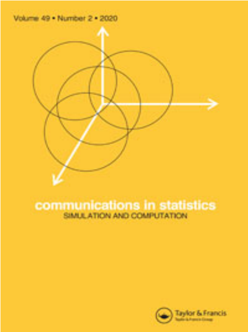

Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Estimation of the mean residual life function in the presence of measurement errors

Hojjat Pourjafar & Vali Zardasht

To cite this article: Hojjat Pourjafar & Vali Zardasht (2020) Estimation of the mean residual life function in the presence of measurement errors, Communications in Statistics - Simulation and Computation, 49:2, 532-555, DOI: 10.1080/03610918.2018.1489054

To link to this article: https://doi.org/10.1080/03610918.2018.1489054

Published online: 09 Dec 2018.

Submit your article to this journal

Article views: 69

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=lssp20

COMMUNICATIONS IN STATISTICS—SIMULATION AND COMPUTATION®
2020, VOL. 49, NO. 2, 532–555
https://doi.org/10.1080/03610918.2018.1489054

Taylor & Francis
Taylor & Francis Group

Check for updates

Estimation of the mean residual life function in the presence of measurement errors

Hojjat Pourjafar and Vali Zardasht

Department of Statistics, Faculty of Sciences, University of Mohaghegh Ardabili, Ardabil, Iran

ABSTRACT
The data available for statistical analysis from many scientific areas often come with measurement error. Ignoring measurement error can bring forth biased estimates and lead to erroneous conclusions to various degrees in a data analysis. This paper considers the problem of estimation of the mean residual life function assuming that observed lifetime random variables are given by a multiplicative measurement error model. The consistency of the estimator are proven under some regularity conditions. It is also shown that the estimator weakly converges to a normal distribution. Finally, numerical examples based on an extensive simulation study and real lifetime data analysis are presented to illustrate the theory and assess performance of the estimator.

ARTICLE HISTORY
Received 26 September 2017
Accepted 6 June 2018

KEYWORDS
Bandwidths; Contaminated lifetime data; Deconvolution kernel density estimator; Measurement error models; Reliability measures

MATHEMATICS SUBJECT CLASSIFICATION
62N02; 62G20

1. Introduction

The mean residual lifetime (MRL) function is an important measure in economics, actuarial relating to life insurance, reliability, and survival analysis. The MRL function computes the expected remaining survival time of a subject given survival up to time t. Let T be a lifetime random variable with continuous distribution function F, survival function $\bar{F} = 1-F$ and density function f. Provided that $E(T)<\infty$, the MRL function is defined as

$$m(t) = E(T-t|T>t) = \frac{\int_t^{\infty} (x-t)f(x)dx}{\bar{F}(t)} = \frac{\int_t^{\infty} \bar{F}(x)dx}{\bar{F}(t)}, \tag{1}$$

and $m(t) = 0$ whenever $\bar{F}(t) = 0$. For more details about the theoretical properties and applications of the MRL function and a comprehensive literature review, we refer the reader to Lai and Xie (2006).

Considering complete and censored failure times data and through different procedures, estimation of the mean residual life function and other statistical inference for that have been taken a great deal of attention by many researchers. For instance, Yang (1978) considered the empirical estimate for the MRL function through replacing the survival function in (1) with its empirical survival function. A natural extension of this estimator to randomly right-censored data is replacing $\bar{F}(t)$ in (1) with its Kaplan-Meier

CONTACT Vali Zardasht zardasht@uma.ac.ir Department of Statistics, Faculty of Sciences, University of Mohaghegh Ardabili, Ardabil, Iran.
© 2018 Taylor & Francis Group, LLC

estimator; see, for example, Yang (1977) and Hall and Wellner (1979) and references therein. Motivated by kernel density estimation, several estimators involving kernels were proposed by various researchers working in the area. For example, Ruiz and Guillamn (1996) estimated the numerator in $m(t)$ by a recursive kernel estimate and used the empirical survival function to estimate the denominator. Chaubey and Sen (1999, 2008), proposed an alternative smoothed estimator based on complete data and right-censored data respectively. Abdous and Berred (2005) used an integrated weighted local linear smoothing technique to smooth the empirical estimator introduced by Yang (1978). Considering the estimation of non-decreasing mean residual life functions problem, Jayasinghe and Zeephongsekul (2013) employed local polynomial regression with fixed design points accompanied by appropriate binning to construct several new estimators for the MRL function.

The data available for statistical analysis from many scientific areas often come with measurement error. For example, many economic data sets are contaminated by the mismeasured variables. In medical image analysis, observable outputs are often blurred images. In astronomy, due to great astronomical distances and atmospheric noise, most data are subject to measurement errors. Ignoring measurement error can bring forth biased estimates and lead to erroneous conclusions to various degrees in a data analysis. Measurement error model is an active and rich research field in statistics. For more examples and details of the measurement error problems, we refer the reader to Buonaccorsi (2010), Meister (2009), and Carroll et al. (2006). In this paper, we consider the problem of point-wise estimation of the mean residual life function when data contain measurement errors. Indeed, considering a non-parametric framework, we propose our estimator for the MRL by replacing the density function in (1) with the standard deconvolution kernel density estimator which was first introduced by Stefanski and Carroll (1990) and Carroll and Hall (1988).

A fundamental problem in measurement error models is to recover an unknown density of a variable when its observed values or data are contaminated with errors. Formally, Let X be the variable of interest, which we cannot observe directly. Instead, based on an observed sample $Y_1, \ldots, Y_n$ drawn independently from the model

$$Y = X + \varepsilon, \tag{2}$$

where the measurement error ε is independent of X, one is interested in estimating f_X, the unknown density function of X. This classical error model is called additive measurement errors model which is commonly considered for contaminated data. It is worth to mention that the circumstances under which the data are collected and their availability determines whether or not an error model is appropriate to use in the data analysis. For the real examples of the classical error model and the model cheing method, we refer the reader to Carroll et al. (2006), chapter 1. Regarding the classical approach to this model, it is assumed that the probability distribution of ε is exactly known; although in many real-life situations this condition cannot be justified. However, in most practical applications, we are able to estimate the error density function of ε from replicated measurements. (cf. Meister 2009, p. 88.)

The well-known deconvolution estimator of f_X is the standard deconvolution kernel density estimator

$$\begin{aligned}\hat{f}_X(x) &= \frac{1}{2\pi}\int \exp(-itx)\phi_K(tb)\frac{\frac{1}{n}\sum_{j=1}^{n}\exp(itY_j)}{\phi_\varepsilon(t)}dt \\ &= \frac{1}{n}\sum_{j=1}^{n}\frac{1}{2\pi}\int e^{it(Y_j-x)}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dt,\end{aligned} \tag{3}$$

where $K:\mathbb{R}\to\mathbb{R}^+$ is kernel function, $b>0$ a bandwidth parameter and ϕ_K, ϕ_ε are characteristic functions of K and ε, respectively. This estimator is well-defined for any non-vanishing ϕ_ε whenever ϕ_K is compactly supported.

From practical point of view, the lifetime variables even observed with measurement errors are positive. It is clear that the additive model may produce a negative value. Thus, the additive model can not be straightly applied for lifetime variables. For this purpose, let T be the lifetime variable which we cannot observe directly. We assume that we observe random variables $Z_1, \ldots, Z_n$ given by multiplicative model

$$Z_i = T_i\delta_i, \quad i = 1, \ldots, n, \tag{4}$$

where δ_i are independent identically distributed random variables, independent of T_is with a known density w.r.t. the Lebesgue measure on the positive real line. This is equivalent (take the natural logarithms of both sides of the equation above) to say that we observe random variables $Y_1 = \ln(Z_1), \ldots, Y_n = \ln(Z_n)$ given by

$$Y_i = X_i + \varepsilon_i, \quad i = 1, \ldots, n, \tag{5}$$

where, $Y_i = \ln(Z_i), X_i = \ln(T_i)$ and $\varepsilon_i = \ln(\delta_i)$. Our objective is to estimate the MRL function of lifetime variable T at any given positive point t from the observations $Z_1, \ldots, Z_n$ or $Y_1, \ldots, Y_n$. Using the estimator (3) and by applying the change of variable technique the density function of T can be estimated by

$$\hat{f}(t) = \frac{1}{t}\hat{f}_X(\ln(t)). \tag{6}$$

Now our proposed estimator for the MRL function $m(t)$ of T can be given by

$$\hat{m}(t) = \frac{\int_t^{\tau_F} x\hat{f}(x)dx}{\int_t^{\tau_F}\hat{f}(x)dx} - t = \frac{\int_t^{\tau_F}\left[1-\hat{F}_X(\ln(x))\right]dx}{1-\hat{F}_X(\ln(t))}, \tag{7}$$

where $\hat{F}_X(t) = \int_{\upsilon_F}^{t}\hat{f}_X(x)dx$ and $\tau_F := \inf\{x : F(x) = 1\}$. The main objective of this article is to investigate some key properties of the proposed estimator defined in (7). It should be mentioned that at the case where the errors are nonrandom and systematic, the above problem can be considered as the estimation of the MRL function under biased samples which has been taken attentions by some researches (cf. Fakoor 2015). The rest of the article is organized as follows. In Sec. 2, we propose and provide proofs of some asymptotic properties of $\hat{m}(t)$. Since a key component of kernel estimating is the selection of an optimal bandwidth, Sec. 3 is devoted to a discussion of bandwidth selection methods. Section 4 looks at a numerical example and finally, some concluding remarks are given in Sec. 5.

