NONPARAMETRIC ESTIMATION IN UNIFORM DECONVOLUTION AND INTERVAL CENSORING

BY PIET GROENEBOOM[a] AND GEURT JONGBLOED[b]

Delft Institute of Applied Mathematics, Mekelweg 4, 2628 CD Delft, The Netherlands.
[a]*P.Groeneboom@tudelft.nl*; [b]*G.Jongbloed@tudelft.nl*

In the uniform deconvolution problem one is interested in estimating the distribution function F_0 of a nonnegative random variable, based on a sample with additive uniform noise. A peculiar and not well understood phenomenon of the nonparametric maximum likelihood estimator in this setting is the dichotomy between the situations where $F_0(1) = 1$ and $F_0(1) < 1$. If $F_0(1) = 1$, the MLE can be computed in a straightforward way and its asymptotic pointwise behavior can be derived using the connection to the so-called current status problem. However, if $F_0(1) < 1$, one needs an iterative procedure to compute it and the asymptotic pointwise behavior of the nonparametric maximum likelihood estimator is not known. In this paper we describe the problem, connect it to interval censoring problems and a more general model studied in [5] to state two competing naturally occurring conjectures for the case $F_0(1) < 1$. Asymptotic arguments related to smooth functional theory and extensive simulations lead us to to bet on one of these two conjectures.

1. Introduction. Consider a random variable U, distributed according to a distribution function F_0 on $\mathbb{R}^+$. Instead of observing U, one observes the sum

$$S = U + V$$

where $V \sim$Unif$(0,1)$, independent of U. The random variable S then has convolution density

$$g_S(s) = \int 1_{[0,1]}(s-v)\,dF_0(v) = \int 1_{[s-1,s]}(v)\,dF_0(v) = F_0(s) - F_0(s-1). \tag{1.1}$$

Statistical inference for F_0 based on n i.i.d. random variables $S_1, \ldots, S_n$ with density g_S, is known as the uniform deconvolution problem. It is not essential that the support of the uniform distribution is $[0,1]$, we can always reduce the deconvolution problem with uniform random variables with another support to this situation. For simplicity we stick in this paper to uniform random variables with support $[0,1]$. As discussed in [8], an interesting application of uniform deconvolution is in the "deblurring" of pictures, blurred by Poisson noise, see, e.g., [13].

The log likelihood of F is then defined as

$$\ell(F) = \sum_{i=1}^{n} \log\left\{F(S_i) - F(S_i - 1)\right\}. \tag{1.2}$$

If F_0 satisfies $F_0(1) = 1$, this model will be seen to be equivalent to the so-called current status (or interval censoring case 1) model. Consequently, there is a one step algorithm to compute the nonparametric MLE of F_0. We explain this further in Section 2.1. We also derive the asymptotic distribution of the nonparametric MLE in this section.

AMS 2000 subject classifications: Primary 62G05, 62N01; secondary 62-04.

Keywords and phrases: deconvolution, current status, interval censoring, nonparametric maximum likelihood, incubation time model, asymptotic distribution.

In section 2.2, we show that if the support of the distribution corresponding to the distribution function F_0 is not contained in $[0,1]$, with upper support point $M > 1$, the model can be shown to be equivalent to an interval censoring, case m, model, where $m = \lceil M \rceil$. Consequently, a one step algorithm to compute the nonparametric MLE is not known, but one can use iterative algorithms in this case to compute the nonparametric MLE, for example the iterative convex minorant algorithm, proposed in [10].

A more general model is studied in [5] and [6]. There the length of the support of the uniform random variable V, say E, is actually random (but observable). This model is used for estimating the distribution of the incubation time of a disease. In this interpretation there is a positive random variable E, the duration of the "exposure time", with (unknown) distribution function F_E. Given E, an independent "infection time" V is drawn, uniformly distributed on the interval $[0,E]$. Moreover, the quantity of interest is the incubation (length) time U with distribution function F_0. Apart from E, the observation is then the time S for getting symptomatic, where $S = U + V$. Here, the random variables U and V are independent, conditionally on E. The observations therefore consist of n i.i.d. pairs of exposure times and times of getting symptomatic

$$(E_i, S_i), \qquad i = 1, \dots, n. \tag{1.3}$$

This model is also considered in [1], [2], [4] and [14].

Under the assumption that E has an absolutely continuous distribution, the asymptotic pointwise distribution of the nonparametric MLE of F_0 is derived in [5]. Remarkably, for this result there is no distinction between different assumptions on the support of the distribution corresponding to F_0, it holds in general. Taking for the distribution of E the point mass at 1 in the statement of Theorem 4.1 in [5], a case certainly not covered by the assumption of absolute continuity in this theorem and coinciding with the uniform deconvolution problem stated above, the correct asymptotic distribution appears in case $F_0(1) = 1$ as derived in our section 2.1. This also provides a natural conjecture (which we do, however, not believe in) for the asymptotic distribution of the nonparametric MLE in the situation where $F_0(1) < 1$. In section 3 we discuss this more general model, used in estimating the incubation time distribution, in more detail and formulate the conjecture.

Before turning to a simulation study and asymptotic considerations to support our actual conjecture in section 5, in section 4 we digress in a discussion of the estimation of other functionals of F_0, rather than the evaluation of the nonparametric MLE itself at a fixed point. For these so-called smooth functionals, asymptotic results are known. Moreover, the idea behind these smooth functionals is needed to understand the asymptotic considerations leading up to our final conjecture for the asymptotic distribution of the nonparametric MLE of $F_0(t_0)$ in the uniform deconvolution setting.

2. Uniform deconvolution and interval censoring. In this section, we establish a relation between the uniform deconvolution problem and various interval censoring models. In those models, also a dichotomy between two cases arises naturally. The uniform deconvolution problem in case $F_0(1) = 1$ will be seen to be related to the interval censoring case 1 (or, current status) problem. This situation will be discussed in section 2.1, where also the asymptotic distribution of the nonparametric MLE of $F_0(t_0)$ is given. In section 2.2, we discuss the relation between the uniform deconvolution model where $F_0(1) < 1$ and interval censoring case m, for $m \geq 2$.

We first briefly describe the interval censoring case 1 (IC-1) problem. Consider a random variable X with distribution function F_0 of interest and independently of this a random variable Y with a distribution function G. Instead of observing X, the pair (Y, Δ) is observed, where Δ indicates whether X is smaller than or equal to Y. So, $\Delta = 1_{\{X \leq Y\}}$. In survival analysis terminology, X is the event time, Y is the inspection time and we observe the "status" of the subject at time Y. The status being whether the event (usually an infection) has already occured at time Y or not. Note that given inspection time $Y = y$, $\Delta \sim$Bernoulli$(F_0(y))$. The problem is to estimate distribution function F_0, based on an i.i.d. sample of (Y_i, Δ_i)'s.

