ESTIMATION OF THE INCUBATION TIME DISTRIBUTION IN THE SINGLY AND DOUBLY INTERVAL CENSORED MODEL

BY PIET GROENEBOOM

Delft University of Technology

Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands.
P.Groeneboom@tudelft.nl

We analyze nonparametric estimators for the distribution function of the incubation time in the singly and doubly interval censoring model. The classical approach is to use parametric families like Weibull, log-normal or gamma distributions in the estimation procedure. We propose nonparametric estimates which stay closer to the data than the classical parametric methods. We also give explicit limit distributions for discrete versions of the models and apply this to compute confidence intervals. The methods complement the analysis of the continuous model in [6] and [7]. R scripts for computation of the estimates are provided in [5].

1. Introduction. In [6] estimation of the incubation time distribution for COVID-19 was considered in the situation where the data are *singly interval censored*. In this case one has an observation interval $[E_L, E_R]$ which is known to contain the time of infection and a time S where the person becomes symptomatic. Following [2] we set the left endpoint of the exposure interval $[E_L, E_R]$ equal to zero ("looking back"). Our observations then consist of the pairs of (lengths of) exposure times and times of getting symptomatic

$$(E, S), \qquad i = 1, \ldots, n,$$

where $S = I + U$ (shifted for taking $E_L = 0$), and I is the (length of) infection time and U (length of the) incubation time.

The times I and U are assumed independent, given E, and are not observable. To ensure that the distribution functin F of the incubation time U is identifiable, we will assume that the time till infection is uniformly distributed on the interval $[0, E]$, conditionally on the length of the exposure time E. The model is for example considered in [18], [2], [1] and [6].

We define the (convolution) density q_F of (E, S) by

$$\begin{aligned} q_F(e,s) &= e^{-1}\{F(s) - F(s-e)\} \\ &= e^{-1}\int_{u=(s-e)_+}^{s} dF(u) = e^{-1}\{F(s) - F(s-e), \qquad e > 0,\ s \geq 0. \end{aligned} \tag{1.1}$$

w.r.t. μ, which is the product of the measure dF_E of the exposure time E and Lebesgue measure. The distribution function F satisfies $F(x) = 0$ for $x \leq 0$. So the underlying measure Q_F for (E, S) is defined by

$$dQ_F(e,s) = q_F(e,s)\, ds\, dF_E(e), \qquad e \in (0, M_2], \qquad s \geq 0, \tag{1.2}$$

where $M_2 < \infty$ is the upper bound of the support of E.

One usually assumes that the underlying distribution function F_0 is absolutely continuous, with density f_0, and that E has no mass on an interval $[0, \epsilon)$, for some $\epsilon > 0$. The probability

AMS 2000 subject classifications: Primary 62G05, 62N01; secondary 62-04.
Keywords and phrases: confidence intervals.

measure Q_F, defined by (1.2) may look somewhat peculiar, but one can verify, for example by taking E uniform on $[1, 20]$ and F_0 uniform on $[0, 20]$ that we have indeed

$$\int dQ_F(e, s) = 1.$$

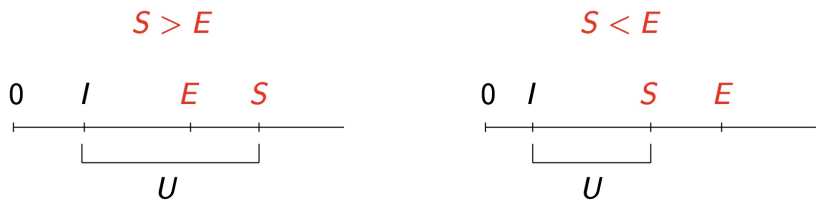


Fig 1: Singly interval censored data. E is the end of the exposure time, S the time of becoming symptomatic, I infection time and U the (length of the) incubation time. We only can observe E and S.

Observations of this type are shown schematically in Figure 1. In practice, however, the variables $S = I + U$ are usually truncated, taking the floor of S, to integers (days), in which case the log likelihood becomes, conditionally on the values of E,

$$\log \int_{s=\lfloor S \rfloor}^{\lceil S \rceil} q_F(E, s)\, ds = \log \left\{ E^{-1} \int_{s=\lfloor S \rfloor}^{\lceil S \rceil} \{F(s) - F(s - E)\}\, ds \right\}, \tag{1.3}$$

where $\lfloor S \rfloor$ and $\lceil S \rceil$ are the floor and ceil of S, respectively. Note that S itself is assumed to have a continuous distribution and that S therefore lies with probability one strictly between two consecutive integers, so we integrate over an interval of length 1.

Usually the exposure times E are also only known as days, so represented by integers, as we will do in the sequel. In principle we can also consider a continuous exposure time E, but we will not do this to avoid an overcomplicated model.

So instead of the parameters $F_0(S)$ and $F_0(S - E)$, we consider estimating the parameters

$$\int_{s=\lfloor S \rfloor}^{\lceil S \rceil} F_0(s)\, ds, \quad \text{and} \quad \int_{s=\lfloor S \rfloor}^{\lceil S \rceil} F_0(s - E)\, ds, \tag{1.4}$$

where $E \in \mathbb{N}$. For a sample $(E_1, S_1), \ldots, (E_n, S_n)$ we get the log likelihood

$$\ell(F) = \sum_{i=1}^{n} \log\{F(S_i) - F(S_i - E_i)\}, \tag{1.5}$$

where we assume that the S_i and E_i are integers.

We can set up a simulation of the model in the following way. We generate the incubation time from a known distribution, for example the truncated Weibull distribution, with parameters a and b, given by the distribution function

$$F_0(x) = \begin{cases} 0, & x < 0, \\ \{1 - \exp(-b\, x^a)\}/\{1 - \exp(-b\, M_1^a)\}, & x \in [0, M_1], \\ 1, & x > M_1. \end{cases} \tag{1.6}$$

Next we define the distribution F_E, for example by the discrete uniform distribution on $\{1, 2, \ldots, M_2\}$. Then we generate random variables (E_i, S_i) by generating E_i from F_E and

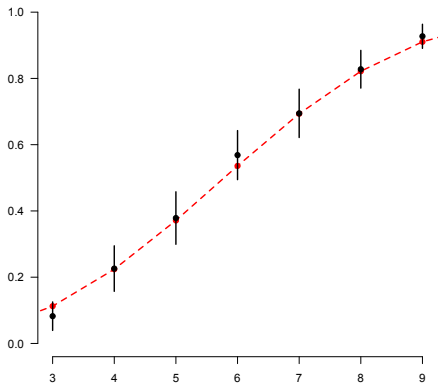


Fig 2: 95% confidence intervals for the values of $\bar{F}_0(i) = \int_i^{i+1} F_0(x)\,dx$ at the points $3, 4, \ldots, 9$ for a sample of size $n = 500$ for the model, where the incubation times are generated from the Weibull distribution, truncated at $M_1 = 15$, with parameters $a = 3.035$ and $b = 0.0026$. The black dots are the values of the MLE $\hat{F}_n$ and the vertical segments denote the symmetric 95% confidence intervals. The dashed red curve gives $\bar{F}_0$, linearly interpolated between the values at the points i and $i + 1$.