2. Asymptotic results

The main purpose of this section is to investigate the asymptotic distribution and consistency of above mentioned $\hat{m}(t)$. In addition, we assume that f the density function of the lifetime random variable T is continuous and bounded. We also assume that the error measurement ε has a non-vanishing characteristic function, ϕ_ε, i.e.,

$$|\phi_\varepsilon(t)|>0, \quad \textit{for all } t \in \mathbb{R}. \tag{8}$$

This holds in many cases of interest, and in particular at the normal model. Let the kernel K be a bounded, even probability density function with characteristic function ϕ_K satisfying (for each fixed positive b)

$$\underset{t\in\mathbb{R}}{Sup}\,|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|<\infty, \qquad \int|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|dt<\infty. \tag{9}$$

The above condition guarantees that $\phi_K^2/|\phi_\varepsilon(./b)|^2, |\phi_K|$ and ϕ_K^2 are all integrable, which in turn implies that ϕ_k is invertible, that is

$$K(x) = (2\pi)^{-1}\int e^{-itx}\phi_K(t)dt.$$

The first theorem gives the consistency of $\hat{m}(t)$.

Theorem 1. *Assume that ϕ_K and ϕ_ε satisfy (8) and (9), and that $\tau_F<\infty$. Assume also that $b \to 0$ and $(nb)^{-1}\int\phi_K^2(t)|\phi_\varepsilon(\frac{t}{b})|^{-2}dt \to 0$ as $n \to \infty$. Then for any t, as $n \to \infty$,*

$$\hat{m}(t)\xrightarrow{p}m(t), \tag{10}$$

where $\xrightarrow{p}$ represents convergence in probability.

Depending on the tail behavior of the characteristic function ϕ_ε, the following two cases are usually distinguished:

a. *ordinary smooth errors*, when the tails of ϕ_ε are polynomial, that is,

$$c_0|t|^{-\beta} \le |\phi_\varepsilon(t)| \le c_1|t|^{-\beta}, \quad \forall t \in \mathbb{R}$$

for some $c_0>0, c_1>0$ and $\beta>0$.

a. *super-smooth errors*, when the tails are exponential, that is,

$$c_0 \exp\left\{-\gamma|t|^\beta\right\} \le |\phi_\varepsilon(t)| \le c_1 \exp\left\{-\gamma|t|^\beta\right\}, \quad \forall t \in \mathbb{R}$$

for some $c_0>0, c_1>0, \gamma>0$ and $\beta>0$.

The Laplace, gamma and double-exponential densities are examples of ordinary smooth error and the normal, mixture normal and Cauchy densities are examples of super-smooth error. Hence Theorem 1 has the following corollaries.

Corollary 1.1. *Assume that the error measurement is ordinary smooth or super-smooth, the kernel K is bounded and even function with bounded characteristic function on a compact and symmetric support. Assume also that $\tau_F<\infty$ and $b \to 0$ as $n \to \infty$. Then for any t (10) holds.*

Corollary 1.2. *Let the error measurement be normally distributed with zero expectation and standard deviation σ_ε, the kernel K be standard normal density and $b > \sigma_\varepsilon$. If $\tau_F < \infty$ and $b \to 0$ as $n \to \infty$, then for any t (10) holds.*

Corollary 1.3. *Suppose that the error measurement is distributed as laplace distribution with zero expectation and that the kernel K is standard normal density. Then for any t (10) holds.*

The next theorem gives the asymptotic distribution of $\hat{m}(t)$.

Theorem 2. *Assume that $\int \frac{\phi_K(t)}{\phi_\varepsilon(t/b)} dt \to c$ as $n \to \infty$, for some constant $c \geq 0$. Then, under the conditions of Theorem 1, as $n \to \infty$,*

$$\sqrt{n}(\hat{m}(t) - m(t)) \xrightarrow{d} N\left(0, \sigma^2(t)\right),$$

($\xrightarrow{d}$ stands for convergence in distribution). $N(0, \sigma^2(t))$ represent normal random variable with mean zero and variance $\sigma^2(t) = \frac{\zeta}{\bar{F}(t)}$ where $\zeta = Var(\int_t^{\tau_F} (1 - L_1(\omega - Y_1))d\omega)$ and L_1 is defined in (17).

Due to their length and being of a somewhat technical nature, we have postponed the proofs of the results to the Appendix.

3. Bandwidth selection

Bandwidth selection in deconvolution problems has been broadly discussed in many papers. Hesse (1999) carried out a theoretical study of the cross-validation (CV) bandwidth selection procedure. Delaigle and Gijbels (2004a) studied a bootstrap procedure to estimate the optimal bandwidth and showed its consistency. Delaigle and Gijbels (2004b) compared several plug-in bandwidth selectors with the CV bandwidth selector and the bootstrap bandwidth selector. Wang and Wang (2010) generalized the plug-in and the bootstrap bandwidth selection methods to the case of heteroscedastic errors.

Here, we recall the rule of thumb and plug-in bandwidth selection methods. In this paper for the normal error we use the rule of thumb according to theorem 1 and theorem 3 of Fan (1991) and for the laplace error we use the plug-in method to obtain optimum bandwidth values.

3.1. Rule of thumb

As in kernel density estimation with error-free data, the criterion of the bandwidth selection in deconvolution problems is the mean integrated squared error (MISE), defined by

$$MISE\left\{\hat{f}\right\} = E\left[\int \left(\hat{f}(x) - f(x)\right)^2 dx\right].$$

In the case of the homoscedastic normal errors, by the definition of super-smooth distribution, the errors have a super-smooth distribution of order $\beta = 2$ with a positive constant $\gamma = 2/\sigma_\varepsilon^2$. Regarding the error distribution, the kernel function, minimizing the approximated MISE (see Fan 1991, Theorem 1), gives the rule-of-thumb optimum bandwidth value as the following:

$$b_{ROT,N} = \left(\frac{4}{\gamma}\right)^{1/\beta} (\log n)^{-1/\beta} = \sqrt{2}\sigma_\varepsilon (\log n)^{-1/2}.$$

In the case of the homoscedastic Laplacian errors (ordinary smooth), the rule-of-thumb bandwidth becomes

$$b_{ROT,L} = \left(\frac{5\sigma_\varepsilon^4}{n}\right)^{1/9}$$

where σ_ε^2 is the variance of the measurement error. In the R package decon the function bw.dnrd applies the rule of thumb methods to get the bandwidth value.

3.2. Plug-in method

The plug-in bandwidth selection method is simply replaces the unknown density function with the normal density function in minimizing the approximated MISE. Stefanski and Carroll (1990) showed that the asymptotic dominating term of the MISE of the deconvolution kernel density estimator (3) can be estimated by

$$\widehat{MISE}(b) = \frac{1}{2\pi nb} \int \frac{|\phi_K(t)|^2}{|\phi_\varepsilon(t/b)|^2} dt + \frac{b^4}{4} R(f_X'') \int x^2 K(x) dx,$$

where $R(f_X'') = \int (f_X''(x))^2 dx$. Evaluating the $M\hat{I}SE(b)$ involves estimating the unknown quantity $R(f_X'')$. If one assumes X to be normal, then $R(\hat{f}_X'') = 0.375\hat{\sigma}_X^{-5}\pi^{-1/2}$, where $\hat{\sigma}_X = \sqrt{\hat{\sigma}_Y^2 - \sigma_\varepsilon^2}$, $\hat{\sigma}_Y^2$ the sample variance of Y and σ_ε^2 is the variance of the measurement error variable. In the R package decon the function bw.dmise has also been utilized to compute the bandwidth value based on the plug-in method.

4. Simulation study

Simulation exercises were undertaken to study the performances of our estimator and assess the performances of this estimator, comparing it with the empirical estimator based on the contaminated random sample $Y_1, Y_2, \ldots, Y_n$, i.e.

$$m_n(t) = \frac{\sum_{i=1}^n (Y_i - t) I_{(0,Y_i)}(t)}{\sum_{i=1}^n I_{(0,Y_i)}(t)}. \tag{11}$$

(In the sequel, the symbol $I_A(t)$ represents the indicator function of the set A.)

We used the following set up in our simulation study. The unobserved distribution f was assumed to be one of the following.

1. Weibull distribution with shape parameter 2 and scale parameter 1.
2. Lognormal distribution with location parameter 0 and scale parameter 1/2.
3. Gamma distribution with shape parameters 3 and scale parameter $\sqrt{3}$.
4. Uniform distribution on (0, 1).

For the measurement error distribution, we considered the following two cases distribution.

1. Normal distribution with zero mean and and standard deviation σ_ε.
2. Laplace distribution with zero mean and standard deviation σ_ε.

The parameter σ_ε was chosen such that we set a specific percent of contamination based on the signal-to-noise ratio $E(X)/\sigma_\varepsilon$, where $X = \ln(T)$. In particular, we let the values of the ratio to be 5 and 2 corresponding to 20% and 50% contamination, respectively.

In kernel density estimation problem under the error-free data, it is well-known that the choice of the kernel function K does not have a big effect on the accuracy of the estimator. However, in deconvolution kernel estimation under contaminated data, the particular structure of the deconvolution estimators require to be satisfied in some circumstances. For normal errors, we used second-order kernel whose characteristic function has a compact and symmetric support (Fan 1992; Delaigle and Gijbels 2004a),

$$K(x) = \frac{48\cos x}{\pi x^4}\left(1 - \frac{15}{x^2}\right) - \frac{144\sin x}{\pi x^5}\left(2 - \frac{5}{x^2}\right) \tag{12}$$

with characteristic function

$$\phi_K(t) = \left(1 - t^2\right)^3 I_{[-1,1]}(t).$$

For Laplacian errors, We considered the standard normal kernel function. In each case, we ran 1,000 simulation trials of different sample sizes n where $n = 50$, $n = 200$ and $n = 500$. We used the R functions bw.dnrd for normal error and bw.dmise for lapalce error (in the decon package of R) to calculate the bandwidth values. For each sample, the values of the estimator (7) and the empirical estimator (11) were obtained for some values of t. This is then repeated for 1,000 samples and their corresponding bias (i.e., the average of the estimates minus $m(t)$), mean-squared errors (i.e., the average of the sum of the squared difference between the estimate and $m(t)$) calculated.