A more general model is the interval censoring case m (IC-m) model, where m stands for the number of distinct inspection times per subject. Formally, $Y = (Y_1, \dots, Y_m)$ is a random vector of distinct inspection times and X a nonnegative random variable independent of these event times. One then observes (Y, Δ), where the $m+1$-vector Δ indicates which of the intervals defined by the vector Y contains X: $\Delta = (\Delta_1, \dots, \Delta_{m+1})$. Here for $1 \le j \le m+1$, $\Delta_j = 1_{(Y_{(j-1)}, Y_{(j)}]}(X)$, denoting by $Y_{(j)}$ the j-th order statistic in the vector Y, and by convention $Y_0 \equiv 0$ and $Y_{(m+1)} = \infty$. Note that given the (ordered) vector of observation times equals $(y_{(1)}, \dots, y_{(m)})$, the vector Δ is multinomially distributed with parameters 1 and probability vector

$$(F_0(y_{(1)}), F_0(y_{(2)}) - F_0(y_{(1)}), \dots, F_0(y_{(m)}) - F_0(y_{(m-1)}), 1 - F_0(y_{(m)})).$$

Given an i.i.d. sample of vectors (Y, Δ), the problem is then to estimate the distribution function F_0.

In the two subsections below, we establish the relation between the uniform deconvolution problem and the interval censoring models and derive results based on that, depending on whether $F_0(1) = 1$ or $F_0(1) < 1$.

2.1. *Case* $F_0(1) = 1$. Following the suggestions of Exercise 2, p. 61, of [10], we set up an equivalence of the interval censoring problem with the current status model, in case $F_0(1) = 1$. Given the observed $S_1, S_2, \dots, S_n$, define

$$\Delta_i = \begin{cases} 1\,, & \text{if } S_i \le 1 \\ 0\,, & \text{if } S_i > 1 \end{cases} \tag{2.1}$$

and let the random variable Y_i be defined by:

$$Y_i = \begin{cases} S_i, & , \text{ if } \Delta_i = 1 \\ S_i - 1\,, & \text{if } \Delta_i = 0, \end{cases} \tag{2.2}$$

translating the S_i's to pairs (Y_i, Δ_i) for $1 \le i \le n$. Note that, for $y \in (0,1)$, considering a single observation (e.g. $i = 1$), we get

$$\begin{aligned} \mathbb{P}(Y \le y) &= \mathbb{P}(Y \le y \wedge \Delta = 0) + \mathbb{P}(Y \le y \wedge \Delta = 1) \\ &= \mathbb{P}(S - 1 \le y \wedge \Delta = 0) + \mathbb{P}(S \le y \wedge \Delta = 1) = \mathbb{P}(1 < S \le y + 1) + \mathbb{P}(S \le y) \\ &= \int_1^{1+y} g_S(s)\,ds + \int_0^y g_S(s)\,ds = \int_0^y F_0(s+1)\,ds = y. \end{aligned}$$

Here we use (1.1) and the fact that $F_0(s+1) = 1$ if $s > 0$, as $F_0(1) = 1$. Hence, the Y_i's are uniformly distributed on $[0,1]$.

Moreover, again using (1.1), for $y \in (0,1)$, the conditional distribution of Δ given $Y = y$ is Bernoulli, with success parameter

$$\mathbb{P}\{\Delta = 1 | Y = y\} = \frac{g_S(y)}{g_S(y) + g_S(y+1)} = F_0(y).$$

This shows the equivalence of the uniform deconvolution model with the current status model in the case $F_0(1) = 1$. More specifically, it is a current status model with event time distrubution F_0 and inspection time distribution uniform on $[0,1]$. This means that we can immediately characterize the nonparametric MLE in this setting and also derive its pointwise asymptotic distribution. Indeed, the MLE is the maximizer of the log likelihood

$$\sum_{i=1}^{n} \left(\Delta_i \log F(Y_i) + (1-\Delta_i)\log(1-F(Y_i))\right),$$

where Δ_i and Y_i are defined by (2.1) and (2.2) as functions of the observations S_i. Being a special instance of a 'generalized isotonic regression problem', the MLE $\hat{F}_n$ is known to coincide with the solution of the isotonic regression problem of minimizing the sum of squares

$$\sum_{i=1}^{n} (\Delta_i - F(Y_i))^2,$$

over all distribution functions F. The one step algorithm for computing the solution, using the so-called cusum diagram of the Δ_i's is described, e.g., on p. 30 of [9]. The solution consists of local averages of the Δ_i and can therefore only have rational values.

Specializing Theorem 3.7 in Section 3.8 of [9] to the setting where the inspection times are uniformly distributed on $[0,1]$, yields the asymptotic distribution of the MLE in the uniform deconvolution model with $F_0(1)=1$.

THEOREM 2.1. *Let F_0 be differentiable on (a,b), $0<a<b<1$ with a continuous positive derivative $f_0(t)$ for $t\in(a,b)$, where $[a,b]$ is the support of f_0. Let $\hat{F}_n$ be the nonparametric MLE of F_0. Then, for $t_0\in(a,b)$:*

(2.3)
$$n^{1/3}\{\hat{F}_n(t_0)-F_0(t_0)\}/\left(4f_0(t_0)F_0(t_0)(1-F_0(t_0))\}\right)^{1/3} \xrightarrow{d} \text{argmin}_{t\in\mathbb{R}}\left\{W(t)+t^2\right\},$$

where W is two-sided Brownian motion on $\mathbb{R}$, originating from zero.

2.2. *Case $F_0(1)<1$.* In line with the previous subsection, we will make a connection of the uniform deconvolution problem to the more general IC-m problem, where $m=\lceil M\rceil$, M denoting the upper support point of the distribution corresponding to distribution function F_0.

We start again by constructing a vector Δ and Y based on the observation S. Note that S now takes values in $[0,M+1]$. Define for $1\le j\le m+1$ the j-th component in Δ by

$$\Delta_j = \begin{cases} 1\,, & \text{if } S\in(j-1,j] \\ 0\,, & \text{else.}\end{cases} \tag{2.4}$$

Moreover, define the components of the vector Y by

$$Y_j = S - \lfloor S\rfloor + j - 1 \tag{2.5}$$

for $1\le j\le m+1$.

Similarly to the IC-1 setting, we can show that Y_1 is uniformly distributed on $[0,1]$. Indeed, for $y\in(0,1)$,

$$\begin{aligned}
\mathbb{P}(Y_1\le y) &= \sum_{j=1}^{m+1}\mathbb{P}(Y_1\le y\wedge\Delta_j=1) = \sum_{j=1}^{m+1}\mathbb{P}(Y_1\le y\wedge j-1<S\le j)\\
&= \sum_{j=1}^{m+1}\mathbb{P}(S-(j-1)\le y\wedge j-1<S\le j) = \sum_{j=1}^{m+1}\mathbb{P}(j-1<S\le y+j-1)\\
&= \sum_{j=1}^{m+1}\int_{j-1}^{j-1+y} g_S(s)\,ds = \sum_{j=1}^{m+1}\int_{j-1}^{j-1+y} F_0(s)-F_0(s-1)\,ds = \int_m^{m+y} F_0(s)\,ds
\end{aligned}$$

Again, the right hand side equals y because $F_0(s) = 1$ for $s \ge m$, as $m \ge M$ and $F_0(M) = 1$. So, the vector Y consists of a uniformly distributed random variable Y_1 on $[0,1]$, followed by $Y_j = Y_1 + j - 1$ for $2 \le j \le m$. It is therefore also automatically ordered.