S_i as $S_i = \lfloor I_i + U_i \rfloor$, where (the infection time) I_i is (continuously) uniform on $[0, E_i]$, conditionally on E_i, and (the incubation time) U_i has the distribution specified by (1.6). For each sample of this type we maximize the log likelihood (1.5) over discrete right-continuous distribution functions F, only having jumps at the integers, to get the MLE $\hat{F}_n$. Note that for such right-continuous distribution functions F, only having jumps at the integers we get:

$$F(i) = \int_i^{i+1} F(t)\,dt.$$

The result of one such a simulation is shown in Figure 2 for sample size $n = 500$, where the parameters of the Weibull distribution are $a = 3.035$ and $b = 0.0026$, which are maximum likelihood estimates of the parameters of the truncated Weibull distribution (1.6) in an analysis of data from travelers from Wuhan, if one assumes that the incubation time distribution is given by (1.6) (see [6]). We took $M_1 = 15$ as an upper bound for the incubation distribution and also $M_2 = 15$ as upper bound for the distribution F_E of E (which we took discretely uniform on $\{1, 2, \ldots, M_2\}$). The confidence intervals were constructed using Theorem 2.1 in Section 2, as further explained in Section 4.

The *doubly interval censored* model is considered in [14]. Although the authors of [14] refer to the treatment of the singly censored model in [1], the model is different. The difference is that it is not assumed that we have (almost) exact data on when a person becomes symptomatic, but only have an interval in time for when this happened. In fact, in [1] there is also an interval, but this is only the interval of one day, whereas in [14] the interval can be 81 days. In [18], the appendix is spent on showing that the models are different. The doubly interval censored model has the following description.

There is an interval $[E_L, E_R]$ for the infection time and an interval $[S_L, S_R]$ for the time of becoming symptomatic. One can, just as in [6], shift the data in such a way that $E_L = 0$, which leaves us with three numbers: the time E ("length of Exposure time") and the times S_L and S_R, adapted for the shifting of E_L to zero. Denoting the time of becoming symptomatic

by S, we have again that S is the sum of the the infection time I and the incubation time U. We also assume, conditionally on the exposure time E, that I and U are independent and that the time of becoming infected is uniformly distributed on the interval $[0, E]$ Typical schematic pictures of the doubly interval censored model are shown in Figure 3 for two different situations for the interval $[S_L, S_R]$, containing S.

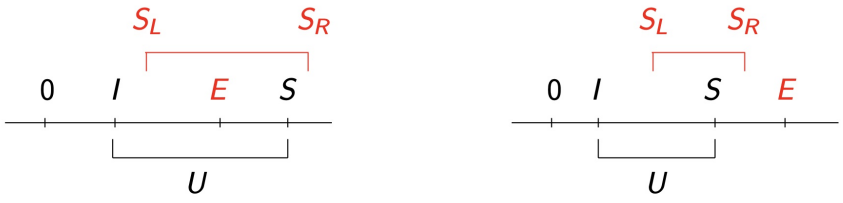


Fig 3: Doubly interval censored data. S the time of becoming symptomatic, I infection time and U the (length of the) incubation time. We can only observe E and the interval $[S_L, S_R]$, containing S.

In this case the log likelihood is of the form

$$\log\left[\mathbb{F}(S_R) - \mathbb{F}(S_L) - \mathbb{F}(S_R - E) + \mathbb{F}(S_L - E)\right], \tag{1.7}$$

neglecting parts not affecting the maximization problem, where $\mathbb{F}(u) = \int_0^u F(x)\,dx$, $u > 0$.

On the topic of the distribution theory for nonparametric estimators for interval censored data there is a large literature in mathematical statistics, but this literature does not seem to have made its way into medical statistics. The analysis in [14] is based on [18] (see [13]) and uses the `R` package `coarseDataTools`, which is created (among others) by Nicholas Reich and Justin Lessler, two of the nine authors of [14]. Distribution theory for nonparametric estimates still seems to be missing, though.

Deriving distribution theory for the nonparametric estimators seems indeed to be extremely hard. It was proved for the continuous model in [7] that the nonparametric maximim likelihood estimator, if not restricted to the set $\mathcal{F}$ of discrete distribution functions only having jumps at the integers, of the incubation time for the singly interval censored interval model converges at cube root n rate to Chernoff's distribution if one would be able to observe the time of getting symptomatic with greater accuracy than just days.

But one can again restrict consideration to the estimation of parameters $\bar{F}_0(i)$ of type (1.4). We discusss this model in Section 2. If one only considers a finite number of parameters of this type, one can specify the asymptotic (normal) distribution, using the Fisher information matrix, for both models. The rate of convergence of the estimates changes to $\sqrt{n}$. The results are given in Theorems 2.1 and 2.2 of Section 2.

Similar changes in convergence rate and asymptotic distribution were observed for the current status model with competing risk in [15], going from the continuous model to a discrete model or to a model for grouped data.

Computation of the nonparametric MLE is non-trivial. We discuss this in Section 3, where the support reduction is introduced for computing the MLE in both models. In section 4 we discuss the construction of confidence intervals for both models. Proofs are given in the appendix.

2. The nonparametric model. In general, if one has or pretends to have 1-dimensional observations that are exactly observable, the (nonparametric) maximum likelihood estimator (MLE) of the distribution function is (in several interpretations we shall not go into here) the

empirical distribution function. If one has right-censored data, the Kaplan-Meier estimator can be seen as the MLE of the distribution function (see, e.g., [12]). In the interval censored situation, the MLE is given by estimates of the type discussed in the present paper.

In the original treatment of interval censored data such as the current status model it was assumed that the nonparametric MLE would converge as a process at $\sqrt{n}$ rate and in particular would be "tight". It was also conjectured that the pointwise limit distribution would be normal ([17], [20], [11]). But it was proved in [4] that in the model where the distribution function is absolutely continuous the process is *not* tight, does *not* pointwise converge at $\sqrt{n}$ rate, and that the actual pointwise limit distribution is also *not* normal, but in fact given by Chernoff's distribution (see [3] and [10]).

