



HAL
open science

On A Deconvolution Problem Under Competing Risks

Christophe Chesneau, Isha Dewan, Hassan Doosti

► **To cite this version:**

Christophe Chesneau, Isha Dewan, Hassan Doosti. On A Deconvolution Problem Under Competing Risks. 2015. hal-01150077v1

HAL Id: hal-01150077

<https://hal.science/hal-01150077v1>

Preprint submitted on 8 May 2015 (v1), last revised 29 Sep 2015 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On A Deconvolution Problem Under Competing Risks

CHRISTOPHE CHESNEAU, ISHA DEWAN AND HASSAN DOOSTI

LMNO, University of Caen, France.

Indian Statistical Institute, New Delhi, India.

Mashhad University of Medical Sciences, Mashhad, Iran.

May 7, 2015

Abstract

Bagai and Prakasa Rao (1992) considered a competing risks model with two dependent risks. The two risks are initially independent but dependence arises because of the additive effect of an independent risk on the two initially independent risks. They showed that the ratio of failure rates are identifiable in the nonparametric set up. In this paper, we consider it as a measurement error/deconvolution problem and suggest a nonparametric kernel type estimator for the ratio of two failure rates. The local error properties of the proposed estimator are studied. Simulation studies show the efficacy of the proposed estimator.

Key words and phrases: competing risks, identifiability, deconvolution, failure rates.

1 Introduction

Consider situations where the outcome consists of paired observation with one component being a continuous random variable and the second a discrete random variable. In lifetime studies the continuous random variable could be the lifetime of the component and the discrete random variable identifying the cause of failure of the component. Such data with competing risks occur in several real life situations - survival analysis, life testing, unemployment studies, sociological studies (occurrence of crime) . . . (see David and Moeschberger (1978), Crowder (2001), Kalbfleisch and Prentice (2002) and Pintilie (2006)).

Bagai and Prakasa Rao (1992) considered an additive model based on two dependent competing risks X and Y , where $X = X_0 + X_1$, $Y = X_0 + X_2$ and X_0, X_1, X_2 are independent random variables. For example, an individual may be exposed to risk of death due to heart and kidney failure. But the presence of diabetes may shorten these theoretical lifetimes. Or change in dietary habits and exercise routines could enhance these theoretical lifetimes.

These latent failure times X, Y are not observable. We can only observe the failure time T and the cause of failure δ , where

$$T = \min(X, Y) = \min(X_0 + X_1, X_0 + X_2) = X_0 + \min(X_1, X_2)$$

and

$$\delta = I(X < Y) = I(X_0 + X_1 < X_0 + X_2) = I(X_1 < X_2).$$

Suppose X_0, X_1, X_2 are independent random variables whose characteristic functions are non-vanishing. Then Kotlarski (1966, 1967) and Prakasa Rao (1968) proved that the distributions of X_0, X_1 and X_2 can be determined uniquely from the joint distribution of X, Y among the class with fixed $E(X_0)$. However, X, Y are not observable in the presence of competing risks. Bagai and Prakasa Rao (1992) assumed that X_0, X_1, X_2 are independent exponential random variables with parameters λ_0, λ_1 and λ_2 , respectively. Consider the class with λ_0 fixed. Then, if only failure time T is observable, $\lambda_1 + \lambda_2$ is identifiable. If both failure time and cause of death (T, δ) are observable then both λ_1 and λ_2 are identifiable.

In the nonparametric case, let X_0, X_1, X_2 be independent random variables with distribution functions G, F_1 and F_2 , density functions g, f_1 and f_2 , survival functions \bar{G}, \bar{F}_1 and \bar{F}_2 and failure rate functions r_G, r_{F_1} and r_{F_2} , respectively (note $r_G(x) = g(x)/\bar{G}(x)$.) Then Bagai and Prakasa Rao (1992) showed that when T, δ are observable, then in the class of distributions with X_0 fixed, individual distributions F_1 and F_2 are not identifiable but the ratio of failure rates H defined by

$$H(t) = \frac{r_{F_1}(t)}{r_{F_2}(t)} = \frac{f_1(t)\bar{F}_2(t)}{f_2(t)\bar{F}_1(t)}$$

is identifiable.

In this paper we propose a kernel based nonparametric estimator for H based on competing risks data $(T_1, \delta_1), (T_2, \delta_2), \dots, (T_n, \delta_n)$.

Let us set $Z = \min(X_1, X_2)$. Let the sub-density functions of $(Z, \delta = 1)$ and $(Z, \delta = 0)$ be given by $f_Z(t, 1)$ and $f_Z(t, 0)$, respectively. Since X_1, X_2 are independent, one can prove that

$$f_Z(t, 1) = f_1(t)\bar{F}_2(t), \quad f_Z(t, 0) = f_2(t)\bar{F}_1(t),$$

Hence we can express H as

$$H(t) = \frac{f_Z(t, 1)}{f_Z(t, 0)}.$$

Let us now discuss the estimations of $f_Z(t, 1)$ and $f_Z(t, 0)$. Remark that the failure time T can be written as

$$T = Z + X_0,$$

where X_0 can be viewed as an error term. Since X_0 and Z are independent, this can be looked upon as additive measurement error model studied first by Carroll and Hall (1988), Stefanski and Carroll (1990) and Zhang (1990). One could also refer to Fan (1991, 1992), Carroll *et al.* (1995), Efromovich (1997), Delaigle and Gijbels (2004 a,b), Meister (2004), Van Es and Uh (2005), Delaigle and Meister (2007, 2008), Hall and Lahiri (2008) and Wang *et al.* (2010). A comprehensive discussion of nonparametric deconvolution methods can be found in Meister (2009).

Extending the deconvolution problem to competing risks, we develop a ratio-type estimator \hat{H} for H . It is based on Fourier-kernel methods and a thresholding technique derived to the approach of Vasiliev (2014). We evaluate its performance under the local risk $\mathcal{R}(\hat{H}, H)(t^*)$ defined by, for a fixed point $t^* \in R$,

$$\mathcal{R}(\hat{H}, H)(t^*) = E(|\hat{H}(t^*) - H(t^*)|). \quad (1.1)$$

We prove that \hat{H} attains fast rates of convergence over Hölder class according to the nature of the density of X_0 (the so-called ordinary smooth and super smooth cases are considered). Moreover, we illustrate its finite sample performance via a simulation study.

The remainder of the paper is structured as follows. In Section 2 we present some notations and formulate some assumptions on the model. Our main estimator is described in Section 3. Section 4 is devoted to its performances in terms of rates of convergence under the local error. The results of a numerical experiment to investigate the properties of our estimator are presented in Sections 5 and 6. The proofs are deferred to Section 8.

2 Notations and Assumptions

For any $p \geq 1$, let

$$L_p(R) = \left\{ h : R \rightarrow R; \int |h(x)|^p dx < \infty \right\}.$$

The convolution product of two functions h and k is given by

$$(h * k)(x) = \int h(x - z)k(z)dz,$$

provided the above integral exists.

The Fourier transform of a function $h \in L_1(R)$ is defined as

$$h^{ft}(x) = \int e^{itx}h(t)dt.$$

Let f_T be the density function of T , g be the density function of X_0 and f_Z be the density function of Z .

We formulate the following boundedness assumptions. Let $t \in R$.

(A1)(t) We suppose that there exists a known constant $c_* > 0$ such that

$$f_Z(t, 0) \geq c_*.$$

(A2)(t) We suppose that there exists a constant $C_* > 0$ such that

$$f_Z(t, j) \leq C_*, \quad j = 0, 1.$$

These boundedness assumptions are satisfied by a wide variety of sub-densities. In our main Theorem, we need that **(A1)**(t) and **(A2)**(t) are satisfied only at the point $t = t^*$ considered in the local error.