Now, consider the conditional distribution of Δ given Y, which only depends on Y_1. For fixed $j \in \{1, 2, \ldots, m+1\}$ and $y_1 \in (0,1)$, using (1.1)

$$\begin{aligned}\mathbb{P}(\Delta_j = 1 | Y_1 = y_1) &= \frac{g_S(j-1+y_1)}{\sum_{k=1}^{m+1} g_S(k-1+y_1)}\\ &= \frac{F_0(j-1+y_1) - F_0(j-2+y_1)}{F_0(m+y_1) - F_0(y_1-1)} = F_0(y_{(j)}) - F_0(y_{(j-1)}),\end{aligned}$$

because $y_1 - 1 < 0$ and $m + y_1 > m \ge M$. In view of the previous discussion on the IC-m model, this shows that translating the sample of S_i's to the sample of (Y_i, Δ_i)'s according to (2.4) and (2.5), yields IC-m data. The distribution of 'inspection times' is quite specific and peculiar, having distance exactly equal to one with a uniformly distributed first.

Having established this relation between the two models, results known for the IC-m model can be used to derive results for our deconvolution model. However, much less results are known for the IC-m model than for IC-1. One indication of the difficulties in going away from the IC-1 model is that the nonparametric MLE can have irrational values in the IC-m, $m > 1$, model. A simple example of this is given in Example 1.3 on p. 48 of [10]. Here the values of the nonparametric maximum likelihood estimate of the distribution function of the hidden variable contain the values $\frac{1}{2} \pm \frac{1}{6}\sqrt{3}$ (the solution can be computed analytically in this simple example). This cannot happen in the IC-1 model.

Let us now turn to the nonparametric MLE. Define

$$m_n = \max_{j: S_j > 1} (S_j - 1). \tag{2.6}$$

Denoting by $\mathbb{Q}_n$ the empirical distribution of $(S_1, \ldots, S_n)$, the nonparametric MLE is defined as maximizer of the log likelihood function

$$\begin{aligned}\ell_n(F) &= \int \log\{F(s) - F(s-1)\}\, d\mathbb{Q}_n(s)\\ &= \int_{s\le 1} \log F(s)\, d\mathbb{Q}_n(s) + \int_{1<s\le m_n} \log\{F(s) - F(s-1)\}\, d\mathbb{Q}_n(s)\\ &\qquad + \int_{s>1\vee m_n} \log\{1 - F(s-1)\}\, d\mathbb{Q}_n(s).\end{aligned} \tag{2.7}$$

Note that for maximizing this log likelihood function, the class of distribution functions can be restricted to those satisfying $F(S_i) = 1$ if $S_i > \max_{j:S_j>1}(S_j - 1)$, and $F(S_i - 1) = 0$ if $S_i - 1 < \min_j S_j$, for different values of F at these points would make the log likelihood smaller. Define the set of distribution functions.

DEFINITION 2.1. $\mathcal{F}_n$ is the set of discrete distribution functions F, which only have mass at the points $\{S_i, S_i - 1 : 1 \le i \le n\}$ and satisfy

$$F(x) = 1 \quad \text{if} \quad x \ge m_n, \tag{2.8}$$

and

$$F(x) = 0 \quad \text{if} \quad x < \min_j S_j. \tag{2.9}$$

We restrict the distribution functions, occurring in the problem of maximizing the likelihood to this set $\mathcal{F}_n$.

The nonparametric MLE can then be characterized using the following process, defined in terms of $F \in \mathcal{F}_n$

$$\begin{aligned} &W_{n,F}(t) \\ &= \int \frac{\{s \le t \wedge 1\}}{F(s)}\, d\mathbb{Q}_n(s) - \int \frac{\{0 < s-1 \le t < s \le m_n\}}{F(s)-F(s-1)}\, d\mathbb{Q}_n(s) \\ &\quad + \int \frac{\{1 < s \le t \wedge m_n\}}{F(s)-F(s-1)}\, d\mathbb{Q}_n(s) - \int \frac{\{1 \vee m_n - 1 < s-1 \le t\}}{1-F(s-1)}\, d\mathbb{Q}_n(s), \quad t \ge 0, \end{aligned} \tag{2.10}$$

where we write the set as shorthand for its indicator function.

Now, let $T_1 < \cdots < T_m$ be the points S_i such that S_i not of type (2.8) and $S_i - 1$ is not of type (2.9). Then $0 < F(T_i) < 1$ for $i = 1, \ldots, m$, if the log likelihood for F is finite, and we can define:

$$\mathcal{Y} = \{\boldsymbol{y} \in (0,1)^m : \boldsymbol{y} = (y_1, \ldots, y_m) = (F(T_1), \ldots, F(T_m), F \in \mathcal{F}_n\}.$$

The following lemma characterizes the MLE and is proved in a completely analogous way as Lemma 2.2 in [5].

LEMMA 2.1. *Let the class of distribution functions $\mathcal{F}_n$ be as defined in Definition 2.1. Then $\hat{F}_n \in \mathcal{F}_n$ maximizes (2.7) over $F \in \mathcal{F}_n$ if and only if*

(i)

$$\int_{u \in [t,\infty)} dW_{n,\hat{F}_n}(u) \le 0, \qquad t \ge 0, \tag{2.11}$$

(ii)

$$\int \hat{F}_n(t)\, dW_{n,\hat{F}_n}(t) = 0. \tag{2.12}$$

where W_{n,F_n} is defined by (2.10). Moreover, $\hat{F}_n \in \mathcal{F}_n$ is uniquely determined by (2.11) and (2.12).

As mentioned above, in order to compute the MLE in this case, an iterative algorithm has to be used. We use the iterative convex minorant algorithm proposed in [10], using line search to determine the step size, as described in [12]. This algorithm was also applied in computing the MLE for the incubation time distribution in [4]. R scripts for computing the MLE in this way are available in [7]. This algorithm uses the characterization of Lemma 2.1.

The asymptotic distribution of the nonparametric MLE in the IC-m model has only partially been established for the case 2 model in the so-called separated case, in which the distance between observation times stays away from zero. However, its proof is given under the additional assumption that the distance between the observation times has an absolutely continuous distribution, a situation we clearly do not have here. For details, see [3]. One conjecture for the asymptotic distribution of the MLE in the uniform deconvolution problem when $F_0(1) < 1$ would be that the statement of Theorem 2.1 is also valid in this setting. In the next section, we consider the more general mixed uniform deconvolution problem for which asymptotic results have been derived leading to a second conjecture for the asymptotic distribution in the general fixed uniform deconvolution problem.