The jump from right-censored data to interval censored data is considerably larger than the jump from completely observable data to right-censored data. For the Kaplan-Meier estimator one still has tightness of the process, square root n convergence and pointwise asymptotic normality. The Kaplan-Meier estimator is widely used in medical statistics, but this can not be said of the nonparametric MLE for interval censored data. In the latter situation one usually falls back on simple parametric estimates, using several classes of well-known distributions, like Weibull, gamma, log normal or Erlang. The disadvantage of this approach is that there is no compelling reason to choose one of these parametric distributions and that these models give inconsistent estimators and cnsiderable bias of the particular parametric model does not hold (see, e.g., [7]). In the case of the estimation of the incubation time distribution, we usually only have discrete observations, corresponding to days when a person become symptomatic. Also, it is clear that an infinite upper bound for the number of days of the incubation time, as assumed in the parametric models, is unrealistic. It seems more realistic to assume an upper bound of 15 or 20 days.

We therefore consider a model where we only try to estimate the parameters $\bar{F}_0(i)$, defined by (1.4). For these parameters we get the following result for the singly interval censored model. Here and the rest of the paper, we assume that the S_i and E_i are integers.

THEOREM 2.1. *Let F_E have support $\{1, 2, \ldots, M_2\}$, with positive mass at each i in this set. Let, given E, the infection time I have a (continuous) uniform distribution on $[0, E]$, and let the time of getting symptomatic S be the sum of the infection time I and the incubation time U, where I and U are independent, given E. Let U have an absolutely continuous distribution function F_0 on $\mathbb{R}$, such that $F_0(0) = 0$ and $F_0(M_1) = 1$, where $M_1 \leq M_2$, and let $\bar{F}_0(i)$ be defined by*

$$\bar{F}_0(i) = \int_i^{i+1} F_0(x)\,dx, \qquad i = 0, 1, 2, \ldots.$$

Moreover, suppose $\mathcal{T} = \{t_0, t_1, \ldots, t_m\} \subset \{0, \ldots, M_1\}$ is a set of points such that $0 = t_0 < t_1 < \cdots < t_m$, $\bar{F}_0(t_m) = 1$ and $p_0(t_i) \stackrel{\text{def}}{=} \bar{F}_0(t_i) - \bar{F}_0(t_{i-1}) > 0$, for each $i = 1, \ldots, m$. Furthermore, let $\mathcal{F}$ be the set of right-continuous discrete distribution functions only having jumps at the positive integers and let $\hat{F}_n \in \mathcal{F}$ maximize the log likelihood

$$\ell(F) = \sum_{i=1}^{n} \log\{F(S_i) - F(S_i - E_i)\}, \tag{2.1}$$

over $F \in \mathcal{F}$ on $\mathcal{T} \setminus \{0\}$. Finally, let $\hat{p}_n(t_i) = \hat{F}_n(t_i) - \hat{F}_n(t_i-)$ be the corresponding point masses. Then:

(i)

$$n^{1/2}\left\{\left(\hat{p}_n(t_1), \ldots, \hat{p}_n(t_{m-1})\right) - \left(p_0(t_1), \ldots, p_0(t_{m-1})\right)\right\} \xrightarrow{\mathcal{D}} N\left(\mathbf{0}, \boldsymbol{\Sigma}^{-1}\right), \tag{2.2}$$

where $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j,=1,\ldots,m-1}$ *is the Fisher information matrix with elements*

$$\sigma_{ij} = \mathbb{E}\,\frac{\big(1_{(S-E,S]}(t_i) - 1_{(S-E,S]}(t_m)\big)\big(1_{(S-E,S]}(t_j) - 1_{(S-E,S]}(t_m)\big)}{\{\bar{F}_0(S) - \bar{F}_0(S-E)\}^2}\,, \tag{2.3}$$

for $i, j = 1, \ldots, m-1$, *and where we assume that* $\boldsymbol{\Sigma}$ *is nonsingular.*

(ii) *Let the covariance matrix* $\boldsymbol{\Sigma}$ *be defined by (2.3) be nonsingular and let* $\hat{F}_n$ *maximize (2.1). Then:*

$$n^{1/2}\Big\{\big(\hat{F}_n(t_1), \ldots, \hat{F}_n(t_{m-1})\big) - \big(\bar{F}_0(t_1), \ldots, \bar{F}_0(t_{m-1})\big)\Big\} \xrightarrow{\mathcal{D}} N\left(\mathbf{0}, \boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^T\right), \tag{2.4}$$

where the matrix $\boldsymbol{A}$ *has rows* $\sum_{j=1}^{t_i} \boldsymbol{e}_j^T$, $i = 1, \ldots, m-1$, *for the unit vectors* $\boldsymbol{e}_j \in \mathbb{R}^{m-1}$.

REMARK 2.1. Note that $p_0(t_m) = 1 - \sum_{i=1}^{m-1} p_0(t_i)$ and $\bar{F}_0(t_m) = 1$ and that we therefore have $m-1$ free parameters, as in the case of the multinomial distribution.

REMARK 2.2. The major difference with the conditions of Theorem 4.1 in [7] is that the maximum likelihood estimators have mass at points of a *a fixed finite set*. Also note that the modest aim to estimate only a fixed finite number of parameters pulls the rate of convergence from cube root n to square root n.

We can prove a similar result for the doubly interval censored case. This time the Fisher information matrix consists of the elements

$$\sigma_{ij} = \mathbb{E}\,\frac{\{\psi(E, S_L, S_R, t_i) - \psi(E, S_L, S_R, t_m)\}\{\psi(E, S_L, S_R, t_j) - \psi(E, S_L, S_R, t_m)\}}{\big\{\int \psi(E, S_L, S_R, t)\, d\bar{F}_0(t)\big\}^2},$$

for $i, j = 1, \ldots, m-1$, where $t_i \in \mathcal{T} \subset \{1, 2, \ldots, M_1\}$, and where

$$\begin{aligned}\psi(e, s_L, s_E, t) &= (s_R - t)1_{\{t\in(0,s_R]\}} - (s_L - t)1_{\{t\in(0,s_L]\}} \\ &\quad - (s_R - e - t)1_{\{t\in(0,s_R-e]\}} + (s_L - e - t)1_{\{t\in(0,s_L-e]\}},\end{aligned} \tag{2.5}$$

Note that

$$\int_{t\in(s_L,s_R]} \{F(t) - F(t-E)\}\, dt = \int \psi(e, s_L, s_E, t)\, dF(t),$$

As noticed in Section 3, exactly the same support reduction algorithm can be used, with the weights $w_i(t_j) = \psi(E_i, S_{L,i}, S_{R,i}, t_j)$ where ψ is defined by (2.5), instead of the weights $w_i(t_j) = 1_{(S_i - E_i, S_i]}(t_j)$ for the singly interval censored model (see (3.1)).

