) - A3) and that \( g(y) \) is continuous, we have for \( 1 - F(x) > 0 \) and any \( \delta > 0 \)
\[ \sqrt{nh} \{ \tilde{\lambda}_n^*(x) - \lambda(x) \} \overset{d}{\to} N \left( \frac{h\lambda''(x)c_2}{6c_1}, \frac{f(x)c_6(\delta)}{[1 - F(x)]^2[1 - G(x)]c_1^2} \right) \]
as \( n \to \infty \), where \( c_6 \) is defined as in (4.2).

5 Simulation study

A Monte Carlo study was conducted to demonstrate the advantage of our variance reduced local smooth estimator \( \tilde{\lambda}_n(x) \) over the kernel smooth estimator \( \bar{\lambda}_n(x) \), under the i.i.d setup. The uniform kernel and Epanechikov kernel were employed for \( \tilde{\lambda}_n(x) \) and \( \bar{\lambda}_n(x) \), respectively. Moreover, value of \( \delta \) in the definition of \( \tilde{\lambda}_n(x) \) was taken as one.

We generated 1000 pseudo-random samples of size \( n = 100 \) from Weibull distribution \( F(x) = 1 - \exp(-x^\alpha), \ x \geq 0 \). We took \( \alpha = 4 \) and compute \( \bar{\lambda}_n(x) \) and \( \tilde{\lambda}_n(x) \) at point \( x \) such that \( F(x) = 0.5 \) for \( h = \frac{\bar{h}_{opt}}{2} + \frac{j}{21}\bar{h}_{opt}, \ j = 0, 1, \cdots, 19 \), where \( \bar{h}_{opt} \) is defined in (3.1). In Figure 1, we plot the mean squared errors of \( \bar{\lambda}_n(x) \) and \( \tilde{\lambda}_n(x) \) against different \( h \). This figure clearly shows that \( \tilde{\lambda}_n(x) \) has a substantially smaller mean squared error than \( \bar{\lambda}_n(x) \). This confirms the asymptotic results.
6 Proofs

Proof of Theorem 1. Let $U_j = F(X_j), j = 1, \ldots, n, G_n(u) = \frac{1}{n} \sum_{i=1}^{n} I(U_i \leq u)$ and $\alpha_n(u) = \sqrt{n}(G_n(u) - u)$. Then, using the result of Komlós, Major and Tusnády (1975), there exists a sequence of Brownian bridges $B_n(u), 0 \leq u \leq 1, n = 1, 2, \ldots$, such that

$$\sup_{0 \leq u \leq 1} |\alpha_n(u) - B_n(u)| = O_p(n^{-1/2} \log n).$$

(6.1)

Note that

$$\begin{align*}
&n^{-1}h^{-1}s_{n,0}(x) = f(x) + O_p(h^2) \\
&n^{-1}h^{-3}s_{n,1}(x) = -f'(x)c_1 + O_p(h) \\
&n^{-1}h^{-3}s_{n,2}(x) = f(x)c_1 + O_p(h^2) \\
&n^{-1}h^{-5}s_{n,3}(x) = -f'(x)c_2 + O_p(h) \\
&n^{-1}h^{-5}s_{n,4}(x) = f(x)c_2 + O_p(h^2).
\end{align*}$$

(6.2)

Using (6.2) we obtain that

$$\begin{align*}
\Delta_{n,1}(x) &= o_p(n^3h^{11}) \\
\Delta_{n,2}(x) &= n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2) + O_p(n^3h^{10}) \\
\Delta_{n,3}(x) &= n^3h^9f(x)f'(x)^2c_1(c_2^2 - c_2) + O_p(n^3h^{10}) \\
\Delta_{n}(x) &= n^4h^{12}f^3(x)f'(x)c_1^2(c_2 - c_1^2) + O_p(n^4h^{13}).
\end{align*}$$

(6.3)