As usual in deconvolution problems in nonparametric estimation, we assume that the density function of g , belongs to one of the two major types of error densities :

(B1) *Ordinary smooth error densities* : There exist constants $C_1 > 0$, $C_2 > 0$ and $\alpha \geq 0$ such that

$$C_1(1 + |x|)^{-\alpha} \leq |g^{ft}(x)| \leq C_2(1 + |x|)^{-\alpha}.$$

(B2) *Super smooth error densities* : There exist constants $C_1 > 0$, $C_2 > 0$, $d_1 > 0$, $d_2 > 0$ and $\gamma \geq 0$ such that

$$C_1 \exp(-d_1|x|^\gamma) \leq |g^{ft}(x)| \leq C_2 \exp(-d_2|x|^\gamma)$$

for some $C_2 > C_1 > 0$, $0 < d_2 < d_1$ and $\gamma > 0$.

Further details are given in Meister (2009). Density function of a normal random variable is an example of a super smooth error density and the Laplace density (also known as double-exponential density) is an example for ordinary smooth error density.

Finally, we say that a function h belongs to the Hölder class $F_{\beta, \lambda; t}(M)$, with $\beta > 0$, $\lambda > 0$ and $M > 0$ if and only if

$$|h^{([\beta])}(y) - h^{([\beta])}(\tilde{y})| \leq M|y - \tilde{y}|^{\beta - [\beta]}, \quad \text{for all } y, \tilde{y} \in [t - \lambda, t + \lambda]$$

and $\sup_{t \in R} |h(t)| \leq M$.

3 The Estimators

In this section we propose an estimator for $H(t)$, the ratio of hazard functions of X_1 and X_2 .

First of all, note that the independence of Z and X_0 implies that the density function of T is given by

$$f_T(x) = (f_Z * g)(x) = \int f_Z(x - z)g(z)dz.$$

An elementary property of the Fourier transform and convolution product gives the following relationship between Fourier transforms of f_T, g and f_Z .

$$f_T^{ft}(x) = g^{ft}(x)f_Z^{ft}(x).$$

For a given g , the estimator \hat{f}_T of f_T is given by $\hat{f}_T^{ft}(x) = g^{ft}(x)\hat{f}_Z^{ft}(x)$ where $\hat{f}_Z^{ft}(x)$ is an estimator of $f_Z^{ft}(x)$. Hence,

$$\hat{f}_Z^{ft}(x) = \frac{\hat{f}_T^{ft}(x)}{g^{ft}(x)}.$$

Then, one can use the inversion techniques to find $\hat{f}_Z(x)$.

We extend these ideas to the competing risks problem. Let the sub-density functions of $(T, \delta = 1)$ and $(T, \delta = 0)$ be given by $f_T(t, 1)$ and $f_T(t, 0)$ respectively. Since Z and X_0 are independent, we have,

$$\begin{aligned} f_T(t, 1) &= \int f_Z(t - z, 1)g(z)dz \\ f_T(t, 0) &= \int f_Z(t - z, 0)g(z)dz \end{aligned}$$

The Fourier transforms of $f_T(t, 1)$ and $f_T(t, 0)$ are given by

$$\begin{aligned} f_T^{ft}(x, 1) &= f_Z^{ft}(x, 1)g^{ft}(x) \\ f_T^{ft}(x, 0) &= f_Z^{ft}(x, 0)g^{ft}(x) \end{aligned}$$

Then, for given g , the estimators for $f_T(t, 1)$ and $f_T(t, 0)$ are given by

$$\begin{aligned} \hat{f}_T^{ft}(x, 1) &= g^{ft}(x)\hat{f}_Z^{ft}(x, 1). \\ \hat{f}_T^{ft}(x, 0) &= g^{ft}(x)\hat{f}_Z^{ft}(x, 0). \end{aligned}$$

Based on paired data $(T_1, \delta_1), (T_2, \delta_2), \dots, (T_n, \delta_n)$, we propose the following estimators for $f_Z(t, 1)$ and $f_Z(t, 0)$,

$$\begin{aligned} \hat{f}_Z(t, 1) &= \frac{1}{2n\pi} \sum_{j=1}^n \delta_j \int e^{-i(t-T_j)x} \frac{K^{ft}(xb)}{g^{ft}(x)} dx, \\ \hat{f}_Z(t, 0) &= \frac{1}{2n\pi} \sum_{j=1}^n (1 - \delta_j) \int e^{-i(t-T_j)x} \frac{K^{ft}(xb)}{g^{ft}(x)} dx. \end{aligned} \tag{3.1}$$

where K^{ft} is Fourier transform of a kernel function, K with bandwidth $b = b_n$ satisfying the following properties :

- (a) $K \in L_1(R) \cap L_2(R)$, K^{ft} has support on $[-1, 1]$,
- (b) $\int z^j K(z) dz = \delta_{j,0}$, for all $j = 0, \dots, [\beta]$, $\delta_{j,\ell}$ is equal to 1, if $j = \ell$, and zero otherwise,
- (c) $\int |K(z)z^\beta| dz < \infty$.

We call these kernel functions β -order kernels (see Meister (2009)).

Let t^* be a fixed point in R . Suppose that **(A1)**(t^*) holds. We consider the following estimator for $H(t^*)$:

$$\hat{H}(t^*) = \frac{\hat{f}_Z(t^*, 1)}{\hat{f}_Z(t^*, 0)} \mathbf{1}_{\{|\hat{f}_Z(t^*, 0)| \geq \frac{c^*}{2}\}}, \quad (3.2)$$

where $\mathbf{1}$ denotes the indicator function.

The proposed thresholding is a technical tool to ensure good theoretical properties of our estimator. It is derived from the general ratio-type estimator studied by Vasiliev (2014).

4 Main Result

In this section we investigate the local error of \hat{H} over Hölder spaces.

Theorem 4.1 *Let t^* be a fixed point in R . Suppose that **(A1)**(t^*) is satisfied and that $f_Z(t^*, 1)$ and $f_Z(t^*, 0)$ belong to $F_{\beta, \lambda; t^*}(M)$.*

- Under **(B1)** : let $\hat{H}(t^*)$ be defined by (3.2) with a β -order kernel and

$$b = O\left(n^{-1/(2\beta+2\alpha+1)}\right). \quad (4.1)$$

Then the local error (1.1) of $\hat{H}(t^*)$ satisfied

$$\mathcal{R}(\hat{H}, H)(t^*) = O\left(n^{-\beta/(2\beta+2\alpha+1)}\right).$$

- Under **(B2)** : let $\hat{H}(t^*)$ be defined by (3.2) with a β -order kernel and

$$b = c_b^{-1/\gamma} (\log(n))^{-1/\gamma}, \quad c_b \in (0, 1/(2d_1)). \quad (4.2)$$

Then the local error (1.1) of $\hat{H}(t^*)$ satisfied

$$\mathcal{R}(\hat{H}, H)(t^*) = O\left(\log(n)^{-\beta/\gamma}\right).$$

The proof of Theorem 4.1 is based on sharp upper bounds and powerful mean L_2 error properties of our estimators $\hat{f}(t, j)$ for $j = 0, 1$.

Let us mention that the rates of convergence in Theorem 4.1 correspond to the optimal ones (in the minimax sense) attained by the standard Fourier-kernel density estimator in a measurement error setting under the local error (see Meister (2009)).

Remark 4.1 *Theorem 4.1 can be extended for the sup error :*

$$\mathcal{R}_o(\hat{H}, H) = \sup_{t \in I} E(|\hat{H}(t) - H(t)|),$$

where $I \subset R$ provided that **(A1)**(t) is satisfied for any $t \in I$.

