On a nonparametric estimator for ruin probability in the classical risk model

Zhimin Zhang[†][*], Hailiang Yang[‡] and Hu Yang[†]

[†]Department of Statistics and Actuarial Science,
Chongqing University, Chongqing, P.R. China

[‡]Department of Statistics and Actuarial Science,
The University of Hong Kong, Pokfulam Road, Hong Kong

Abstract

In this paper, we present a nonparametric estimator for ruin probability in the classical risk model with unknown claim size distribution. We construct the estimator by Fourier inversion and kernel density estimation method. Under some conditions imposed on the kernel, bandwidth and claim size density, we present some large sample properties of the estimator. Some simulation studies are also given to show the finite sample performance of the estimator.
Keywords: Ruin probability; Fourier transform; Kernel; Bias; Variance; Asymptotic normality.

1   Introduction

Consider the following classical compound Poisson risk model

$$U_t = u + ct - \sum_{j=1}^{N_t} X_j,$$

where $u \geq 0$ is the initial surplus, $c > 0$ is the constant premium rate. The claim number process $\{N_t, t \geq 0\}$ is a Poisson process with intensity $\lambda > 0$. The i.i.d. random variables $\{X_j, j = 1, 2, \cdots\}$, independent of $\{N_t\}$, have the same distribution as that of X with density f and distribution function $F = 1 - \overline{F}$. The ruin time defined by

$$\tau = \inf\{t > 0 : U_t < 0\}$$

is the first time when the surplus becomes negative. Given the initial surplus u the ruin probability is defined by

$$\psi(u) = \mathbb{P}(\tau < \infty | U_0 = u).$$

The calculation, approximation as well as estimation of the ruin probability have been hot topics in risk theory for a long time (see for instance Rolski et al. (1999)

---

[*]Corresponding author. E-mail address: cquzzm@163.com, zmzhang@cqu.edu.cn

and Asmussen (2000)). As is known to all, closed-form calculation formulas for ruin probability are usually not easy to obtain, and approximation can only provide some rough information on ruin. Furthermore, because these theoretical results are usually dependent on the parametric assumptions imposed on the risk models, they are somewhat restrictive in practical situations. However, statistical estimation and inference do not depend heavily on the specific models, and usually some satisfactory results can be obtained when some sample information on the risk model is available.

In the last few decades, a lot of contributions have been made to the nonparametric estimation of the ruin probability, see, e.g. Grandell (1979), Frees (1986), Hipp (1989), Croux and Veraverbeke (1990), Pitts (1994), and Politis (2003) to name a few. The claim size density is usually assumed to be unknown and it is even not imposed on any parametric assumptions, while the Poisson intensity is sometimes assumed to be unkown. Most of the nonparametric estimators of the ruin probability are proposed on the ground of empirical type estimation of the claim size distribution. Mnatsakanov et al. (2008) considered an empirical-type estimator of the Laplace transform of the ruin probability, and recovered the ruin probability by a regularized Laplace inversion technique. This method has also been used by Shimizu (2010) to study the Gerber-Shiu discounted penalty function. Instead of using Laplace transform method, we shall present an estimator by Fourier (inversion) transform.

As one of the most important and efficient method, kernel method is very popular in nonparametric statistics literature, see e.g. Prakasa Rao (1983) and Wand and Jones (1995). Croux and Veraverbeke (1990) propose an estimator for the ruin probability from the Pollaczeck-Khinchine formula by using U-statistics. However, their estimator is not unbiased if the the mean of the claim size is not known. Kernel method has also been considered by Croux and Veraverbeke (1990) to estimate the claim size density. In this paper, we assume that the Poisson intensity λ is known but the density function of the individual claim size is unknown. Instead, we assume that a random sample of the individual claim sizes is available and we estimate the claim size density by kernel method. Different from Croux and Veraverbeke (1990), methodology employed in this paper is Fourier (inversion) transform. Note that one advantage of Fourier transform is that it can transform a geometric series function to a single function. We first start from a Fourier inversion transform representation for ruin probability that can be obtained by Pollaczeck-Khinchine formula, then propose an estimator for ruin probability by some plug-in techniques. Fourier transform method has also been applied by Embrechts et al. (1993) to risk theory, but their attention was paid to the empirical-type estimator. Kernel smoothing has an appealing advantage that the order of the variance of the estimator can be effectively reduced by choosing an appropriate smoothing parameter. This is the reason why kernel method is so popular in nonparametric statistics literature.

The reminder of this paper is organized as follows. In Section 2, we illustrate the construction procedure of the estimator. In Section 3, we give some assumptions on the kernel, bandwidth and claim size density, and present some large sample properties of the estimator. The finite sample performance of the estimator is examined by some simulation studies in Section 4. Finally, some concluding remarks are given in Section 5. All the technical arguments are given in Appendix A and B.

2   The estimator

Throughout this paper, all integrals without an indicated domain of integration are taken over the whole real line. We will apply Fourier inversion to propose our estimator. In the sequel of this paper, for an absolutely integrable function v we denote its Fourier transform by

$$\phi_v(s) = \int e^{isx} v(x) dx.$$

Under some mild integrable conditions, Fourier inversion formula gives

$$v(x) = \frac{1}{2\pi} \int e^{-isx} \phi_v(s) ds,$$

where x is a continuous point of v.

We start from the Pollaczeck-Khinchine type formula for ruin probability. Let μ denote the mean of individual claim amount, and let $f_e(x) = \overline{F}(x)/\mu$ for $x \geq 0$ and be identical to zero for $x < 0$. If the net profit condition $c > \lambda\mu$ holds, then Pollaczeck-Khinchine formula states that

$$
\begin{aligned}
\psi(u) &= \left(1 - \frac{\lambda\mu}{c}\right) \sum_{j=1}^{\infty} \left(\frac{\lambda\mu}{c}\right)^j \left(1 - \int_0^u f_e^{*j}(x) dx\right) \\
&= \frac{\lambda\mu}{c} - \left(1 - \frac{\lambda\mu}{c}\right) \sum_{j=1}^{\infty} \left(\frac{\lambda\mu}{c}\right)^j \int_0^u f_e^{*j}(x) dx, \quad u \geq 0,
\end{aligned}
\tag{2.1}
$$

where f_e^{*j} is the j-fold convolution of the density f_e with itself.

Employing integration by part it is easy to verify that

$$\phi_{f_e}(s) = \frac{\phi_f(s) - 1}{i\mu s},$$

and then noting the fact that the Fourier transform of convolution is equal to the product of the transforms, we obtain

$$f_e^{*j}(x) = \frac{1}{2\pi} \int e^{-isx} \left(\frac{\phi_f(s) - 1}{i\mu s}\right)^j ds, \quad j = 1, 2, \dots.$$

We remark that the above inversion formula does not hold at the discontinuous point $x = 0$. However, because we are only interested in the integral of f_e^{*j} (see (2.1)), this discontinuous point will not impact the following plug-in argument.

Plugging the above expression for f_e^{*j} into (2.1) and using Fubini's theorem, we have

$$
\begin{aligned}
\psi(u) &= \frac{\lambda\mu}{c} - \left(1 - \frac{\lambda\mu}{c}\right) \frac{1}{2\pi} \sum_{j=1}^{\infty} \left(\frac{\lambda\mu}{c}\right)^j \int_0^u \int e^{-isx} \left(\frac{\phi_f(s) - 1}{i\mu s}\right)^j ds dx \\
&= \frac{\lambda\mu}{c} - \left(1 - \frac{\lambda\mu}{c}\right) \frac{1}{2\pi} \int \int_0^u e^{-isx} \frac{\lambda\phi_f(s) - \lambda}{ics - \lambda\phi_f(s) + \lambda} dx ds \\
&= \frac{\lambda\mu}{c} - \left(1 - \frac{\lambda\mu}{c}\right) \varphi(u),
\end{aligned}
\tag{2.2}
$$

3

where

$$\varphi(u) = \frac{1}{2\pi} \int \frac{1 - e^{-isu}}{is} \frac{\frac{\lambda}{c} \frac{\phi_f(s)-1}{is}}{1 - \frac{\lambda}{c} \frac{\phi_f(s)-1}{is}} ds.$$

There exist two unknown quantities in (2.2), i.e. the mean μ and the function $\varphi(u)$. Hence, an estimator for $\psi(u)$ can be obtained if we can find estimators for μ and φ.

Suppose that an i.i.d. sample $X_1, X_2, \ldots, X_n$ of the individual claim sizes can be obtained. We construct the estimators for μ and $\varphi(u)$ based on this observed sample. Firstly, we estimate μ by the sample mean $\hat{\mu}_n$ that is defined by

$$\hat{\mu}_n = \frac{1}{n} \sum_{j=1}^{n} X_j. \tag{2.3}$$

Next, we employ kernel method to get an estimator for $\varphi(u)$. The kernel density estimator for the claim size density f, given the sample $X_1, X_2, \ldots, X_n$, is defined as

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right),$$

where K is a kernel function, and $h_n > 0$ is a smoothing parameter, usually called bandwidth, such that $h_n \to 0$ as $n \to \infty$. With the understanding that the bandwidth sequence $\{h_n\}$ depends on the sample size, we will omit the subscript n for notational convenience. Let $K_h(\cdot) = \frac{1}{h} K(\frac{\cdot}{h})$. We can calculate the Fourier transform of $\hat{f}_n$ as follows,

$$\phi_{\hat{f}_n}(s) = \int e^{isx} \frac{1}{n} \sum_{j=1}^{n} K_h(x - X_j) dx = \phi_{emp}(s)\phi_K(sh),$$

where $\phi_{emp}(s) = \frac{1}{n} \sum_{j=1}^{n} e^{isX_j}$ is the empirical characteristic function. Although $\phi_{emp}(s)$ is an unbiased estimate, it is not a good estimate at high frequencies. For this reason, it is usual to incorporate the damping factor $\phi_K(sh)$.