Since

$$\{\hat{\lambda}_n(x) - \lambda(x)\} \Delta_n(x)$$

$$\begin{align*}
&= \sum_{j=1}^{n} \{\Lambda_n(X_j) - \Lambda(X_j)\} k\left(\frac{x - X_j}{h}\right)\{\Delta_{n,1}(x) + (X_j - x)\Delta_{n,2}(x) + (X_j - x)^2\Delta_{n,3}(x)\} \\
&\quad + \sum_{j=1}^{n} \{\Lambda(X_j) - \lambda(x)(X_j - x) - \frac{1}{2}\lambda(x)(X_j - x)^2\} k\left(\frac{x - X_j}{h}\right) \\
&\quad \times \{\Delta_{n,1}(x) + (X_j - x)\Delta_{n,2}(x) + (X_j - x)^2\Delta_{n,3}(x)\} \\
&= I + II,
\end{align*}$$

12
it is easy to see that Theorem 1 holds if we show that

$$\sqrt{n}h^{-1}n^{-1}I \overset{d}{\to} N\left(0, \frac{f^2(x)[f'(x)]^2c_1(c_2-c_1)^2c_3}{[1-F(x)]^2}\right) \quad (6.4)$$

and

$$n^{-4}h^{-14}II \overset{p}{\to} \frac{1}{6}X''(x)f^2(x)f'(x)c_2c_1(c_2-c_1) \quad (6.5)$$
as $n \to \infty$. The proof of (6.5) is straightforward, hence it is omitted. To prove (6.4), we decompose $I$ as follows:

$$I = -\sum_{j=1}^{n} \log \frac{1-F_{n}(X_j)}{1-F(X_j)} \{\Delta_{n,1}(x) + (X_j-x)\Delta_{n,2}(x) + (X_j-x)^2\Delta_{n,3}(x)\}$$

$$= \sum_{j=1}^{n} \frac{F_n(X_j) - F(X_j)}{1 - F(X_j)} \{1 + O_p(n^{-1/2} \log n)\} \times k\left(\frac{x-X_j}{h}\right) \{\Delta_{n,1}(x) + (X_j-x)\Delta_{n,2}(x) + (X_j-x)^2\Delta_{n,3}(x)\}$$

$$= \left[\frac{1+O_p(h)+O_p(n^{-1/2} \log n)}{1-F(x)} \Delta_{n,1} \sum_{j=1}^{n} \{F_n(X_j) - F(X_j)\} k\left(\frac{x-X_j}{h}\right) \right]$$

$$+ \left[\frac{1+O_p(h)+O_p(n^{-1/2} \log n)}{1-F(x)} \{\Delta_{n,2} - n^3h^9f^2(x)f'(x)c_1(c_2-c_1)^2\} \right]$$

$$\times \sum_{j=1}^{n} \{F_n(X_j) - F(X_j)\} (X_j-x) k\left(\frac{x-X_j}{h}\right)$$

$$+ \left[\frac{1+O_p(h)+O_p(n^{-1/2} \log n)}{1-F(x)} \{\Delta_{n,3}(x) - n^3h^9f(x)[f'(x)]^2c_1(c_2-c_1)\} \right]$$

$$\times \sum_{j=1}^{n} \{F_n(X_j) - F(X_j)\} (X_j-x)^2 k\left(\frac{x-X_j}{h}\right)$$

$$+ \left[\frac{1+O_p(h)+O_p(n^{-1/2} \log n)}{1-F(x)} \sum_{j=1}^{n} \{F_n(X_j) - F(X_j)\} k\left(\frac{x-X_j}{h}\right) \right]$$

$$\times \{(X_j-x)n^3h^9f^2(x)f'(x)c_1(c_2-c_1) + (X_j-x)^2n^3h^9f(x)[f'(x)]^2c_1(c_2-c_1)\}$$

$$= I_1 + I_2 + I_3 + I_4. \quad (6.6)$$
Using (6.1)–(6.3), the terms $I_1$–$I_3$ in (6.6) satisfy that
\[
\sqrt{n}hn^{-1}h^{-12}I_j = o_p(1), \; j = 1, 2, 3.
\]