Remark 4.2 *If n is such that $c_* \geq 1/\log(\log(n))$, one can define \hat{H} (3.2) with $1/\log(\log(n))$ instead of c_* . Then Theorem 4.1 is still valid with a penalty of $\log(\log(n))$ in the rates of convergence.*

5 Numerical Study

In this section we carry out a numerical experiment to study the performance of our proposed estimator \hat{H} given by (3.2) . We employed a smooth version of estimators for ratio of hazard rates, H after local linear regression (see, e.g., Fan (1992) and Chesneau *et al.* (2014)). The competing estimator for ratio of hazard rates is the ordinary kernel type estimator for hazard function under random censorship. We also plot the estimators for sub-density functions given by (3.1) and the kernel-type estimators for sub-densities. We show that our proposed estimators perform better especially when the error term is with non-zero mean.

Here we consider two cases - when both X_1 and X_2 follow Weibull distribution or Gamma distribution. The error density of X_0 is considered Normal or Laplace. In all cases the error terms have nonzero mean, $\mu = \min(sd(X_1), sd(X_2))/2$ and standard deviation $\sigma = \min(sd(X_1), sd(X_2))/20$. Graphs for zero mean error term are given in supplementary files. In each example we plot three figures, the left figure shows the real ratio of failure rates, *i.e.* H with blue curves and our proposed \hat{H} with solid lines and ordinary kernel based estimator of H with red lines. The two other figures shows the real sub-densities $f_Z(t, 1)$ and $f_Z(t, 0)$ with blue lines versus two estimators: our proposed estimators for sub-densities and the classical kernel-type estimator for sub-density functions with solid lines and red lines, respectively. The codes are written *R* and we employed different packages: *decon*, *OOmisc*, *muhaz*, *KernSmooth*, *logspline*. We used excellent software package **decon** for *R*, which was presented by Wang and Wang (2011).

Example 1 : We generate $n = 250$, X_1, X_2 observations from two independent Weibull distributions. Figure 1 shows the results for Weibull distributions with parameters (3,1) and (4,2), respectively. Figure 2 shows the results for Weibull distributions with parameters (6,1) and (3,2), respectively.

Example 2 : We generate $n = 250$ X_1, X_2 observations from two independent Gamma distributions - Figure 3 is for Gamma with parameters (6,1) and (3,2), respectively and Figure 4 is for Gamma with parameters (3,2) and (2,1), respectively.

On the basis of Figures 1-4 we conclude

- (i) The proposed estimator of $H(t)$, the ratio of hazard functions is very close to the original function for sample size 250.
- (ii) There are problems at the boundary with \hat{H} showing a drop.
- (iii) The alternative kernel type estimator does not perform well in all cases.
- (iv) The performance of \hat{H} is similar for both Normal and Laplace error densities.
- (v) Our estimators of the sub-density functions are smoother than the original sub-densities.
- (vi) The estimators of sub-density functions based on kernel estimators are not performing well.
- (vii) The pattern of results is similar for both Weibull and Gamma densities.

The performance of our estimators with non-zero error mean (both Normal and Laplace) is significantly better than the case of zero error mean. See Supplementary files for illustrations. The results for higher sample sizes ($n = 500, 1000$) show similar pattern.

In all Figures the function H is represented by the curve in blue and its estimator \hat{H} by curve in black. It is clear that the estimator \hat{H} is very close to the original ratio of the hazard rates $H(t)$. The results are similar for error distributions Normal and Laplace . The estimator is not too good at the tails when the error density is Laplace.

Similar conclusions hold in Figures 3 and 4 when we have Gamma variables with parameter above. Note that the pictures in Figures 3 and 4 look more apart because of the scale chosen for the Y-axis.

6 Real Data

We consider the follicular cell lymphoma data given in Pintilie (2006). There were 541 patients diagnosed with follicular cell lymphoma (I or II). The two treatments were radiation alone and a combination of radiation and chemo. The two competing risks were disease

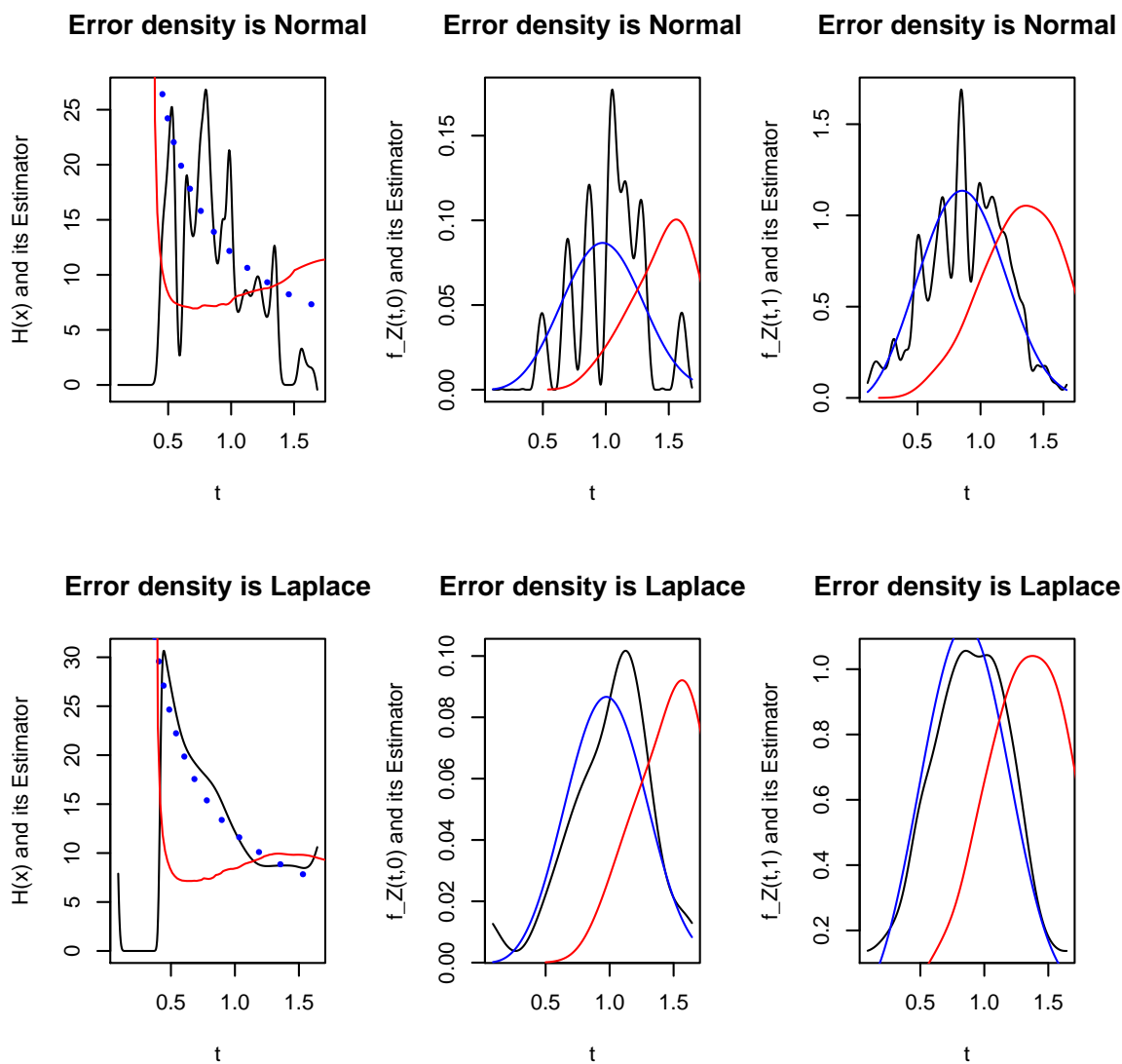


Figure 1 Ratio of hazard rates and sub-densities with their estimators. Real curves are blue and proposed estimators are solid and competitor estimators are red curves. The top array gives the results for Weibulls with parameters (3,1) and (4,2), respectively when the error density is normal whereas the the bottom one corresponds to error density Laplace.