Tables 1–8 summarize the results of the 1,000 simulation trials for the different experimental set ups. In the tables in some cases when t has large value, the NaN value was recorded for $m_n(t)$. This is because for some replications we had not any value greater than given t and the denominator of (11) was vanished. As is evident from these tables, the performance of our proposed estimator is generally better than that of the empirical estimator in terms of the both of empirical mean square error and bias. From a practical point of view, the main objective of the simulation is to show that the performance of our proposed estimator is comparable or exceeds that of the empirical estimator based on contaminated data. From the results obtained, we may conclude that this is indeed the case here. The results revealed that with large sample size, the improvement of the proposed estimator over empirical estimator would be further enhanced.

4.1. Real life data analysis

In this section, we use a couple of real data set coming from medical and epidemiologic studies to examine the behavior of our proposed estimator.

Table 1. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have Weibull distribution with shape 2 and scale 1 with normal contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.25	0.5	1.00	2.00
		Normal errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.005	0.003	0.004	0.005	0.028
		(0.041)	(0.003)	(0.007)	(0.014)	(−0.030)
	$m_n(t)$	12.10	0.273	0.019	0.008	NaN
		(3.478)	(0.517)	(0.126)	(−0.077)	(NaN)
n = 200	$\hat{m}(t)$	0.000	0.001	0.001	0.001	0.009
		(0.001)	(0.004)	(0.006)	(0.012)	(0.012)
	$m_n(t)$	3.517	0.271	0.016	0.006	NaN
		(1.875)	(0.520)	(0.125)	(−0.077)	(NaN)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.000	0.003
		(0.002)	(0.003)	(0.004)	(0.009)	(0.016)
	$m_n(t)$	3.522	0.270	0.016	0.006	0.015
		(1.876)	(0.519)	(0.125)	(−0.078)	(−0.119)
		Normal errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.008	0.005	0.006	0.013	0.029
		(0.065)	(0.031)	(0.051)	(0.092)	(0.106)
	$m_n(t)$	12.12	0.275	0.022	0.006	NaN
		(3.481)	(0.520)	(0.136)	(−0.061)	(NaN)
n = 200	$\hat{m}(t)$	0.001	0.001	0.002	0.006	0.016
		(0.019)	(0.024)	(0.039)	(0.073)	(0.104)
	$m_n(t)$	3.526	0.273	0.019	0.004	NaN
		(1.877)	(0.521)	(0.136)	(−0.060)	(NaN)
n = 500	$\hat{m}(t)$	0.000	0.000	0.001	0.004	0.012
		(0.015)	(0.021)	(0.031)	(0.061)	(0.098)
	$m_n(t)$	3.518	0.273	0.018	0.004	0.010
		(1.875)	(0.522)	(0.134)	(−0.061)	(−0.096)

Example 1. *The first data set consists of the Framingham Heart Study (*Carroll et al. 2006*). This study consists of a series of exams taken two years apart. Following Carroll et al. (2006), we use systolic blood pressure (SBP) measurements of 1,615 men aged 31−65, from Exam two and Exam three and treat the* $Y = \log(SBP-50)$ *values of each individual j for the two exams* $(Y_{j,1}, Y_{j,2})$ *as repeated measures of the long-term average SBP, which is denoted by* X_j*:*

$$Y_{j,1} = X_j + \varepsilon_{j,1},$$
$$Y_{j,2} = X_j + \varepsilon_{j,2},$$

for individuals $j = 1, \ldots, n$*. Furthermore, we illustrate the analysis using the mean of the two exams,* $Y_j = (Y_{j,1} + Y_{j,2})/2$*, so that the model in our case is*

$$Y_j = X_j + \varepsilon_j$$

where $\varepsilon_j = (\varepsilon_{j,1} + \varepsilon_{j,2})/2$ *and we are interesting in the estimation of MRL from the data* Y_j*,* $j = 1, \ldots, 1615$*. Following Carroll et al. (2006), we obtain* $\hat{\sigma}^2_\varepsilon = 0.0125$*. Assuming* $\varepsilon \sim N(0, 0.125)$*, Figure 1 depicts the two estimators (7) and (11) of the MRL function for these data.*

Example 2. *The second data set consists of the NHANES-I Epidemiologic Study Cohort data set (Jones et al. 1987) is a cohort study originally consisting of 8,596 women who were interviewed about their nutrition habits and later examined for evidence of cancer. Carroll et al. (2006) restricted attention to a subcohort of 3,145 women aged 25-50 who*

Table 2. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have log-normal distribution with shape 0 and scale 1/2 with normal contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.25	0.5	1.00	2.00
		Normal errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.007	0.006	0.007	0.011	0.071
		(0.006)	(0.001)	(0.002)	(0.009)	(−0.033)
	$m_n(t)$	3.671	0.256	0.009	0.031	NaN
		(1.914)	(0.501)	(0.074)	(−0.167)	(NaN)
n = 200	$\hat{m}(t)$	0.001	0.001	0.001	0.003	0.020
		(0.004)	(0.004)	(0.005)	(0.009)	(0.006)
	$m_n(t)$	3.664	0.256	0.007	0.028	0.118
		(1.914)	(0.505)	(0.079)	(−0.165)	(−0.339)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.001	0.008
		(0.002)	(0.003)	(0.005)	(0.008)	(0.005)
	$m_n(t)$	3.657	0.255	0.006	0.027	0.115
		(1.912)	(0.505)	(0.079)	(−0.166)	(−0.338)
		Normal errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.008	0.008	0.010	0.018	0.064
		(0.024)	(0.029)	(0.048)	(0.073)	(0.076)
	$m_n(t)$	3.655	0.261	0.011	0.028	NaN
		(1.910)	(0.506)	(0.085)	(−0.155)	(NaN)
n = 200	$\hat{m}(t)$	0.002	0.002	0.003	0.005	0.020
		(0.018)	(0.023)	(0.033)	(0.053)	(0.068)
	$m_n(t)$	3.653	0.257	0.008	0.025	0.109
		(1.911)	(0.506)	(0.084)	(−0.156)	(−0.326)
n = 500	$\hat{m}(t)$	0.001	0.001	0.001	0.003	0.009
		(0.016)	(0.018)	(0.031)	(0.049)	(0.054)
	$m_n(t)$	3.656	0.257	0.007	0.024	0.110
		(1.912)	(0.506)	(0.086)	(−0.155)	(−0.330)

Table 3. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have gamma distribution with shape 3 and scale $\sqrt{3}$ with normal contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.5	1.0	2.5	3.5
		Normal errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.019	0.022	0.021	0.058	0.136
		(0.009)	(0.021)	(0.031)	(0.057)	(0.014)
	$m_n(t)$	2.843	0.031	0.187	0.313	NaN
		(1.684)	(−0.157)	(−0.428)	(−0.555)	(NaN)
n = 200	$\hat{m}(t)$	0.005	0.004	0.005	0.015	0.037
		(0.008)	(0.013)	(0.021)	(0.043)	(0.037)
	$m_n(t)$	2.844	0.026	0.185	0.312	0.329
		(1.686)	(−0.158)	(−0.429)	(−0.558)	(−0.571)
n = 500	$\hat{m}(t)$	0.002	0.002	0.002	0.007	0.016
		(0.010)	(0.013)	(0.017)	(0.040)	(0.040)
	$m_n(t)$	2.850	0.025	0.185	0.310	0.328
		(1.688)	(−0.158)	(−0.429)	(−0.557)	(−0.572)
		Normal errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.033	0.039	0.064	0.144	0.184
		(0.097)	(0.125)	(0.196)	(0.310)	(0.329)
	$m_n(t)$	2.856	0.029	0.173	0.293	NaN
		(1.687)	(−0.151)	(−0.410)	(−0.536)	(NaN)
n = 200	$\hat{m}(t)$	0.010	0.013	0.024	0.079	0.107
		(0.063)	(0.089)	(0.137)	(0.256)	(0.286)
	$m_n(t)$	2.841	0.024	0.172	0.285	0.299
		(1.685)	(−0.151)	(−0.414)	(−0.533)	(−0.545)
n = 500	$\hat{m}(t)$	0.006	0.008	0.017	0.054	0.077
		(0.058)	(0.077)	(0.121)	(0.219)	(0.256)
	$m_n(t)$	2.848	0.023	0.170	0.285	0.298
		(1.687)	(−0.151)	(−0.412)	(−0.533)	(−0.545)

Table 4. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have uniform distribution with shape 0 and scale 1 with normal contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.2	0.4	0.6	0.8
		Normal errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.001	0.001	0.000	0.002	0.002
		(−0.004)	(0.008)	(0.021)	(0.043)	(0.050)
	$m_n(t)$	9.922	0.397	0.064	0.015	0.010
		(3.147)	(0.625)	(0.247)	(0.115)	(0.092)
n = 200	$\hat{m}(t)$	0.000	0.000	0.000	0.001	0.002
		(0.001)	(0.006)	(0.014)	(0.034)	(0.045)
	$m_n(t)$	9.995	0.390	0.061	0.014	0.008
		(3.160)	(0.623)	(0.246)	(0.116)	(0.090)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.000	0.001
		(0.000)	(0.004)	(0.012)	(0.030)	(0.042)
	$m_n(t)$	9.959	0.386	0.061	0.013	0.008
		(3.155)	(0.621)	(0.247)	(0.116)	(0.091)
		Normal errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.001	0.004	0.011	0.012	0.005
		(0.001)	(0.060)	(0.104)	(0.111)	(0.077)
	$m_n(t)$	10.01	0.474	0.139	0.089	0.089
		(3.160)	(0.682)	(0.364)	(0.288)	(0.284)
n = 200	$\hat{m}(t)$	0.000	0.002	0.007	0.010	0.005
		(0.001)	(0.043)	(0.087)	(0.100)	(0.073)
	$m_n(t)$	10.00	0.464	0.134	0.081	0.084
		(3.162)	(0.679)	(0.364)	(0.283)	(0.286)
n = 500	$\hat{m}(t)$	0.000	0.001	0.006	0.009	0.005
		(0.000)	(0.037)	(0.079)	(0.094)	(0.071)
	$m_n(t)$	9.992	0.465	0.131	0.080	0.082
		(3.160)	(0.681)	(0.362)	(0.283)	(0.286)