Next, the term $I_4$ in (6.6) can be estimated as follows:
\[
I_4 = \frac{1-F(x)}{1+O_p(n^{-1/2} \log n)}
\]
\[
= n \int \{F_n(s) - F(s)\}k \left( \frac{x-a}{h} \right) \{(s-x)n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2)
\]
\[
+ (s-x)^2n^3h^9f(x)f'(x)[f'(x)]^2c_1(c_1^2 - c_2) \} dF_n(s)
\]
\[
= \frac{n}{2} \int k \left( \frac{x-a}{h} \right) \{(s-x)n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2)
\]
\[
+ (s-x)^2n^3h^9f(x)f'(x)[f'(x)]^2c_1(c_1^2 - c_2) \} d[F_n(s) - F(s)]^2
\]
\[
+ n \int \{F_n(s) - F(s)\}k \left( \frac{x-a}{h} \right) \{(s-x)n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2)
\]
\[
+ (s-x)^2n^3h^9f(x)f'(x)[f'(x)]^2c_1(c_1^2 - c_2) \} f(s) \, ds
\]
\[
= \left[ -\frac{n}{2} \int [F_n(s) - F(s)]^2 d\{k \left( \frac{x-a}{h} \right) \}
\]
\[
\times [(s-x)n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2) + (s-x)^2n^3h^9f(x)[f'(x)]^2c_1(c_1^2 - c_2)] \right]
\]
\[
+ \left[ \sqrt{n} \int \{\sqrt{n}[F_n(s) - F(s)] - B_n(F(s))\}k \left( \frac{x-a}{h} \right)
\]
\[
\times \{(s-x)n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2) + (s-x)^2n^3h^9f(x)[f'(x)]^2c_1(c_1^2 - c_2) \} f(s) \, ds \right]
\]
\[
+ \left[ \sqrt{n} \int B_n(F(s))k \left( \frac{x-a}{h} \right)
\]
\[
\times \{(s-x)n^3h^9f^2(x)f'(x)c_1(c_2 - c_1^2) + (s-x)^2n^3h^9f(x)[f'(x)]^2c_1(c_1^2 - c_2) \} f(s) \, ds \right]
\]
\[
= III_1 + III_2 + III_3.
\]
\[EIII^2_3 = 2nh^2 \int_{-1}^{1} \int_{-1}^{t} \{F(x - th) - F(x - th)F(x - sh)\}k(s)k(t)\]
\[\times \{-sn^3h^{10}f^2(x)f'(x)c_1(c_2 - c_1^2) + s^2n^3h^{11}f(x)[f'(x)]^2c_1(c_1^2 - c_2)\}\]
\[\times \{tr^3h^{10}f^2(x)f'(x)c_1(c_2 - c_1^2) + t^2n^3h^{11}f(x)[f'(x)]^2c_1(c_1^2 - c_2)\}\]
\[\times f(x - sh)f(x - th) \, dsdt\]
\[= nh^2 \int_{-1}^{1} \int_{-1}^{1} F(x)[1 - F(x)]k(s)k(t)\]
\[\times \{-sn^3h^{10}f^2(x)f'(x)c_1(c_2 - c_1^2) + s^2n^3h^{11}f(x)[f'(x)]^2c_1(c_1^2 - c_2)\}\]
\[\times \{tr^3h^{10}f^2(x)f'(x)c_1(c_2 - c_1^2) + t^2n^3h^{11}f(x)[f'(x)]^2c_1(c_1^2 - c_2)\}\]
\[\times f(x - sh)f(x - th) \, dsdt\]
\[+ 2nh^2 \int_{-1}^{1} \int_{-1}^{t} \{-thf(x) + thF(x)f(x) + shF(x)f(x) + O(h^2)\}k(s)k(t)\]
\[\times \{-sn^3h^{10}f^2(x)f'(x)c_1(c_2 - c_1^2) + s^2n^3h^{11}f(x)[f'(x)]^2c_1(c_1^2 - c_2)\}\]
\[\times \{tr^3h^{10}f^2(x)f'(x)c_1(c_2 - c_1^2) + t^2n^3h^{11}f(x)[f'(x)]^2c_1(c_1^2 - c_2)\}\]
\[\times f(x - sh)f(x - th) \, dsdt\]
\[= O(n^7h^{24}) + 2n^7h^{23}f^7(x)[f'(x)]^2c_1^2(c_2 - c_1^2)^2\]
\[\times \int_{-1}^{1} \int_{-1}^{t} \{-t + tF(x) + sF(x)\}k(s)k(t)st \, dsdt\]
\[= O(n^7h^{24}) + 2n^7h^{23}f^7(x)[f'(x)]^2c_1^2(c_2 - c_1^2)^2[F(x) - 1]\]
\[\times \int_{-1}^{1} \int_{-1}^{t} k(s)k(t)(s + t)st \, dsdt\]
\[+ 2n^7h^{23}f^7(x)[f'(x)]^2c_1^2(c_2 - c_1^2)^2 \int_{-1}^{1} \int_{-1}^{t} k(s)k(t)s^2t \, dsdt\]
\[= O(n^7h^{24}) + n^7h^{23}f^7(x)[f'(x)]^2c_1^2(c_2 - c_1^2)^2[F(x) - 1]\]
\[\times \int_{-1}^{1} \int_{-1}^{t} k(s)k(t)(s + t)st \, dsdt\]
\[+ n^7h^{23}f^7(x)[f'(x)]^2c_1^2(c_2 - c_1^2)^2c_3\]
\[= O(n^7h^{24}) + n^7h^{23}f^7(x)[f'(x)]^2c_1^2(c_2 - c_1^2)^2c_3,\]
we have
\[ \sqrt{nh}n^{-4}h^{-12}III_3 \overset{d}{\to} N\left(0, f^2(x)|f'(x)|^2c_1^2(c_2 - c_1^2)^2c_3\right). \] (6.9)

