Journal of Statistical Research *ISSN 0256 - 422 X*
2010, Vol. 44, No. 2, pp. 207-217
Bangladesh

BOOTSTRAP BANDWIDTH SELECTION FOR A SMOOTH SURVIVAL FUNCTION ESTIMATOR FROM CENSORED DATA

Arthur Berg
Pennsylvania State University, 500 University Dr., Hershey, PA, 17033, USA
Email: berg@psu.edu

Kagba Suaray
California State University
Long Beach, 1250 Bellflower Blvd., Long Beach, CA, 90840-1001, USA
Email: ksuaray@csulb.edu

summary

Since the seminal paper by Efron [5], the bootstrap technique has been applied extensively to problems in the analysis of censored data. In this paper, the general problem is discussed, and a brief literature review is given. A new smooth estimator for the survival function is proposed, and its strong consistency proved. A bandwidth selection procedure based on Efron's censored bootstrap that incorporates a plug-in principle is proposed to determine a choice for the optimal bandwidth. Simulations indicates favorable performance of the proposed estimator when compared to the original Kaplan-Meier estimator.

Keywords and phrases: Censoring, Bootstrap, Kernel Estimation

AMS Classification: 62G07, 62N02

1 Introduction

The bootstrap was created by Efron [4] as a generalization of the jackknife. It quickly emerged as a reliable and mathematically solid technique for approximating measurements of accuracy for many useful but elusive quantities, such as the median. This utility translated to applications involving a variety of data structures, one of the first being censored data.

Gonzalez-Manteiga et al. [9] invoked the smoothed bootstrap for censored data for use in improving hazard function estimation. In this manuscript, we focus on improving estimation of the survival function by smoothing Peterson's representation of the survival function [17] while introducing two smoothing parameters. Through reformulation of the estimator, we motivate estimating the smoothing parameters independently in which one parameter is estimated with the smoothed bootstrap as in [9] and the other parameter is estimated with the plug-in principle, as described in [1], but generalized here to the censored data context.

The seminal paper of Kaplan and Meier [10] introduced the product-limit estimator – the nonparametric maximum likelihood estimator of the cumulative distribution function (CDF) in censored data – which, analogous to the empirical distribution function, is a $\sqrt{n}$-convergent estimator with several other desirable properties. But there is still room for improvement, especially when the sample size is small. Traditional kernel smoothing of the empirical distribution function and Kaplan-Meier estimator, have shown to be effective in improving finite-sample performance in the nonparametric estimation of the CDF [18, 6] and survival function [8, 13]. Kim et al. [12] recently suggested a much different type of kernel estimator of the CDF that combines three kernel estimates of the survival function, and finite-sample improvements were obtained.

The approach we take in the manuscript is much different from directly smoothing the Kaplan-Meier estimator itself. Peterson provided a representation of the Kaplan-Meier estimator as an explicit function of two empirical sub-survival functions [17]. Based on this alternative representation of the Kaplan-Meier estimator, a new smoothed estimator of the survival function is proposed. Utilizing similar techniques as originally considered in [17], strong consistency of the smooth estimator is established.

In the next section, we setup the notation and provide well known properties of estimators involving censored data. Peterson's formulation of the survival function is described in Section 3 along with the proposed smoothed form of the estimator with the corresponding theorem for strong consistency stated. The bootstrap method for selecting the optimal bandwidth is described in Section 4. Simulations provided in Section 5 indicate considerable improvements in performance with the proposed estimator in comparison with the Kaplan-Meier estimator. A proof of the strong consistency theorem stated at the end of Section 3 is provided in the Appendix.

2 Notation and Preliminaries

The study of lifetime variables highlights the importance of a number of related functions. The methods described later will enable automatic nonparametric estimation of these functions. The first is the survival function, $S(x) = P(X > x)$, i.e. the probability of non-occurrence of some event of interest by time x, with $x > 0$. Moreover, we will be interested in the hazard function, $\lambda(x)$, defined as the instantaneous rate of death or failure given survival up to time x:

$$\lambda(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x \mid X \geq x)}{\Delta x},$$

which is equal to

$$\frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)}.$$

It will also be useful to define the cumulative hazard function:

$$\Lambda(x) = \int_0^x \lambda(t)dt.$$

These functions are all mathematically equivalent, as is readily apparent, by the following relationship:

$$\lambda(x) = -\frac{d}{dx} \ln S(x)$$

or equivalently,

$$S(x) = \exp(-\Lambda(x)). \tag{2.1}$$

Extensive treatment and application of these functions in the parametric, as well as semi- and non- parametric settings can be found in many texts, such as Cox and Oakes [3], Lawless [14], and Miller [16].