Table 5. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have Weibull distribution with shape 2 and scale 1 with Laplace contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.25	0.5	1.00	2.00
		Laplace errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.004	0.004	0.004	0.005	0.033
		(0.004)	(0.005)	(0.007)	(0.012)	(−0.044)
	$m_n(t)$	3.544	0.276	0.019	0.008	NaN
		(1.880)	(0.520)	(0.128)	(−0.076)	(NaN)
n = 200	$\hat{m}(t)$	0.001	0.001	0.000	0.001	0.009
		(0.004)	(0.004)	(0.005)	(0.009)	(0.003)
	$m_n(t)$	3.533	0.272	0.017	0.006	NaN
		(1.879)	(0.520)	(0.127)	(−0.077)	(NaN)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.000	0.005
		(0.001)	(0.002)	(0.004)	(0.007)	(0.159)
	$m_n(t)$	3.520	0.269	0.016	0.006	0.005
		(1.875)	(0.519)	(0.126)	(−0.078)	(0.074)
		Laplace errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.004	0.004	0.005	0.009	0.240
		(0.014)	(0.017)	(0.029)	(0.054)	(0.009)
	$m_n(t)$	3.534	0.278	0.022	0.006	NaN
		(1.877)	(0.522)	(0.136)	(−0.060)	(NaN)
n = 200	$\hat{m}(t)$	0.001	0.001	0.001	0.002	0.019
		(0.009)	(0.010)	(0.018)	(0.035)	(0.044)
	$m_n(t)$	3.527	0.273	0.019	0.004	NaN
		(1.877)	(0.521)	(0.134)	(−0.061)	(NaN)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.001	0.009
		(0.006)	(0.008)	(0.012)	(0.024)	(0.044)
	$m_n(t)$	3.523	0.274	0.018	0.004	0.009
		(1.876)	(0.523)	(0.133)	(−0.062)	(−0.093)

Table 6. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have log-normal distribution with shape 0 and scale 1/2 with Laplace contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.25	0.5	1.00	2.00
		Laplace errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.007	0.008	0.007	0.012	0.089
		(0.004)	(0.005)	(0.006)	(0.014)	(−0.007)
	$m_n(t)$	3.669	0.261	0.010	0.030	NaN
		(1.914)	(0.506)	(0.078)	(−0.162)	(NaN)
n = 200	$\hat{m}(t)$	0.001	0.001	0.001	0.003	0.019
		(0.005)	(0.001)	(0.004)	(0.009)	(0.008)
	$m_n(t)$	3.664	0.255	0.007	0.028	0.115
		(1.913)	(0.504)	(0.078)	(−0.165)	(−0.336)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.001	0.008
		(0.002)	(0.000)	(0.003)	(0.005)	(0.012)
	$m_n(t)$	3.658	0.253	0.006	0.028	0.114
		(1.912)	(0.503)	(0.078)	(−0.166)	(−0.337)
		Laplace errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.008	0.008	0.009	0.015	0.088
		(0.010)	(0.011)	(0.024)	(0.029)	(0.010)
	$m_n(t)$	3.658	0.260	0.012	0.028	NaN
		(1.911)	(0.505)	(0.086)	(−0.157)	(NaN)
n = 200	$\hat{m}(t)$	0.002	0.002	0.002	0.003	0.022
		(0.006)	(0.009)	(0.010)	(0.021)	(0.016)
	$m_n(t)$	3.655	0.258	0.008	0.024	0.111
		(1.911)	(0.507)	(0.083)	(−0.154)	(−0.330)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.001	0.009
		(0.004)	(0.007)	(0.009)	(0.015)	(0.016)
	$m_n(t)$	3.658	0.258	0.007	0.024	0.109
		(1.912)	(0.507)	(0.085)	(−0.155)	(−0.329)

Table 7. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have gamma distribution with shape 3 and scale $\sqrt{3}$ with Laplace contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.5	1.0	2.5	3.5
		Laplace errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.021	0.020	0.023	0.058	0.177
		(0.013)	(0.007)	(0.020)	(0.022)	(−0.020)
	$m_n(t)$	2.859	0.031	0.189	0.322	NaN
		(1.688)	(−0.161)	(−0.429)	(−0.563)	(NaN)
n = 200	$\hat{m}(t)$	0.005	0.005	0.005	0.015	0.040
		(0.004)	(0.010)	(0.013)	(0.030)	(0.031)
	$m_n(t)$	2.842	0.026	0.186	0.313	0.328
		(1.685)	(−0.158)	(−0.430)	(−0.558)	(−0.571)
n = 500	$\hat{m}(t)$	0.002	0.002	0.002	0.006	0.016
		(0.006)	(0.008)	(0.015)	(0.028)	(0.029)
	$m_n(t)$	2.847	0.025	0.183	0.311	0.327
		(1.687)	(−0.158)	(−0.428)	(−0.557)	(−0.571)
		Laplace errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.025	0.025	0.030	0.101	0.289
		(0.027)	(0.045)	(0.054)	(0.116)	(0.085)
	$m_n(t)$	2.855	0.028	0.176	0.289	NaN
		(1.687)	(−0.148)	(−0.414)	(−0.532)	(NaN)
n = 200	$\hat{m}(t)$	0.006	0.006	0.008	0.027	0.076
		(0.019)	(0.026)	(0.037)	(0.070)	(0.074)
	$m_n(t)$	2.853	0.024	0.171	0.286	0.298
		(1.688)	(−0.151)	(−0.412)	(−0.534)	(−0.543)
n = 500	$\hat{m}(t)$	0.002	0.002	0.003	0.012	0.034
		(0.012)	(0.017)	(0.026)	(0.052)	(0.062)
	$m_n(t)$	2.845	0.023	0.171	0.285	0.296
		(1.686)	(−0.151)	(−0.413)	(−0.533)	(−0.543)

Table 8. Empirical mean square error and bias(in parenthesis) for estimating $m(t)$ when actual data have uniform distribution with shape 0 and scale 1 with Laplace contamination. $\hat{m}(t)$ is our suggested estimator (7) and $m_n(t)$ is empirical estimator (11).

		t				
		0.05	0.2	0.4	0.6	0.8
		Laplace errors with zero expectation and 20% contamination				
n = 50	$\hat{m}(t)$	0.002	0.001	0.000	0.000	0.001
		(−0.019)	(0.004)	(0.011)	(0.024)	(0.036)
	$m_n(t)$	2.768	0.394	0.065	0.015	0.009
		(1.660)	(0.623)	(0.249)	(0.115)	(0.084)
n = 200	$\hat{m}(t)$	0.000	0.000	0.000	0.000	0.000
		(−0.019)	(0.003)	(0.007)	(0.017)	(0.030)
	$m_n(t)$	2.768	0.390	0.062	0.013	0.007
		(1.662)	(0.623)	(0.247)	(0.113)	(0.084)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.000	0.000
		(−0.019)	(0.002)	(0.006)	(0.013)	(0.026)
	$m_n(t)$	2.763	0.387	0.062	0.013	0.007
		(1.662)	(0.622)	(0.248)	(0.115)	(0.084)
		Laplace errors with zero expectation and 50% contamination				
n = 50	$\hat{m}(t)$	0.002	0.002	0.002	0.002	0.003
		(−0.021)	(0.007)	(0.017)	(0.033)	(0.038)
	$m_n(t)$	2.818	0.473	0.126	0.076	0.084
		(1.674)	(0.681)	(0.346)	(0.260)	(0.268)
n = 200	$\hat{m}(t)$	0.000	0.000	0.000	0.0001	0.001
		(−0.018)	(0.005)	(0.011)	(0.024)	(0.036)
	$m_n(t)$	2.837	0.468	0.124	0.067	0.076
		(1.683)	(0.682)	(0.350)	(0.256)	(0.270)
n = 500	$\hat{m}(t)$	0.000	0.000	0.000	0.000	0.001
		(−0.019)	(0.003)	(0.008)	(0.018)	(0.032)
	$m_n(t)$	2.831	0.464	0.123	0.067	0.075
		(1.682)	(0.680)	(0.349)	(0.257)	(0.272)