Using (6.1)–(6.3), we notice that terms \( III_1 \) and \( III_2 \) in (6.8) satisfy
\[ \sqrt{nh}n^{-4}h^{-12}III_j = o_p(1), \ j = 1, 2. \] (6.10)

Thus, (6.4) follows from (6.7), (6.9) and (6.10). This completes the proof of Theorem 1.

**Proof of Theorem 2.** Define
\[ H(u) = H_1(u) + H_2(u), \]
with \( H_1(u) = P(Z_j \leq u, \delta_j = 1) \) and \( H_2(u) = P(Z_j \leq u, \delta_j = 0) \), and let
\[ H_n(u) = H_{n,1}(u) + H_{n,2}(u) \]
with \( H_{n,1}(u) = \frac{1}{n}\sum_{j=1}^{n}I(Z_j \leq u, \delta_j = 1) \) and \( H_{n,2}(u) = \frac{1}{n}\sum_{j=1}^{n}I(Z_j \leq u, \delta_j = 0) \). Let \( T \) be such that \( 1 - H(T) > d \) with some \( d > 0 \) and \( M, \lambda \) denote generic positive constants. Then it follows from Major and Rejto (1988) that the process \( \{F_n^*(u) - F(u), -\infty < u < \infty, 1 - H(u) > 0\} \) can be represented as
\[ F_n^*(u) - F(u) = (1 - F(u))[B_1(n, u) + B_2(n, u)] + R(n, u), \]
where
\[ B_1(n, u) = \frac{H_{n,1}(u) - H_1(u)}{1 - H(u)} - \int_{-\infty}^{u} \frac{H_{n,1}(y) - H_1(y)}{[1 - H(y)]^2} dH(y), \]
\[ B_2(n, u) = \int_{-\infty}^{u} \frac{H_n(y) - H(u)}{[1 - H(y)]^2} dH_2(y) \]
and for any \( \delta_0 > 0 \)
\[ P(A_1) \leq Me^{-\lambda h^{-\delta_0}} \]
where \( A_1 = \{\sup_{u \leq T} n|R(n, u)| > h^{-\delta_0}\} \). Moreover, there exists a Gaussian Process \( W(u), -\infty < u < \infty, \) with \( E(W(u)) = 0 \) and covariance
\[ E(W(s)W(t)) = \gamma(s) = \int_{-\infty}^{\delta} [1 - G(u)]^{-1}[1 - F(u)]^{-2} dF(u) \] (6.11)
for $-\infty < s \leq t < \infty$, such that
\[ P(A_2) \leq Me^{-\lambda x} \quad \text{and} \quad P(A_3) \leq Me^{-\lambda x}, \] (6.12)