In survival analysis, the data collected are lifetimes, modeled as $X_1, \ldots, X_n$, positive i.i.d. random variables with survival function S. We will use the word "lifetime" as a generic term for the observed time until death or failure, or any event of interest. In addition, we have independent censoring variables $Y_1, \ldots, Y_n$, positive i.i.d. random variables with survival function H (we will take S and H absolutely continuous with densities f_X and f_Y, respectively). For a pair (X_i, Y_i), we will say that the *ith* participant in the study has a realized failure time if $X_i \leq Y_i$, and that they are censored otherwise. Thus, the observed data can be summarized by the pairs (Z_i, Δ_i), where $Z_i = \min(X_i, Y_i)$, and $\Delta_i = 1_{[X_i \leq Y_i]}$, for $i = 1, \ldots, n$. Moreover, we define G, the survival function of the observed data, i.e. $G(x) \equiv P(Z \geq x)$. Independence of X and Y yields the relationship $G(x) = S(x)H(x)$. We will keep the notation of this paragraph throughout the remainder of this paper.

In addition to the aforementioned, the subsurvival functions $G_i(x), i = 0, 1$ will play a central role in the analysis and results contained herein. We define them as

$$G_i(x) = P(Z \geq x, \Delta = i).$$

These functions are directly proportional to the conditional survival function on the (non)censored observations to the degree that censoring occurs. Note also that

$$G_0(x) + G_1(x) = G(x). \tag{2.2}$$

3 Estimating the Survivor Function

In the absence of parametric assumptions, the most widely used estimator for the survival function is the Kaplan Meier estimator (KME), defined as

$$\hat{S}_n(x) = 1 - \hat{F}_n(x) = \prod_{i=1}^{n} \left(1 - \frac{\Delta_i}{n-i+1}\right)^{1_{[Z_{(i)} < x]}},$$

where $Z_{(i)}$, $i = 1, \ldots, n$ are the ordered Z_i's, and $\Delta_{(i)}$ the corresponding indicators. Here, and for the remainder of this paper, we take $x < \tau_S \leq \tau_H$, where $\tau_R = \inf\{x : R(x) > 0\} < \infty$.

3.1 Peterson's Representation of the KME

An alternative approach to estimation of the survival function stems from Peterson's representation of $S(t)$ as a functional of the subsurvival functions of the lifetime and censoring variables (see [17]). For S and H continuous, we have

$$S(x) \triangleq \Psi(G_1, G_0; x) = \exp\left[\int_0^x \frac{dG_1(u)}{G_0(u) + G_1(u)}\right]. \tag{3.1}$$

Note that this is essentially an alternate expression of (2.1) above. Peterson showed that the Kaplan Meier estimator can analogously be constructed as a functional of the empirical subsurvival functions, i.e. $\hat{S}_n(x) = \Psi(G_{1,n}, G_{0,n}; x)$, where

$$G_{i,n}(u) = \frac{1}{n}\sum_{j=1}^{n} 1_{[Z_j > u, \Delta_j = i]}$$

for $i = 0, 1$. Peterson's representation of the KME is key to proving its strong consistency. It is important to point out that the denominator of the integrand in (3.1) is simply $G(u)$, as in (2.2). Because the numerator of the integrand is a function of only the observed lifetimes, this representation has the intriguing property of accounting for censoring without explicitly expressing it.