3. The mixed model. The *mixed uniform deconvolution model* as described in the introduction (so with the random length E of the support of the uniform noise V) is clearly a generalization of the what we from now will call the *fixed uniform deconvolution model*. The first reducing to the latter in case E has the degenerate distribution at the point 1. Let M_1 and M_2 be the upper support points of the distributions corresponding to E and U respectively, so that $M = M_1 + M_2$ is the upper support point of the distribution of $S = U + V$. Denoting by μ the product of the measure dF_E and Lebesgue measure on $[0, M]$, (E_i, S_i) has (convolution) density q_F with respect to μ

$$\begin{aligned} q_F(e,s) &= e^{-1}\{F(s) - F(s-e)\} \\ &= e^{-1}\int_{u=(s-e)_+}^{s} dF(u), \qquad e > 0,\, s \in [0, M]. \end{aligned} \tag{3.1}$$

We define the underlying measure Q_0 for (E_i, S_i) by

$$dQ_0(e,s) = q_F(e,s)\,ds\,dF_E(e), \qquad s \in [0,M], \qquad e \in (0, M_2]. \tag{3.2}$$

For estimating the distribution function F_0 of U, usually parametric distributions are used, like the Weibull, log-normal or gamma distribution. However, in [4] the nonparametric MLE is used. This maximum likelihood estimator $\hat{F}_n$ maximizes the function

$$\ell(F) = n^{-1}\sum_{i=1}^{n} \log\left(F(S_i) - F(S_i - E_i)\right) \tag{3.3}$$

over *all* distribution functions F on $\mathbb{R}$ which satisfy $F(x) = 0$, $x \le 0$, see [4]. The monotonicity and boundedness of F (between 0 and 1) ensures that this maximization problem has a solution. Note that the random variable E_i now replaces the value 1 in $F(S_i - 1)$ compared to the fixed uniform deconvolution problem, but that for the rest the model is the same.

The asymptotic distribution of the nonparametric MLE in the mixed uniform deconvolution problem is given as Theorem 4.1 in [5].

THEOREM 3.1. *Let the conditions of Theorem 4.1 in [5] be satisfied. In particular, let E_i stay away from zero and have an absolutely continuous distribution. Then*

$$n^{1/3}\{\hat{F}_n(t_0) - F_0(t_0)\}/(4f_0(t_0)/c_E)^{1/3} \xrightarrow{d} \text{argmin}\left\{W(t) + t^2\right\}, \tag{3.4}$$

where W is two-sided Brownian motion on $\mathbb{R}$, originating from zero, and where the constant c_E is given by:

$$c_E = \int e^{-1}\left[\frac{1}{F_0(t_0) - F_0(t_0 - e)} + \frac{1}{F_0(t_0+e) - F_0(t_0)}\right] dF_E(e). \tag{3.5}$$

This result is valid, irrespective of whether the support of distribution corresponding to F_0 is contained in $[0, 1]$ or not.

Here we see a condition for that result, namely that E has an absolutely continuous distribution, a condition clearly not met by the degenerate distribution at 1. Ignoring this for the moment, the statement of the theorem would yield the asymptotic distribution with variance

$$(4f_0(t_0)/c_E)^{2/3}\,\text{Var}\left(\text{argmin}_{t\in\mathbb{R}}\left\{W(t) + t^2\right\}\right),$$

where,

$$\begin{aligned} c_E &= \int e^{-1}\left[\frac{1}{F_0(t_0) - F_0(t_0 - e)} + \frac{1}{F_0(t_0+e) - F_0(t_0)}\right] dF_E(e) \\ &= \frac{1}{F_0(t_0) - F_0(t_0 - 1)} + \frac{1}{F_0(t_0+1) - F_0(t_0)}. \end{aligned} \tag{3.6}$$

Now, if $F_0(1)=1$, this expression (for $t_0\in(0,1)$) reduces to

$$c_E=\frac{1}{F_0(t_0)}+\frac{1}{1-F_0(t_0)}=\frac{1}{F_0(t_0)(1-F_0(t))}. \tag{3.7}$$

This is precisely the constant in the asymptotic variance of Theorem 2.1, proved in the case $F_0(1)=1$. Now two natural conjectures come up. The first is that also for the setting where $F_0(1)<1$ for c_E will equal (3.7). The second being that it equals (3.6). In Section 5 we will further investigate this asymptotic variance, using a simulation study as well as asymptotic calculations involving so-called smooth functionals. In section 4, we discuss asymptotic results for this type of functionals.

4. Interlude; examples of applications of smooth functional theory in the fixed model.

Though estimating the distribution function at a fixed point in the uniform deconvolution function is intrinsically more complicated than based on a direct sample from the unknown distribution, reflected in a slower rate of estimation than the 'parametric rate' $\sqrt{n}$, there are functionals of F_0 that can be estimated at rate $\sqrt{n}$ in the uniform deconvolution model. For the functionals discussed in this section, we have known (normal) asymptotic behavior of the functions of the nonparametric MLE. The theory does not depend on whether the support of the distribution, corresponding to F_0, is contained in $[0,1]$ or not. We illustrate the theory with F_0 the truncated exponential distribution function on $[0,2]$, given by

$$F_0(x)=\frac{1-\exp\{-x\}}{1-\exp\{-2\}}1_{[0,2]}(x)+1_{(2,\infty)}(x). \tag{4.1}$$

Note that in this case the support of f_0 is not contained in $[0,1]$.

By "smooth functionals" we mean both global smooth functionals such as the "mean functional" and local smooth functionals such as smooth approximations of the distribution function and density, which can be estimated by applying kernel smoothing to the nonparametric MLE. These estimates are asymptotically normal. For deriving the asymptotic behavior, we establish a representation in the "observation space" of the functional in the "hidden space", using the score function. The setup is briefly described in the appendix (section 7.1). For a more detailed discussion of the smooth functional theory, see [9].

EXAMPLE 4.1. Using the theory of Section 7 for the functional $F\mapsto\int x\,dF(x)$, we get from Lemma 7.1 for the score function θ_F on $[0,3]$:

$$\theta_F(x)=\begin{cases}-\{1-F(x)\}-\{1-F(x+1)\}, & x\in[0,1],\\ F(x-1)-\{1-F(x)\}, & x\in(1,2],\\ F(x-2)+F(x-1), & x\in(2,3].\end{cases} \tag{4.2}$$

For $F=F_0$ as in (4.1), this becomes:

$$\theta_F(x)=\begin{cases}\bigl(2e^x-e(1+e)\bigr)/\bigl((e^2-1)e^x\bigr), & x\in[0,1],\\ \bigl((1+e^2)e^x-e^2(1+e)\bigr)/\bigl((e^2-1)e^x\bigr), & x\in(1,2],\\ (2e^2-(1+e)e^{3-x})/(e^2-1), & x\in(2,3].\end{cases} \tag{4.3}$$

A picture of this function is given in Figure 1.