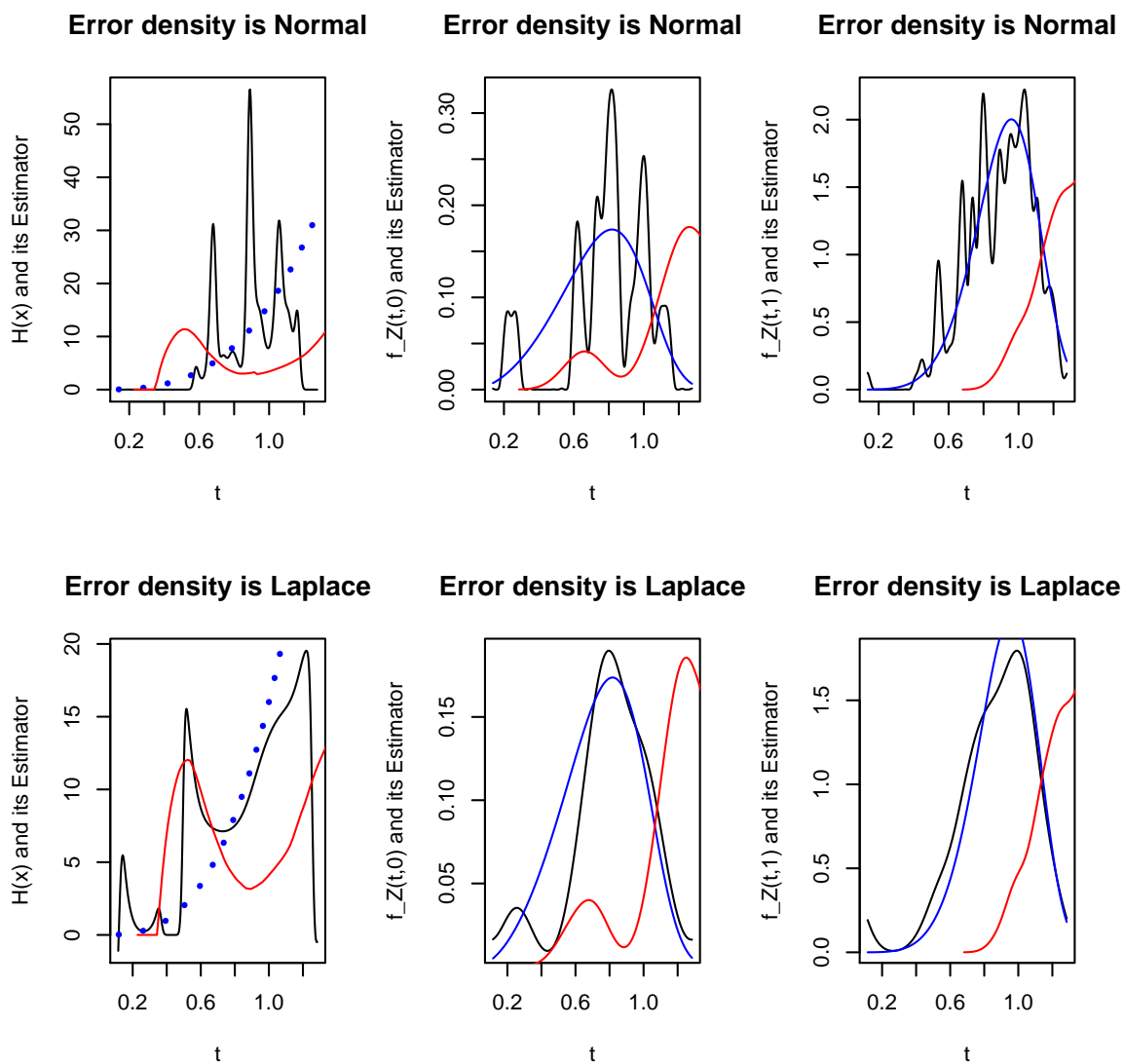


Figure 2 Ratio of hazard rates and sub-densities with their two estimators. Curve types are as in Figure 1. Densities considered are Weibull with parameters (6,1) and (3,2), respectively.

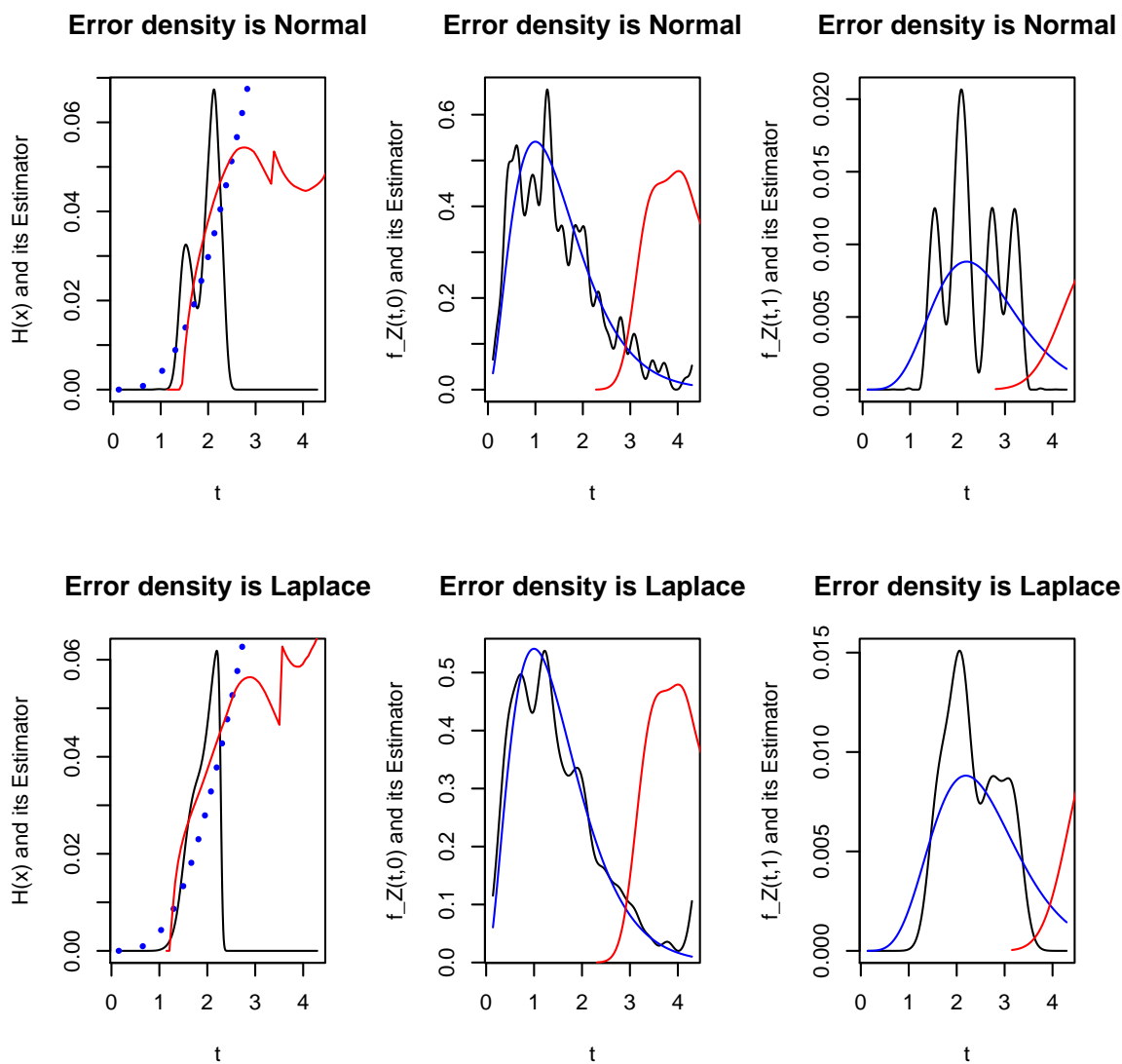


Figure 3 Ratio of hazard rates and sub-densities with two estimators. Curve types are as in Figure 1. Densities considered are Gamma with parameters (6,1) and (3,2), respectively.