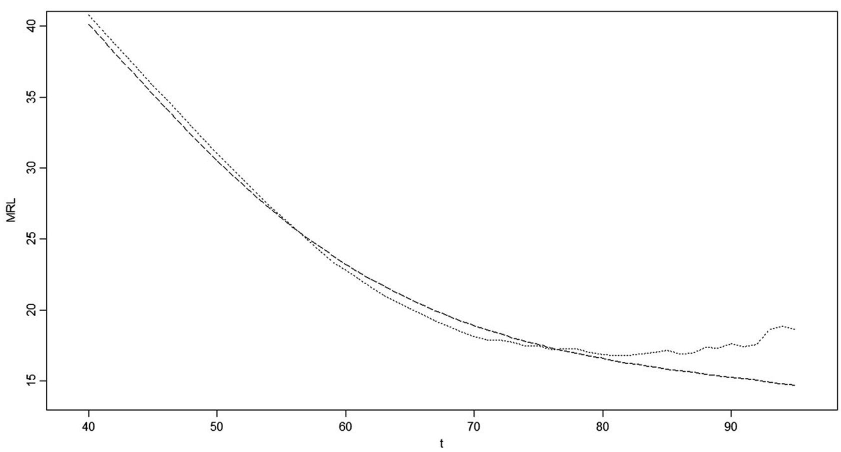

Figure 1. Estimation of the MRL function for the Framingham Heart Study data (dashed line: our proposed estimator (7), dotted line: the estimator (11).).

have no missing data on the variables of interest. In the NHANES data, because of both difficult and expensive to measure long-term diet in a large cohort, instead of observing Y long-term diet, the measured X the amount of intake of saturated fat (in grams) in the previous 24 hours. That the measurement error is large in 24-hour recalls has been documented previously (Beaton et al. 1979; Wu, Whittemore, and Jung 1986). Indeed, there is evidence to support the conclusion that more than half of the variability in the observed

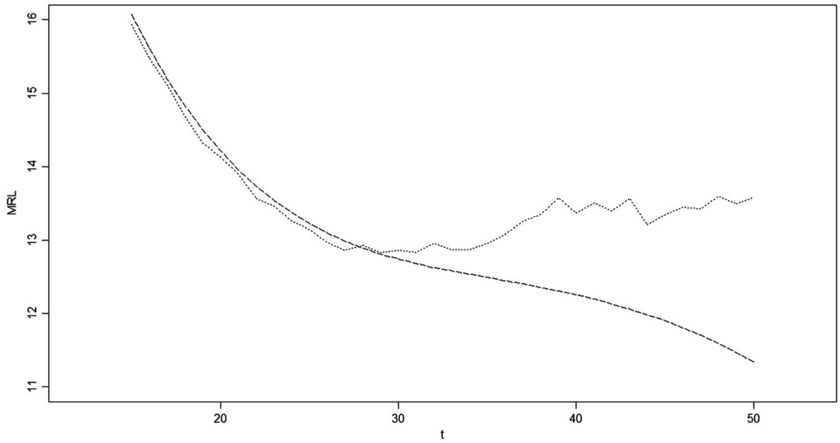

Figure 2. Estimation of the MRL function for the NHANES-I Epidemiologic Study Cohort data. (dashed line: our proposed estimator (7), dotted line: the estimator (11).).

data is due to measurement error. To analyze these data Carroll et al. (2006) used logarithmic transformation log(5 + saturated fat). The transformation was chosen for illustrative purposes and because it makes the observed values nearly normally distributed. By using data from the Continuing Survey of Food Intake by Individuals (CSFII, see Thompson et al. 1992), they estimated that over 75% of the variance of a single 24-hour recall is made up of measurement error. They simply took the estimate as given, namely, that the observed sample variance of Y is 0.233, and for the additive measurement error model, the measurement error variance is estimated as $\hat{\sigma}_{\varepsilon}^{2} = 0.171$. *Here we investigate the behavior of our estimator under 5% normal contamination. Figure 2 plots the estimates.*

5. Conclusion

In this article, we considered the problem of point-wise estimation of $m(t)$ in the presence of measurement errors. Our proposed estimator was based on a multiplicative model and the standard deconvolution kernel density function estimation using the logarithmic transformation of the contaminated data. The estimator under some conditions was shown to be a consistent estimator. In addition, for famous error distributions, we mentioned some conditions under which our proposed estimator is consistent. This estimator was shown to be asymptotically unbiased and consistent, and also to converge in distribution to a normal random variable. An extensive simulation exercise was undertaken to compare between the performance of this estimator and one which uses directly the empirical distributions of the true underlying distributions based on the contaminated data and the results argue favorably for the proposed estimator. Finally, a couple of real data set was used to illustrate the estimator of $m(t)$ under contaminated data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

APPENDIX

Proof of Theorem 1. According to the Eq. (6), the limiting behavior of f and f_X is the same, then throughout the proofs we use f instead of f_X. Under the assumptions of the theorem we want to proof

$$\hat{m}(t) = \frac{\int_t^{\tau_F} \omega \hat{f}(\omega) d\omega}{\int_t^{\tau_F} \hat{f}(\omega) d\omega} - t \xrightarrow{p} \frac{\int_t^{\infty} \omega f(\omega) d\omega}{\int_t^{\infty} f(\omega) d\omega} - t = m(t)$$

One can see that both of the numerator and the denominator of $\hat{m}(t)$ are in the following form

$$\int \rho(\omega) \hat{f}(\omega) d\omega, \tag{13}$$

where, $\rho(\omega) = \omega 1_{[t,\tau_f]}(\omega)$ gives the numerator and $\rho(\omega) = 1_{[t,\tau_f]}(\omega)$ gives the denominator. Thus, it is sufficient to prove that

$$\int \rho(\omega) \hat{f}(\omega) d\omega \xrightarrow{p} \int \rho(\omega) f(\omega) d\omega \tag{14}$$

The result then follows from the fact that if $a_n \rightarrow_p a$ and $b_n \rightarrow_p b$ then $(a_n/b_n) - t \rightarrow_p (a/b) - t$ provided $b \neq 0$(cf. Resnick 1999, p. 174).

First of all, note that the estimator (3) can be rewritten as

$$\begin{aligned}\hat{f}(x) &= \frac{1}{n} \sum_{j=1}^{n} \frac{1}{2\pi} \int e^{it(Y_j - x)} \frac{\phi_K(tb)}{\phi_\varepsilon(t)} dt \\ &= \frac{1}{nb} \sum_{j=1}^{n} L\left(\frac{x - Y_j}{b}\right),\end{aligned}$$

where

$$L(x) = \frac{1}{2\pi} \int e^{-itx} \frac{\phi_K(t)}{\phi_\varepsilon(t/b)} dt.$$

Hence

$$\begin{aligned}\int \rho(\omega) \hat{f}(\omega) d\omega &= \int \rho(\omega) \frac{1}{nb} \sum_{j=1}^{n} L\left(\frac{\omega - Y_j}{b}\right) d\omega \\ &= \frac{1}{n} \sum_{j=1}^{n} \frac{1}{b} \int \rho(\omega) L\left(\frac{\omega - Y_j}{b}\right) d\omega \\ &= \frac{1}{n} \sum_{j=1}^{n} \xi_j,\end{aligned} \tag{15}$$

where

$$\begin{aligned}\xi_j &= \frac{1}{b} \int \rho(\omega) L\left(\frac{\omega - Y_j}{b}\right) d\omega \\ &= \frac{1}{b} \int \rho(\omega) \frac{1}{2\pi} \int e^{-it\left(\frac{\omega - Y_j}{b}\right)} \frac{\phi_K(tb)}{\phi_\varepsilon(t)} dt d\omega \\ &= \frac{1}{2\pi} \int \int \rho(\omega) e^{-it\omega} \phi_K(tb) \frac{e^{itY_j}}{\phi_\varepsilon(t)} dt d\omega\end{aligned} \tag{16}$$

Regarding the Chebyshev inequality, to prove (14), it is enough to show that under the given conditions the bias and variance of (13) converge to zero. Using (15) we have

$$
\begin{aligned}
E\left(\int \rho(\omega)\hat{f}(\omega)d\omega\right) &= E(\xi) \\
&= E\left(\frac{1}{2\pi}\int\int \rho(\omega)e^{-it\omega}\phi_K(tb)\frac{e^{itY}}{\phi_\varepsilon(t)}dtd\omega\right) \\
&= \frac{1}{2\pi}\int\int \rho(\omega)e^{-it\omega}\phi_K(tb)\frac{E(e^{itY})}{\phi_\varepsilon(t)}dtd\omega \\
&= \frac{1}{2\pi}\int\int \rho(\omega)e^{-it\omega}\phi_K(tb)\frac{\phi_Y(t)}{\phi_\varepsilon(t)}dtd\omega \\
&= \frac{1}{2\pi}\int\int \rho(\omega)e^{-it\omega}\phi_K(tb)\phi_X(t)dtd\omega \\
&\to \frac{1}{2\pi}\int\int \rho(\omega)e^{-it\omega}\phi_X(t)dtd\omega \quad \left(\phi_K(tb)\to 1 as b \to 0\right) \\
&= \int \rho(\omega)\left(\frac{1}{2\pi}\int e^{-it\omega}\phi_X(t)dt\right)d\omega \\
&= \int \rho(\omega)f(\omega)d\omega,
\end{aligned}
$$

where, in the line five, we use the fact that under the additive measurement error model and the assumption that X and ε are independence we have $\phi_Y(t) = \phi_X(t).\phi_\varepsilon(t)$.