3.2 Smoothing Peterson's Representation

The crux of our approach is to construct smooth global estimates for the survival function by taking the convolution of the $G_{i,n}$ with a kernel function, $K_h(u) = \frac{1}{h}K(u/h)$, that satisfies the following:

$(K1)$ $K(u) \geq 0$ $(K2)$ $K \in C[-1, 1]$

$(K3)$ $\int_{\infty}^{\infty} K(u)du = 1$ $(K4)$ $K(-u) = K(u)$

$(K5)$ K is of bounded variation

Thus, K is a symmetric sufficiently smooth, even density function with compact support. The parameter h is called the bandwidth, and controls the amount of smoothing applied to our estimate. Conditions on bandwidth parameters will be stated below. For the density f we assume

$$(D1) \;\; f'' \text{ exists} \quad \text{and} \;\; (D2) \;\; \int |f''| < \infty.$$

We define the so derived estimator as $\breve{S}(x) = \Psi(\breve{G}_1, \breve{G}_0; x)$ where

$$\breve{G}_1(u) = \frac{1}{nh_1}\sum_{j=1}^{n} \bar{K}\left(\frac{u - Z_j}{h_1}\right)\Delta_i,$$

where

$$\bar{K} = \int_x^\infty K(u)\,du.$$

and $\breve{G}_0$ is the above with Δ_i replaced with $1 - \Delta_i$. These are essentially smoothed versions of the respective subsurvival functions. We observe that smoothing causes the terms in the quotient in (3.1) to take an interesting form. By the fundamental theorem of calculus, the numerator becomes

$$\frac{1}{nh_1}\sum_{j=1}^{n} K\left(\frac{u - Z_j}{h_1}\right)\Delta_i$$

which is what Blum and Susarla [2] call $fH_n(u)$, essentially an estimator of $f \cdot H$, the joint sub-density of $(Z_i, \Delta_i = 1)$. As for the denominator, it is easily seen to be equal to

$$\frac{1}{nh_2}\sum_{j=1}^{n} \bar{K}\left(\frac{u - Z_j}{h_2}\right). \tag{3.2}$$

Note that (3.2) is clearly a smoothed version of the empirical survival function of the Z_i, i.e. the observed data. We will call this estimator $\breve{G}(u)$.

Now as for the bandwidths h_1 and h_2 appearing above, we assume

$(B1)$ $h_1 \to 0$ $\qquad$ $(B2)$ $nh_1 \to \infty$

$(B3)$ $\dfrac{nh_1}{\log\log n} \to \infty$ *almost surely* $\qquad$ $(B4)$ $h_2 \sim O(n^{-1/3})$

We establish the strong consistency of our estimator in the following

Theorem 1. *Let conditions $(K1) - (K5)$, $(D1), (D2)$, and $(B1) - (B3)$ hold, and let x be a continuity point of both $f \cdot H$ and f_Y. Then $\breve{S}_n(x) \to S(x)$ almost surely.*

Remark 1. It is now well understood how to analyze the MSE asymptotics of the smoothed KME, but the Peterson representation was designed to analyze the strong consistency of the KME, not the MSE asymptotics. The proposed smoothed Peterson representation is expected to improve the MSE asymptotics of the estimator and achieve $\sqrt{n}$ consistency, but the form of the Peterson estimator is not conducive to demonstrating MSE performance. This will be the topic of further research.

4 Bootstrapping the MISE

As typical in nonparametric estimation problems, we develop a global error criterion that will enable us to automatically determine the value of the smoothing parameter that is optimal in the prescribed sense. One of the most frequently used criteria in this context is that of mean integrated squared error, or MISE. The MISE of $\breve{S}_n(x)$, is defined as

$$R\left((Z_1, \Delta_1), \ldots, (Z_n, \Delta_n), h_1, S\right) = E \int \left(\breve{S}_n(x) - S(x)\right)^2 dx \tag{4.1}$$

Conspicuously absent is reference to the smoothing parameter h_2. It is a well established fact that optimal estimation of rational functionals occurs when the numerator is optimized. With this in mind, we choose to use a plug in bandwidth for the function in the denominator. We discuss this further in the next section.

Our approach will be to generate a bootstrap estimate of the MISE, which we will call MISE*, which we will minimize with respect to the bandwidth h_1. To do this, we employ the smooth censored resampling plan of Gonzalez-Manteiga et al. [9], which we now state for completeness. In what follows, g_1 and g_2 represent pilot bandwidths necessary to initiate the resampling process, and the symbol $\star$ refers to the convolution operation.

Step 1: Generate bootstrap resamples $X_1^*, \ldots, X_n^*$ from the smoothed lifetime KME, $\hat{S}_n(x) \star K_{g1}$. This entails numerically inverting the smoothed lifetime CDF, $\hat{F}^{-1}$, and computing $X_i^* = \hat{F}^{-1}(u_i)$ where $u_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ for $i = 1, \ldots, n$.