The asymptotic variance of $\sqrt{n}\int x\,d\hat F_n(x)$ is given by:

$$\sigma_{F_0}^2=\int_0^3\theta_{F_0}(x)^2\{F_0(x)-F_0(x-1)\}\,dx\approx 0.357915,$$

and we have:

$$\sqrt{n}\left\{\int x\,d\hat F_n(x)-\int x\,dF_0(x)\right\}\xrightarrow{\mathcal{D}}N(0,\sigma_{F_0}^2),$$

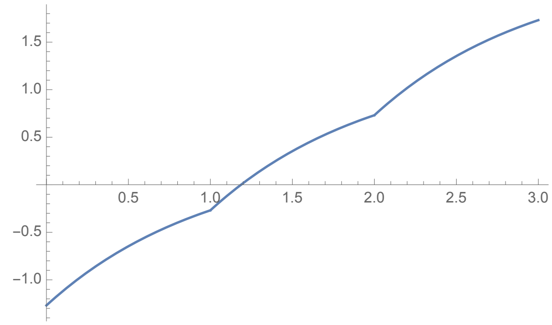


Fig 1: The score function θ_{F_0} for the truncated exponential distribution, given by (4.3).

where $N(0,\sigma_{F_0}^2)$ is a normal distribution with expectation zero and variance $\sigma_{F_0}^2$. One can in fact consistently estimate the asymptotic variance $\sigma_{F_0}^2$ by

$$\int \theta_{\hat F_n}(x)^2\{\hat F_n(x)-\hat F_n(x-1)\}\,dx,$$

where one takes $F=\hat F_n$ in (4.2). Note that the rate of convergence is $\sqrt{n}$ instead of the cube root n rate, expected for $\hat F_n(t)$ itself.

The asymptotic efficiency of this estimate of the first moment is proved in [15], see in particular example 11.2.3e on p. 230 of [15]. It beats obvious moment estimates like $\bar S_n - 1/2$, because of the particular form of the characteristic function of the Uniform distribution (having zeroes). But these matters are not the main concern of our paper.

EXAMPLE 4.2. We can estimate the density f_0 by the density estimator

$$\hat f_{nh}(t)=\int K_h(t-x)\,d\hat F_n(x)$$

where $h>0$ is a bandwidth and $K_h(u)=h^{-1}K(u/h)$ for a symmetric smooth kernel K, for example the triweight kernel

$$K(u)=\tfrac{35}{32}(1-u^2)^3 1_{[-1,1]}(u).$$

This estimator was considered in [8]. It was shown that, under the condition given there:

$$\left(nh^3\right)^{1/2}\left\{\hat f_{nh}(t)-\int K_h(t-x)\,dF_0(x)\right\}\xrightarrow{\mathcal{D}}N(0,\sigma_t^2),$$

where

$$\sigma_t^2=\lim_{h\downarrow 0}h^3\int \theta_{h,t,F_0}^2(s)\{F_0(s)-F_0(s-1)\}\,ds, \tag{4.4}$$

and θ_{h,t,F_0} is given by

$$\theta_{h,t,F_0}(x)=\sum_{i=0}^{m-1}\{1-F_0(x+i)\}K_h'(t-(x+i)),\qquad x\in[0,1].$$

Here $\theta_{h,t,F_0}(x+i)=\theta_{h,t,F_0}(x+i-1)-K_h'(t-(x+i-1))$, $i=1\dots,m$ and $m=\lceil M\rceil$, where M is the upper bound of the support of f_0.

We show that σ_t^2, given by (4.4) can be simplified to

$$\sigma_t^2=F_0(t)\{1-F_0(t)\}\int K'(u)^2\,du. \tag{4.5}$$

We have, for $t \in (0,1)$:

$$
\begin{aligned}
&\int \theta_{h,t,F_0}^2(x)\{F_0(x)-F_0(x-1)\}\,dx\\
&=\int_{x=t-h}^{t+h} K_h'(t-x)^2\{1-F_0(x)\}^2F_0(x)\,dx\\
&\qquad+\int_{x=t-h}^{t+h} K_h'(t-x)^2F_0(x)^2\{F_0(x+1)-F_0(x)\}\,dx\\
&\qquad+\dots\\
&\qquad+\int_{x=t-h}^{t+h} K_h'(t-x)^2F_0(x)^2\{1-F_0(x+m-1)\}\,dx\\
&=\int_{x=t-h}^{t+h} K_h'(t-x)^2\{1-F_0(x)\}^2F_0(x)\,dx\\
&\qquad\qquad+\int_{x=t-h}^{t+h} K_h'(t-x)^2F_0(x)^2\{1-F_0(x)\}\,dx\\
&=\int_{x=t-h}^{t+h} K_h'(t-x)^2F_0(x)\{1-F_0(x)\}\,dx\\
&\sim h^{-3}F_0(t)\{1-F_0(t)\}\int K'(u)^2\,du,\qquad h\downarrow 0,
\end{aligned}
$$

using the telescoping sum $F_0(x+1)-F_0(x)+\dots+1-F_0(x+m-1)$.

For $t \in (i,i+1)$, $i \ge 1$ we get a similar argument and for $t=i$, $i=1,2,\dots$ we get the result from the contributions from both sides of i, which each contribute one half to the total mass.

EXAMPLE 4.3. In an entirely similar way, the asymptotic distribution of the estimator

$$\tilde F_{nh}(t)=\int I\!\!K_h(t-x)\,d\hat F_n(x)$$

of the distribution function F_0 itself can be derived, where $I\!\!K_h(u)=I\!\!K(u/h)$ and $I\!\!K$ is the integrated kernel K, defined by

$$I\!\!K(u)=\int_{-\infty}^u K(x)\,dx.$$

We have the following result.

THEOREM 4.1. *Let F_0 have a continuous positive density f_0 on $(0,M)$, where $M>0$, and let $t_0 \in (0,M)$. Then, if $nh\to\infty$ and $h\downarrow 0$:*

$$\bigl(nh\bigr)^{1/2}\left\{\tilde F_{nh}(t)-\int I\!\!K_h(t-x)\,dF_0(x)\right\}\xrightarrow{\mathcal{D}} N(0,\tilde\sigma_t^2),$$

where

$$\tilde\sigma_t^2=F_0(t)\{1-F_0(t)\}\int K(u)^2\,du.$$

The proof proceeds along a similar path as the proof of the analogous result for the density in the preceding example. The truncated exponential distribution function, defined by (4.1) satisfies the conditions of the theorem, for $M=2$.

5. Simulations and asymptotic considerations. Having two competing conjectures for the asymptotic distribution of $\hat{F}_n(t_0)$ in the fixed uniform deconvolution problem, related to the constants (3.6) and (3.7) in the asymptotic variance, in this section we will first present a simulation study supporting (3.7). One could say that "plugging in" the point mass at 1 as distribution for E should be possible in the mixed case setting in case $F_0(1) = 1$. In the case $F_0(1) < 1$, however, this does not seem to lead to the right result. Diving into the proof of Theorem 4.1 in [5], we identify a term that in case $F_0(1) = 1$ is identically zero. But if $F_0(1) < 1$ this term is not zero and has order $O_p(n^{-5/6})$ in the mixed model, but order $O_p(n^{-2/3})$ in the fixed model, which means that it is negligible in the mixed model but not negligible in the asymptotic expansion for the fixed model.