For the variance we have

$$
\begin{aligned}
Var(\xi) &\le E\left(\xi^2\right) \\
&= \int\int \xi^2 f(x)dxg(\varepsilon)d\varepsilon \\
&= \frac{1}{b^2}\int\int\left(\int \rho(\omega)L\left(\frac{\omega-y}{b}\right)d\omega\right)^2 f(x)dxg(\varepsilon)d\varepsilon \\
&= \frac{1}{b^2}\int\int H_b^2\left(\frac{\omega-y}{b}\right)f(x)dxg(\varepsilon)d\varepsilon,
\end{aligned}
$$

where

$$
\begin{aligned}
H_b(u) &= \int \rho(\omega)L(u)d\omega \\
&= \frac{1}{2\pi}\int\int \rho(\omega)e^{-itu}\frac{\phi_K(t)}{\phi_\varepsilon(t/b)}L(u)d\omega.
\end{aligned}
$$

Also we have

$$
\begin{aligned}
\int H_b^2(u)du &= \left(\tfrac{1}{2\pi}\right)^2\int\left(\int\int \rho(\omega)e^{-itu}\tfrac{\phi_K(t)}{\phi_\varepsilon(t/b)}d\omega\right)^2 du \\
&\le \left(\tfrac{1}{2\pi}\right)^2\int\int\left(\int \rho(\omega)e^{-itu}\tfrac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)^2 d\omega du \\
&= \left(\tfrac{1}{2\pi}\right)^2\int \rho^2(\omega)\int\left(\int e^{-itu}\tfrac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)^2 dud\omega \\
&= \frac{1}{2\pi}\int \rho^2(\omega)\int \phi_K^2(t)|\phi_\varepsilon(t/b)|^{-2}dtd\omega \\
&= \frac{1}{2\pi}\left(\int \rho^2(\omega)d\omega\right)\left(\int \phi_K^2(t)|\phi_\varepsilon(t/b)|^{-2}dt,\right)
\end{aligned}
$$

where, in line four we use the Parseval theorem and the square integrability of $|\phi_K/\phi_\varepsilon(./b)|$. Thus

$$
\begin{aligned}
Var(\xi) &\leq \frac{1}{b^2}\int\int H_b^2\left(\frac{\omega - y}{b}\right) f(x)dxg(\varepsilon)d\varepsilon \\
&= \frac{1}{b^2}\int\int H_b^2\left(\frac{\omega - x - \varepsilon}{b}\right) f(x)dxg(\varepsilon)d\varepsilon \\
&= \frac{1}{b}\int\int H_b^2(-u) f(x)dxg(\varepsilon)d\varepsilon \\
&= \frac{1}{b}\left(\int A(b;\omega-\varepsilon)g(\varepsilon)d\varepsilon\right)\left(\int H_b^2(-u)du\right) \\
&\leq \frac{1}{b}B_f\int H_b^2(u)du, \left(\int H_b^2(-u)du = \int H_b^2(u)du\right) \\
&\leq \frac{1}{b}B_f\frac{1}{2\pi}\left(\int \rho^2(\omega)d\omega\right)\left(\int \phi_K^2(t)|\phi_\varepsilon(t/b)|^{-2}dt,\right)
\end{aligned}
$$

where

$$
A(b;a) := \frac{\int H_b^2(u)f(a+bu)du}{\int H_b^2(u)du},
$$

which is bounded by $B_f = Sup_x f(x)$. Hence

$$
\begin{aligned}
Var\left(\int \rho(\omega)\hat{f}(\omega)d\omega\right) &= \frac{1}{n}Var(\xi) \\
&\leq B_f\frac{1}{2\pi}\left(\int \rho^2(\omega)d\omega\right)\left(\frac{1}{nb}\int \phi_K^2(t)|\phi_\varepsilon(t/b)|^{-2}dt\right)
\end{aligned}
$$

Under the assumption that $\tau_F<\infty$, the integral $\int \rho^2(\omega)d\omega$ is finite. Then, the assumption (iii) implies that $Var(\int \rho(\omega)\hat{f}(\omega)d\omega) \to 0$ which completes the proof.

Proof of corollary 1. Under the assumptions, f is continuous and bounded, $0<|\phi_\varepsilon(t)|$ and the kernel K is bounded and even. Let $[-a,a]$ be the support of $\phi_K(t)$ and L be the upper bound of $|\phi_K(t)|$. We have

$$
\begin{aligned}
\underset{t\in\mathbb{R}}{Sup}\,|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)| &= \underset{-a\leq t\leq a}{Sup}\,|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)| \\
&= \left(\underset{-a\leq t\leq a}{Sup}\,|\phi_K(t)|\right)\left(\underset{-a\leq t\leq a}{Sup}\,|\phi_\varepsilon^{-1}\left(\frac{t}{b}\right)|\right) \\
&= L\times\left(\underset{-a\leq t\leq a}{Sup}\,|\phi_\varepsilon^{-1}\left(\frac{t}{b}\right)|\right) \\
&\leq \begin{cases} L\times\left(Sup_{-a\leq t\leq a}\, c_0^{-1}|\frac{t}{b}|^\beta\right), & \textit{for ordinary smooth error} \\ L\times\left(Sup_{-a\leq t\leq a}\, c_0^{-1}\exp\left\{\gamma|\frac{t}{b}|^\beta\right\}\right), & \textit{for supersmooth error} \end{cases} \\
&= \begin{cases} L\times\left(c_0^{-1}\left(\frac{a}{b}\right)^\beta\right) \\ L\times\left(c_0^{-1}\exp\left\{\gamma\left(\frac{a}{b}\right)^\beta\right\}\right) \end{cases} \\
&< \infty.
\end{aligned}
$$

Also, for the ordinary smooth error, we have

$$
\begin{aligned}
\int |\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|dt &= \int_{-a}^{a} |\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|dt \\
&< \left(\underset{-a\le t\le a}{Sup} |\phi_K(t)| \right) \left(\int_{-a}^{a} |\phi_\varepsilon^{-1}\left(\frac{t}{b}\right)|dt \right) \\
&\le L \times \left(\int_{-a}^{a} c_0^{-1} |\frac{t}{b}|^\beta dt \right) \\
&= L \times \left(2c_0^{-1} \int_0^a \left(\frac{t}{b}\right)^\beta dt \right) \\
&= L \times \left(2c_0^{-1} \frac{a}{\beta+1} \left(\frac{a}{b}\right)^\beta \right) \\
&< \infty,
\end{aligned}
$$

and, for supersmooth error

$$
\begin{aligned}
\int |\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|dt &= \int_{-a}^{a} |\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|dt \\
&\le \left(\underset{-a\le t\le a}{Sup} |\phi_K(t)| \right) \left(\int_{-a}^{a} |\phi_\varepsilon^{-1}\left(\frac{t}{b}\right)|dt \right) \\
&\le L \times \left(\int_{-a}^{a} c_0^{-1} \exp\left\{ \gamma |\frac{t}{b}|^\beta \right\} dt \right) \\
&= L \times \left(2c_0^{-1} \int_0^a \exp\left\{ \gamma \left(\frac{t}{b}\right)^\beta \right\} dt \right) \\
&< L \times \left(2c_0^{-1} \int_0^a t^{\beta-1} \exp\left\{ \gamma \left(\frac{t}{b}\right)^\beta \right\} dt \right) \\
&= 2Lc_0^{-1} \frac{b^\beta}{\gamma\beta} \exp\left\{ \gamma \left(\frac{t}{b}\right)^\beta \right\} \Big|_0^a \\
&= 2Lc_0^{-1} \frac{b^\beta}{\gamma\beta} \left(\exp\left\{ \gamma \left(\frac{a}{b}\right)^\beta \right\} - 1 \right) \\
&< \infty.
\end{aligned}
$$

In addition, for ordinary smooth error we have

$$
\begin{aligned}
\frac{1}{nb}\int \phi_K^2(t)|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt &= \frac{1}{nb}\int_{-a}^{a} \phi_K^2(t)|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt \\
&\le \frac{1}{nb}\left(\underset{-a\le t\le a}{Sup}|\phi_K(t)|\right)^2\left(\int_{-a}^{a}|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt\right) \\
&\le \frac{1}{nb}\times L^2\times\left(\int_{-a}^{a} c_0^{-2}|\frac{t}{b}|^{2\beta}dt\right) \\
&= \frac{L^2}{nb}\times\left(2c_0^{-2}\int_0^a\left(\frac{t}{b}\right)^{2\beta}dt\right) \\
&= \frac{L^2}{nb}\times\left(2c_0^{-2}\frac{a}{2\beta+1}\left(\frac{a}{b}\right)^{2\beta}\right) \\
&\to 0, \qquad as\, n\to\infty,
\end{aligned}
$$

and for the supersmooth error

$$
\begin{aligned}
\frac{1}{nb}\int \phi_K^2(t)|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt &= \frac{1}{nb}\int_{-a}^{a} \phi_K^2(t)|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt \\
&\le \frac{1}{nb}\left(\underset{-a\le t\le a}{Sup}|\phi_K(t)|\right)^2\left(\int_{-a}^{a}|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt\right) \\
&\le \frac{1}{nb}\times L^2\times\left(\int_{-a}^{a} c_0^{-2}\exp\left\{2\gamma|\frac{t}{b}|^{\beta}\right\}dt\right) \\
&= \frac{L^2}{nb}\times\left(2c_0^{-2}\int_0^a\exp\left\{2\gamma\left(\frac{t}{b}\right)^{\beta}\right\}dt\right) \\
&< \frac{L^2}{nb}\times\left(2c_0^{-2}\int_0^a t^{\beta-1}\exp\left\{2\gamma\left(\frac{t}{b}\right)^{\beta}\right\}dt\right) \\
&= \frac{c_0^{-1}L^2}{nb}\frac{b^\beta}{\gamma\beta}\exp\left\{\gamma\left(\frac{t}{b}\right)^{\beta}\right\}\Bigg|_0^a \\
&= \frac{c_0^{-1}L^2}{nb}\frac{b^\beta}{\gamma\beta}\left(\exp\left\{\gamma\left(\frac{a}{b}\right)^{\beta}\right\}-1\right) \\
&\to 0, \qquad as\, n\to\infty.
\end{aligned}
$$

Thus, all the conditions of Theorem 1 satisfy and the result holds.