By a plug-in device we obtain the following estimator for $\varphi(u)$,

$$\tilde{\varphi}_n(u) = \frac{1}{2\pi} \int \frac{1 - e^{-isu}}{is} \frac{\frac{\lambda}{c} \frac{\phi_{emp}(s)\phi_K(sh)-1}{is}}{1 - \frac{\lambda}{c} \frac{\phi_{emp}(s)\phi_K(sh)-1}{is}} ds. \tag{2.4}$$

Note that there is no guarantee that $\left|1 - \frac{\lambda}{c} \frac{\phi_{emp}(s)\phi_K(sh)-1}{is}\right|$ is strictly positive, although the probability of this event tends to one. Thus, the integral in (2.4) is possibly infinite. To get a finite-valued estimate, we consider the following modification,

$$\hat{\varphi}_n(u) = (M \wedge \tilde{\varphi}_n(u)) \vee (-M), \tag{2.5}$$

where M is a large constant. There is no need to assume that M is a function of n such that it tends to infinity as $n \to \infty$. To see this, we obtain from (2.2) that

$$\varphi(u) \leq \frac{c + \lambda\mu}{c - \lambda\mu}.$$

Note that the right hand side of the above inequality is unknown due to the unknown mean μ. However, in practical applications we can set M to be a large enough constant.

By (2.2), (2.3), (2.5) and a plug-in device we can get the following estimator for the ruin probability,

$$\hat{\psi}_n(u) = \frac{\lambda\hat{\mu}_n}{c} - \left(1 - \frac{\lambda\hat{\mu}_n}{c}\right)\hat{\varphi}_n(u). \tag{2.6}$$

Remark 1 *When the Poison claim arrival intensity is also unknown, we can construct an estimator for λ if a sample on the process $\{N_t\}$ is available. Assume that we can observe the total number of claims that arrive prior to some lattice time points, i.e. for $d > 0$ a sample $\{N_d, N_{2d}, N_{3d}, \ldots\}$ can be observed. Then an unbiased estimator for λ is*

$$\hat{\lambda}_n = \frac{N_{nd}}{nd}.$$

Hence, we can estimate $\varphi(u)$ by

$$\tilde{\varphi}_n^*(u) = \frac{1}{2\pi}\int \frac{1 - e^{-isu}}{is} \frac{\frac{\hat{\lambda}_n}{c}\frac{\phi_{emp}(s)\phi_K(sh)-1}{is}}{1 - \frac{\hat{\lambda}_n}{c}\frac{\phi_{emp}(s)\phi_K(sh)-1}{is}} ds. \tag{2.7}$$

Similar to (2.5), it is more preferable to consider the following modification

$$\hat{\varphi}_n^*(u) = (M \wedge \tilde{\varphi}_n^*(u)) \vee (-M), \tag{2.8}$$

By a plug-in device, we obtain the following estimator for ruin probability,

$$\hat{\psi}_n^*(u) = \frac{\hat{\lambda}_n\hat{\mu}_n}{c} - \left(1 - \frac{\hat{\lambda}_n\hat{\mu}_n}{c}\right)\hat{\varphi}_n^*(u). \tag{2.9}$$

In the reminder of this paper, we will only focus our attention to the estimator $\hat{\psi}_n(u)$. We would like to assert that the estimator $\hat{\psi}_n^*(u)$ also has some consistency properties, but to see this we need more involved technical arguments.

3 Assumptions and asymptotic properties

To study the estimator, we present some assumptions imposed on the kernel K, bandwidth h and density f. For a function v, define $\mu_j(v) = \int |x|^j v(x)dx$ and $R_j(v) = \int v(x)^j dx$ for $j = 0, 1, 2, \cdots$. In the sequel, all the limits are taken as $n \to \infty$. Let $C > 0$ denote a generic finite-valued constant that can take different values at different steps.

Assumption K K is a probability kernel, symmetric about zero, $\mu_3(K) < \infty$ and $R_2(K) < \infty$.

Assumption H $h = Cn^{-\gamma}$ for some $C > 0$ and $0 < \gamma < 1$.

Assumption F

(F1) $f(x) \equiv 0$ for $x < 0$, f is continuously differentiable in $(0, \infty)$, right-continuous at zero, and f'' exists almost everywhere;

(F2) $\int |f'(x)| dx < \infty$, $\int |f''(x)| dx < \infty$, $R_2(f') < \infty$;

(F3) $\mathbb{E} X^4 < \infty$, and for some $0 < \alpha < \frac{1}{2}$, $x^{\alpha+1} \overline{F}(x^\alpha) \to 0$ as $x \to \infty$;

(F4) there exists a constant $\rho > 0$ such that $1 - \frac{\lambda \mu}{c} \geq \rho$.

Remark 2 *Assumption **K** is very standard and widely used in the nonparametric kernel estimate literature. Assumption **H** implies that the standard assumptions on the bandwidth, $h \to 0$ and $nh \to \infty$, hold. Assumptions (F1), (F2) are satisfied by many continuous densities that are widely used in risk theory. Assumption (F3) is satisfied by all the light-tailed distributions, and some heavy-tailed distributions such as Lognormal, Weibull as well as Pareto with large shape parameter. Finally, assumption (F4) is the net profit condition.*

Proposition 1 *Suppose that assumptions **K**, **H**, **F** hold, then the bias of $\hat{\varphi}_n(u)$, defined by $Bias(\hat{\varphi}_n(u) = \mathbb{E}\hat{\varphi}_n(u) - \varphi(u)$, has the following order,*

$$Bias(\hat{\varphi}_n(u)) = O(n^{-1} + h^2). \tag{3.1}$$

Proposition 2 *Suppose that assumptions **K**, **H**, **F** hold, then the variance of $\hat{\varphi}_n(u)$ has the following order,*

$$Var(\hat{\varphi}_n(u)) = O(n^{-1}) + O((n^2 h)^{-1} + h^4). \tag{3.2}$$

It follows from Proposition 1 and 2 that the mean squared error (MSE) of $\hat{\varphi}_n(u)$ has the following order,

$$
\begin{aligned}
\text{MSE}(\hat{\varphi}_n(u)) &= [\text{Bias}(\hat{\varphi}_n(u)]^2 + \text{Var}(\hat{\varphi}_n(u)) \\
&= O(\frac{1}{n} + \frac{h^2}{n} + \frac{1}{n^2 h} + h^4).
\end{aligned} \tag{3.3}
$$

From (3.3) it is not easy to see the order of optimal bandwidth that minimizes the asymptotic MSE. From (3.1) it seems that the smaller the bandwidth, the smaller the bias. While from (3.2) it seems that smaller bandwidth may cause a larger variance. This phenomena also appears in the kernel density estimate literature. Instead of finding a balance between bias and variance, we can minimize the term $\frac{1}{n^2 h} + h^4$ in (3.2) to obtain a plausible order of the optimal bandwidth that minimizes the variance, which is $O(n^{-\frac{2}{5}})$. If we choose $h = Cn^{-\frac{2}{5}}$ for some $C > 0$, then from (3.3) we know that the order of $\text{MSE}(\hat{\varphi}_n(u))$ is $O(n^{-1})$. Thus, $\hat{\varphi}_n(u)$ is a $\sqrt{n}$-consistent estimator.

Proposition 3 *Suppose that assumptions **K**, **H**, **F** hold and $nh^4 \to 0$, then*

$$\frac{\hat{\varphi}_n(u) - \varphi(u)}{\sqrt{Var(\hat{\varphi}_n(u))}} \xrightarrow{\mathfrak{D}} \mathcal{N}(0, 1),$$

where $\mathcal{N}(0, 1)$ is the standard normal distribution.

Remark 3 *Now suppose that the assumptions in Proposition 3 hold. From the proof of Proposition 3 we know that $Var(\hat{\varphi}_n(u)) = O(n^{-1})$. Then by Proposition 1 we have $Bias(\hat{\varphi}_n(u))/\sqrt{Var(\hat{\varphi}_n(u))} \to 0$. Consequently, by Proposition 3 we have*

$$\frac{\hat{\varphi}_n(u) - \mathbb{E}\hat{\varphi}_n(u)}{\sqrt{Var(\hat{\varphi}_n(u))}} = \frac{\hat{\varphi}_n(u) - \varphi(u)}{\sqrt{Var(\hat{\varphi}_n(u))}} - \frac{Bias(\hat{\varphi}_n(u))}{\sqrt{Var(\hat{\varphi}_n(u))}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Now we study the order of the bias and variance of $\hat{\psi}_n(u)$.

Theorem 1 *Suppose that assumptions $\boldsymbol{K}$, $\boldsymbol{H}$, $\boldsymbol{F}$ hold, then*

$$Bias(\hat{\psi}_n(u)) = O(n^{-\frac{1}{2}} + h^2). \tag{3.4}$$

Proof. By (2.6) we have

$$\hat{\psi}_n(u) - \psi(u) = \frac{\lambda}{c}(\hat{\mu}_n - \mu)(1 + \hat{\varphi}_n(u)) - \left(1 - \frac{\lambda\mu}{c}\right)(\hat{\varphi}_n(u) - \varphi(u)). \tag{3.5}$$

Note that $\hat{\varphi}_n(u)$ is bounded. By Cauchy-Schwarz inequality we have

$$|\mathbb{E}(\hat{\mu}_n - \mu)(1 + \hat{\varphi}_n(u))| \leq C[\mathbb{E}(\hat{\mu}_n - \mu)^2]^{\frac{1}{2}} = O(n^{-\frac{1}{2}}).$$

By Proposition 1 we have

$$\mathbb{E}\hat{\varphi}_n(u) - \varphi(u) = O(n^{-1} + h^2).$$

Then (3.4) follows from the above results. $\qquad\square$

Theorem 2 *Suppose that assumptions $\boldsymbol{K}$, $\boldsymbol{H}$, $\boldsymbol{F}$ hold, then*

$$Var(\hat{\psi}_n(u)) = O(n^{-1}) + O((n^2 h)^{-1} + h^4). \tag{3.6}$$

Proof. By (3.5) and Cauchy-Schwarz inequality we have

$$\text{Var}(\hat{\psi}_n(u)) \leq 2\frac{\lambda^2}{c^2}\text{Var}((\hat{\mu}_n - \mu)(1 + \hat{\varphi}_n(u))) + 2\left(1 - \frac{\lambda}{c}\right)^2 \text{Var}(\hat{\varphi}_n(u)). \tag{3.7}$$

Noting that $\hat{\varphi}_n(u)$ is bounded, we have

$$\text{Var}((\hat{\mu}_n - \mu)(1 + \hat{\varphi}_n(u))) \leq \mathbb{E}((\hat{\mu}_n - \mu)(1 + \hat{\varphi}_n(u)))^2 \leq C\mathbb{E}(\hat{\mu}_n - \mu)^2 = O(n^{-1}),$$

which together with (3.7) and Proposition 2 gives (3.6). $\qquad\square$

By Theorem 1 and 2, the mean squared error of $\hat{\psi}_n(u)$ is given by

$$\text{MSE}(\hat{\psi}_n(u)) = O(\frac{1}{n} + \frac{h^2}{\sqrt{n}} + \frac{1}{n^2 h} + h^4). \tag{3.8}$$

If $n^2 h^5 \to 0$, then we can find a balance between bias and variance by minimizing the term $\frac{h^2}{\sqrt{n}} + \frac{1}{n^2 h}$ in (3.8). In this case, the plausible optimal bandwidth is $h_{opt} = Cn^{-\frac{1}{2}}$ for some $C > 0$, and from (3.8) we know that the order of $\text{MSE}(\hat{\psi}_n(u))$ is also $O(n^{-1})$ under such bandwidth setting. Consequently, $\hat{\psi}_n(u)$ is a $\sqrt{n}$-consistent estimator.