To simulate the mixed model, we first generate the random pairs

$$(U_i, E_i), \qquad U_i \sim F_0, \qquad E_i \sim F_E = (\cdot - 0.5)\, 1_{[0.5,1.5)} + 1_{[1.5,\infty)}, \qquad i = 1, \ldots, n.$$

Then, conditionally on E_i a Uniform$(0, E_i)$ random variable V_i is drawn. Our observations then consist of

$$(E_i, S_i), \quad \text{where} \quad S_i = U_i + V_i, \qquad i = 1, \ldots, n.$$

This example satisfies the conditions of Theorem 4.1 in [5], which means that the asymptotic variance of the nonparametric maximum likelihood estimator $\hat{F}_n$ of F_0 evaluated at $t_0 \in (0, 2)$ is given by:

$$\sigma_{t_0}^2 = (4f_0(t_0)/c_E)^{-2/3}\, \mathrm{var}\left(\mathrm{argmin}_{t\in\mathbb{R}} \left\{W(t) + t^2\right\}\right), \tag{5.1}$$

where c_E is given by (3.5). It was computed numerically in [11], using Airy functions, that

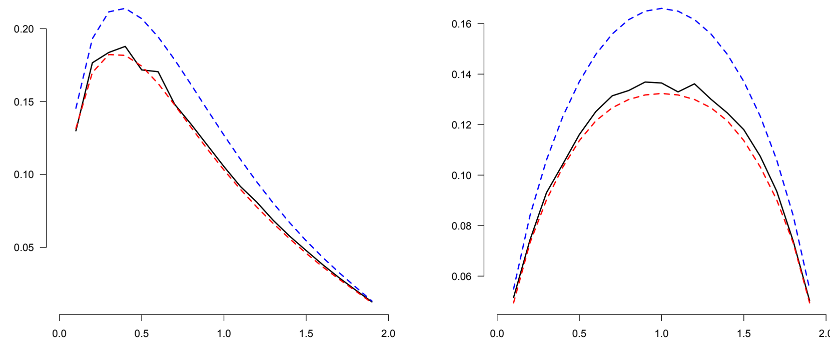
$$\mathrm{var}\left(\mathrm{argmin}_{t\in\mathbb{R}} \left\{W(t) + t^2\right\}\right) \approx 0.263555964.$$

To illustrate Theorem 4.1 in [5] we can simulate from the model and compute the variances of $\hat{F}_n(t_i)$ for, say 10,000, samples for $t_i = i \cdot 0.1$, $i = 1, \ldots, 19$ and compare the variances of $\hat{F}_n(t_i)$ with the asymptotic variances $\sigma_{t_i}^2$, taking $t_0 = t_i$ in (5.1). A comparison of simulated and asymptotic values is shown in Figure 2, where we take $n = 10{,}000$, and F_0 the truncated standard exponential and the uniform distribution function on $[0, 2]$, respectively.

If $M > 1$, as in this example where $M = 2$, the values of the asymptotic variances are no longer of the form (5.1) in the fixed model. We shall now explain the reason of this discrepancy. We start with the characterization of Lemma 2.1. The MLE is determined by the process $W_{\hat{F}_n}$ and, just as in [5], we now investigate the terms in a local expansion of $W_{\hat{F}_n}$. To this end, we first define the process X_n.

LEMMA 5.1. *Let F_0 have a continuous positive density f_0 on $(0, M)$, where $M > 1$, let $t_0 \in (0, M)$ and let the process X_n be defined by:*

$$\begin{aligned} X_n(t) = &\int_{t_0 < s \le t_0 + n^{-1/3}t} \frac{\delta_1(s)}{\hat{F}_n(s)}\, d(\mathbb{Q}_n - Q_0)(s) \\ &- \int_{t_0 < s-1 \le t_0 + n^{-1/3}t} \frac{\delta_2(s)}{\hat{F}_n(s) - \hat{F}_n(s-1)}\, d(\mathbb{Q}_n - Q_0)(s) \\ &+ \int_{t_0 < s \le t_0 + n^{-1/3}t} \frac{\delta_2(s)}{\hat{F}_n(s) - \hat{F}_n(s-1)}\, d(\mathbb{Q}_n - Q_0)(s) \\ &- \int_{t_0 < s-1 \le t_0 + n^{-1/3}t} \frac{\delta_3(s)}{1 - \hat{F}_n(s-1)}\, d(\mathbb{Q}_n - Q_0)(s), \end{aligned} \tag{5.2}$$



(a) (b)

Fig 2: Simulated variances, times $n^{2/3}$, for the mixed model of $\hat F_n(t_i)$, for $t_i = 0.1, 0.2, \dots, 1.9$ (blue solid curve, linearly interpolated between values at the t_i), compared with the asymptotic values (5.1) of Theorem 4.1 in [5] (red, dashed) for E_i uniform on $[0.5, 1.5]$. The simulated variances are based on $10,000$ simulations of samples of size $n = 10,000$ for (a) F_0 the truncated exponential distribution function on $[0, 2]$ and (b) F_0 the uniform distribution function on $[0, 2]$. The blue dashed curves are the corresponding asymptotic variance curves of Conjecture 5.1 for the fixed model.

where $\delta_1 = 1_{[0,1]}$, $\delta_2 = 1_{(1,m_n]}$ *and* $\delta_3 = 1_{(m_n,\infty)}$ *and* $\mathbb{Q}_n$ *is the empirical distribution measure of the* S_i *with corresponding underlying measure* Q_0. *Then* $n^{2/3}X_n$ *converges in distribution, in the Skorohod topology, to the process*

$$t \mapsto \sqrt{c}W(t), \qquad t \in \mathbb{R},$$

where c *is defined by*

$$c = \frac{1}{F_0(t_0) - F_0(t_0 - 1)} + \frac{1}{F_0(t_0 + 1) - F_0(t_0)}\,.$$

PROOF. The proof follows the proof of Lemma 9.4 in [5]. □

We now have:

LEMMA 5.2.

$$W_{n,\hat F_n}(t_0 + n^{-1/3}t) - W_{n,\hat F_n}(t_0) = X_n(t) + Y_n(t),$$

where X_n *is defined by (5.2) and* Y_n *by*

$$Y_n(t) = \int_{t_0 < s \le t_0 + n^{-1/3}t} \left\{ \frac{F_0(s) - F_0(s-1)}{\hat F_n(s) - \hat F_n(s-1)} - \frac{F_0(s+1) - F_0(s)}{\hat F_n(s+1) - \hat F_n(s)} \right\} ds. \tag{5.3}$$

We can write

$$Y_n(t) = A_n(t) + B_n(t),$$

where

$$A_n(t) = -\int_{s \in [t_0, t_0 + n^{-1/3}t)} \left\{ \hat F_n(s) - F_0(s) \right\} \left\{ \frac{1}{\hat F_n(s) - \hat F_n(s-1)} + \frac{1}{\hat F_n(s+1) - \hat F_n(s)} \right\} ds,$$

and

(5.4)
$$B_n(t) = \int_{s\in[t_0,t_0+n^{-1/3}t)} \left\{ \frac{\hat F_n(s-1) - F_0(s-1)}{\hat F_n(s) - \hat F_n(s-1)} + \frac{\hat F_n(s+1) - F_0(s+1)}{\hat F_n(s+1) - \hat F_n(s)} \right\} ds.$$