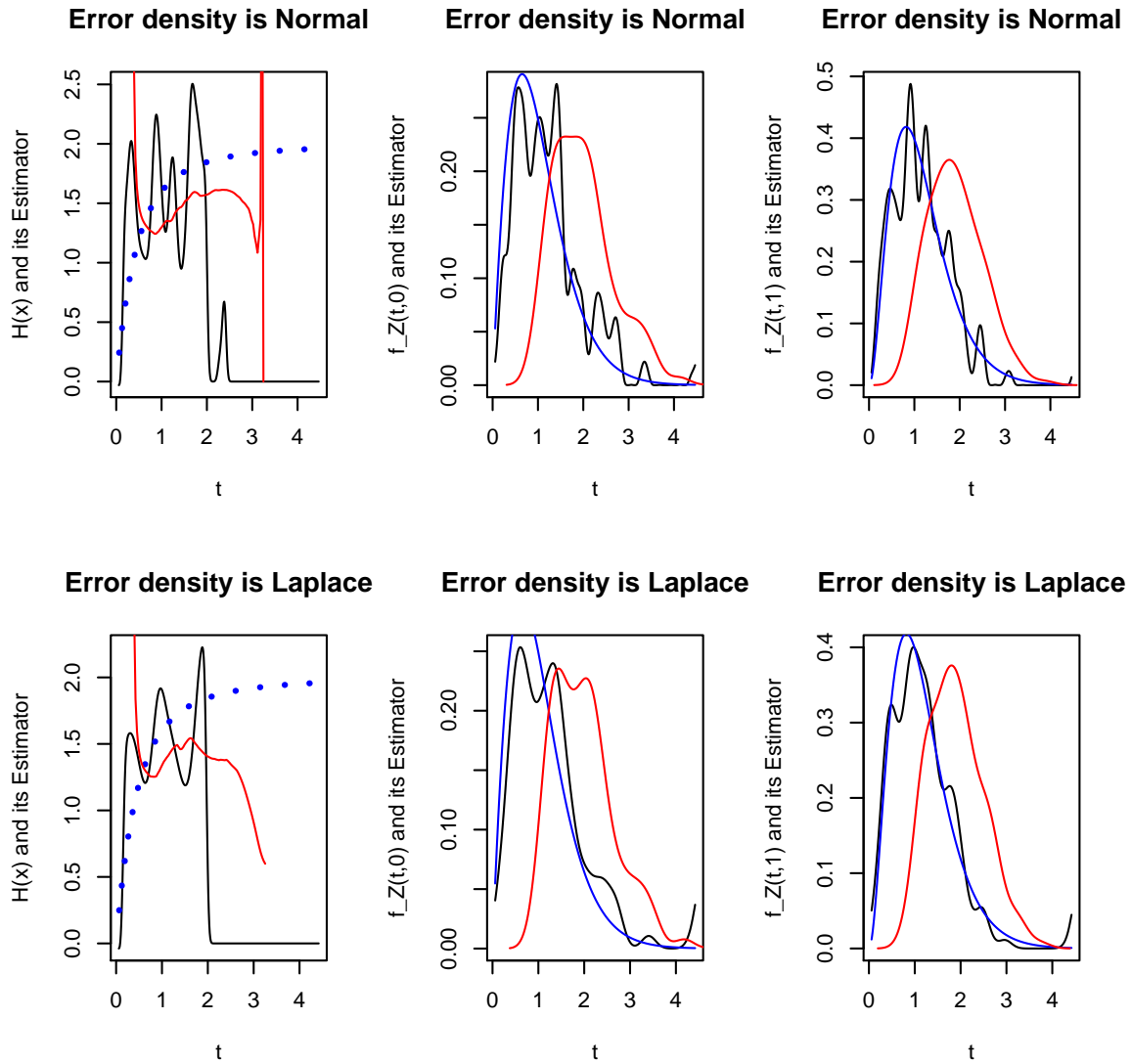


Figure 4 Ratio of hazard rates and sub-densities with their two estimators. Curve types are as in Figure 1. Densities considered are Gamma with parameters (3,2) and (2,1), respectively.

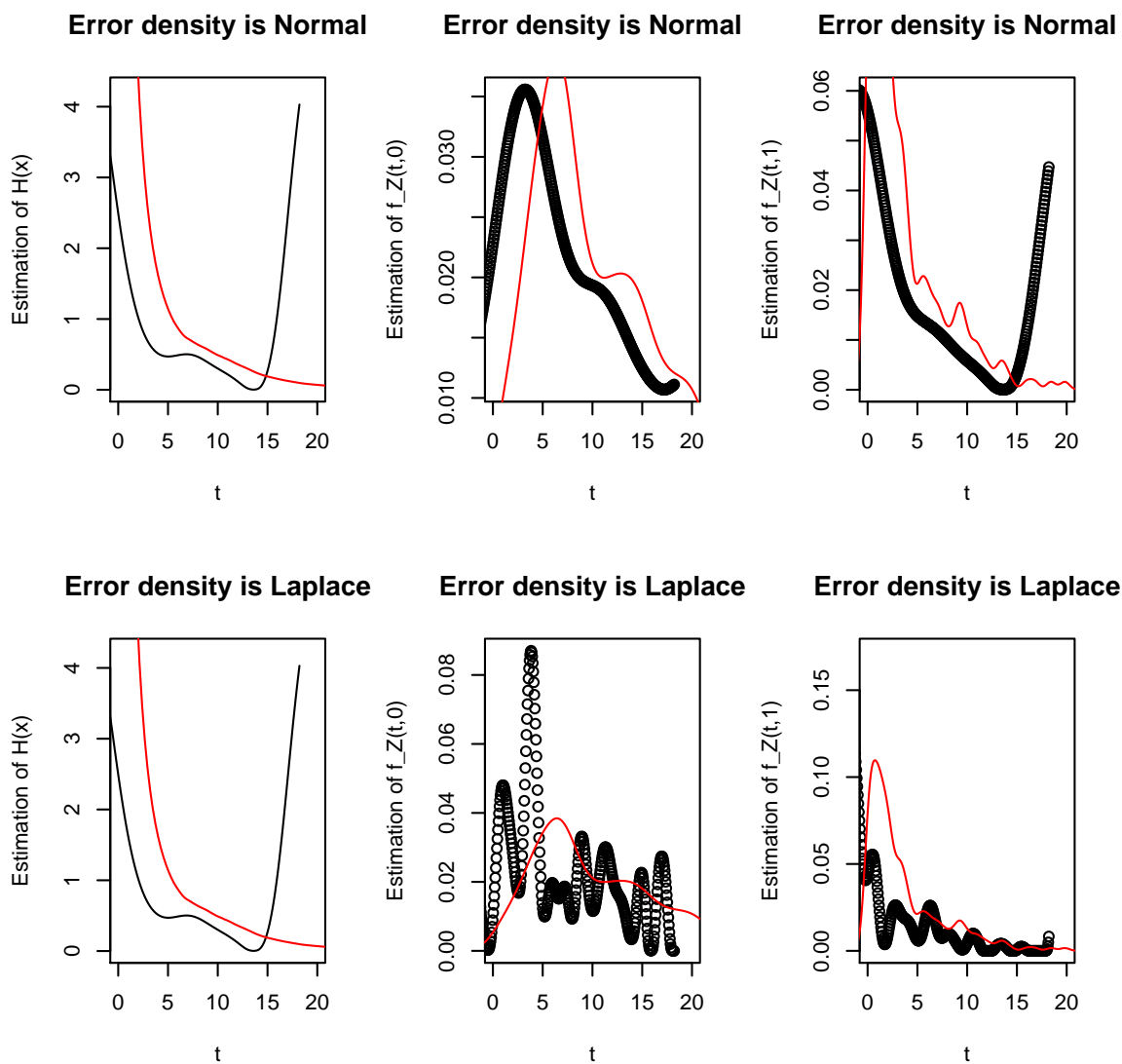


Figure 5 Two estimators for ratio of hazard rates and sub-densities. Proposed estimators are solid lines and the competitors are red lines. The expectation of error term is 3.