Remark 4 *Consider the case when μ is known. Estimator (2.6) reduces to the following semi-parametric estimator*

$$\hat{\psi}_{n,sem}(u) = \frac{\lambda\mu}{c} - \left(1 - \frac{\lambda\mu}{c}\right)\hat{\varphi}_n(u).$$

It is easily seen that the bias and variance of $\hat{\psi}_{n,sem}(u)$ has the same order as that of $\hat{\varphi}_n(u)$. Also, from Proposition 3 we know that $\hat{\psi}_{n,sem}(u)$ has asymptotic normality property.

4   Simulation studies

In this section the finite sample performance of the estimator is examined by some examples. In these simulation studies, we assume that the claim size densities are respectively exponential, Gamma and Pareto. For exponential and Gamma claim size densities, explicit formula for ruin probability can be obtained by inverting the Laplace transform of the ruin probability (see formula 3 in Albrecher and Kortschak (2009)). In the case of Pareto claim size density, we use the integral expression for ruin probability presented by Ramsay (2003).

$K(\cdot)$ is chosen to be the Gaussian kernel, i.e.

$$K(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad \phi_K(s) = e^{-\frac{s^2}{2}}.$$

Note that only the function $\hat{\varphi}_n(u)$ depends on the bandwidth h. Thus, instead of using h_{opt}, we prefer to use the bandwidth that minimizes the order of $\text{Var}(\hat{\varphi}_n(u))$, i.e. $h = Cn^{-\frac{2}{5}}$. We set $C = 0.95$ in our simulation studies by hand.

Example 1 *Premium rate $c = 1.5$; Poisson intensity $\lambda = 1$; Exp(1) claim size density with*

$$f(x) = e^{-x}, \quad x \geq 0.$$

Example 2 *Premium rate $c = 8$; Poisson intensity $\lambda = 1.5$; Gamma(4,1) claim size density with*

$$f(x) = \frac{1}{6}x^3e^{-x}, \quad x \geq 0.$$

Example 3 *Premium rate $c = 1.1$; Poisson intensity $\lambda = 1$; Pareto claim size density with*

$$f(x) = \frac{6}{5}\left(1 + \frac{x}{5}\right)^{-7}$$

Figure 1 shows the behavior of the estimator with sample sizes $n = 50, 100, 200$ and the parameter setting in Example 1. As is expected, the estimator is improved as the sample size n increases. In particular, when the initial surplus u is small, the estimators perform better than when u is large. In order to overcome the "curse" of
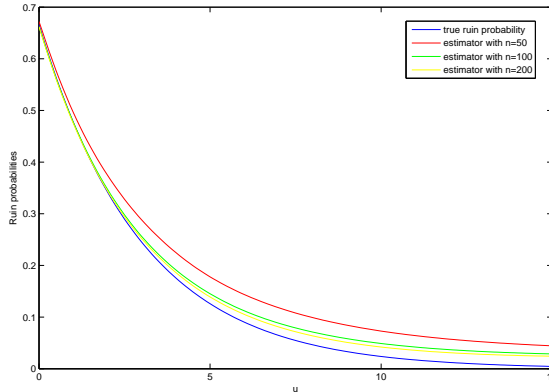
Figure 1: Comparison of the true ruin probability and estimators when f is Exp(1), sample sizes $n = 50; 100; 200$.

large initial surplus, we adopt a scale transform approach. Note that

$$
\begin{aligned}
\psi(u) &= \mathbb{P}\left(\inf_{t}\left(u + ct - \sum_{j=1}^{N(t)} X_j\right) < 0\right) \\
&= \mathbb{P}\left(\inf_{t}\left(\frac{u}{\xi} + \frac{c}{\xi}t - \sum_{j=1}^{N(t)} \frac{X_j}{\xi}\right) < 0\right),
\end{aligned}
$$

where $\xi > 0$ is a scale parameter. Thus, for any initial surplus, the ruin probability can be reduced to the case with small initial surplus if we choose the scale parameter large enough. In Figure 2, we show the comparison results of the true ruin probability, original estimator and the scaled estimator. In a sense, scale transform indeed improves the performance of the original estimator for a fixed sample size. Also, the performance of the scaled estimator is improved as the sample size increases. In figure 3-4, we show the performance of the estimators when the claim size density is Gamma(4,1), and obtain almost the same conclusions as that in the exponential claim size situation. We believe that the estimator has better performance if the sample size is sufficiently large.

Both exponential and Gamma distributions belong to the light-tailed class. In Example 3, we consider a heavy-tailed case, i.e. the claim sizes follow a Pareto distribution. Compared with light-tailed claim size distributions, simulation study shows that larger sample sizes are needed to guarantee good results for heavy-tailed distribution (see Figure 5). Also, we would like to assert that the scale transform fail to improve the estimate when the initial surplus is too large in the heavy-tailed case. Thus, we need to find other approach to solve this problem. For example, instead of estimating the ruin probability, we can estimate its asymptotic behavior for large initial surplus. For this purpose, we should first check the tail behavior of the claim size distribution, and this will lead to some new problems. We leave this work for further research.
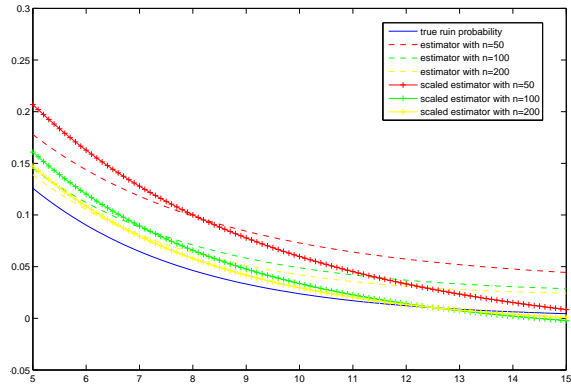
9

Figure 2: Comparison of the true ruin probability and (scaled) estimators when f is Exp(1), sample sizes $n = 50; 100; 200$ and scale parameter $\xi = 5$.
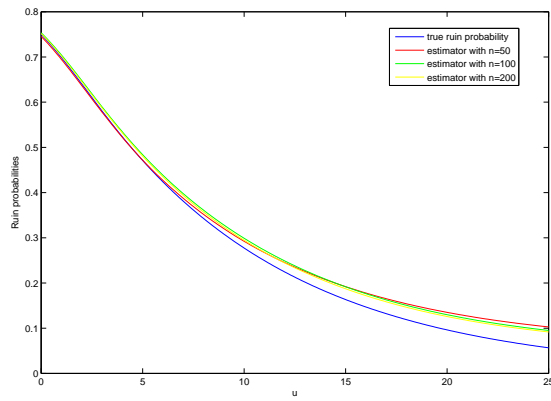

Figure 3: Comparison of the true ruin probability and estimators when f is Gamma(4,1), sample sizes $n = 50; 100; 200$.

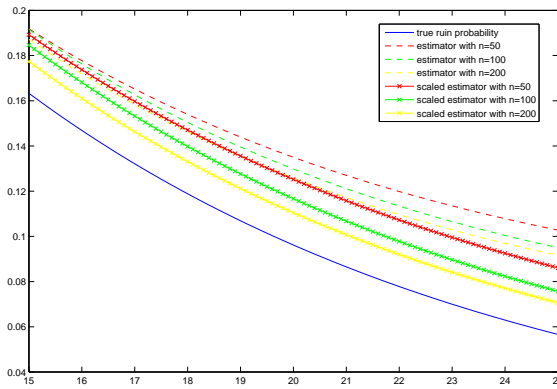Figure 4: Comparison of the true ruin probability and (scaled) estimators when f is Gamma(4,1), sample sizes $n = 50; 100; 200$ and scale parameter $\xi = 5$.
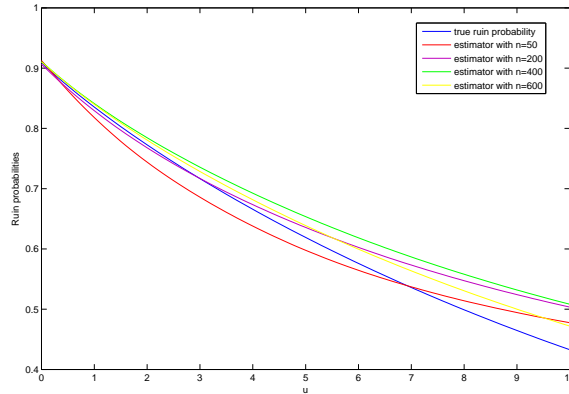
Figure 5: Comparison of the true ruin probability and estimators when f is Pareto, sample sizes $n = 50; 200; 400; 600$.

5 Concluding remarks

In this paper, we have presented a nonparametric estimator for ruin probability by Fourier inversion and kernel method. The consistent property of the estimator is analyzed by MSE, and the finite sample size performance is studied by some simulations.

The methodology presented in this paper can be used to estimate other ruin related quantities, such as the distributions and moments of the surplus before and after ruin, or the Gerber-Shiu function with general penalty function. However, when estimating the Gerber-Shiu function with a positive discount parameter, we should estimate more parameters or functions, and consequently the argument may be more involved.

A Proofs of Proposition 1-3

In this appendix, we derive the order of bias and variance of $\hat{\varphi}_n(u)$, and show that it has asymptotic normality. To begin with, we present some preliminaries.