The crucial difference between the mixed model and the fixed model is the term $B_n(t)$ in the expansion of $W_{n,\hat F_n}$ in Lemma 5.2. In the mixed model the corresponding term

(5.5)
$$\tilde B_n(t) = \int_{s\in[t_0,t_0+n^{-1/3}t)} \int e^{-1} \left\{ \frac{\hat F_n(s-e) - F_0(s-e)}{\hat F_n(s) - \hat F_n(s-e)} + \frac{\hat F_n(s+e) - F_0(s+e)}{\hat F_n(s+e) - \hat F_n(s)} \right\} dF_E(e)\, ds$$

is shown in [5] to be of order $O_p(n^{-5/6})$ in the leading expansion, but (5.4) is of order $O_p(n^{-2/3})$. In fact, the terms $A_n(t)$ and $B_n(t)$ are both of order $O_p(n^{-2/3})$ in the fixed model, whereas the corresponding terms are of order $O_p(n^{-2/3})$ and $O_p(n^{-5/6})$, respectively, in the mixed model, see the proof of Lemma 9.8 and Lemma 9.9 in [5], respectively. Integration w.r.t. dF_E, where F_E is absolutely continuous, causes the lower order of the term. Showing that (5.5) is of order $O_p(n^{-5/6})$, however, is not at all easy and is in fact the main effort in [5]. We need to apply the smooth functional theory here, as discussed in Section 4.

We have the following conjecture on the asymptotic behavior of $\hat F_n$ if V_i is Uniform$(0,1)$.

CONJECTURE 5.1. *Let F_0 have a continuous positive density f_0 on $(0,M)$, where $M > 0$, and let $t_0 \in (0,M)$. Let $\hat F_n$ be the nonparametric MLE of F_0. Then, for $t_0 \in (a,b)$:*

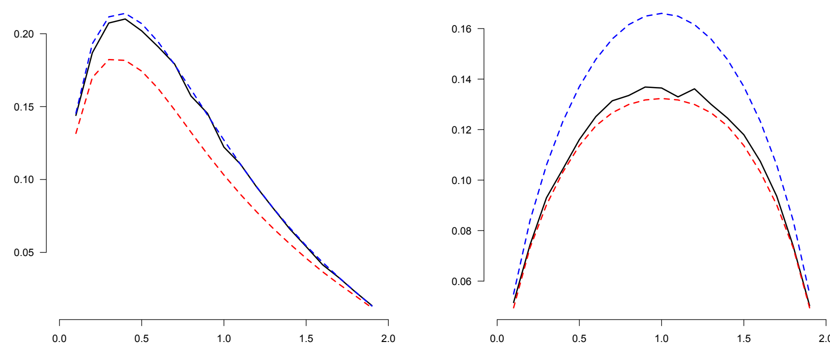
(5.6)
$$n^{1/3}\{\hat F_n(t_0) - F_0(t_0)\}/\left(4f_0(t_0)F_0(t_0)(1-F_0(t_0))\}\right)^{1/3} \xrightarrow{d} \text{argmin}_{t\in\mathbb{R}}\left\{W(t)+t^2\right\},$$

where W is two-sided Brownian motion on $\mathbb{R}$, originating from zero.

So in this case we expect the MLE to have exactly the same limit behavior as in the case that the support of the distribution is contained in $[0,1]$ (see Theorem 2.1). The reason we believe the conjecture might be true partly relies on simulations and partly on the behavior of the smooth functionals in Section 4, where also the $F(1-F)$ "Bernoulli factor" enters into the variance. If the conjecture is true, we think it enters via telescoping sums, as we demonstrated for the density estimator, based on the MLE, but we have not been able to prove this.

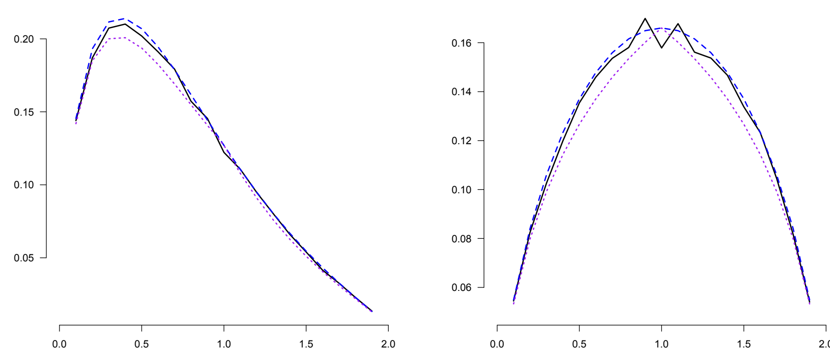
We finally show in Figure 4 a picture of the variance curve for 10,000 samples of size $n = 10,000$, where the blue dashed curve shows the theoretical curve of Conjecture 5.1 and the purple dotted curve the theoretical curve if we would apply Theorem 4.1 in [5] with E degenerate at 1 (ignoring the conditions of Theorem 4.1 in [5]).

The "dip" of the empirical variance curves at the point 1 if F_0 is the uniform distribution function on $[0,2]$, which was also visible for the mixed model in Figure 2, is remarkable, but we do not have an explanation for it. Note that the theoretical curves touch at this point, because the theoretical variance at the point $t=1$ contains $F_0(t)\{1-F_0(t)\}$ for both conjectures there.



(a) (b)

Fig 3: Simulated variances, times $n^{2/3}$, for the fixed model of $\hat{F}_n(t_i)$, for $t_i = 0.1, 0.2, \ldots, 1.9$ (blue solid curve, linearly interpolated between values at the t_i), compared with the asymptotic values of Conjecture 5.1 (blue, dashed). The simulated variances are based on $10,000$ simulations of samples of size $n = 10,000$ for (a) F_0 the truncated exponential distribution function on $[0, 2]$ and (b) F_0 the uniform distribution function on $[0, 2]$. The red dashed curves are the corresponding asymptotic variance curves of Theorem 4.1 in [5] for the mixed model and E_i uniform on $[0.5, 1.5]$.



(a) (b)

Fig 4: Simulated variances, times $n^{2/3}$, for the fixed model of $\hat{F}_n(t_i)$, for $t_i = 0.1, 0.2, \ldots, 1.9$ (black solid curve, linearly interpolated between values at the t_i), compared with the asymptotic values of Conjecture 5.1 (blue, dashed) and the conjecture that would follow from Theorem 4.1 in [5], ignoring the conditions (purple, dotted), for (a) F_0 the truncated exponential distribution function on $[0, 2]$ and (b) F_0 the uniform distribution function on $[0, 2]$. The simulated variances are based on $10,000$ simulations of samples of size $n = 10,000$.