Step 2: Independently generate bootstrap resamples $Y_1^*, \ldots, Y_n^*$ from the smoothed censoring KME, $\hat{H}_n(x) \star K_{g1}$. This is performed as in Step 1, but on the smoothed censoring CDF.

Step 3: The smooth censored bootstrap sample is determined by $(Z_1^*, \Delta_1^*), \ldots, (Z_n^*, \Delta_n^*)$, where $Z_i^* = \min(X_i^*, Y_i^*)$, and $\Delta_i^* = 1_{[X_i^* \leq Y_i^*]}$, for $i = 1, \ldots, n$.

Step 4: The MISE (4.1) is approximated by the bootstrapped resamples for different bandwidths and the bandwidth yielding the smallest estimated MISE is selected.

As noted by Sanchez-Sellero et al. [19], in this context, obtaining a closed form for the bootstrap MISE is quite difficult to obtain, and so we proceed with Monte Carlo simulations to approximate the optimal bandwidth.

5 Simulations

To measure finite sample performance of the proposed estimator, simulations are performed. The proposed smoothed estimator of the survival function is somewhat computationally intense and invoking a bootstrap routine forces us to limit the number of simulations performed. Since the Kaplan-Meier estimator is already a $\sqrt{n}$-consistent estimator, we focus on smaller sample sizes where the benefit of smoothing is greatest; we consider sample sizes of $n = 30$ and $n = 100$. For simulating the censored data, lifetime and censoring variables are independently and identically simulated from a lognormal distribution with mean 4 and standard deviation .6 on the log scale. Therefore 50% censoring is observed on average. Two bandwidth parameters, h_1 and h_2, are first estimated. The bootstrap procedure described in the previous section is used to estimate h_1. The the denominator in the smoothed survival function estimator represents a smoothed version of the empirical survival function of the observed data, we utilize a modified version of Azzalini's plug-in bandwidth estimate [1] to censored data. This will provide us with an estimate of the h_2. Details of the modified Azzalini procedure are provided below.

5.1 The Plug-in Bandwidth

As noted above, the asymptotic performance of a quotient of statistics is more heavily dependent upon the numerator than the denominator. With this in mind, we sought to streamline the bootstrap process with the use of a plug in selection method (see Silverman [20] for the complete sample case) for h_2 in the kernel estimate (3.2). The usual normal reference rule is inadequate for lifetime data, so instead we choose a lognormal reference. Because the denominator is essentially a kernel CDF on the observed data, we may use the method of Azzalini [1] to designate our plug in rule, namely $h = (u/4vn)^{1/3}$,where

$$u = \left[1 - \int_{-1}^{1} \bar{K}^2(t)dt\right] \int_0^\infty f(x)dx \quad \text{and} \quad v = \left[\int_{-1}^{1} t^2 K(t)dt\right]^2 \left[\frac{1}{4}\int_0^\infty [f'(x)]^2 dx\right].$$

In the above we take f to be the lognormal density, and choose for K the Epanechnikov kernel. These selections yield

$$h = \left[\frac{72\sigma^3\sqrt{5\pi}}{35\exp\left(-3\mu + \frac{9\sigma^2}{4}(2+\sigma^2)\right)n}\right]^{1/3}.$$

It is clear that the parameters μ and σ must be estimated. Given the expressions for the mean and variance of the lognormal distribution, it can easily be shown that if $X \sim LogN(\mu, \sigma^2)$, then

$$\sigma = \left[\log\left(\frac{E(X^2)}{E(X)^2}\right)\right] \quad \text{and} \quad \mu = \log\left[E(X)\right] - \frac{\sigma^2}{2}.$$

To estimate the moments of f, we use the Kaplan-Meier weighted sample moments

$$\hat{\mu}_k = \int_0^\infty x^k d\hat{F}_n(x).$$

5.2 Results

As the bootstrap procedure is computationally intensive, the bootstrap replications were limited to 100 replications per realization. Performing 50 realizations per simulation, MISE performance is provided it Table 1.

Here we see the proposed smooth estimator provides much better MISE performance when compared with the original Kaplan-Meier estimator. Again, we focus on relatively small sample sizes where the effect of smoothing can be observed. Indeed we see the diminishing effect of smoothing with the larger sample size, but the proposed smooth survival function estimator is still seen to considerably outperform the Kaplan-Meier estimator.