Firstly, let

$$A_n(s) = \frac{\lambda}{c} \frac{\phi_f(s) - \phi_{emp}(s)\phi_K(sh)}{is}.$$

Using the inequalities $|e^{ix} - 1| \leq |x| \wedge 2$, $|\phi_f(s)| \leq 1$ and $|\phi_K(sh)| \leq 1$, we have

$$
\begin{aligned}
\frac{c}{\lambda}|A_n(s)| &\leq |\phi_f(s)| \left| \frac{1 - \phi_K(sh)}{is} \right| + |\phi_K(sh)| \left(\left| \frac{1 - \phi_f(s)}{is} \right| + \left| \frac{1 - \phi_{emp}(s)}{is} \right| \right) \\
&\leq \left(h\mu_1(K) \wedge \frac{2}{|s|} + \mu \wedge \frac{2}{|s|} + \hat{\mu}_n \wedge \frac{2}{|s|} \right) (|\phi_f(s)| \vee |\phi_K(sh)|) \\
&\leq (h\mu_1(K) + \mu + \hat{\mu}_n) \wedge \left(\frac{2}{|s|} (|\phi_f(s)| \vee |\phi_K(sh)|) \right).
\end{aligned}
$$

For any constant $L > 0$, by Cauchy-Schwarz inequality and Parseval's identity, we have

$$
\begin{aligned}
\int_{|s| \geq L} \frac{1}{|s|} |\phi_f(s)| ds &\leq \left(\int_{|s| \geq L} \frac{1}{|s|^2} ds \right)^{\frac{1}{2}} \left(\int |\phi_f(s)|^2 ds \right)^{\frac{1}{2}} \\
&\leq C \left(\int_{|s| \geq L} \frac{1}{|s|^2} ds \right)^{\frac{1}{2}} \left(\int f(x)^2 dx \right)^{\frac{1}{2}} \\
&< \infty.
\end{aligned}
$$

Similarly, for fixed n we can show that $\frac{1}{|s|} |\phi_K(sh)|$ is also absolutely integrable on $\{s : |s| > L\}$. Whence, for a given sample we conclude that $A_n(s)$ is absolutely integrable with respect to s.

Secondly, let $\mathbf{1}_{(A)}$ denote the indicator function of event A, and let $\hat{F}_n(x) = \int_{-\infty}^{x} \hat{f}_n(y) dy$, $\overline{\hat{F}}(x) = \int_x^{\infty} \hat{f}_n(y) dy$. By integration by part, it is easy to check that

$$
\begin{aligned}
&\frac{\phi_f(s) - \phi_{emp}(s)\phi_K(sh)}{is} \\
&= \int e^{isx} \left(\mathbf{1}_{(x \geq 0)}(\overline{F}(x) - \overline{\hat{F}}_n(x)) + \mathbf{1}_{(x<0)} \hat{F}_n(x) \right) dx \\
&= \int e^{isx} \left(\mathbf{1}_{(x \geq 0)}(\overline{F}(x) - \overline{\hat{F}}_n(x)) - \mathbf{1}_{(x<0)}(F(x) - \hat{F}_n(x)) \right) dx, \qquad \text{(A.1)}
\end{aligned}
$$

where in the last step we used the fact that $F(x) = 0$ for $x < 0$. From (A.1) we know that $G_n(x) := \mathbf{1}_{(x \geq 0)}(\overline{F}(x) - \overline{\hat{F}}_n(x)) - \mathbf{1}_{(x<0)}(F(x) - \hat{F}_n(x))$ is the original function corresponding to the Fourier transform $\frac{\phi_f(s) - \phi_{emp}(s)\phi_K(sh)}{is}$.

Finally, we need the following result,

$$
\begin{aligned}
\frac{1}{1 - \frac{\lambda}{c} \frac{\phi_{emp}(s)\phi_K(sh) - 1}{is}} &= \frac{1}{1 - \frac{\lambda}{c} \frac{\phi_f(s) - 1}{is} + A_n(s)} \\
&= \frac{1}{1 - \frac{\lambda}{c} \frac{\phi_f(s) - 1}{is}} - \frac{A_n(s)}{\left(1 - \frac{\lambda}{c} \frac{\phi_f(s) - 1}{is}\right)^2} \\
&\quad + \frac{A_n^2(s)}{\left(1 - \frac{\lambda}{c} \frac{\phi_f(s) - 1}{is} + A_n(s)\right)\left(1 - \frac{\lambda}{c} \frac{\phi_f(s) - 1}{is}\right)^2},
\end{aligned}
$$

which implies that

$$
\begin{aligned}
\tilde{\varphi}_n(u) &= \frac{1}{2\pi} \int \frac{1 - e^{-isu}}{is} \left(\frac{1}{1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is} + A_n(s)} - 1 \right) ds \\
&= \varphi(u) - \frac{1}{2\pi} \int \frac{1 - e^{-isu}}{is} \frac{A_n(s)}{\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is}\right)^2} ds \\
&\quad + \frac{1}{2\pi} \int \frac{1 - e^{-isu}}{is} \frac{A_n^2(s)}{\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is} + A_n(s)\right)\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is}\right)^2} ds \\
&:= \varphi(u) - \mathrm{I}_n(u) + \mathrm{II}_n(u). \quad\quad\quad (A.2)
\end{aligned}
$$

It is readily seen that $\mathrm{I}_n(u)$ and $\mathrm{II}_n(u)$ are real functions. Using the inequality $|\phi_f(s) - 1| \leq |i\mu s|$, we have

$$
\left| 1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is} \right|^2 \geq \left(1 - \frac{\lambda}{c}\left|\frac{\phi_f(s)-1}{is}\right| \right)^2 \geq \left(1 - \frac{\lambda\mu}{c} \right)^2 > \rho
$$

thanks to assumption (F4).

Hence,

$$
\begin{aligned}
|\mathrm{I}_n(u)| &\leq \frac{1}{2\pi} \int \left| \frac{1 - e^{-isu}}{is} \right| \cdot \frac{|A_n(s)|}{\left| 1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is} \right|^2} ds \\
&\leq C \int \left| \frac{1 - e^{-isu}}{is} \right| |A_n(s)| ds \\
&\leq C \left(\int |A_n(s)|^4 ds \right)^{\frac{1}{4}}, \quad\quad\quad (A.3)
\end{aligned}
$$

where the last step follows from Hölder's inequality and the integrability of $\left|\frac{1-e^{-isu}}{is}\right|^{\frac{4}{3}}$. For $0 < \delta < \rho$, define the following set

$$
B_{n,\delta} = \left\{ \sup_s |A_n(s)| \leq \delta \right\}
$$

and let $\overline{B}_{n,\delta}$ be its complementary set. On the set $B_{n,\delta}$, we have

$$
\left| 1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is} + A_n(s) \right| \geq 1 - \frac{\lambda}{c}\left|\frac{\phi_f(s)-1}{is}\right| - |A_n(s)| \geq 1 - \frac{\lambda\mu}{c} - \delta \geq \rho - \delta,
$$

and consequently,

$$
\begin{aligned}
|\mathrm{II}_n(u)| &\leq \frac{1}{2\pi} \int \left| \frac{1 - e^{-isu}}{is} \frac{A_n^2(s)}{\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is} + A_n(s)\right)\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is}\right)^2} \right| ds \\
&\leq C \int \left| \frac{1 - e^{-isu}}{is} \right| |A_n(s)|^2 ds,
\end{aligned}
$$

13

which leads to

$$|\text{II}_n(u)| \leq C \int |A_n(s)|^2 ds \qquad \text{(A.4)}$$

and

$$|\text{II}_n(u)| \leq C \left(\int |A_n(s)|^4 ds \right)^{\frac{1}{2}} \qquad \text{(A.5)}$$

thanks to Cauchy-Schwarz inequality. Note that (A.4) and (A.5) hold on the set $B_{n,\delta}$.

For notational convenience in the following arguments, we define

$$q(x) = \sum_{j=1}^{\infty} \left(\frac{\lambda\mu}{c} \right)^j f_e^{*j}(x).$$

Proof of Proposition 1. Let ϵ_1 and ϵ_2 be two small constants such that

$$\varphi(u) + \epsilon_1 + \epsilon_2 < M.$$

Let

$$V_n = \{|\text{I}_n(u)| < \epsilon_1, \ |\text{II}_n(u)| < \epsilon_2\}$$

and $\overline{V}_n$ be its complementary set. Then it follows from (A.2) that $|\tilde{\varphi}_n(u)| < M$ on the set V_n. Consequently, on the set V_n the truncation in (2.5) is neglectable. With this in hand, we have

$$
\begin{aligned}
&\hat{\varphi}_n(u) - \varphi(u) \\
=\ &(\hat{\varphi}_n(u) - \varphi(u))\, \mathbf{1}_{(V_n)} + (\hat{\varphi}_n(u) - \varphi(u))\mathbf{1}_{(\overline{V}_n)} \\
=\ &(\tilde{\varphi}_n(u) - \varphi(u))\, \mathbf{1}_{(V_n)} + (\hat{\varphi}_n(u) - \varphi(u))\, \mathbf{1}_{(\overline{V}_n)} \\
=\ &\sum_{i=1}^{4} H_{n,i}(u), \qquad\qquad\qquad\qquad\qquad \text{(A.6)}
\end{aligned}
$$

where

$$H_{n,1}(u) = -\text{I}_n(u), \quad H_{n,2}(u) = \text{II}_n(u)\mathbf{1}_{(V_n)}, \quad H_{n,3}(u) = \text{I}_n(u)\mathbf{1}_{(\overline{V}_n)},$$
$$H_{n,4}(u) = (\hat{\varphi}_n(u) - \varphi(u))\mathbf{1}_{(\overline{V}_n)}.$$

Thus, the bias can be expressed as

$$\text{Bias}\,(\hat{\varphi}_n(u)) = \sum_{i=1}^{4} \mathbb{E}[H_{n,i}(u)]. \qquad \text{(A.7)}$$

We treat the expectations in (A.7) respectively.