Proof of corollary 2. Here, also it is enough to show that the conditions of Theorem 1 satisfy. Again, under the assumptions, f is continuous and bounded and $\phi_\varepsilon(t) = e^{-\frac{1}{2}\sigma_\varepsilon^2 t^2}>0$, for all $t\in\mathbb{R}$. Here the kernel is the standard normal density $K(x) = 1/\sqrt{2\pi}e^{-x^2/2}$ which is bounded and even. For this kernel we have

$$\underset{t\in\mathbb{R}}{Sup}\left|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)\right| = \underset{t\in\mathbb{R}}{Sup}\left|e^{-\frac{t^2}{2}}/e^{-\frac{1}{2}\sigma_\varepsilon^2\frac{t^2}{b^2}}\right|$$
$$= \underset{t\in\mathbb{R}}{Sup}\left|e^{-\frac{t^2}{2}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}\right|.$$

For $b>\sigma_\varepsilon$, $(1-\sigma_\varepsilon^2/b^2)>0$ and therefore the supreme equals to 1. Also

$$\begin{aligned}\int\left|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)\right|dt &= \int\left|e^{-\frac{t^2}{2}}/e^{-\frac{1}{2}\sigma_\varepsilon^2\frac{t^2}{b^2}}\right|dt\\ &= \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}dt\\ &= 2\int_0^{+\infty} e^{-\frac{t^2}{2}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}dt\\ &< 2\int_0^{+\infty} te^{-\frac{t^2}{2}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}dt\\ &= -2\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)^{-1} e^{-\frac{t^2}{2}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}\Bigg|_0^{+\infty}\\ &= 2\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)^{-1}(b>\sigma_\varepsilon)\\ &<\infty.\end{aligned}$$

For $b>\sigma_\varepsilon$, the standard normal kernel satisfies in assumption (iii). Finally,

$$\begin{aligned}\frac{1}{nb}\int\phi_K^2(t)\left|\phi_\varepsilon\left(\frac{t}{b}\right)\right|^{-2}dt &= \frac{1}{nb}\int e^{-t^2}\left|e^{-\frac{1}{2}\sigma_\varepsilon^2\frac{t^2}{b^2}}\right|^{-2}dt\\ &= \frac{1}{nb}\int_{-\infty}^{+\infty} e^{-t^2}e^{\sigma_\varepsilon^2\frac{t^2}{b^2}}dt\\ &= \frac{2}{nb}\int_0^{+\infty} e^{-t^2\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}dt\\ &< \frac{2}{nb}\int_0^{+\infty} te^{-t^2\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}dt\\ &= -\frac{1}{nb}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)^{-1}e^{-t^2\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)}\Bigg|_0^{+\infty}\\ &= \frac{1}{nb}\left(1-\frac{\sigma_\varepsilon^2}{b^2}\right)^{-1} \qquad (b>\sigma_\varepsilon)\\ &\to 0, \qquad as\, n\to\infty.\end{aligned}$$

Thus,the standard normal kernel with $b>\sigma_\varepsilon$ meets the last condition of Theorem 1.

Proof of corollary 3. Similar to the above proofs, we have

$E(e^{it\varepsilon}) = (1+\beta^2 t^2)^{-1} > 0$, for all $t \in \mathbb{R}$. For standard normal kernel we get

$$\underset{t\in\mathbb{R}}{Sup}\left|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)\right| = \max_{t\in\mathbb{R}}\left|e^{-\frac{1}{2}t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)\right|$$
$$= \begin{cases} e^{-\frac{1}{2}t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)\Big|_{t=\pm\sqrt{2-\left(b^2/\beta^2\right)}}, & if 2\beta^2 > b \\ e^{-\frac{1}{2}t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)\Big|_{t=0}, & if 2\beta^2 \le b \end{cases}$$
$$< \infty.$$

Also

$$\int|\phi_K(t)/\phi_\varepsilon\left(\frac{t}{b}\right)|dt = \int|e^{-\frac{1}{2}t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)|dt$$
$$= 2\int_0^\infty e^{-\frac{1}{2}t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)dt$$
$$= \sqrt{2\pi}\left(1+\frac{\beta^2}{b^2}\right)$$
$$< \infty,$$

and

$$\frac{1}{nb}\int\phi_K^2(t)|\phi_\varepsilon\left(\frac{t}{b}\right)|^{-2}dt = \frac{1}{nb}\int e^{-t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)^2 dt$$
$$= \frac{2}{nb}\int_0^\infty e^{-t^2}\left(1+\beta^2\left(\frac{t}{b}\right)^2\right)^2 dt$$
$$= \sqrt{\pi}\frac{4b^4+4b^2\beta^2+3\beta^4}{4nb^5}$$
$$\to 0, \qquad as n \to \infty.$$

Proof of Theorem 2. First note that the estimator (3) can be written as

$$\hat{f}(x) = \frac{1}{nb}\sum_{j=1}^n L\left(\frac{x-Y_j}{b}\right),$$

where

$$L(z) = \frac{1}{2\pi}\int e^{-itz}\frac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt.$$

Thus

$$\hat{F}(\omega) = \int_{\upsilon_F}^{\omega}\hat{f}(x)dx$$
$$= \frac{1}{nb}\sum_{j=1}^n\int_{\upsilon_F}^{\omega}L\left(\frac{x-Y_j}{b}\right)dx$$
$$= \frac{1}{n}\sum_{j=1}^n\int_{\upsilon_F}^{\frac{\omega-Y_j}{b}}L(v)dv,$$

where, $\upsilon_F := \sup\{x : F(x) = 0\}$. Hall and Lahiri (2008) showed that $\int_{-\infty}^{u}L(v)dv = L_1(bu)$, where

$$L_1(z) = \frac{1}{2} + \frac{1}{2\pi}\int \frac{\sin(tz)}{t}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dt. \tag{17}$$

Hence

$$\begin{aligned}\hat{F}(\omega) &= \frac{1}{n}\sum_{j=1}^{n} L_1(\omega - Y_j)\\ &= \frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{2} + \frac{1}{2\pi}\int \frac{\sin(t(\omega - Y_j))}{t}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dt\right)\\ &= \frac{1}{2} + \frac{1}{2n\pi}\sum_{j=1}^{n}\int \frac{\sin(t(\omega - Y_j))}{t}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dt.\end{aligned} \tag{18}$$

At the proof of Theorem, 1 we obtained that

$$1-\hat{F}(t) = \int_t^{\tau_F}\hat{f}(\omega)d\omega \xrightarrow{p} \int_t^{\infty} f(\omega)d\omega = 1-F(t).$$

Thus, if we show that the numerator of $\hat{m}(t)$ convergences in distribution to a normal random variable, the result will follow from the Slutsky theorem.

Using (18) the numerator of $\hat{m}(t)$ can be given by

$$\begin{aligned}\int_t^{\tau_F}\left(1 - \hat{F}(\omega)\right)d\omega &= \int_t^{\tau_F}\left(1 - \frac{1}{n}\sum_{j=1}^{n} L_1(\omega - Y_j)\right)d\omega\\ &= \frac{1}{n}\sum_{j=1}^{n}\int_t^{\tau_F}(1 - L_1(\omega - Y_j))d\omega\\ &= \frac{1}{n}\sum_{j=1}^{n}\mathcal{W}_j,\end{aligned}$$

where $\mathcal{W}_j := \int_t^{\tau_F}(1 - L_1(\omega - Y_j))d\omega$. To apply the central limit theorem, we first show that the Lyapounov condition (cf. Shao 2003, p. 69)

$$\underset{n\to\infty}{Lim}\frac{1}{\sigma_n^{2+\delta}}\sum_{j=1}^{n} E|\mathcal{W}_j - E\mathcal{W}_j|^{2+\delta} = 0, \quad for\ some\ \delta>0 \tag{19}$$

holds, where $\sigma_n^2 = Var(\sum_{j=1}^{n}\mathcal{W}_j) = nVar(\mathcal{W}_1) = n\zeta$ with $\zeta := Var(\mathcal{W}_1) = Var(\int_t^{\tau_F}(1 - L_1(\omega - Y_1))d\omega)$. We have

$$\begin{aligned}\mathcal{W}_1^2 &= \left(\int_t^{\tau_F}(1 - L_1(\omega - Y_1))d\omega\right)^2\\ &= \left(\int_t^{\tau_F}\left(1 - \int_{\upsilon_F}^{\omega} L\left(\tfrac{x-Y_1}{b}\right)dx\right)d\omega\right)^2\\ &= \left(\int_t^{\tau_F}\left(1 - \tfrac{b}{2\pi}\int_{\upsilon_F}^{\omega}\int_{-\infty}^{+\infty} e^{-it(x-Y_1)}\tfrac{\phi_K(tb)}{\phi_\varepsilon(t)}dtdx\right)d\omega\right)^2\\ &\le \left(\int_t^{\tau_F}\left(1 + \tfrac{b}{2\pi}(\omega - \upsilon_F)\int_{-\infty}^{+\infty}\tfrac{\phi_K(tb)}{\phi_\varepsilon(t)}dt\right)d\omega\right)^2\\ &= \left((\tau_F - t) + \tfrac{b}{2\pi}\left(\int_{-\infty}^{+\infty}\tfrac{\phi_K(tb)}{\phi_\varepsilon(t)}dt\right)\underbrace{\int_t^{\tau_F}(\omega - \upsilon_F)d\omega}_{A}\right)^2\\ &= \left((\tau_F - t) + \tfrac{A}{2\pi}\left(\int_{-\infty}^{+\infty}\tfrac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)\right)^2\\ &= \left((\tau_F - t) + \tfrac{Ac}{2\pi}(1 + o(1))\right)^2\end{aligned}$$

where the last line follows from assumption (i). Then

$$\zeta = Var(\mathcal{W}_1) \le E(\mathcal{W}_1^2) \le \left((\tau_F - t) + \frac{Ac}{2\pi}(1+o(1))\right)^2 < \infty.$$