Note that the "singly censored interval censoring model", where the terms of the log likelihood are

$$\int_{\lfloor S\rfloor}^{\lceil S\rceil} \{F(s) - F(s_E)\}\, ds = \int_{\lfloor S\rfloor}^{\lfloor S\rfloor+1} \{F(s) - F(s_E)\}\, ds$$

also satisfies

$$\int_{\lfloor S\rfloor}^{\lfloor S\rfloor+1} \{F(s) - F(s_E)\}\, ds = \int \psi(e, s_L, s_E, t)\, dF(t),$$

if F only has jumps at the integers, and $s_L = s_L(S) = \lfloor S\rfloor$ and $s_R = s_R(S) = \lceil S\rceil = \lfloor S\rfloor + 1$ (here we assume that S itself is continuously distributed!) and that the singly interval censored model is therefore a special case of the doubly interval censored model, if the observations are truncated to integers.

We get the following analogue of Theorem 2.1 for the doubly interval censored model.

THEOREM 2.2. *Let F_0 and F_E be distributions with the properties defined in Theorem 2.1, and let the time of getting symptomatic S satisfy $S \in [S_L, S_R]$. Moreover, suppose $\mathcal{T} = \{t_0, t_1, \ldots, t_m\} \subset \{0, 1, \ldots, M_1\}$ is a set of points such that $0 = t_0 < t_1 < \cdots < t_m$, $\bar{F}_0(t_m) = 1$ and $p_0(t_i) \stackrel{\text{def}}{=} \bar{F}_0(t_i) - \bar{F}_0(t_{i-1}) > 0$, for each $i = 1, \ldots, m$. Finally, let $\hat{F}_n$ maximize the log likelihood*

$$\ell(F) = \sum_{i=1}^{n} \log \int \psi(E, S_L, S_R, t)\, dF(t) \tag{2.6}$$

on $\mathcal{T} \setminus \{0\}$ where ψ is defined by (2.5), over the set of discrete distribution functions $\mathcal{F}$, defined in Theorem 2.1, and let $\hat{p}_n(t_i) = \hat{F}_n(t_i) - \hat{F}_n(t_i-)$ be the corresponding point masses. Then:

(i)

$$n^{1/2} \left\{ \left(\hat{p}_n(t_1), \ldots, \hat{p}_n(t_{m-1})\right) - \left(p_0(t_1), \ldots, p_0(t_{m-1})\right) \right\} \xrightarrow{\mathcal{D}} N\left(\mathbf{0}, \boldsymbol{\Sigma}^{-1}\right), \tag{2.7}$$

where $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j,=1,\ldots,m-1}$ is the Fisher information matrix with elements

$$\sigma_{ij} = \mathbb{E} \frac{\{\psi(E, S_L, S_E, t_i) - \psi(E, S_L, S_E, t_m)\}\{\psi(E, S_L, S_E, t_j) - \psi(E, S_L, S_E, t_m)\}}{\left\{\int \psi(E, S_L, S_R, t)\, d\bar{F}_0(t)\right\}^2}, \tag{2.8}$$

for $i, j = 1, \ldots, m-1$, where ψ isdefined by (2.5).

(ii) *Let the covariance matrix $\boldsymbol{\Sigma}$ be defined by (2.3). Then:*

$$n^{1/2} \left\{ \left(\hat{F}_n(t_1), \ldots, \hat{F}_n(t_{m-1})\right) - \left(\bar{F}_0(t_1), \ldots, \bar{F}_0(t_{m-1})\right) \right\} \xrightarrow{\mathcal{D}} N\left(\mathbf{0}, \boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}^T\right), \tag{2.9}$$

where the matrix $\boldsymbol{A}$ has rows $\sum_{j=1}^{t_i} \boldsymbol{e}_j^T$, $i = 1, \ldots, m-1$, for the unit vectors $\boldsymbol{e}_j \in \mathbb{R}^{m-1}$.

For the data set analyzed in [14] we get the picture in Figure 4 for the nonparametric MLE for the doubly interval censored data, considered in [14].

3. Computation of the nonparametric maximum likelihood estimators. In [6] two methods were discussed to compute the nonparametric MLE in the singly interval censored model: the EM algorithm and the iterative convex minorant algorithm. The EM algorithm is excruciatingly slow for this model and for this reason the iterative convex minorant algorithm was used in the simulations. But in the case of doubly interval censored data it is less clear how the iterative convex minorant algorithm should be used, although we could think of ways to apply it in this situation too. However, we will turn to a third method of computing the nonparametric MLE, the *support reduction algorithm*, see [8]. This method can be applied with equal ease to the two models.

We first discuss the support reduction algorithm for the singly interval censored model. The support reduction algorithm starts by specifying a grid of points $\mathcal{S} = \{t_1, \ldots, t_m\}$ which could be points of mass of the MLE. As an example, for the data set analyzed in [6], one could take set of points $\mathcal{S} = \{1, 2, \ldots, 20\}$, where $i \in \mathcal{S}$ would be the number of days of the incubation time. We can also take the points S_i and $(S_i - E_i)1_{\{S_i - E_i > 0\}}$ because these are the points appearing in the log likelihood.

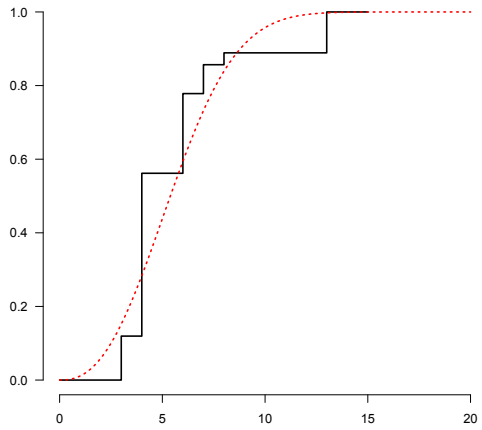


Fig 4: The nonparametric MLE (black solid curve) for the doubly interval censored data, considered in [14]. The dotted red curve is the Weibull curve, as computed by the function dic.fit in the `R` package `coarseDataTools`. Sample size is $n = 181$.