By Fubini's theorem and Fourier inversion, we have

$$
\begin{aligned}
\mathrm{I}_n(u) &= \frac{1}{2\pi} \int \frac{1 - e^{-isu}}{is} \frac{A_n(s)}{\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is}\right)^2} ds \\
&= \frac{1}{2\pi} \int \int_0^u e^{-isx} dx \frac{A_n(s)}{\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is}\right)^2} ds \\
&= \int_0^u \frac{1}{2\pi} \int e^{-isx} \frac{A_n(s)}{\left(1 - \frac{\lambda}{c}\frac{\phi_f(s)-1}{is}\right)^2} ds dx \\
&= \frac{\lambda}{c} \int_0^u \left(\sum_{j=0}^\infty \left(\frac{\lambda\mu}{c}\right)^j f_e^{*j}\right)^{*2} * G_n(x) dx \\
&= \frac{\lambda}{c} \left(\int_0^u G_n(x)dx + 2\int_0^u q * G_n(x)dx + \int_0^u q^{*2} * G_n(x)dx\right) \quad (A.8)
\end{aligned}
$$

Then using Lemma 8 in Appendix B, we obtain

$$
\mathbb{E}[H_{n,1}(u)] = O(h^2). \tag{A.9}
$$

Note that

$$
\mathbb{E}|H_{n,2}(u)| = \mathbb{E}\left[|\mathrm{II}_n(u)|; V_n \cap B_{n,\delta}\right] + \mathbb{E}\left[|\mathrm{II}_n(u)|; V_n \cap \overline{B}_{n,\delta}\right]. \tag{A.10}
$$

By (A.4) and (B.2) we have

$$
\begin{aligned}
\mathbb{E}[|\mathrm{II}_n(u)|; V_n \cap B_{n,\delta}] &\leq C \cdot \mathbb{E}\int |A_n(s)|^2 ds \\
&= \begin{cases} O(n^{-1} + h^3), & \text{if } f(0) > 0, \\ O(n^{-1} + h^4), & \text{if } f(0) = 0. \end{cases}
\end{aligned} \tag{A.11}
$$

By Lemma 2 in Appendix B we have

$$
\mathbb{E}\left[|\mathrm{II}_n(u)|; V_n \cap \overline{B}_{n,\delta}\right] \leq \epsilon_2 P(\overline{B}_{n,\delta}) = o(n^{-1}). \tag{A.12}
$$

Thus, by (A.10)-(A.12) we have

$$
\mathbb{E}|H_{n,2}(u)| = \begin{cases} O(n^{-1} + h^3), & \text{if } f(0) > 0, \\ O(n^{-1} + h^4), & \text{if } f(0) = 0. \end{cases} \tag{A.13}
$$

In order to study the third and fourth expectations in (A.7), we need bound the probability $\mathbb{P}(\overline{V}_n)$. Obviously, we have

$$
\begin{aligned}
\mathbb{P}(\overline{V}_n) &\leq \mathbb{P}(|\mathrm{I}_n(u)| \geq \epsilon_1) + \mathbb{P}(|\mathrm{II}_n(u)| \geq \epsilon_2) \\
&\leq \mathbb{P}(|\mathrm{I}_n(u)| \geq \epsilon_1) + \mathbb{P}(\{|\mathrm{II}_n(u)| \geq \epsilon_2\} \cap B_{n,\delta}) + \mathbb{P}(\overline{B}_{n,\delta}).
\end{aligned}
$$

By (A.3), Markov's inequality and (B.3), we have

$$
\mathbb{P}(|\mathrm{I}_n(u)| \geq \epsilon_1) \leq \frac{1}{\epsilon_1^4}\mathbb{E}|\mathrm{I}_n(u)|^4 \leq C \cdot \mathbb{E}\int |A_n(s)|^4 ds = O(h^4 + \frac{1}{n^2 h}). \tag{A.14}
$$

15

Note that inequality (A.5) holds on the set $B_{n,\delta}$. Thus, by Markov's inequality and (B.3) we have

$$\mathbb{P}\left(\{|\mathrm{II}_n(u)| \geq \epsilon_2\} \cap B_{n,\delta}\right) \leq \frac{1}{\epsilon_2^2}\mathbb{E}[|\mathrm{II}_n(u)|^2; B_{n,\delta}] \leq C \cdot \mathbb{E}\int |A_n(s)|^4 ds = O(h^4 + \frac{1}{n^2 h}),$$

which together with (A.14) and Lemma 2 in Appendix B gives

$$\mathbb{P}(\overline{V}_n) = o(n^{-1}) + O(h^4 + \frac{1}{n^2 h}). \tag{A.15}$$

By Hölder's inequality, (A.3), (A.15) and (B.3), we have

$$\begin{aligned}
\mathbb{E}\,|H_{n,3}(u)| &= \mathbb{E}\left[|\mathrm{I}_n(u)|\mathbf{1}_{(\overline{V}_n)}\right] \\
&\leq \left(\mathbb{E}|\mathrm{I}_n(u)|^4\right)^{\frac{1}{4}}\left(\mathbb{P}(\overline{V}_n)\right)^{\frac{3}{4}} \\
&= \left[O(h^4 + \frac{1}{n^2 h})\right]^{\frac{1}{4}}\left[o(n^{-1}) + O(h^4 + \frac{1}{n^2 h})\right]^{\frac{3}{4}} \\
&= o(n^{-1}) + O(h^4 + \frac{1}{n^2 h}). \tag{A.16}
\end{aligned}$$

Also, by (A.15) we have

$$\mathbb{E}\,|H_{n,4}(u)| \leq (M + \varphi(u))\mathbb{P}(\overline{V}_n) = o(n^{-1}) + O(h^4 + \frac{1}{n^2 h}). \tag{A.17}$$

Finally, by (A.7), (A.9), (A.13), (A.16) and (A.17) we obtain the desired result. This completes the proof.

Proof of Proposition 2. By (A.6) we obtain

$$\begin{aligned}
\mathrm{Var}(\hat{\varphi}_n(u)) &= \sum_{i=1}^{4}\mathrm{Var}(H_{n,i}(u)) + 2\sum_{i=1}^{4}\sum_{j=i+1}^{4}\mathrm{Cov}(H_{n,i}(u), H_{n,j}(u)) \\
&\leq 4\sum_{i=1}^{4}\mathrm{Var}(H_{n,i}(u)), \tag{A.18}
\end{aligned}$$

where we used Cauchy-Schwarz inequality and the inequality $2ab \leq a^2 + b^2$ in the second step.

We analyze the variances in (A.18) respectively. Firstly, by (A.8) and Lemma 9 , we have

$$\begin{aligned}
&\mathrm{Var}(H_{n,1}(u)) \\
&= \frac{\lambda^2}{c^2}\mathrm{Var}\left(\int_0^u G_n(x)dx + 2\int_0^u q * G_n(x)dx + \int_0^u q^{*2} * G_n(x)dx\right) \\
&= \frac{\lambda^2}{c^2}\frac{\sum_{j=1}^{3}Q_j(u) + 2\sum_{j=4}^{6}Q_j(u)}{n} + o(n^{-1}). \tag{A.19}
\end{aligned}$$

Secondly, by Cauchy-Schwarz inequality we have

$$\mathrm{Var}(H_{n,2}(u)) \leq 2\mathrm{Var}(\mathrm{II}_n(u)\mathbf{1}_{(V_n \cap B_{n,\delta})}) + 2\mathrm{Var}(\mathrm{II}_n(u)\mathbf{1}_{(V_n \cap \overline{B}_{n,\delta})}).$$

16

By (A.5) and (B.3) we have

$$
\begin{aligned}
\mathrm{Var}(\mathrm{II}_n(u)\mathbf{1}_{(V_n \cap B_{n,\delta})}) &\leq \mathbb{E}\left[|\mathrm{II}_n(u)|^2; V_n \cap B_{n,\delta}\right] \\
&\leq C \cdot \mathbb{E}\int |A_n(s)|^4 ds \\
&= O(h^4 + \frac{1}{n^2 h}),
\end{aligned}
$$

which, together with $\mathrm{Var}(\mathrm{II}_n(u)\mathbf{1}_{(V_n \cap \overline{B}_{n,\delta})}) \leq C \cdot \mathbb{P}(\overline{B}_{n,\delta}) = o(n^{-1})$, gives

$$
\mathrm{Var}(H_{n,2}(u)) = o(n^{-1}) + O(h^4 + \frac{1}{n^2 h}). \tag{A.20}
$$

Thirdly, by Cauchy-Schwarz inequality, (A.3), (A.15) and (B.3), we have

$$
\begin{aligned}
\mathrm{Var}(H_{n,3}(u)) &\leq \mathbb{E}\left[|\mathrm{I}_n(u)|^2; \overline{V}_n\right] \\
&\leq \left(\mathbb{E}|\mathrm{I}_n(u)|^4\right)^{\frac{1}{2}} \left(\mathbb{P}(\overline{V}_n)\right)^{\frac{1}{2}} \\
&= \left[O(h^4 + \frac{1}{n^2 h})\right]^{\frac{1}{2}} \left[o(n^{-1}) + O(h^4 + \frac{1}{n^2 h})\right]^{\frac{1}{2}} \\
&= o(n^{-1}) + O(h^4 + \frac{1}{n^2 h}). \tag{A.21}
\end{aligned}
$$

Finally, by (A.15) we have

$$
\begin{aligned}
\mathrm{Var}(H_{n,4}(u)) &\leq \mathbb{E}[(\hat{\varphi}_n(u) - \varphi(u))^2; \overline{V}_n] \\
&\leq (M + \varphi(u))^2 \mathbb{P}(\overline{V}_n) \\
&= o(n^{-1}) + O(h^4 + \frac{1}{n^2 h}). \tag{A.22}
\end{aligned}
$$

Hence, by (A.18)-(A.22) we obtain the desired result. This completes the proof.

Proof of Proposition 3. It follows from Cauchy-Schwarz inequality that for two random variables ξ_1 and ξ_2, if $\mathrm{Var}(\xi_2) = o(\mathrm{Var}(\xi_1))$, then $\mathrm{Var}(\xi_1 + \xi_2) \sim \mathrm{Var}(\xi_1)$. Whence, by (A.6) and (A.19)-(A.22) we have

$$
\mathrm{Var}\left(\hat{\varphi}_n(u)\right) \sim \mathrm{Var}(H_{n,1}(u)) \sim \frac{\lambda^2}{c^2} \frac{\sum_{j=1}^{3} Q_j(u) + 2\sum_{j=4}^{6} Q_j(u)}{n}. \tag{A.23}
$$

Employing (A.6) we have

$$
\frac{\hat{\varphi}_n(u) - \varphi(u)}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}} = \frac{\sum_{i=1}^{4} H_{n,i}(u)}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}}.
$$

It follows from Markov's inequality that for any $\epsilon > 0$,

$$
\mathbb{P}\left(\frac{\left|\sum_{i=2}^{4} H_{n,i}(u)\right|}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}} > \epsilon\right) \leq \frac{\sum_{i=2}^{4} \mathbb{E}|H_{n,i}(u)|}{\epsilon\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}},
$$

which together with (A.13), (A.16), (A.17) and (A.23) implies that $\frac{\sum_{i=2}^{4} H_{n,i}(u)}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}}$ converges to zero in probability, and by Slutsky's theorem it can be neglected. Whence, we only need to study $\frac{H_{n,1}(u)}{\sqrt{\mathrm{Var}(\hat{\varphi}(u))}}$, which can be rewritten as

$$\frac{H_{n,1}(u) - \mathbb{E}[H_{n,1}(u)]}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}} + \frac{\mathbb{E}[H_{n,1}(u)]}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}}.$$

By (A.9) and (A.23) we have $\frac{\mathbb{E}[H_{n,1}(u)]}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}} \to 0$. Thus, we only need to verify

$$\frac{H_{n,1}(u) - \mathbb{E}[H_{n,1}(u)]}{\sqrt{\mathrm{Var}(\hat{\varphi}_n(u))}} \xrightarrow{\mathfrak{D}} \mathcal{N}(0,1),$$

which can be proved by using (A.23) and checking some sufficient conditions for central limit theorem.