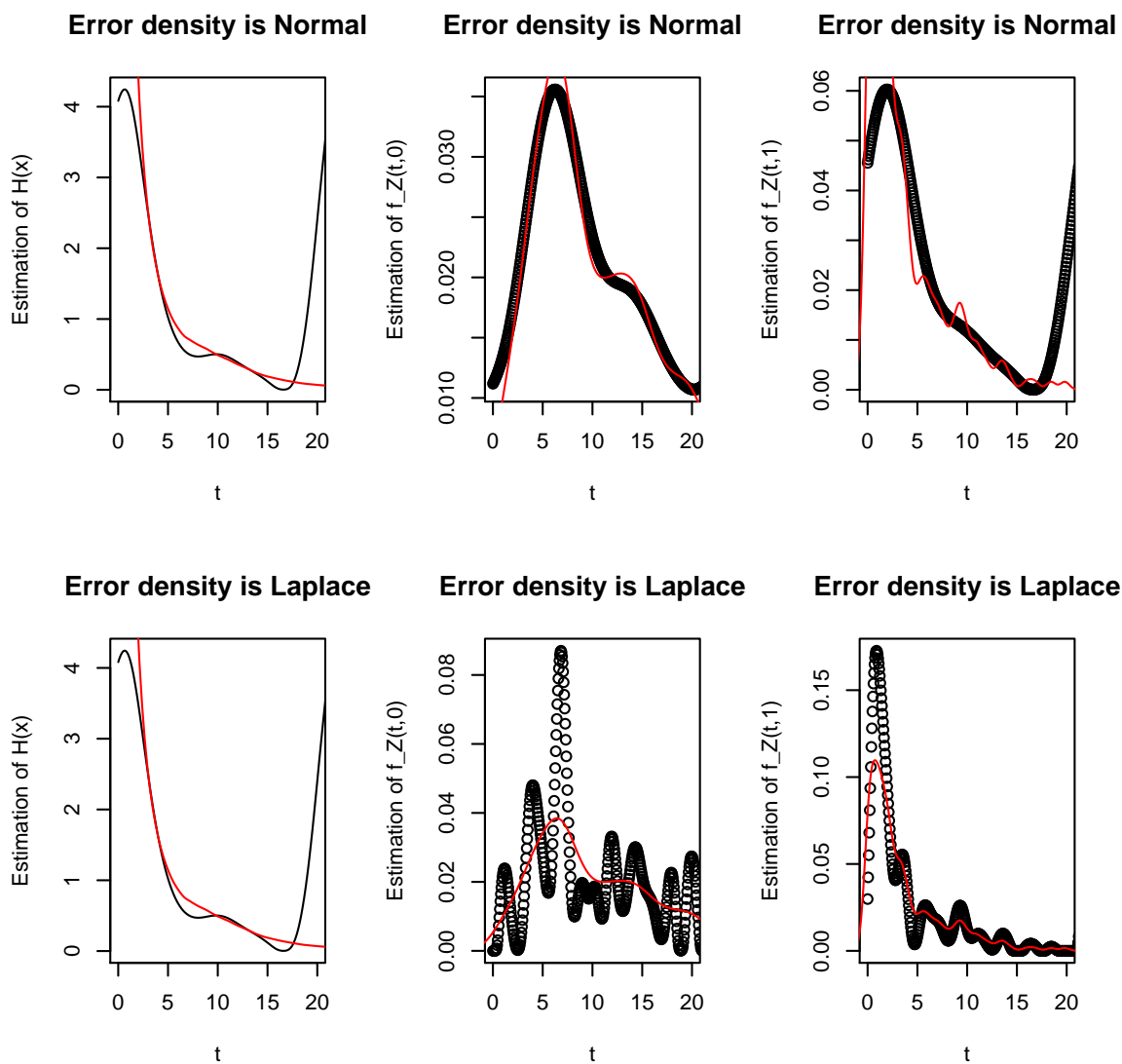


Figure 6 Ratio of hazard rates and sub-densities with two estimators. Proposed estimators are solid lines and the competitors are red lines. The expectation of error term is zero.

relapse (and no response) and death in remission. The average age of patients was 57 with $sd=14$. The haemoglobin levels of the patient had mean =138 and $sd=15$. The study had a median follow-up time of 5.5 years. Out of the 541 patients, 272 patients were in the no treatment response or relapse category, 76 patients died without relapse and 193 patients were censored.

The data was analyzed by Scheike and Zhang (2011) using Timereg Package R. Using both semi-parametric and nonparametric models they concluded that age and stage of the disease were significant factors. For our analysis, death in remission corresponds to $\delta = 1$ and disease relapse with no response and censored observations have been pooled to correspond to $\delta = 0$. We assume that the observed death times ($\delta = 1$) and relapse/censored times ($\delta = 0$) are sum of the actual times which are unobserved and common additive random effect of age and stage of the disease. We assume that this additive effect has either normal distribution or Laplace distribution.

Figure 5 shows the estimators of the ratio of hazard functions H and the two sub-density functions $f_Z(t, 1)$ and $f_Z(t, 0)$ for Normal and Laplace errors with $\mu = 3$. In each of the figures 5 and 6 the black curves corresponds to the proposed estimator for follicular data. The red curve for the ratio estimators corresponds to ordinary kernel type estimators of the hazard functions. the red curve for sub-density functions correspond to classical kernel-type estimators for sub-density functions.

Figure 6 corresponds to the case when both error densities have mean $\mu = 0$.

It is clear that if the mean of the error density is 0, then proposed estimators are closer to the estimators already available in literature. However, if mean effect of the additive error is different from zero, then one should make use of the newly proposed estimators 3.2 and 3.1. From the data analysis done previously, there is significant affect of age and stage of diseases. Hence one should use 3.2 and 3.1.

7 Conclusions

The initial motivation of the paper was to find nonparametric estimator for the ratio of failure rates $H(t)$. It turned out to be the extension of measurement error problem to competing risks model. We have proposed estimators for $H(t)$ and studied their properties. In the process we have also proposed kernel type estimators for the sub-density function under measurement error model. These estimators are not restricted to competing risks model but are applicable to all real situations giving rise to paired data where one component is a continuous random variable and the other is a discrete random variable.

8 Proofs

8.1 Auxiliary results

Let us introduce Lemmas 8.1 and 8.2 below before proving Theorem 4.1. Note that the main difference in the proof from other similar results in the literature (see Meister(2009)) is that we are working with sub-density functions and hence there is another factor involving δ in the estimators.