The log likelihood, divided by n, for this set of points can be written

$$n^{-1} \sum_{i=1}^{n} \log \left\{ \sum_{j=1}^{m} p_j 1_{\{t_j \in (S_i - E_i, S_i]\}} \right\},$$

where the $p_j \geq 0$, $\sum_{j=1}^{m} p_j = 1$, and where the subset of locations t_j of *strictly* positive mass p_j have to be estimated. Introducing the notation

$$w_i(t_j) = 1_{\{t_j \in (S_i - E_i, S_i]\}}, \tag{3.1}$$

we can write the log likelihood, divided by n, for $\{p_1, \ldots, p_m\}$ at $\{t_1, \ldots, t_m\}$

$$\ell(p_1, \ldots, p_m) = n^{-1} \sum_{i=1}^{n} \log \left\{ \sum_{j=1}^{m} p_j w_i(t_j) \right\} \tag{3.2}$$

Turning the maximization problem into a minimization problem on the cone $\mathbb{R}^m_+$, with a Lagrange term to ensure that the solution satisfies $\sum_{i=1}^{m} p_i = 1$, we get as our criterion function

$$\phi(p_1, \ldots, p_m) = -n^{-1} \sum_{i=1}^{n} \log \left\{ \sum_{j=1}^{m} p_j w_i(t_j) \right\} + \sum_{i=1}^{m} p_i - 1 \tag{3.3}$$

For this function have the following lemma.

LEMMA 3.1 (Fenchel duality conditions for minimization on a cone). *The function ϕ in (3.3) is minimized on $\mathbb{R}^m_+$ if and only if*

(i)

$$\frac{\partial}{\partial p_j} \phi(p_1, \ldots, p_m) \geq 0, \qquad j = 0, \ldots, m, \tag{3.4}$$

and

(ii)

$$\sum_{j=1}^{m} p_j \frac{\partial}{\partial p_j}\phi(p_1,\ldots,p_m) = 0. \tag{3.5}$$

The algorithms mentioned (EM, convex minorant algorithm and support reduction algorithm) all three try to make the conditions of Lemma 1 satisfied. The EM algorithm tries to do this by simple iteration:

$$p_j' = p_j\left\{1 - \frac{\partial}{\partial p_j}\phi(p_1,\ldots,p_m)\right\}, \qquad j = 1,\ldots,m,$$

(see (7) in [6]), the iterative convex minorant algorithm by introducing a quadratic approximation, parametrizing with the values of the distribution function instead of the point masses (see [6]).

As in the iterative convex minorant algorithm, the support reduction algorithm employs quadratic approximation, but the parametrization uses the point masses. Expanding the first two terms of the log likelihood at a fixed vector $\boldsymbol{p}^{(0)} = (p_1^{(0)},\ldots,p_m^{(0)})$, we get:

$$\begin{aligned} &n^{-1}\sum_{i=1}^{n}\log\left\{\sum_{j=1}^{m} p_j w_i(t_j)\right\} - n^{-1}\sum_{i=1}^{n}\log\left\{\sum_{j=1}^{m} p_j^{(0)} w_i(t_j)\right\} \\ &\approx n^{-1}\sum_{i=1}^{n}\frac{\sum_{j=1}^{m}\left(p_j - p_j^{(0)}\right)w_i(t_j)}{\sum_{j=1}^{m} p_j^{(0)} w_i(t_j)} - \tfrac{1}{2}n^{-1}\sum_{i=1}^{n}\frac{\left\{\sum_{j=1}^{m}\left(p_j - p_j^{(0)}\right)w_i(t_j)\right\}^2}{\left\{\sum_{j=1}^{m_0} p_j^{(0)} w_i(t_j)\right\}^2}. \end{aligned}$$

We now use iterative minimization. For fixed $m_0 < m$ and a subset $\{t_{j_1},\ldots,t_{j_{m_0}}\} \subset \{t_1,\ldots,t_m\}$, we minimize

$$\begin{aligned} &\frac{1}{2n}\sum_{i=1}^{n}\frac{\sum_{k=1}^{m_0} p_{j_k}^2 w_i(t_{j_k})^2 + 2\sum_{k<\ell} p_{j_k}p_{j_\ell} w_i(t_{j_k})w_i(t_{j_\ell})}{\left\{\sum_{j=1}^{m} p_j^{(0)} w_i(t_j)\right\}^2} \\ &\qquad - \frac{2}{n}\sum_{i=1}^{n}\frac{\sum_{k=1}^{m_0} p_{j_k} w_i(t_{j_k})}{\sum_{j=1}^{m} p_j^{(0)} w_i(t_j)} + \sum_{k=1}^{m_0} p_{j_k}. \end{aligned} \tag{3.6}$$

as a function of $p_1,\ldots,p_{m_0}$, where, at the start of the iterations

$$m_0 = 1 \qquad \text{and} \qquad p_j^{(0)} = \frac{1}{m}, \qquad j = 1,\ldots,m,$$

and where the double sum in the first numerator vanishes if $m_0 = 1$.

Then we investigate if adding a point $t_{j_{m_0+1}}$ not in the set $\{t_{j_1},\ldots,t_{j_{m_0}}\}$ and minimizing (3.6) over $\{p_{j_1},\ldots,p_{j_{m_0+1}}\}$, with m_0 replaced by m_0+1, leads to a smaller value of (2.1) with the extra point. This may lead to a solution with negative p_i's. In that case we remove the point t_k with the smallest value of $p_k < 0$ and solve the least squares minimization problem for (3.6) again. It can be proved that this procedure does not remove the point just added again (see, e.g., [16] and [8]). If this solution gives again values $p_i < 0$, we reduce the set further to a subset of $m_0 - 1$ points and solve the least squares minimization problem for (3.6) again with m_0 replaced by $m_0 - 1$, continuing until we find a solution with only positive p_i's.

Then we repeat the whole procedure, starting by investigating whether adding a point $t_{j_{m_0'+1}}$ not in the set $\{t_{j_1},\ldots,t_{j_{m_0'}}\}$ leads to a smaller value of the criterion for the new subset $\{t_{j_1},\ldots,t_{j_{m_0'}}\} \subset \{t_1,\ldots,t_m\}$. Continuing in this way we find a subset $\{t_{j_1},\ldots,t_{j_{m_0}}\}$ and

corresponding $p_{j_1}, \ldots, p_{j_{m_0}}$ which solves the least squares problem for all possible subsets $\{t_{j_1}, \ldots, t_{j_{m_0}}\} \subset \{t_1, \ldots, t_m\}$.