B   Lemmas and Proofs

We present some lemmas that are useful in deriving the order of $\mathrm{Bias}(\hat{\varphi}_n(u))$ and $\mathrm{Var}(\hat{\varphi}_n(u))$.

Lemma 1 *Suppose that assumption (F3) holds. Then for any $\epsilon > 0$*

$$\mathbb{P}\left(\int_0^\infty |F(x) - F_{emp}(x)| dx > \epsilon\right) = o(n^{-1}),$$

where $F_{emp}(x) = \frac{1}{n}\sum_{j=1}^{n} I_{(X_j \le x)}$ is the empirical distribution function.

Proof.   For $\alpha > 0$ we can find some positive integer $n_{\epsilon,\alpha}$ such that for all $n \ge n_{\epsilon,\alpha}$, $\int_{n^\alpha}^\infty (1 - F(x)) dx < \frac{\epsilon}{2}$. Then for $n \ge n_{\epsilon,\alpha}$, we have

$$\mathbb{P}\left(\int_0^\infty |F(x) - F_{emp}(x)| dx > \epsilon\right)$$
$$\le \mathbb{P}\left(\int_0^{n^\alpha} |F(x) - F_{emp}(x)| dx + \int_{n^\alpha}^\infty (1 - F(x)) dx + \int_{n^\alpha}^\infty (1 - F_{emp}(x)) dx > \epsilon\right)$$
$$\le \mathbb{P}\left(\int_0^{n^\alpha} |F(x) - F_{emp}(x)| dx + \int_{n^\alpha}^\infty (1 - F_{emp}(x)) dx > \frac{\epsilon}{2}\right)$$
$$\le \mathbb{P}\left(\int_0^{n^\alpha} |F(x) - F_{emp}(x)| dx > \frac{\epsilon}{4}\right) + \mathbb{P}\left(\int_{n^\alpha}^\infty (1 - F_{emp}(x)) dx > \frac{\epsilon}{4}\right).$$

By Lemma 2.1.1 in Prakasa Rao (1983), we have

$$\mathbb{P}\left(\int_0^{n^\alpha} |F(x) - F_{emp}(x)| dx > \frac{\epsilon}{4}\right) \le \mathbb{P}\left(\sup_{x>0} |F(x) - F_{emp}(x)| > \frac{\epsilon}{4n^\alpha}\right)$$
$$\le C\exp\left(-\frac{\epsilon^2}{8}n^{1-2\alpha}\right)$$

for some constant $C > 0$, which implies that for $0 < \alpha < \frac{1}{2}$

$$\mathbb{P}\left(\int_0^{n^\alpha} |F(x) - F_{emp}(x)|dx > \frac{\epsilon}{4}\right) = o(n^{-1}). \tag{B.1}$$

By Markov's inequality, we have

$$
\begin{aligned}
\mathbb{P}\left(\int_{n^\alpha}^\infty (1 - F_{emp}(x))dx > \frac{\epsilon}{4}\right) &= \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n (X_i - n^\alpha)I(X_i > n^\alpha) > \frac{\epsilon}{4}\right) \\
&\leq \frac{4}{\epsilon}\mathbb{E}[(X_i - n^\alpha)I(X_i > n^\alpha)] \\
&= \frac{4}{\epsilon}\int_{n^\alpha}^\infty (x - n^\alpha)dF(x) \\
&= \frac{4}{\epsilon}\int_{n^\alpha}^\infty \overline{F}(x)dx.
\end{aligned}
$$

By the above result, L'Hôpital's rule and assumption (F3), we have

$$\mathbb{P}\left(\int_{n^\alpha}^\infty (1 - F_{emp}(x))dx > \frac{\epsilon}{4}\right) = o(n^{-1}),$$

which together with (B1) gives the desired result. $\qquad\square$

Lemma 2 *Suppose that $h \to 0$, $\mu_1(K) < \infty$, and assumption (F3) holds. Then for any $\epsilon > 0$*

$$\mathbb{P}\left(\sup_s \left|\frac{\phi_f(s) - \phi_{emp}(s)\phi_K(sh)}{is}\right| > \epsilon\right) = o(n^{-1}).$$

Proof. Using the inequality $|e^{isx} - 1| \leq |sx|$ and the following result

$$
\begin{aligned}
\left|\frac{\phi_f(s) - \phi_{emp}(s)}{is}\right| &= \left|\frac{\int_0^\infty (e^{isx} - 1)dF(x) - \int_0^\infty (e^{isx} - 1)dF_{emp}(x)}{is}\right| \\
&= \left|\int_0^\infty \int_0^x e^{isy}dydF(x) - \int_0^\infty \int_0^x e^{isy}dydF_{emp}(x)\right| \\
&= \left|\int_0^\infty (1 - F(x))e^{isx}dx - \int_0^\infty (1 - F_{emp}(x))e^{isx}dx\right| \\
&\leq \int_0^\infty |F(x) - F_{emp}(x)|dx,
\end{aligned}
$$

we have

$$
\begin{aligned}
&\left|\frac{\phi_f(s) - \phi_{emp}(s)\phi_K(sh)}{is}\right| \\
&\leq |\phi_f(s)|\left|\frac{1 - \phi_K(sh)}{is}\right| + |\phi_K(sh)|\left|\frac{\phi_f(s) - \phi_{emp}(s)}{is}\right| \\
&\leq h\mu_1(K) + \int_0^\infty |F(x) - F_{emp}(x)|dx.
\end{aligned}
$$

19

For $\epsilon > 0$, there exists some positive integer n_0 such that for $n \geq n_0$, $h\mu_1(K) < \frac{\epsilon}{2}$. Then for $n \geq n_0$, we have

$$\mathbb{P}\left(\sup_s \left|\frac{\phi_f(s) - \phi_{emp}(s)\phi_K(s)}{is}\right| > \epsilon\right)$$

$$\leq \mathbb{P}\left(h\mu_1(K) + \int_0^\infty |F(x) - F_{emp}(x)|dx > \epsilon\right)$$

$$\leq \mathbb{P}\left(\int_0^\infty |F(x) - F_{emp}(x)|dx > \frac{\epsilon}{2}\right)$$

$$= o(n^{-1})$$

thanks to Lemma 1. This completes the proof. $\qquad \square$

Lemma 3 *Suppose that assumptions **K**, (F1), (F2) and (F3) hold. Then we have*

$$\mathbb{E}\int |A_n(s)|^2 ds = \begin{cases} O(n^{-1} + h^3), & \text{if } f(0) > 0, \\ O(n^{-1} + h^4), & \text{if } f(0) = 0, \end{cases} \tag{B.2}$$

and

$$\mathbb{E}\int |A_n(s)|^4 ds = O(h^4 + \frac{1}{n^2 h}). \tag{B.3}$$

Proof. By Parseval's identity and Bias-Variance decomposition, we have

$$\mathbb{E}\int |A_n(s)|^2 ds \leq C \cdot \mathbb{E}\frac{1}{2\pi}\int \left|\frac{\phi_f(s) - \phi_{emp}(s)\phi_K(sh)}{is}\right|^2 ds$$

$$= C \cdot \mathbb{E}\int G_n(x)^2 dx$$

$$= C \cdot \int [\mathbb{E}G_n(x)]^2 dx + C \cdot \int \text{Var}(G_n(x))dx,$$

which together with Lemma 6 and Lemma 7 gives (B.2).

By C_r-inequality, we have

$$\mathbb{E}\int |A_n(s)|^4 ds \leq C \cdot \mathbb{E}\int \left|\frac{(\phi_{emp}(s) - \phi_f(s))\phi_K(sh)}{is}\right|^4 ds$$

$$+ C \int \left|\frac{\phi_f(s)(\phi_K(sh) - 1)}{is}\right|^4 ds. \tag{B.4}$$

Using the symmetry of the kernel K and the inequality $|e^{ix} - 1 - ix| \leq \frac{x^2}{2}$, we have

$$|\phi_K(sh) - 1| \leq \int |e^{ishx} - 1 - ishx|K(x)dx$$

$$\leq \frac{1}{2}s^2 h^2 \mu_2(K).$$

Consequently,

$$\int \left|\frac{\phi_f(s)(\phi_K(sh) - 1)}{is}\right|^4 ds \leq C \int \left|\frac{\phi_f(s)}{is}\right|^4 |\phi_K(sh) - 1|^2 ds \leq Ch^4 \tag{B.5}$$

20

thanks to $\int |\phi_f(s)|^4 ds < \infty$. Using the inequality $|e^{ix} - 1| \le |x|$ and C_r-inequality, we have

$$\mathbb{E}\left|\frac{\phi_f(s) - e^{isX_j}}{is}\right|^4 \le \mathbb{E}(\mu + X_j)^4 \le C(\mu^4 + \mathbb{E}X_j^4) < \infty.$$

Thus, we can use Rosenthal's inequality (see Theorem 2.12 in Hall and Heyde (1980)) to obtain

$$\mathbb{E}\left|\frac{\phi_f(s) - \phi_{emp}(s)}{is}\right|^4 \le \frac{C}{n^2},$$

which holds uniformly in s. By the above inequality we have

$$\mathbb{E}\int\left|\frac{(\phi_{emp}(s) - \phi_f(s))\phi_K(sh)}{is}\right|^4 ds \le \frac{C}{n^2}\int|\phi_K(sh)|^4 ds \le \frac{C}{n^2 h},$$

which, together with (B.4) and (B.5), gives (B.3). $\qquad\square$

We need the following Taylor's expansions that are special cases of Lemma 4.3 and 4.4 in van Eeden (1985).