Lemma 8.1 *Let $t \in R$. Consider the estimator $\hat{f}_Z(t, j)$ (3.1) for $j = 0, 1$. Set $K_b(t) = (1/b)K(t/b)$ and $f_{Z_j}(t) = f_Z(t, j)$ for $j = 0, 1$. Then, under **(A2)**(t), for $j = 0, 1$,*

- we have

$$E(\hat{f}_Z(t, j)) = (K_b * f_{Z_j})(t).$$

- we have

$$\text{Var}(\hat{f}_Z(t, j)) \leq \frac{1}{2\pi n} C_* \int \left| \frac{K^{ft}(xb)}{g^{ft}(x)} \right|^2 dx.$$

Proof of Lemma 8.1 : We have

$$E(\delta_j e^{ixT_j}) = \int \int \delta_j e^{ixt_j} dF_T(t_j, \delta_j) = f_T^{ft}(x, 1) = g^{ft}(x) f_Z^{ft}(x, 1). \quad (8.1)$$

Similarly,

$$E((1 - \delta_j) e^{ixT_j}) = g^{ft}(x) f_Z^{ft}(x, 0).$$

Now we obtain the expectation of $\hat{f}_Z(t, 1)$: The Fubini theorem, the equality (8.1) and the Fourier inversion give

$$\begin{aligned} E(\hat{f}_Z(t, 1)) &= \frac{1}{2\pi n} \int e^{-itx} K^{ft}(xb) \frac{1}{g^{ft}(x)} \sum_{j=1}^n E(\delta_j e^{ixT_j}) dx \\ &= \frac{1}{2\pi} \int e^{-itx} K^{ft}(xb) f_Z^{ft}(x, 1) dx \\ &= \frac{1}{2\pi} \int e^{-itx} (K_b * f_{Z_1})^{ft}(x) dx = (K_b * f_{Z_1})(t). \end{aligned}$$

Expectation of $\hat{f}_Z(t, 0)$ can be obtained in a similar manner.

Using the change of variable $z = t - t_j$, $\sup_{t \in R} |(g * f_{Z_1})(t)| \leq C_*$ by **(A2)**(t), the inequality : $Var(B) \leq E(B^2)$ and the Parseval identity, we obtain

$$\begin{aligned}
Var(\hat{f}_Z(t, 1)) &= \frac{1}{(2\pi n)^2} \sum_{j=1}^n Var \left(\delta_j \int e^{-i(t-T_j)x} \frac{K^{ft}(xb)}{g^{ft}(x)} dx \right) \\
&\leq \frac{1}{(2\pi)^2 n} E \left(\left| \delta_j \int e^{-i(t-T_j)x} \frac{K^{ft}(xb)}{g^{ft}(x)} dx \right|^2 \right) \\
&= \frac{1}{(2\pi)^2 n} \int \left| \int e^{-i(t-t_j)x} \frac{K^{ft}(xb)}{g^{ft}(x)} dx \right|^2 f_T(t_j, 1) dt_j \\
&= \frac{1}{(2\pi)^2 n} \int \left| \int e^{-izx} \frac{K^{ft}(xb)}{g^{ft}(x)} dx \right|^2 (g * f_{Z_1})(t - z) dz \\
&\leq C_* \frac{1}{(2\pi)^2 n} \int \left| \int e^{-izx} \frac{K^{ft}(xb)}{g^{ft}(x)} dx \right|^2 dz \\
&= C_* \frac{1}{2\pi n} \int \left| \frac{K^{ft}(xb)}{g^{ft}(x)} \right|^2 dx.
\end{aligned}$$

A similar result for variance of $\hat{f}_Z(t, 0)$ can be obtained.

This ends the proof of Lemma 8.1. □

Lemma 8.2 *Let $t \in R$. Consider the estimator $\hat{f}_Z(t, j)$ (3.1) for $j = 0, 1$ and suppose that **(A2)**(t) is satisfied. Then*

- under **(B1)** :

$$E \left(|\hat{f}_Z(t, j) - f_Z(t, j)|^2 \right) = O \left(\frac{b^{-1-2\alpha}}{n} + b^{2\beta} \right),$$

- under **(B2)** :

$$E \left(|\hat{f}_Z(t, j) - f_Z(t, j)|^2 \right) = O \left(\frac{e^{2d_1 b^{-\gamma}}}{bn} + b^{2\beta} \right).$$

Moreover :

- under **(B1)** : choosing b as in (4.1), we obtain

$$E \left(|\hat{f}_Z(t, j) - f_Z(t, j)|^2 \right) = O \left(n^{-2\beta/(2\beta+2\alpha+1)} \right),$$

- under **(B2)** : choosing b as in (4.2), we obtain

$$E \left(|\hat{f}_Z(t, j) - f_Z(t, j)|^2 \right) = O \left(\log(n)^{-2\beta/\gamma} \right).$$

Proof of Lemma 8.2 :

- under **(B1)** : for any $t \in R$, thanks to Lemma 8.1,

$$\begin{aligned} E \left(|\hat{f}_Z(t, j) - f_Z(t, j)|^2 \right) &= V \left(\hat{f}_Z(t, j) \right) + \left| E(\hat{f}_Z(t, j)) - f_Z(t, j) \right|^2 \\ &\leq C_* \frac{1}{2\pi n} \int \left| \frac{K^{ft}(xb)}{g^{ft}(x)} \right|^2 dx + |(K_b * f_{Z_j})(t) - f_Z(t, j)|^2. \end{aligned}$$

Since K is a β -orders kernel and f_{Z_j} belong to $F_{\beta, \lambda; t}(M)$, similar arguments to those developed in Meister (2009) (pages 37-38) yield

$$|(K_b * f_{Z_j})(t) - f_Z(t, j)|^2 = O(b^{2\beta}).$$

Moreover, by **(B1)**,

$$\frac{1}{2\pi n} \int \left| \frac{K^{ft}(xb)}{g^{ft}(x)} \right|^2 dx = O \left(\frac{1}{nb} \min_{|t| \leq 1/b} |g^{ft}(t)|^{-2} \right) = O \left(\frac{b^{-1-2\alpha}}{n} \right).$$

It follows from the inequalities above that

$$E \left(|\hat{f}_Z(t, j) - f_Z(t, j)|^2 \right) = O \left(\frac{b^{-1-2\alpha}}{n} + b^{2\beta} \right),$$

- under **(B2)**, the proof is similar to the first point. It is enough to observe that

$$\frac{1}{2\pi n} \int \left| \frac{K^{ft}(xb)}{g^{ft}(x)} \right|^2 dx = O \left(\frac{e^{2d_1 b^{-\gamma}}}{bn} \right).$$

The considered bandwidths are chosen to minimize the right terms.

The proof of Lemma 8.2 is complete. □

8.2 Proof of Theorem 4.1

Proof of Theorem 4.1 : First of all, note that **(A2)**(t^*) is satisfied with $C_* = M$ since $f_Z(t^*, 1)$ and $f_Z(t^*, 0)$ belong to $F_{\beta, \lambda; t^*}(M)$. An elementary decomposition yields

$$|\hat{H}(t^*) - H(t^*)| = \left| \frac{\hat{f}_Z(t^*, 1)}{\hat{f}_Z(t^*, 0)} - \frac{f_Z(t^*, 1)}{f_Z(t^*, 0)} \right| \mathbf{1}_{\{|\hat{f}_Z(t^*, 0)| \geq \frac{c_*}{2}\}} + \frac{f_Z(t^*, 1)}{f_Z(t^*, 0)} \mathbf{1}_{\{|\hat{f}_Z(t^*, 0)| < \frac{c_*}{2}\}}.$$

It follows from **(A1)**(t^*) and **(A2)**(t^*), the embeddings

$$\begin{aligned} \left\{ |\hat{f}_Z(t^*, 0)| < \frac{c_*}{2} \right\} &\subseteq \left\{ f_Z(t^*, 0) - |\hat{f}_Z(t^*, 0)| > \frac{c_*}{2} \right\} \\ &\subseteq \left\{ |\hat{f}_Z(t^*, 0) - f_Z(t^*, 0)| > \frac{c_*}{2} \right\} \end{aligned}$$

and the Markov inequality that

$$\begin{aligned} \frac{f_Z(t^*, 1)}{f_Z(t^*, 0)} \mathbf{1}_{\{|\hat{f}_Z(t^*, 0)| < \frac{c_*}{2}\}} &\leq \frac{M}{c_*} \mathbf{1}_{\{|\hat{f}_Z(t^*, 0)| < \frac{c_*}{2}\}} \leq \frac{M}{c_*} \mathbf{1}_{\{|f_Z(t^*, 0) - \hat{f}_Z(t^*, 0)| > \frac{c_*}{2}\}} \\ &\leq \frac{2M}{(c_*)^2} |f_Z(t^*, 0) - \hat{f}_Z(t^*, 0)|. \end{aligned}$$