Next we change the values $p_j^{(0)}$ in the denominators of (3.6). Let

$$\boldsymbol{p} = (p_1, \ldots, p_m), \qquad \boldsymbol{p}^{(0)} = \left(p_1^{(0)}, \ldots, p_m^{(0)}\right),$$

where $\boldsymbol{p}$ consists of the values p_{j_k} found in the iterative least squares minimization procedure and zeroes for indices j_k not corresponding to indices of the subset $\{t_{j_1}, \ldots, t_{j_{m_0}}\}$. Using a line search procedure, for which we use Armijo's rule, we look for a convex combination

$$\boldsymbol{p}' = \alpha \boldsymbol{p} + (1-\alpha)\boldsymbol{p}^{(0)}, \qquad \alpha \in (0,1),$$

such that $\phi(\boldsymbol{p}') < \phi(\boldsymbol{p}^{(0)})$, where ϕ is defined by (3.3). Then we set $\boldsymbol{p}^{(0)} := \boldsymbol{p}'$, and repeat the iterative least squares minimization procedure, described above.

We repeat these inner and outer iterations until conditions (i) and (ii) of Lemma 3.1 are satisfied up to a certain tolerance, for which we took 10^{-10}.

As an example, the algorithm is applied to the data set, given in [6], starting with $p_i^{(0)} = 1/20$ at the points $\{1, \ldots, 20\}$ and $p_1 = 1$ at $t_1 = 10$.

iteration	criterion	$\min_{j:p_j>0} \frac{\partial}{\partial p_j}\phi(\boldsymbol{p})$	$\lvert\langle \boldsymbol{p}, \phi'(\boldsymbol{p})\rangle\rvert$	$\#\{j : p_j > 0\}$
1	1.5042265478	−0.1222076701	0.0650411285	7
2	1.4607858577	−0.0245250080	0.0650411285	7
3	1.4528857033	−0.0016619636	0.0008604701	7
4	1.4523204985	−0.0000347676	0.0000174294	7
5	1.4522988585	−0.0000003963	0.0000001969	7
6	1.4522974627	−0.0000000040	0.0000000020	7
7	1.4522973319	−0.0000000000	0.0000000000	7

It is seen that after the first least squares iteration run, the algorithm has found 7 points of strictly positive mass (the points 3 to 9) and that this number does not change in the following outer iterations. It is also clear that the outer iteration are of (quadratic) Newton type. The end solution coincides in all 10 decimals with the result of the iterative convex minorant algorithm in [6]. It can be reproduced by running the `R` script for the support reduction algorithm in [5].

The support reduction algorithm can be run in exactly the same way for the *doubly interval censored* data. The only change concerns the $w_j(t_i)$. This time $w_j(t_i)$ is defined by (2.5). The log likelihood is again given by (3.2), but with the new definition of the $w_j(t_i)$. The solution of the maximization problem on the set of points $\mathcal{S} = \{t_1, \ldots, t_m\}$ is again characterized by Lemma 3.1.

If the data are given in integers, the singly interval censored model can be treated as a doubly interval censored model by replacing the S_i by the interval $[S_i, S_i + 1]$.

4. Confidence intervals. To construct confidence intervals for the distribution function of the incubation time $\bar{F}_0$, discretized on the integers as in (1.4), based on the nonparametric MLE for the singly and doubly interval censoring models, we can use Theorems 2.1 and 2.2. Because the Weibull distribution is a popular tool for modeling the incubation time distribution in medical statistics, we use simulations from this distribution as our examples.

Since we have square root n convergence and asymptotic normality, we can also use bootstrap confidence intervals. We do not run into the inconsistency difficulties from which the classical nonparametric bootstrap suffers in the continuous model (see [6] and [7]). Bootstrapping has the advantage that we do not have to estimate the asymptotic variances.

We start by considering asymptotic confidence intervals, using Theorems 2.1 and 2.2 for estimating the variance. If we want a 95% confidence intervals for the distribution function at a fixed point t, we can use the interval

$$[\hat{F}_n(t) - 1.96\,\hat{\sigma}_n(t)/\sqrt{n}, \hat{F}_n(t) + 1.96\,\hat{\sigma}_n(t)/\sqrt{n}], \tag{4.1}$$

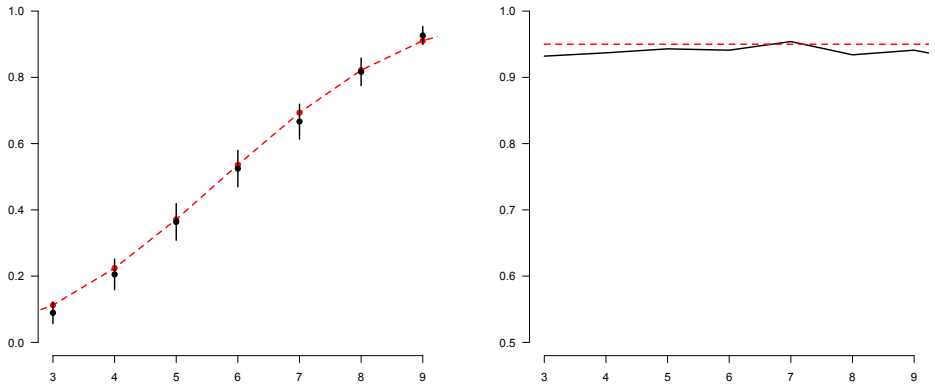
where $\hat{F}_n$ is the nonparametric MLE and $\hat{\sigma}_n(t)$ is the square root of a diagonal element of the inverse *observed* Fisher information matrix, corresponding to the Fisher information matrix of Theorems 2.1 and 2.2.

The observed Fisher information matrix is defined by $\boldsymbol{F} = (f_{jk})_{j,k=1,\ldots,m-1}$, where

$$f_{jk} = n^{-1} \sum_{i=1}^{n} \frac{\big(1_{(S_i-E_i,S_i]}(t_j) - 1_{(S_i-E_i,S_i]}(t_m)\big)\big(1_{(S_i-E_i,S_i]}(t_k) - 1_{(S_i-E_i,S_i]}(t_m)\big)}{\{\hat{F}_n(S_i) - \hat{F}_n(S_i - E_i)\}^2}, \tag{4.2}$$

for $k, j = 1, \ldots, m-1$ and where the t_j are points of mass of the MLE $\hat{F}_n$. The diagonal elements of the matrix $\boldsymbol{A}\boldsymbol{F}^{-1}\boldsymbol{A}^T$ can be used as the estimates of the asymptotic variances of $\hat{F}_n(t_1), \ldots, \hat{F}_n(t_{m-1})$, where the matrix $\boldsymbol{A}$ has rows $\sum_{j=1}^{t_i} \boldsymbol{e}_j^T$, $i = 1, \ldots, m-1$, for the unit vectors $\boldsymbol{e}_j \in \mathbb{R}^{m-1}$ (see (ii) of Theorem 2.1).