Lemma 4 *Suppose that Assumption (F1) and (F2) hold. If $x > 0$ and $x - uh < 0$, then*

$$f(x - uh) - f(x) + uhf'(x) - h^2\int_0^u (u - s)f''(x - sh)ds = -f'(0+)(x - uh) - f(0).$$

If $x < 0$ and $x - uh > 0$, then

$$f(x - uh) - f(x) + uhf'(x) - h^2\int_0^u (u - s)f''(x - sh)ds = f'(0+)(x - uh) + f(0).$$

If $x(x - uh) > 0$, then

$$f(x - uh) - f(x) + uhf'(x) - h^2\int_0^u (u - s)f''(x - sh)ds = 0.$$

The following inequality is given by Lemma A.1 in Tsybakov (2009).

Lemma 5 *(Generalized Minkowski inequality) For any Borel function p on $\mathbb{R} \times \mathbb{R}$, we have*

$$\int\left(\int p(y, x)dy\right)^2 dx \le \left(\int\left(\int (p(y, x))^2 dx\right)^{\frac{1}{2}} dy\right)^2.$$

Lemma 6 *Suppose that assumptions $\boldsymbol{K}$, (F1) and (F2) hold. Then*

$$\int [\mathbb{E}G_n(x)]^2\, dx = \begin{cases} O(h^3), & \text{if } f(0) > 0, \\ O(h^4), & \text{if } f(0) = 0. \end{cases}$$

Proof. For $x \geq 0$, by Taylor's expansions given in Lemma 4 we have

$$
\begin{aligned}
\mathbb{E}G_n(x) &= \mathbb{E} \int_x^\infty \left(f(y) - \hat{f}_n(y) \right) dx \\
&= \int_x^\infty \int_0^\infty (f(y) - K_h(y-z)) f(z) dz dy \\
&= \int_x^\infty \int K(s)(f(y) - f(y-sh)) ds dy \\
&= \int_x^\infty \int K(s) \left(f(0) + f'(0+)(y-sh) \right) \mathbf{1}_{(y<sh)} ds dy \\
&\quad -h^2 \int_x^\infty \int K(s) \int_0^s (s-t) f''(y-th) dt ds dy \\
&= f(0) \int K(s)(sh-x) \mathbf{1}_{(x \leq sh)} ds \\
&\quad -\frac{1}{2} f'(0+) \int K(s)(x-sh)^2 \mathbf{1}_{(x \leq sh)} ds \\
&\quad +h^2 \int K(s) \int_0^s (s-t)[f'(0+)\mathbf{1}_{(x \leq th)} + f'(x-th)\mathbf{1}_{(x>th)}] dt ds \\
&:= \chi_{+,1}(x) + \chi_{+,2}(x) + \chi_{+,3}(x).
\end{aligned}
$$

Similarly, for $x < 0$ we have

$$
\begin{aligned}
\mathbb{E}G_n(x) &= \mathbb{E} \int_{-\infty}^x (\hat{f}_n(y) - f(y)) dy \\
&= f(0) \int K(s)(x-sh) \mathbf{1}_{(x \geq sh)} ds \\
&\quad +\frac{1}{2} f'(0+) \int K(s)(x-sh)^2 \mathbf{1}_{(x \geq sh)} ds \\
&\quad +h^2 \int K(s) \int_0^s (s-t)[f'(x-th) - f'(0+)]\mathbf{1}_{(x>th)} dt ds \\
&= \chi_{-,1}(x) + \chi_{-,2}(x) + \chi_{-,3}(x).
\end{aligned}
$$

Then

$$
\begin{aligned}
\int [\mathbb{E}G_n(x)]^2 dx &= \int \left(\sum_{j=1}^3 (\chi_{+,j}(x) + \chi_{-,j}(x)) \right)^2 dx \\
&\leq 3 \sum_{j=1}^3 \int_0^\infty \chi_{+,j}(x)^2 dx + 3 \sum_{j=1}^3 \int_{-\infty}^0 \chi_{-,j}(x)^2 dx. \quad \text{(B.6)}
\end{aligned}
$$

By Lemma 5 we have

$$
\begin{aligned}
\int_0^\infty \chi_{+,1}(x)^2 dx &= f(0)^2 \int_0^\infty \left(\int K(s)(sh-x)\mathbf{1}_{(x\le sh)} ds \right)^2 dx \\
&\le f(0)^2 \left(\int K(s) \left(\int_0^\infty (sh-x)^2 \mathbf{1}_{(x\le sh)} dx \right)^{\frac{1}{2}} ds \right)^2 \\
&\le \frac{1}{3} h^3 f(0)^2 \left(\int K(s)|s|^{\frac{3}{2}} ds \right)^2 \\
&\le Ch^3 f(0)^2,
\end{aligned}
\tag{B.7}
$$

and similarly,

$$
\int_0^\infty \chi_{+,2}(x)^2 dx \le \frac{1}{20} h^5 f'(0+)^2 \left(\int K(s)|s|^{\frac{5}{2}} ds \right)^2 \le Ch^5.
\tag{B.8}
$$

Let $m_+(x,t) = f'(0+)\mathbf{1}_{(x\le th)} + f'(x-th)\mathbf{1}_{(x>th)}$, and write $\chi_{+,3}(x) = \chi_{+,3,1}(x) + \chi_{+,3,2}(x)$, where for $x \ge 0$

$$
\begin{aligned}
\chi_{+,3,1}(x) &= h^2 \int_0^\infty K(s) \int_0^s (s-t)m_+(x,t)dt\,ds, \\
\chi_{+,3,2}(x) &= h^2 \int_{-\infty}^0 K(s) \int_s^0 (t-s)f'(x-th)dt\,ds.
\end{aligned}
$$

By Cauchy-Schwarz inequality, for $s \ge 0$ we have

$$
\begin{aligned}
&\left(\int_0^s (s-t)m_+(x,t)dt \right)^2 \\
&\le \left(\int_0^s (s-t)dt \right) \left(\int_0^s (s-t)(m_+(x,t))^2 dt \right) \\
&\le s^2 \int_0^s (s-t) \left[f'(0+)^2 \mathbf{1}_{(x\le th)} + f'(x-th)^2 \mathbf{1}_{(x>th)} \right] dt.
\end{aligned}
$$

By this inequality, Lemma 5 and Fubini's theorem, we have

$$
\begin{aligned}
&\int_0^\infty \chi_{+,3,1}(x)^2 dx \\
&= h^4 \int_0^\infty \left(\int_0^\infty K(s) \int_0^s (s-t)m_+(x,t)dt\,ds \right)^2 dx \\
&\le h^4 \left(\int_0^\infty K(s) \left(\int_0^\infty \left(\int_0^s (s-t)m_+(x,t)dt \right)^2 dx \right)^{\frac{1}{2}} ds \right)^2 \\
&\le Ch^4 \left(\int_0^\infty K(s) \left(f'(0+)^2 hs^5 + R(f')s^4 \right)^{\frac{1}{2}} ds \right)^2 \\
&\le Ch^4 \left(\int_0^\infty K(s)(s^{\frac{5}{2}} \vee s^2) ds \right)^2 \\
&\le Ch^4,
\end{aligned}
$$

and similarly,

$$\int_0^\infty \chi_{+,3,2}(x)^2 dx \leq Ch^4 \left(\int_{-\infty}^0 K(s)s^2 ds \right)^2 \leq Ch^4.$$

Thus, we have

$$
\begin{aligned}
\int_0^\infty \chi_{+,3}(x)^2 dx &= \int_0^\infty (\chi_{+,3,1}(x) + \chi_{+,3,2}(x))^2 dx \\
&\leq 2\int_0^\infty (\chi_{+,3,1}(x)^2 + \chi_{+,3,2}(x)^2) dx \\
&\leq Ch^4.
\end{aligned}
\tag{B.9}
$$

By exactly the same arguments as above, we can obtain

$$\int_{-\infty}^0 \chi_{-,1}(x)^2 dx \leq Ch^3 f(0)^2, \tag{B.10}$$

$$\int_{-\infty}^0 \chi_{-,2}(x)^2 dx \leq Ch^5, \tag{B.11}$$

$$\int_{-\infty}^0 \chi_{-,3}(x)^2 dx \leq Ch^4. \tag{B.12}$$

Finally, by (B.6)-(B.12) we have

$$\int [\mathbb{E}G_n(x)]^2 \, dx \leq Ch^3 f(0)^2 + Ch^5 + Ch^4.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Lemma 7 *Suppose that assumptions* **K**, *(F1) and (F2) hold. Then*

$$\int Var(G_n(x))^2 dx = O(n^{-1}).$$

Proof. Firstly, we have

$$\int Var(G_n(x))dx = \int_0^\infty Var\left(\overline{\hat{F}_n}(x)\right) dx + \int_{-\infty}^0 Var\left(\hat{F}_n(x)\right) dx.$$

We only treat the first integral on the right hand side of the above equation, because the other one can be analyzed similarly. By some straightforward calculations we obtain

$$
\begin{aligned}
\int_0^\infty Var\left(\overline{\hat{F}_n}(x)\right) dx &= \frac{1}{n}\int_0^\infty Var\left(\int_x^\infty K_h(y-X)dy\right) dx \\
&\leq \frac{1}{n}\int_0^\infty \mathbb{E}\left(\int_x^\infty K_h(y-X)dy\right)^2 dx \\
&= \frac{h}{n}\int_{-\infty}^0 \int_0^\infty \left(\int_t^\infty K(s)ds\right)^2 f(x-th)dxdt \\
&\quad + \frac{h}{n}\int_0^\infty \left(\int_t^\infty K(s)ds\right)^2 dt,
\end{aligned}
$$

Note that $\int_t^\infty K(s)ds \leq 1$. Then

$$\frac{h}{n}\int_{-\infty}^0\int_0^\infty \left(\int_t^\infty K(s)ds\right)^2 f(x-th)dxdt \leq \frac{h}{n}\int_{-\infty}^0\int_0^\infty f(x-th)dxdt = \frac{\mu}{n}.$$

and

$$\frac{h}{n}\int_0^\infty \left(\int_t^\infty K(s)ds\right)^2 dt \leq \frac{h}{n}\int_0^\infty \int_t^\infty K(s)dsdt = \frac{\mu_1(K)h}{2n}.$$

This completes the proof. $\qquad\square$

The following two lemmas can be obtained by exactly the same arguments as that of Zhang et al. (2011). We only give the outline of the derivation of the covariance.