On the other hand, using again **(A1)**(t^*), we obtain

$$\begin{aligned} &\left| \frac{\hat{f}_Z(t^*, 1)}{\hat{f}_Z(t^*, 0)} - \frac{f_Z(t^*, 1)}{f_Z(t^*, 0)} \right| \mathbf{1}_{\{|\hat{f}_Z(t^*, 0)| \geq \frac{c_*}{2}\}} \\ &\leq \frac{2}{(c_*)^2} \left| \hat{f}_Z(t^*, 1) f_Z(t^*, 0) - \hat{f}_Z(t^*, 0) f_Z(t^*, 1) \right| \\ &\leq \frac{2}{(c_*)^2} \left(f_Z(t^*, 1) |\hat{f}_Z(t^*, 0) - f_Z(t^*, 0)| + f_Z(t^*, 0) |\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)| \right) \\ &\quad + \left| \hat{f}_Z(t^*, 0) - f_Z(t^*, 0) \right| |\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)| \\ &\leq \frac{2}{(c_*)^2} \left(M |\hat{f}_Z(t^*, 0) - f_Z(t^*, 0)| + M |\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)| \right) \\ &\quad + \left| \hat{f}_Z(t^*, 0) - f_Z(t^*, 0) \right| |\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)|. \end{aligned}$$

Therefore, there exists a constant $C > 0$ such that

$$\begin{aligned} &|\hat{H}(t^*) - H(t^*)| \\ &\leq C \left(|\hat{f}_Z(t^*, 0) - f_Z(t^*, 0)| + |\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)| \right) \\ &\quad + \left| \hat{f}_Z(t^*, 0) - f_Z(t^*, 0) \right| |\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)| \end{aligned}$$

The Cauchy-Schwarz inequality yields

$$\begin{aligned} \mathcal{R}(\hat{H}, H)(t^*) &\leq C \left(\sqrt{E(|\hat{f}_Z(t^*, 0) - f_Z(t^*, 0)|^2)} + \sqrt{E(|\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)|^2)} \right) \\ &+ \sqrt{E(|\hat{f}_Z(t^*, 0) - f_Z(t^*, 0)|^2)} \sqrt{E(|\hat{f}_Z(t^*, 1) - f_Z(t^*, 1)|^2)}. \end{aligned}$$

It follows from Lemma 8.2 at the point $t = t^*$ that

- under **(B1)** : with b as (4.1), we have

$$\mathcal{R}(\hat{H}, H)(t^*) = O \left(n^{-\beta/(2\beta+2\alpha+1)} \right),$$

- under **(B2)** : with b as (4.2), we have

$$\mathcal{R}(\hat{H}, H)(t^*) = O \left(\log(n)^{-\beta/\gamma} \right).$$

This concludes the proof. □

References

- [1] Bagai, I. and Prakasa Rao, B.L.S. (1992). Analysis of Survival Data with Two Dependent Competing Risks. *Biometrical Journal*, **7**, 801-814.
- [2] Carroll, R. J. and Hall, P. (1988). Optimal Rates of Convergence for Deconvolving a Density. *Journal of the American Statistical Associations*, **83**, 1184-1186.
- [3] Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*. Chapman Hall, London.
- [4] Chesneau, C., Dewan, I. And Doosti, H. (2014). Nonparametric estimation of two dimensional continuous- discrete density function by wavelets. *Statistical Methodology*, **18**, 64-78.
- [5] Crowder, M. (2001). *Classical Competing Risks*. Chapman and Hall, London.
- [6] David, H. A. and Moeschberger, M. L. (1978). *The Theory of Competing Risks*. Macmillan, New York.
- [7] Delaigle, A. and Gijbels, I. (2004a). Bootstrap Bandwidth Selection in Kernel Density Estimation from a Contaminated Sample. *Annals of the Institute of Statistical Mathematics*, **56**, 19-47.

- [8] Delaigle, A. and Gijbels, I. (2004b). Practical Bandwidth Selection in Deconvolution Kernel Density Estimation. *Computational Statistics and Data Analysis*, **45**, 249-267.
- [9] Delaigle, A. and Meister, A. (2007). Nonparametric Regression Estimation in the Heteroscedastic Errors-in-Variables Problem. *Journal of the American Statistical Association*, **102**, 1416-1426.
- [10] Delaigle, A. and Meister, A. (2008). Density Estimation with Heteroscedastic Error. *Bernoulli*, **14**, 562-579.
- [11] Fan, J. (1992). Design-adaptative nonparametric regression. *Journal of the American Statistical Association*, **87**, 998-1004.
- [12] Efromovich, S. (1997). Density Estimation for the Case of Supersmooth Measurement Error. *Journal of the American Statistical Association*, **92**, 526-535.
- [13] Fan, J. (1991). On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems. *The Annals of Statistics*, **19**, 1257-1272.
- [14] Fan, J. (1992). Deconvolution with Supersmooth Distributions. *Canadian Journal of Statistics*, **20**, 155-169.
- [15] Hall, P. and Lahiri, S. N. (2008). Estimation of Distributions, Moments and Quantiles in Deconvolution Problems. *The Annals of Statistics*, **36**, 2110-2134.
- [16] Kalbfleisch, J. D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data*. Wiley, New York.
- [17] Kotlarski, I. (1966). On Characterizing the Normal Distribution by Students' law. *Biometrika*, **53**, 603-606.
- [18] Kotlarski, I. (1967). On Characterizing the Gamma and the Normal Distribution. *Pacific Journal Mathematics*, **20**, 69-76.
- [19] Meister, A. (2004). On the Effect of Misspecifying the Error Density in a Deconvolution Problem. *Canadian Journal of Statistics*, **32**, 439-449.
- [20] Meister, A. (2009). Deconvolution Problems in Nonparametric Statistics. Lecture Notes in Statistics, **193**, Springer.
- [21] Pintilie, M. (2006) . Competing Risks - A Practical Perspective. Wiley.
- [22] Prakasa Rao, B.L.S. (1968). On a Characterization of Probability Distribution on Locally Compact Abelian groups. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **9**, 98-100.

- [23] Rosenblatt, M. (1991). *Stochastic Curve Estimation*. NSF-CBMS Regional Conference Series in Probability and Statistics, 3, Institute of Mathematical Statistics, Hayward, California.
- [24] Scheike, T.H. and Zhang, M. (2011). Analyzing competing risk data using the R timereg package. *Journal of Statistical Software*, **38**, 1-15.
- [25] Stefanski, L. A. and Carroll, R. J. (1990). Deconvoluting Kernel Density Estimators. *Statistics*, **21**, 169-184.
- [26] Van Es, B. and Uh, H. W. (2005). Asymptotic Normality of Kernel-Type Deconvolution Estimators. *Scandinavian Journal of Statistics*, **32**(3), 467-483.
- [27] Vasiliev, V. (2014). A truncated estimation method with guaranteed accuracy, *Ann. Inst. Stat. Math*, 66, 1, 141-163.
- [28] Wang X, F. and Wang, B. (2011). Deconvolution Estimation in Measurement Error Models: The R Package decon. *Journal of Statistical Software*, **39**, (10), 1-24.
- [29] Zhang, CH. (1990). Fourier Methods for Estimating Mixing Densities and Distributions. *The Annals of Statistics*, **18**, 806-830.