The resulting confidence intervals for a sample of size $n = 1000$ is shown in Figure 5 at the points 3 to 9, where the MLE puts most of its mass. The coverage of these intervals is also shown. Here we generated 1000 samples of size $n = 1000$, and computed the fraction of times the parameters $\bar{F}_0(i) = \int_i^{i+1} F_0(t)\,dt$ were inside the intervals (4.1) at the points 3 to 9.



(a) (b)

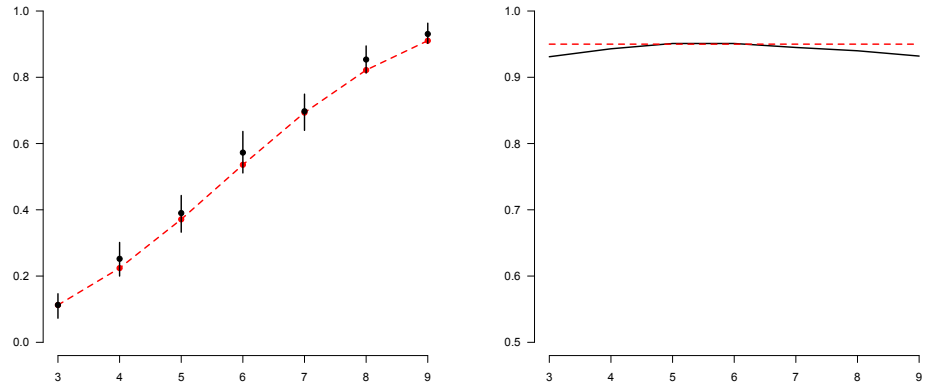
Fig 5: (a) 95% confidence intervals in the singly interval censored model, using (4.1), for the values of $\bar{F}_0(i) = \int_i^{i+1} iF_0(x)\,dx$ (red dots and linearly interpolated dashed red curve) at the points $3, 4, \ldots, 9$ for a sample of size $n = 1000$, where F_0 is the Weibull distribution function, , with parameters $a = 3.035$ and $b = 0.0026$, truncated at $M_1 = 15$. The black dots are the values of $\hat{F}_n$ at these points. (b) Coverage percentages of the 95% confidence intervals at the points $3, 4, \ldots, 9$, using (4.1), for sample size $n = 1000$.

We can run a bootstrap experiment to generate confidence intervals of this type in the following way. We resample with replacement from the data (E_i, S_i) 1000 samples of the

same size n and compute for each of these bootstrap samples the MLE. This gives 1000 bootstrap values $\hat{F}_n^*(t) - \hat{F}_n(t)$. For these bootstrap values of $\hat{F}_n^*(t) - \hat{F}_n(t)$ we compute the 0.025 and 0.975 quantiles $Q^*_{0.025}(t)$ and $Q^*_{0.975}(t)$, respectively. This gives the bootstrap 95% confidence intervals

$$[\hat{F}_n(t) - Q^*_{0.975}(t), \hat{F}_n(t) - Q^*_{0.025}(t)]. \tag{4.3}$$

The results are shown in Figure 6. It is seen that the results are similar to the results of the method, using the inverse observed Fisher information matrix for generating the confidence intervals.



(a) (b)

Fig 6: (a) 95% bootstrap confidence intervals in the singly interval censored model, using (4.3), for the values of $\bar{F}_0(i) = \int_i^{i+1} F_0(x)\,dx$ at the points $3, 4, \ldots, 9$ (red dots and linearly interpolated dashed red curve) for a sample of size $n = 1000$, where F_0 is the Weibull distribution function, truncated at $M_1 = 15$. The black dots are the values of $\hat{F}_n$ at these points. (b) Coverage percentages of the bootstrap 95% confidence intervals at the points $3, 4, \ldots, 9$, using (4.3), for sample size $n = 1000$.

We also simulated data for the doubly censored model. Here we took S_R discretely uniform on $\{\lceil S\rceil, \ldots, \lceil S\rceil + 3\}$ and S_L discretely uniform on $\{\lfloor S\rfloor, \ldots, \lfloor S\rfloor - 3\}$ (replacing $\lfloor S\rfloor - i$ by 0 if $\lfloor S\rfloor - i < 0$). This time the observed Fisher information matrix is defined by

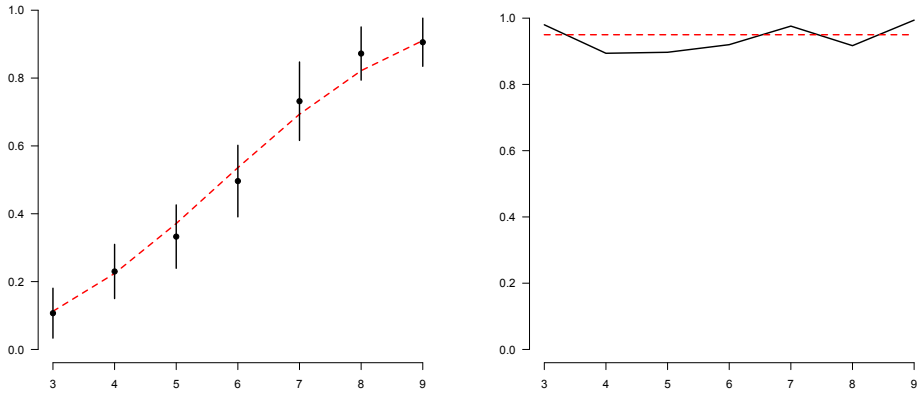
$$f_{jk} = n^{-1} \sum_{i=1}^{n} \frac{\tilde{\psi}(E_i, S_{L,i}, S_{R,i}, t_j)\tilde{\psi}(E_i, S_{L,i}, S_{R,i}, t_k)}{\left\{\int \psi(E_i, S_{L,i}, S_{R,i}, t)\, d\hat{F}_n(t)\right\}^2}, \qquad j, k = 1, \ldots, m-1, \tag{4.4}$$

where

$$\tilde{\psi}(E_i, S_{L,i}, S_{R,i}, t) = \psi(E_i, S_{L,i}, S_{R,i}, t) - \psi(E_i, S_{L,i}, S_{R,i}, t_m),$$

and where the t_j are points of mass of the MLE $\hat{F}_n$ and ψ is defined by (2.5). The diagonal elements of the matrix $\boldsymbol{A}\boldsymbol{F}^{-1}\boldsymbol{A}^T$ can be used as the estimates of the asymptotic variances of $\hat{F}_n(t_1), \ldots, \hat{F}_n(t_{m-1})$, where the matrix $\boldsymbol{A}$ has rows $\sum_{j=1}^{t_i} \boldsymbol{e}_j^T$, $i = 1, \ldots, m-1$, for the unit vectors $\boldsymbol{e}_j \in \mathbb{R}^{m-1}$ (see (ii) of Theorem 2.2).