Lemma 8 *Suppose that assumptions **K**, (F1) and (F2) hold. Then*

$$\mathbb{E}\int_0^u G_n(x)dx = O(h^2),$$

$$\mathbb{E}\int_0^u q*G_n(x)dx = O(h^2),$$

$$\mathbb{E}\int_0^u q^{*2}*G_n(x)dx = O(h^2).$$

Lemma 9 *Suppose that assumptions **K**, (F1) and (F2) hold. Then*

$$Var\left(\int_0^u G_n(x)dx\right) = \frac{1}{n}Q_1(u) + o(n^{-1}),$$

$$Var\left(2\int_0^u q*G_n(x)dx\right) = \frac{1}{n}Q_2(u) + o(n^{-1}),$$

$$Var\left(\int_0^u q^{*2}*G_n(x)dx\right) = \frac{1}{n}Q_3(u) + o(n^{-1}),$$

$$Cov\left(\int_0^u G_n(x)dx, 2\int_0^u q*G_n(x)dx\right) = \frac{1}{n}Q_4(u) + o(n^{-1}),$$

$$Cov\left(\int_0^u G_n(x)dx, \int_0^u q^{*2}*G_n(x)dx\right) = \frac{1}{n}Q_5(u) + o(n^{-1}),$$

$$Cov\left(2\int_0^u q*G_n(x)dx, \int_0^u q^{*2}*G_n(x)dx\right) = \frac{1}{n}Q_6(u) + o(n^{-1}),$$

where

$$Q_1(u) = \int_0^u \int_0^u \overline{F}(x_1 \vee x_2) dx_2 dx_1 - \left(\int_0^u \overline{F}(x) dx \right)^2,$$

$$Q_2(u) = 4 \int_0^u \int_0^u \int_0^{x_1} \int_0^{x_2} q(x_1 - z_1) q(x_2 - z_2) \overline{F}(z_1 \vee z_2) dz_2 dz_1 dx_2 dx_1$$
$$-4 \left(\int_0^u \int_0^x q(x - y) \overline{F}(y) dy dx \right)^2,$$

$$Q_3(u) = \int_0^u \int_0^u \int_0^{x_1} \int_0^{x_2} q^{*2}(x_1 - z_1) q^{*2}(x_2 - z_2) \overline{F}(z_1 \vee z_2) dz_2 dz_1 dx_2 dx_1$$
$$- \left(\int_0^u \int_0^x q^{*2}(x - y) \overline{F}(y) dy dx \right)^2,$$

$$Q_4(u) = 2 \int_0^u \int_0^u \int_{x_1}^\infty \int_0^{x_2 \wedge z_1} q(x_2 - z_2) f(z_1) dz_2 dz_1 dx_2 dx_1$$
$$-2 \left(\int_0^u \overline{F}(x) dx \right) \left(\int_0^u \int_0^x q(x - y) \overline{F}(y) dy dx \right),$$

$$Q_5(u) = \int_0^u \int_0^u \int_{x_1}^\infty \int_0^{x_2 \wedge z_1} q^{*2}(x_2 - z_2) f(z_1) dz_2 dz_1 dx_2 dx_1$$
$$- \left(\int_0^u \overline{F}(x) dx \right) \left(\int_0^u \int_0^x q^{*2}(x - y) \overline{F}(y) dy dx \right),$$

$$Q_6(u) = 2 \int_0^u \int_0^u \int_0^{x_1} \int_0^{x_2} q(x_1 - z_1) q^{*2}(x_2 - z_2) \overline{F}(z_1 \vee z_2) dz_2 dz_1 dx_2 dx_1$$
$$-2 \left(\int_0^u \int_0^x q(x - y) \overline{F}(y) dy dx \right) \left(\int_0^u \int_0^x q^{*2}(x - y) \overline{F}(y) dy dx \right).$$

Proof. We only derive the fourth formula, since the others can be obtained similarly. From the definition of G_n, we have

$$\text{Cov}\left(\int_0^u G_n(x)dx, \int_0^u q * G_n(x)dx\right)$$

$$= \text{Cov}\left(\int_0^u \overline{\hat{F}}_n(x)dx, \int_0^u \int_0^x q(x-y)\overline{\hat{F}}_n(y)dydx\right)$$

$$-\text{Cov}\left(\int_0^u \overline{\hat{F}}_n(x)dx, \int_0^u \int_{-\infty}^0 q(x-y)\hat{F}_n(y)dydx\right)$$

$$= \frac{1}{n}\mathbb{E}\left[\left(\int_0^u \int_x^\infty K_h(z-X)dzdx\right)\right.$$

$$\left.\times\left(\int_0^u \int_0^x \int_y^\infty q(x-y)K_h(z-X)dzdydx\right)\right]$$

$$-\frac{1}{n}\mathbb{E}\left[\left(\int_0^u \int_x^\infty K_h(z-X)dzdx\right)\right.$$

$$\left.\times\left(\int_0^u \int_{-\infty}^0 \int_{-\infty}^y q(x-y)K_h(z-X)dzdydx\right)\right]$$

$$-\frac{1}{n}\left(\int_0^u \overline{F}(x)dx\right)\left(\int_0^u \int_0^x q(x-y)\overline{F}(y)dydx\right) + o(n^{-1}). \qquad \text{(B.13)}$$

where the last step follows from Lemma 8. We treat the above two expectations respectively. Firstly, by some changes of variables and Fubini's theorem we have

$$\mathbb{E}\left[\left(\int_0^u \int_x^\infty K_h(z-X)dzdx\right)\right.$$

$$\left.\times\left(\int_0^u \int_0^x \int_y^\infty q(x-y)K_h(z-X)dzdydx\right)\right]$$

$$= \int\int\int_0^u \int_0^u \int_{x_1}^\infty \int_0^{x_2} K(t_1)K(t_2)q(x_2-z_2)f(z_1-t_1h)$$

$$\times \mathbf{1}_{(z_1 \geq t_1 h, z_1 + t_2 h - t_1 h \geq z_2)}dz_2dz_1dx_2dx_1dt_2dt_1.$$

Next, by the Taylor's expansion formulas given in Lemma 4, it is easy to show that

$$\mathbb{E}\left[\left(\int_0^u \int_x^\infty K_h(z-X)dzdx\right)\right.$$

$$\left.\times\left(\int_0^u \int_0^x \int_y^\infty q(x-y)K_h(z-X)dzdydx\right)\right]$$

$$= \int_0^u \int_0^u \int_{x_1}^\infty \int_0^{x_2 \wedge z_1} q(x_2-z_2)f(z_1)dz_2dz_1dx_2dx_1 + o(1). \qquad \text{(B.14)}$$

Similarly, we can obtain

$$\mathbb{E}\left[\left(\int_0^u \int_x^\infty K_h(z-X)dzdx\right)\right.$$

$$\left.\times\left(\int_0^u \int_{-\infty}^0 \int_{-\infty}^y q(x-y)K_h(z-X)dzdydx\right)\right] = o(1). \qquad \text{(B.15)}$$

27

By (B.13)-(B.15) we obtain the desired result. $\qquad\square$

References

[1] ALBRECHER, H., KORTSCHAK, D. (2009). On ruin probability and aggregate claim representations for Pareto claim size distributions. *Insurance: Mathemtics and Economics* **45**, 362-373.

[2] ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.

[3] CROUX, K., VERVAERBEKE, N. (1990). Nonparametric estimators for the probability of ruin. *Insurance: Mathematics and Economics* **9**, 127-130.

[4] EMBRECHTS P., GRÜBEL R., PITTS, S.M. (1993). Some applications of the fast Fourier transform algorithm in insurance mathematics. *Statistica Neerlandica* **47**, 59-75.

[5] FREES, E. W. (1986). Nonparametric estimation of the probability of ruin. *ASTIN Bulletin* **16**, 81-90.

[6] GERBER, H.U. (1979). *An introduction to mathematical risk theory*. University of Pennsylvania, Philadelphia.

[7] GRANDELL, J. (1979). Empirical bounds for ruin probabilities. *Stochastic Processes and their Applications* **8**, 243-255.

[8] HALL, P., HEYDE, C.C. (1980). *Martingale Limit Theory and its Applications*. Academic Press, New York.

[9] HIPP, C. (1989). Estimators and bootstrap confidence intervals for ruin probabilities. *ASTIN Bulletin* **19**, 57-70.

[10] MNATSAKANOV, R., RUYMGAART, L.L., RUYMGAART, F.H. (2008). Nonparametric estimation of ruin probabilities given a random sample of claims. *Mathematical Methods of Statistics* **17**, 35-43.

[11] PITTS, S.M. (1994). Non-parametric estimation of compound distributions with applications in insurance.*Annals of the Institute of Statistical Mathematics* **46**, 537-555.

[12] POLITIS, K. (2003). Semiparametric Estimation for Non-Ruin Probabilities. *Scandinavian Actuarial Journal* **2003(1)**, 75-96.

[13] PRAKASA RAO, B.L.S. (1983). *Nonparametric Functional Estimation*. Academic Process, New York.

[14] RAMSAY, C.M. (2003). A solution to the ruin problem for Pareto distributions *Insurance: Mathematics and Economics* **33**, 109-116.

[15] ROLSKI, T., SCHMIDLI, H., SCHMIDT, V., TEUGELS, J. (1999). *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, Ltd, Chichester.

[16] TSYBAKOV, A.B. (2009). *Introduction to Nonparametric Estimation*. Springer, York.

[17] VAN EEDEN, C. (1985). Mean integrated squared error of kernel estimators when the density and its derivative are not necessarily continuous. *Annals of Institute of Statistical Mathematics* **37**, 461-472.

[18] WAND, M.P., JONES, M.C. (1995). *Kernel Smoothing*, Chapman & Hall, London.

[19] YASUTAKA S. (2012). Non-parametric estimation of the Gerber-Shiu function for the Wiener-Poisson risk model. *Scandinavian Actuarial Journal* **2012 (1)** 56-69.

[20] ZHANG, Z., YANG, H., YANG, H. (2011). On a nonparametric estimator for the finite time survival probability with zero initial surplus, *submitted for publication*.