where $A_2 = \{ \sup_{-\infty < u \leq T} \sqrt{n} | \sqrt{n} | B_1(n, u) + B_2(n, u) | W(u) | > h^{x/\lambda} \}$ and $A_3 = \{ \sup_{-\infty < u < \infty} \sqrt{n} | H_1(n) - H(u) | > h^{x/\lambda} \}$. It is easy to check that
\[
\begin{align*}
n^{-1}h^{-1}s_{n,0}(x) &= H'_1(x) + O_P(h^2) \\
n^{-1}h^{-3}s_{n,1}(x) &= -H''_1(x)c_1 + O_P(h) \\
n^{-1}h^{-3}s_{n,2}(x) &= H'_1(x)c_1 + O_P(h^2) \\
n^{-1}h^{-5}s_{n,3}(x) &= -H''_1(x)c_2 + O_P(h) \\
n^{-1}h^{-5}s_{n,4}(x) &= H'_1(x)c_2 + O_P(h^2).
\end{align*}
\] (6.13)
By (6.13), we obtain that
\[
\Delta_n(x) = o_p(n^3h^{11})
\]
\[
\begin{align*}
\Delta_{n,2}(x) &= n^3h^9[H'_1(x)]^2H''_1(x)c_1(c_2 - c_1^2) + O_p(n^3h^{10}) \\
\Delta_{n,3}(x) &= n^3h^9[H'_1(x)][H''_1(x)]^2c_1(c_2 - c_1^2) + O_p(n^3h^{10}) \\
\Delta_n(x) &= n^4h^{12}[H'_1(x)]^3H''_1(x)c_1^2(c_2 - c_1^2) + O_p(n^4h^{13}).
\end{align*}
\] (6.14)
Since
\[
\{ \lambda_n(x) - \lambda(x) \} \Delta_n(x)
\]
\[
= \sum_{j=1}^{n} \Lambda_n(Z_j) - \Lambda(Z_j) \} k(\frac{x - Z_j}{h}) \{ \Delta_{n,1}(x) + (Z_j - x)\Delta_{n,2}(x) + (Z_j - x)^2\Delta_{n,3}(x) \} \\
+ \sum_{j=1}^{n} \Lambda(Z_j) - \Lambda_n(x) - \lambda(x)\lambda(Z_j - x) - \frac{1}{2} \lambda'(x)(Z_j - x)^2 \} k(\frac{x - Z_j}{h}) \\
\times \{ \Delta_{n,1}(x) + (Z_j - x)\Delta_{n,2}(x) + (Z_j - x)^2\Delta_{n,3}(x) \} \\
= \ I + \ II,
\] (6.15)
Theorem 2 holds if we can show that
\[
\sqrt{n}\lambda^{-1}h^{-12} I \Rightarrow N \left( 0, \frac{f(x)[H'_1(x)]^6[H''_1(x)]^2c_1^2(c_2 - c_1^2)c_3}{[1 - F(x)]^2[1 - G(x)]} \right)
\] (6.16)
and
\[
n^{-1}h^{-14} I \overset{p}{\rightarrow} \frac{1}{6} \lambda''(x) [H'_1(x)]^3 H''_1(x) c_2 c_1 (c_2 - c_1^2) \tag{6.17}
\]
as \( n \to \infty \), where \( H'_1(x) = [1 - G(x)] f(x) \). It is easy to check that (6.17) holds. To prove (6.16), we decompose the term \( I \) in (6.15) as follows:

\[
I = \sum_{j=1}^{n} \log \left( \frac{1 - F_n(Z_j)}{1 - F(Z_j)} \right) k\left( \frac{x - Z_j}{h} \right) \times \left\{ \Delta_{n,1}(x) + (Z_j - x) \Delta_{n,2}(x) + (Z_j - x)^2 \Delta_{n,3}(x) \right\}
\]

\[
= \sum_{j=1}^{n} \frac{F_n(Z_j) - F(Z_j)}{1 - F(Z_j)} (1 + O_p(n^{-1/2} \log n)) k\left( \frac{x - Z_j}{h} \right)
\]

\[
\times \left\{ \Delta_{n,1}(x) + (Z_j - x) \Delta_{n,2}(x) + (Z_j - x)^2 \Delta_{n,3}(x) \right\}
\]

\[
= \left[ 1 + O_p(h) + O_p(n^{-1/2} \log n) \right] \sum_{j=1}^{n} \frac{F_n(Z_j) - F(Z_j)}{1 - F(Z_j)} k\left( \frac{x - Z_j}{h} \right) \delta_j
\]

\[
\times \left\{ \Delta_{n,1}(x) + (Z_j - x) \Delta_{n,2}(x) + (Z_j - x)^2 \Delta_{n,3}(x) \right\}
\]

\[
= I_1 + I_2 + I_3 + I_4. \tag{6.18}
\]
Using (6.12)-(6.14) the terms $I_1-I_3$ can be estimated as follows:

$$\sqrt{n}h_n^{-1} h^{-12} I_j = o_p(1), \ j = 1, 2, 3. \tag{6.19}$$

Next, we estimate the term $I_4$ in (6.18). Note that

$$dH_{n,1}(u) = [1 - H(u)]d[B_1(n, u) + B_2(n, u)] + H'_1(u) du - \frac{H_n(u) - h(u)}{1 - H(u)} H'_2(u) du.$$ 

We obtain that

$$I_4 \equiv \frac{1 - F(x)}{1 + o_p(h) + o_p(n^{-1/2} h - o)}$$

$$= \sum_{i=1}^{n} \delta k \left( \frac{x - Z_i}{h} \right) R(n, Z_i) \left\{ \left( Z_i - x \right) n^3 h^9 [H'_1(x)]^2 H''_1(x) c_1 (c_2 - c_1^2) \right. \right.$$ 

$$+ \left. \left( Z_i - x \right)^2 n^3 h^9 H'_1(x) [H''_1(x)]^2 c_1 (c_2^2 - c_2) \right\}$$

$$+ n \int [1 - F(s)] [B_1(n, s) + B_2(n, s)] k \left( \frac{s - x}{h} \right) \left\{ (s - x) n^3 h^9 [H'_1(x)]^2 H''_1(x) c_1 (c_2 - c_2^2) \right. \right.$$ 

$$\left. + (s - x)^2 n^3 h^9 H'_1(x) [H''_1(x)]^2 c_1 (c_2^2 - c_2) \right\} dH_{n,1}(s)$$

$$= \left[ \sum_{i=1}^{n} \delta k \left( \frac{x - Z_i}{h} \right) R(n, Z_i) \left\{ \left( Z_i - x \right) n^3 h^9 [H'_1(x)]^2 H''_1(x) c_1 (c_2 - c_1^2) \right. \right.$$ 

$$\left. + \left( Z_i - x \right)^2 n^3 h^9 H'_1(x) [H''_1(x)]^2 c_1 (c_2^2 - c_2) \right\} \right]$$

$$+ \left[ n \int [1 - F(s)] \left\{ \sqrt{n} [B_1(n, s) + B_2(n, s)] - W(s) \right\} k \left( \frac{s - x}{h} \right) \right.$$ 

$$\left. \times \{(s - x) n^3 h^9 [H'_1(x)]^2 H''_1(x) c_1 (c_2 - c_1^2) \right. \right.$$ 

$$\left. + (s - x)^2 n^3 h^9 H'_1(x) [H''_1(x)]^2 c_1 (c_2^2 - c_2) \right\} H'_1(s) ds \right]$$