On the other hand,

$$\begin{aligned}
0 &\le |W_j - EW_j| \\
&= |\textstyle\int_t^{\tau_F} (1 - L_1(\omega - Y_j))d\omega - E\int_t^{\tau_F} (1 - L_1(\omega - Y_j))d\omega| \\
&= |\textstyle\int_t^{\tau_F} (EL_1(\omega - Y_j) - L_1(\omega - Y_j))d\omega| \\
&\le \textstyle\int_t^{\tau_F} |EL_1(\omega - Y_j) - L_1(\omega - Y_j)|d\omega \\
&= \textstyle\int_t^{\tau_F} |E\int_{\upsilon_F}^{\omega} L\left(\frac{x-Y_j}{b}\right)dx - \int_{\upsilon_F}^{\omega} L\left(\frac{x-Y_j}{b}\right)dx|d\omega \\
&= \textstyle\int_t^{\tau_F} |\frac{b}{2\pi}E\displaystyle\int_{\upsilon_F}^{\omega}\int_{-\infty}^{+\infty} e^{-it(x-Y_j)}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dtdx - \frac{b}{2\pi}\int_{\upsilon_F}^{\omega}\int_{-\infty}^{+\infty} e^{-it(x-Y_j)}\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dtdx|d\omega \\
&= \textstyle\int_t^{\tau_F} |\frac{b}{2\pi}\displaystyle\int_{\upsilon_F}^{\omega}\int_{-\infty}^{+\infty} (Ee^{-it(x-Y_j)} - e^{-it(x-Y_j)})\frac{\phi_K(tb)}{\phi_\varepsilon(t)}dtdx|d\omega \\
&\le \textstyle\int_t^{\tau_F} |\frac{b}{\pi}\displaystyle\int_{\upsilon_F}^{\omega}\int_{-\infty}^{+\infty} \frac{\phi_K(tb)}{\phi_\varepsilon(t)}dtdx|d\omega \\
&= \frac{1}{\pi}\left(\int_{-\infty}^{+\infty}\frac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)\underbrace{\left(\int_t^{\tau_F}(\omega - \upsilon_F)d\omega\right)}_{\text{A}} \\
&= \frac{A}{\pi}\left(\int_{-\infty}^{+\infty}\frac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)
\end{aligned}$$

Hence (19) follows from

$$\begin{aligned}
\frac{1}{\sigma_n^{2+\delta}}\sum_{j=1}^{n}E|W_j - EW_j|^{2+\delta} &\le \frac{n}{(n\zeta)^{\delta}}\left(\frac{A}{\pi}\right)^{2+\delta}\left(\int_{-\infty}^{+\infty}\frac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)^{2+\delta} \\
&= \frac{1}{n\zeta^2}\left(\frac{A}{\pi}\right)^4\left(\int_{-\infty}^{+\infty}\frac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)^4, \quad (\delta = 2) \\
&= \left(\tfrac{A}{\pi}\right)^4\left(\tfrac{1}{n^{1/4}\zeta^{1/2}}\textstyle\int_{-\infty}^{+\infty}\tfrac{\phi_K(t)}{\phi_\varepsilon(t/b)}dt\right)^4 \\
&= \left(\tfrac{A}{\pi}\right)^4\left(\tfrac{c}{n^{1/4}\zeta^{1/2}}(1+o(1))\right)^4 \\
&\to 0
\end{aligned}$$

Thus, the Lindeberg central limit theorem (cf. Shao 2003, p. 67) implies that as $n \to \infty$,

$$\frac{1}{\sqrt{n\zeta}}\sum_{j=1}^{n}(\mathcal{W}_j - E\mathcal{W}_j) = \sqrt{\frac{n}{\zeta}}\left(\int_t^{\tau_F}\left(1 - \hat{F}_X(\omega)\right)d\omega - E\mathcal{W}_1\right)\xrightarrow{d}N(0,1).$$

From the proof of Theorem 1 we have

$$\begin{aligned}
E\mathcal{W}_1 &= E\textstyle\int_t^{\tau_F}(1 - L_1(\omega - Y_1))d\omega \\
&= \textstyle\int_t^{\tau_F}(1 - EL_1(\omega - Y_1))d\omega \\
&= \textstyle\int_t^{\tau_F}\left(1 - E\hat{F}(\omega)\right)d\omega \\
&= \textstyle\int_t^{\tau_F}E\left(1 - \hat{F}(\omega)\right)d\omega \\
&\to \textstyle\int_t^{\tau_F}(1 - F(\omega))d\omega,
\end{aligned}$$

as $n \to \infty$. Then

$$\sqrt{\frac{n}{\zeta}}\left(\int_t^{\tau_F}\left(1-\hat{F}(\omega)\right)d\omega-\int_t^{\tau_F}\left(1-F(\omega)\right)d\omega\right)\xrightarrow{d}N(0,1)$$

Now applying the Slutsky theorem we get that as $n\to\infty$

$$\sqrt{n}(\hat{m}(t)-m(t))\xrightarrow{d}N\left(0,\frac{\zeta}{\bar{F}_X(t)}\right),$$

which completes the proof.

References

Abdous, B., and A. Berred. 2005. Mean residual life estimation. *Journal of Statistical Planning and Inference* 132 (1–2):3–19.

Beaton, G. H., J. Milner, P. Corey, V. McGuire, M. Cousins, E. Stewart, M. de Ramos, D. Hewitt, P. V. Grambsch, N. Kassim, et al. 1979. Sources of variation in 24-hour dietary recall data: Implications for nutrition study design and interpretation. *The American Journal of Clinical Nutrition* 32 (12):2546–59.

Buonaccorsi, J. P. 2010. *Measurement error: Models, methods, and applications*. Boca Raton, FL: Chapman and Hall/CRC.

Carroll, R. J., and P. Hall. 1988. Optimal rates of convergence for deconvolving a density. *Journal of American Statistics Association* 83 (404):1184–6.

Carroll, R. J., D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu. 2006. *Measurement error in nonlinear models: A modern perspective*. 2nd ed. London: Chapman and Hall.

Chaubey, Y. P., and P. K. Sen. 1999. On smooth estimation of residual life. *Journal of Statistical Planning and Inference* 75 (2):223–36.

Chaubey, Y. P., and A. Sen. 2008. Smooth estimation of mean residual life under random censoring. *Institute of Mathematical Statistics Collections* 1:35–49.

Delaigle, A., and I. Gijbels. 2004a. Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *Annals of the Institute of Statistical Mathematics* 56 (1):19–47.

Delaigle, A., and I. Gijbels. 2004b. Practical bandwidth selection in deconvolution kernel density estimation. *Computational Statistics and Data Analysis* 45 (2):249–67.

Fakoor, V. 2015. On the nonparametric mean residual life estimator in length-biased sampling. *Communications in Statistics - Theory and Methods* 44 (3):512–9.

Fan, J. 1991. On the optimal rates of convergence for nonparametric deconvolution problems. *The Annals of Statistics* 19 (3):1257–72.

Fan, J. 1992. Deconvolution with supersmooth distributions. *Canadian Journal of Statistics* 20 (2):155–69.

Hall, P., and S. N. Lahiri. 2008. Estimation of distributions, moments and quantiles in deconvolution problems. *The Annals of Statistics* 36 (5):2110–34.

Hall, W. J., and J. A. Wellner. 1979. *Estimation of mean residual life*. Unpublished manuscript.

Hesse, C. H. 1999. Data-driven deconvolution. *Journal of Nonparametric Statistics* 10 (4):343–73.

Jayasinghe, C., and P. Zeephongsekul. 2013. Local polynomial fitting of the mean residual life function. *IEEE Transactions on Reliability* 62 (2):317–28.

Jones, D. Y., A. Schatzkin, S. B. Green, G. Block, L. A. Brinton, R. G. Ziegler, R. Hoover, and P. R. Taylor. 1987. Dietary fat and breast cancer in the national health and nutrition survey I: Epidemiologic follow-up study. *Journal of the National Cancer Institute* 79:465–471.

Lai, C. D., and M. Xie. 2006. *Stochastic ageing and dependence for reliability*. Berlin, Germany: Springer.

Meister, A. 2009. *Deconvolution problems in nonparametric statistics*. Berlin, Germany: Springer.

Resnick, S. I. 1999. *A probability path*. Boston: Birkhauser.

Ruiz, J. M., and A. Guillamón. 1996. Nonparametrie recursive estimator for mean residual life and vitality function under dependence conditions. *Communication in Statistics- Theory and Methods* 25 (9):1997–2011.

Shao, J. 2003. *Mathematical statistics*. New York: Springer.

Stefanski, L. A., and R. J. Carroll. 1990. Deconvoluting kernel density estimators. *Statistics: A Journal of Theoretical and Applied Statistics* 21 (2):169–84.

Thompson, F. E., M. F. Sowers, E. A. Frongillo, and B. J. Parpia. 1992. Sources of fiber and fat in diets of U.S. women aged 19-50: Implications for nutrition education and policy. *American Journal of Public Health* 82 (5):695–718.

Wang, X. F., and B. Wang. 2010. Simultaneous confidence bands and bootstrap bandwidth selection in deconvolution with heteroscedastic error. *Computational Statistics and Data Analysis*, 54: 25–36.

Wu, M. L., A. S. Whittemore, and D. L. Jung. 1986. Errors in reported dietary intakes. I. Short-term recall. *American Journal of Epidemiology* 124 (5):826–35.

Yang, G. L. 1977. Life expectancy under random censorship. *Stochastic Processes and Their Applications* 6 (1):33–9.

Yang, G. L. 1978. Estimation of a biometric function. *The Annals of Statistics* 6 (1):112–6.