Table 1: Estimated MISE performance of the smooth estimator compared to the non-smooth Kaplan-Meier estimator with standard errors given in parentheses.

Sample Size	Smooth Estimator	Kaplan-Meier
$n = 30$	**.212**$_{(.077)}$	.547$_{(.020)}$
$n = 100$	**.169**$_{(.053)}$	.357$_{(.030)}$

A Proof of Theorem 1

Lemma A.1. *Let $(K1)-(K5)$, $(B1)-(B3)$, and $(D1),(D2)$ hold. Then*

$$\sup_x \mid fH_n(x) - fH(x) \mid = O\left[\left(\frac{\log\log n}{nh}\right)^{\frac{1}{2}}\right] + O(h)$$

with probability 1 for every x that is a continuity point of f and H.

Proof. Define

$$\hat{f}_n(x) = \int K_h(x-y)\, d\hat{F}_n(y) \tag{A.1}$$

where $\hat{F}_n(y) = 1 - \hat{S}_n(y)$. Then (A.1) is the well studied convolution estimator introduced in [2] and [7]. The triangle inequality yields

$$\sup_x \mid fH_n(x) - fH(x) \mid \leq \sup_x \mid H(x) \mid \sup_x \mid \hat{f}_n(x) - f(x) \mid + \sup_x \mid \hat{f}_n(x)H(x) - fH_n(x) \mid \tag{A.2}$$

Under $(K3),(K5)$ and $(D1),(D2)$, the first term on RHS of (A.2) can be shown to be $O\left[\left(\frac{\log\log n}{nh}\right)^{\frac{1}{2}}\right] + O(h)$ w.p. 1 (see Theorem 2.2 (ii) in [11]). The second term is clearly $O\left[\left(\frac{\log\log n}{n}\right)^{\frac{1}{2}}\right] + O(h)$ by Theorem 1(ii) in [15]. □

This next lemma addresses the strong consistency of the denominator of the quotient in (3.1).

Lemma A.2. *If G is continuous and $G(\tau_F) < 1$, then*

$$\sup_{x\in\mathbf{R}} \mid \breve{G}(x) - G(x) \mid \to 0$$

with probability 1.

Proof. The proof follows standard lines of argument for strong convergence of kernel estimators. We first develop an expression of the estimator that will be more conducive to the desired result. As expressed in (3.2),

$$\check{G}(x) = \int_{-\infty}^{x} \hat{g}(u)\, du \tag{A.3}$$

where $\hat{g}_n(x)$ is just the kernel estimator constructed from the observed data. As discussed in [7], G is not required to be absolutely continuous in order to construct a density based estimator as in (A.3). An application of Fubini's theorem establishes the equivalence

$$\int_{-\infty}^{x} \int K_h(u-y)\, dG_n(y)\, du = \int \int_{-\infty}^{x} K_h(u-y)\, du\, dG_n(y) = \int \bar{K}_h(u-y)\, dG_n(y),$$

where $G_n(x)$ represents the empirical distribution function of the Z_i. Finally, the integration by parts technique of Földes, Rejtö and Winter yields the equality

$$\int \bar{K}_h(u-y) dG_n(y) = \int G_n(y) d\bar{K}_h(u-y).$$

Now define

$$\breve{G}(x) = \int G(y) d\bar{K}_h(u-y).$$

We have that

$$\mid \check{G}(x) - G(x) \mid \leq \mid \check{G}(x) - \breve{G}(x) \mid + \mid \breve{G}(x) - G(x) \mid . \tag{A.4}$$

We proceed by showing that the supremum over the real line of each of the terms on the right hand side of (A.4) is sufficiently small. For the first term, we have that

$$\mid \check{G}(x) - \breve{G}(x) \mid \leq \int \mid G(y) - G_n(y) \mid d\bar{K}_h(u-y) \leq \sup_y \mid G(y) - G_n(y) \mid V(K) \tag{A.5}$$

where $V(K)$ is the total variation of the kernel K. Condition $(K5)$, along with the Glivelo-Canteli theorem imply that the last term in (A.5) tends to zero almost surely.