$$+ \left[ n \int [1 - F(s)] W(s) k \left( \frac{s - x}{h} \right) \left\{ (s - x) n^3 h^9 [H'_1(x)]^2 H''_1(x) c_1 (c_2 - c_2^2) \right. \right.$$ 

$$\left. + (s - x)^2 n^3 h^9 H'_1(x) [H''_1(x)]^2 c_1 (c_2^2 - c_2) \right\} H'_1(s) ds \right]$$

$$- \left[ n \int [1 - F(s)] [B_1(n, s) + B_2(n, s)] k \left( \frac{s - x}{h} \right) \left\{ (s - x) n^3 h^9 [H'_1(x)]^2 H''_1(x) \right. \right.$$ 

$$\times c_1 (c_2 - c_1^2) + (s - x)^2 n^3 h^9 H'_1(x) [H''_1(x)]^2 c_1 (c_2^2 - c_2) \right\} \frac{H_n(u) - h(u)}{1 - H(u)} H'_2(u) du \right]$$

$$= III_1 + III_2 + III_3 + III_4 + III_5.$$
Using (6.12)-(6.14), the terms $III_1, III_2, III_3$ and $III_5$ satisfy

$$
\sqrt{\eta h n^{\frac{1}{2}}} III_j = o_p(1), \quad j = 1, 2, 3, 5.
$$

(6.21)

Since

$$
EIII_4^n
= 2n^2 \int_1^t \int_1^t [1 - F(x - st)][1 - F(x - th)] \gamma(x - th)k(s)k(t)
$$

$$
\times \{-sn^3h^{10}[H''(x)]^2H''(x)c_1(c_2 - c_1^2) + s^2n^3h^{11}H''(x)[H''(x)]^2c_1(c_1^2 - c_2)\}H'(x - sh)
$$

$$
\times \{-tn^3h^{10}[H''(x)]^2H''(x)c_1(c_2 - c_1^2) + t^2n^3h^{11}H''(x)[H''(x)]^2c_1(c_1^2 - c_2)\}H'(x - th) \, ds \, dt
$$

$$
= nh^2 \int_1^t \int_1^t [1 - F(x - sh)][1 - F(x - th)] \gamma(x - th)k(s)k(t)
$$

$$
\times \{-sn^3h^{10}[H''(x)]^2H''(x)c_1(c_2 - c_1^2) + s^2n^3h^{11}H''(x)[H''(x)]^2c_1(c_1^2 - c_2)\}H'(x - sh)
$$

$$
\times \{-tn^3h^{10}[H''(x)]^2H''(x)c_1(c_2 - c_1^2) + t^2n^3h^{11}H''(x)[H''(x)]^2c_1(c_1^2 - c_2)\}H'(x - th) \, ds \, dt
$$

$$
+ 2n^2 \int_1^t \int_1^t [1 - F(x - th)][1 - F(x - sh)][-th\gamma'(x) + O(h^2)]k(s)k(t)
$$

$$
\times \{-sn^3h^{10}[H''(x)]^2H''(x)c_1(c_2 - c_1^2) + s^2n^3h^{11}H''(x)[H''(x)]^2c_1(c_1^2 - c_2)\}H'(x - sh)
$$

$$
\times \{-tn^3h^{10}[H''(x)]^2H''(x)c_1(c_2 - c_1^2) + t^2n^3h^{11}H''(x)[H''(x)]^2c_1(c_1^2 - c_2)\}H'(x - th) \, ds \, dt
$$

$$
= O(n^7h^{21}) + 2n^7h^{23} \int_1^t \int_1^t [1 - F(x)]^2[-t\gamma'(x)]k(s)k(t)st[H''(x)]^6[H''(x)]^2c_1^2(c_2 - c_1^2)^2 \, ds \, dt
$$

$$
= O(n^7h^{21}) - n^7h^{23} \int_1^t \int_1^t [1 - F(x)]^2\gamma'(x)k(s)k(t)st[H''(x)]^6[H''(x)]^2c_1^2(c_2 - c_1^2)^2 \, ds \, dt
$$