6. Concluding remarks. We showed that the asymptotic distribution of the nonparametric maximum likelihood estimator (MLE) in the uniform deconvolution model can de derived from a corresponding result for the current status model in the case that the support of the distribution is contained in the unit interval. If the support of the unknown distribution is not contained in the unit interval, the asymptotic distribution is unknown. But also in this setting, the model is shown to be related to an interval censoring model, but now case m for $m \ge 2$.

The asymptotics in the mixed uniform deconvolution model (where the length of the support of the uniform variable is random) was studied in [5] under a smoothness condition on the distribution of the interval length. In that setting there is no distinction depending on the support of the distribution corresponding to F_0. In case $F_0(1) = 1$, the statement of the theorem reduces to that for the current status model if the (non-smooth) degenerate distribution is substituted. A natural question is whether the same can be done in case $F_0(1) < 1$.

Simulations indicate this cannot be done. Moreover, an important term in the asymptotic analysis of the nonparametric MLE in the mixed uniform deconvolution problem turns out to behave essentially differently, depending on whether $F_0(1) = 1$ or $F_0(1) < 1$. This discrepancy is explained from smooth functional theory, the strength of which is also demonstrated using more easily understandable functionals in Section 4.

`R` scripts for computing the MLE and producing the pictures in this paper are available in [7].

7. Appendix.

7.1. *Score equations.* As explained in the Appendix of [4], the theory of the estimation of smooth functionals is based on certain score equations. We define the score function θ_F by:

$$\theta_F(s) = E\left\{a(X)|S=s\right\} = \frac{\int_{x\in(s-1,s]} a(x)\,dF(x)}{F(s)-F(s-1)}\,,$$

(compare to (A.1) in [4]). This is the conditional expectation of $a(X)$ in the "hidden" space of the variable of interest, given our observation S. Note that we changed the notation somewhat w.r.t. [4], and denote the distribution function of the incubation time by F instead of G.

Defining

$$\phi_F(t) = \int_{x\in[0,t]} a(x)\,dF(x),$$

we get the following representation for the score function, conditioned on $X = x$:

$$E\{\theta_F(S)|X=x\} = \int_{s\in(x,x+1]} \frac{\phi_F(s)-\phi_F(s-1)}{F(s)-F(s-1)}\,ds. \tag{7.1}$$

By differentiation w.r.t. x we get the following equation, with on the right the derivative w.r.t. x of the functional we want to estimate, denoted by ψ:

$$\frac{\phi_F(x+1)-\phi_F(x)}{F(x+1)-F(x)} - \frac{\phi_F(x)-\phi_F(x-1)}{F(x)-F(x-1)} = \psi(x), \qquad x\in[0,M]. \tag{7.2}$$

Here and in the sequel, we assume that distribution functions are right-continuous.

We have the following lemma.

LEMMA 7.1. *The solution of the system (7.2) is found in the following way. Let $m = \lceil M \rceil$ and let $\theta_F(x)$ be given by*

$$\theta_F(x) = -\sum_{i=0}^{m-1}\{1-F(x+i)\}\psi(x+i), \qquad x\in[0,1].$$

and $\theta_F(x+i) = \theta_F(x+i-1) + \psi(x+i-1)$, $i = 1\dots,m+1$.

Then ϕ_F *is given by*

$$\phi_F(x) = F(x)\theta_F(x), \qquad x \in [0,1],$$

and

$$\phi_F(x+i) - \phi_F(x+i-1) = \{F(x+i) - \{F(x+i-1)\}\theta_F(x+i), \qquad i = 1,\dots,m.$$

REMARK 7.1. This is in accordance with Example 11.2.3e on p. 230 of [15], where $b_{F_0} = \theta_{F_0}$. It is also in line with formula (12) in Theorem 2 in [8].

REFERENCES

[1] BACKER, J. A., KLINKENBERG, D. and WALLINGA, J. (2020). Incubation period of 2019 novel coronavirus (2019-nCov) infections among travellers from Wuhan, China, 20-28 January 2020. *Euro Surveill.* **25**.

[2] BRITTON, T. and SCALIA TOMBA, G. (2019). Estimation in emerging epidemics: bases and remedies. *J. R. Soc. Interface* **16**.

[3] GROENEBOOM, P. (1996). Lectures on inverse problems. In *Lectures on probability theory and statistics (Saint-Flour, 1994)*. *Lecture Notes in Math.* **1648** 67–164. Springer, Berlin. https://doi.org/10.1007/BFb0095675 MR1600884 (99c:62092)

[4] GROENEBOOM, P. (2021). Estimation of the incubation time distribution for COVID-19. *Stat. Neerl.* **75** 161–179. https://doi.org/10.1111/stan.12231 MR4245907

[5] GROENEBOOM, P. (2024a). Nonparametric estimation of the incubation time distribution. *Electron. J. Stat.* **18** 1917–1969. https://doi.org/10.1214/24-ejs2243 MR4736274

[6] GROENEBOOM, P. (2024b). Estimation of the incubation time distribution in the singly and doubly interval censored model. *Stat. Neerl.* **78** 617–635. https://doi.org/10.1111/stan.12335 MR4827421

[7] GROENEBOOM, P. (2025). Deconvolution. https://github.com/pietg/deconvolution.

[8] GROENEBOOM, P. and JONGBLOED, G. (2003). Density estimation in the uniform deconvolution model. *Statist. Neerlandica* **57** 136–157. https://doi.org/10.1111/1467-9574.00225 MR2035863 (2004k:62121)

[9] GROENEBOOM, P. and JONGBLOED, G. (2014). *Nonparametric Estimation under Shape Constraints*. Cambridge Univ. Press, Cambridge.

[10] GROENEBOOM, P. and WELLNER, J. A. (1992). *Information bounds and nonparametric maximum likelihood estimation*. *DMV Seminar* **19**. Birkhäuser Verlag, Basel. MR1180321 (94k:62056)

[11] GROENEBOOM, P. and WELLNER, J. A. (2001). Computing Chernoff's distribution. *J. Comput. Graph. Statist.* **10** 388–400. https://doi.org/10.1198/10618600152627997 MR1939706

[12] JONGBLOED, G. (1998). The iterative convex minorant algorithm for nonparametric estimation. *J. Comput. Graph. Statist.* **7** 310–321. https://doi.org/10.2307/1390706 MR1646718

[13] O'SULLIVAN, F. and ROY CHOUDHURY, K. (2001). An analysis of the role of positivity and mixture model constraints in Poisson deconvolution problems. *J. Comput. Graph. Statist.* **10** 673–696. https://doi.org/10.1198/106186001317243395 MR1938974

[14] REICH, N. G., LESSLER, J., CUMMINGS, D. A. T. and BROOKMEYER, R. (2009). Estimating incubation period distributions with coarse data. *Stat. Med.* **28** 2769–2784. https://doi.org/10.1002/sim.3659 MR2750164

[15] VAN DE GEER, S. A. (2000). *Applications of empirical process theory*. *Cambridge Series in Statistical and Probabilistic Mathematics* **6**. Cambridge University Press, Cambridge. MR1739079 (2001h:62002)