As for the second term in the right hand side of (A.4), we have that

$$\mid \breve{G}(x) - G(x) \mid = \left| \int G(x-vh)K(v)dv - \int G(x)K(v)\, dv \right|,$$

which we can see is less than

$$\int_{|vh|\leq\delta} \mid G(x-vh) - G(x) \mid K(v)\, dv + \int_{|vh|>\delta} \mid G(x-vh) - G(x) \mid K(v)\, dv. \tag{A.6}$$

The first term of (A.6) is evidently less than

$$\sup_{|vh|\leq\delta} \mid G(x-vh) - G(x) \mid \int_{|vh|\leq\delta} \mid K(v) \mid dv.$$

Holding h constant, and letting δ tend to zero, this term approaches zero by the continuity of G at x. Finally, the second term of (A.6) is less than

$$2\int_{|vh|\leq\delta} K(v)\,dv,$$

since the distribution function G is bounded above by 1. Now letting n tend to ∞, the result follows from $(K4)$. □

Proof. (**Theorem 1**) The theorem now follows from Property 3.1 of [17], which we paraphrase here for completeness:

> *The function $\Psi(R_1(\cdot), R_2(\cdot), x)$ is a continuous function of the arguments $R_1(\cdot)$ and $R_2(\cdot)$, where the metric on the space of continuous functions is the supremum metric.*

This fact, along with the above lemmas and the continuous mapping theorem prove the theorem. □

References

[1] A. Azzalini. A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, pages 326–328, 1981.

[2] J.R. Blum and V. Susarla. Maximal deviation theory of density and failure rate function estimates based on censored data. In *Multivariate Analysis–V: Proceedings of the Fifth International Symposium on Multivariate Analysis*, page 213. North-Holland, 1980.

[3] D.R. Cox and D. Oakes. *Analysis of survival data.* Chapman & Hall/CRC, 1984.

[4] B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[5] B. Efron. Censored data and the bootstrap. *Journal of the American Statistical Association*, 76(374):312–319, 1981.

[6] M. Falk. Relative deficiency of kernel type estimators of quantiles. *The Annals of Statistics*, 12(1):261–268, 1984.

[7] A. Földes, L. Rejtő, and BB Winter. Strong consistency properties of nonparametric estimators for randomly censored data, II: estimation of density and failure rate. *Periodica Mathematica Hungarica*, 12(1):15–29, 1981.

[8] J.K. Ghorai. Deficiency of the mle of a smooth survival function under the proportional hazard model. *Communications in statistics. Theory and methods*, 18(8):3047–3056, 1989.

[9] W. Gonzalez-Manteiga, J.S. Marron, and R. Cao. Bootstrap selection of the smoothing parameter in nonparametric hazard rate estimation. *Journal of the American Statistical Association*, 91(435), 1996.

[10] E.L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, pages 457–481, 1958.

[11] R.J. Karunamuni and S. Yang. Weak and strong uniform consistency rates of kernel density estimates for randomly censored data. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, pages 349–359, 1991.

[12] C. Kim, S. Kim, M. Park, and H. Lee. A bias reducing technique in kernel distribution function estimation. *Computational Statistics*, 21(3):589–601, 2006.

[13] K.B. Kulasekera, C.L. Williams, M. Coffin, and A. Manatunga. Smooth estimation of the reliability function. *Lifetime Data Analysis*, 7(4):415–433, 2001.

[14] J.F. Lawless. *Statistical models and methods for lifetime data.* Wiley New York, 1982.

[15] J. Mielniczuk. Some asymptotic properties of kernel estimators of a density function in case of censored data. *The Annals of Statistics*, 14(2):766–773, 1986.

[16] R.G. Miller, G. Gong, and A. Muñoz. *Survival analysis.* Wiley New York, 1981.

[17] A.V. Peterson. Expressing the kaplan-meier estimator as a function of empirical subsurvival functions. *Journal of the American Statistical Association*, 72(360):854–858, 1977.

[18] R.D. Reiss. Nonparametric estimation of smooth distribution functions. *Scandinavian Journal of Statistics*, pages 116–119, 1981.

[19] C. Sánchez-Sellero, W. González-Manteiga, and R. Cao. Bandwidth selection in density estimation with truncated and censored data. *Annals of the Institute of Statistical Mathematics*, 51(1):51–70, 1999.

[20] B.W. Silverman. *Density estimation for statistics and data analysis.* Chapman & Hall/CRC, 1986.