$$
+ 2n^7h^{23} \int_1^t \int_1^t [1 - f(x)]^2s\gamma'(x)k^2(s)k(t)st[H''(x)]^6[H''(x)]^2c_1^2(c_2 - c_1^2)^2 \, ds \, dt
$$

$$
= O(n^7h^{21}) + n^7h^{23}[1 - G(x)]^{-1}f(x)[H''(x)]^6[H''(x)]^2c_1^2(c_2 - c_1^2)^2c_3,
$$

we have

$$
\sqrt{\eta h n^{-4}} h^{-12} III_3 \xrightarrow{d} N\left(0, [1 - G(x)]^{-1}f(x)[H''(x)]^6[H''(x)]^2c_1^2(c_2 - c_1^2)^2c_3 \right).
$$

(6.22)

Hence, (6.15) follows from (6.19), (6.21) and (6.22). This completes the proof of Theorem 2.
Proof of Theorem 3. It follows from the proof of Theorem 1 that
\[
\sqrt{n}h\{\hat{\lambda}_n(x) - \lambda(x)\} = \frac{1}{4}\sqrt{n}h\{\hat{\lambda}_n(x - (\sqrt{1/2} + 1)\delta h) - \lambda(x - (\sqrt{1/2} + 1)\delta h)\}
\]
\[+ \frac{1}{2}\sqrt{n}h\{\hat{\lambda}_n(x - \sqrt{1/2}\delta h) - \lambda(x - \sqrt{1/2}\delta h)\}
\]
\[+ \frac{1}{4}\sqrt{n}h\{\hat{\lambda}_n(x - (\sqrt{1/2} - 1)\delta h) - \lambda(x - (\sqrt{1/2} - 1)\delta h)\}
\]
\[+ \frac{1}{4}\sqrt{n}h\{\lambda(x - (\sqrt{1/2} + 1)\delta h) - \lambda(x)\}
\]
\[+ \frac{1}{2}\sqrt{n}h\{\lambda(x - \sqrt{1/2}\delta h) - \lambda(x)\}
\]
\[+ \frac{1}{4}\sqrt{n}h\{\lambda(x - (\sqrt{1/2} - 1)\delta h) - \lambda(x)\}
\]
\[= \left\{\frac{1}{4}\sqrt{n}h + \frac{1}{2} + \frac{1}{4}\right\} B_n(F(x - (\sqrt{1/2} + 1)\delta h + sh))k(s) ds
\]
\[+ \frac{1}{4}\sqrt{n}h \int_{-1}^{1} B_n(F(x - \sqrt{1/2}\delta h + sh))k(s) ds
\]
\[+ \frac{1}{4}\sqrt{n}h \int_{-1}^{1} B_n(F(x - (\sqrt{1/2} - 1)\delta h + sh))k(s) ds + o_p(1).
\]

Note that
\[
\frac{B_n(F(x + sh)) - B_n(F(x))}{\sqrt{h}} \overset{D}{\to} \sqrt{f(x)}W(s)
\]
in $D([0, T])$, where $T > 0$ and $W(s)$ is a Wiener process, and
\[
E\left\{\left(\int_{-1}^{1} W(s - a)k(s) ds\right)\left(\int_{-1}^{1} W(s - b)k(s) ds\right)\right\} = c_5(a, b),
\]
where $c_5(a, b)$ is defined as in (4.1). We prove Theorem 3 by the above equations.

Proof of Theorem 4. Similar to the proof of Theorem 3.

References


Figure 1: Mean squared errors. The mean squared errors of $\tilde{\lambda}_n(x)$ (dotted line) and $\tilde{\lambda}_n(x)$ (solid line) are plotted against $h = \bar{h}_{opt} + \frac{j}{31} h_{opt}$, $j = 0, 1, \cdots, 19$, where $\bar{h}_{opt}$ is defined in (3.1). We take $x$ such that $F(x) = 0.5$. 