The coverages, based on the Fisher information matrix from one sample were not very good this time, but were rather satisfactry if we estimate the Fisher information matrix by the mean of these matrices over 1000 bootstrap samples from one sample. See Figure 7.



(a) (b)

Fig 7: 95% confidence intervals for a doubly interval censored simulation sample at the points $3, \ldots, 9$, using (4.1), with the variances estimated by the means of the diagonal of the inverse observed Fisher information matrices over 1000 bootstrap samples from one sample. (b) Coverage percentages for the same example over 1000 samples, using the same estimates of the variances. Sample size is $n = 1000$.

5. Conclusion. We studied the nonparametric maximum likelihood estimator of the incubation time, which was such an important parameter in the Covid-19 pandemic. The incubation time in our model is part of the sum of the infection time I and the incubation time U. Usually the only really (partly) observable quantity is the time of becoming symptomatic, which is given by $S = I + U$. So our variable of interest U has to be pulled out of this sum by deconvolution.

An additional problem is that most of the time S is not directly observable, but only an interval $[S_L, S_R]$ is available, which we know to contain S. If this interval only consists of one day, one usually acts as if S itself is observable, and one calls this the *singly interval censored model* (the beginning of the incubation time lies in an interval which represents the exposure time and this is the singly interval censored part).

It seems most reasonable to assume that the time till infection and the incubation time have continuous distributions, and one can analyze this model under the assumption that S is exactly observable. In that case the MLE of the distribution function of the incubation time converges, under some conditions, at cube root n rate to Chernoff's distribution, see [7].

However, taking into account that the observations are usually rounded to days, one can also restrict oneself to estimating the means over one day, represented by

$$\int_i^{i+1} F_0(t)\, dt, \tag{5.1}$$

if F_0 is the distribution function of the incubation time. Since this gives us a fixed bounded number of parameters (we considered 15 parameters of this type in the present paper), we can use classical theory, connecting maximum likelihood with the Fisher information, to derive asymptotic distribution theory for the maximum likelihood estimators of these parameters.

If the interval $[S_L, S_R]$, known to contain S, is larger that one day, the model is called the *doubly interval censored model* for which we still can estimate the parameters (5.1) by maximum likelihood. One can use the Fisher information to characterize the (normal) limiting distribution.

We applied the theory, developed in Section 2 to construct confidence intervals, either by using Theorems 2.1 and 2.2 directly, or by using a bootstrap method. In a sense, this purely nonparametric method lies at the other extreme of the parametric methods. If one wants to estimate the density, one will have to use some kind of smoothing, as was done in [6], which is an intermediate method that is still nonparametric and avoids the need to choose between several parametric models.

The support reduction algorithm seems at present to be the most stable method to estimate the parameters. The computing of the MLE and the confidence intervals is implemented in [5] and discussed in Section 3.

6. Appendix. Proofs. In contrast with the difficulties of the continuous model in [7], the estimation of the parameters $\int_i^{i+1} F_0(t)\,dt$ seems rather standard, once we have figured out the relation between the continuous model and the discrete observations of the times of becoming symptomatic. We still have to deal with a deconvolution problem, which we do by using nonparametric maximum likelihood.

PROOF OF LEMMA 3.1. The key step is to reduce the maximization on the simplex $\{\boldsymbol{p} = (p_1,\ldots,p_m) \in \mathbb{R}_+^m : \sum_{i=1}^m p_i = 1\}$ to maximization on the cone $\mathbb{R}_+^m$. One can check that minimizing minus the log likelihood (3.2) under the restriction $\sum_{i=1}^m p_i = 1$ is equivalent to minimizing (3.3), which is the criterion function + a Lagrange term with Lagrange multiplier $\lambda = 1$, on $\mathbb{R}_+^m$. Then the necessary and sufficient conditions for the minimum follow from Fenchel's duality theorem, see [19], Theorem 31.4. □

PROOF OF THEOREM 2.1. The log likelihood is of the form

$$\ell(p_1,\ldots,p_m) = \sum_{i=1}^n \log \sum_{j=1}^m p_j 1_{((S_i-E_i)_+,S_i]}(t_j),$$

so we count the number of times the point of mass t_j belongs to an interval $(S_i - E_i)_+, S_i]$, where S_i and E_i are integers, and multiply this with the probability p_j. We have

$$\begin{aligned}
&\frac{\partial}{\partial p_j}\ell\Bigl(p_1,\ldots,1-\sum_{k=1}^{m-1} p_k\Bigr)\\
&= \sum_{i=1}^n \frac{1_{((S_i-E_i)_+,S_i]}(t_j) - 1_{((S_i-E_i)_+,S_i]}(t_m)}{\sum_{k=1}^m p_j 1_{((S_i-E_i)_+,S_i]}(t_k)}, \qquad j = 1,\ldots,m-1,
\end{aligned} \tag{6.1}$$

and

$$\begin{aligned}
&\frac{\partial^2}{\partial p_j \partial p_l}\ell\Bigl(p_1,\ldots,1-\sum_{k=1}^{m-1} p_k\Bigr)\\
&= -\sum_{i=1}^n \frac{\bigl\{1_{((S_i-E_i)_+,S_i]}(t_j) - 1_{((S_i-E_i)_+,S_i]}(t_m)\bigr\}\bigl\{1_{((S_i-E_i)_+,S_i]}(t_l) - 1_{((S_i-E_i)_+,S_i]}(t_m)\bigr\}}{\bigl\{\sum_{k=1}^m p_j 1_{((S_i-E_i)_+,S_i]}(t_k)\bigr\}^2},
\end{aligned}$$

for $j,l = 1,\ldots,m-1$, using the convention $0/0 = 0$.

By the assumptions on F_0 and F_E, the variables S_i and $(S_i - E_i)_+$ will be such that for large n, the score functions (6.1) will be zero for $p_j = \hat{p}_j$, where $\hat{\boldsymbol{p}} = (\hat{p}_1,\ldots,\hat{p}_m)$ is the MLE of $\boldsymbol{p}_0 = (p_0(t_1),\ldots,p_0(t_m))$ (that is: no isotonization is needed), and the result now follows from standard theory. □

PROOF OF THEOREM 2.2. The proof is entirely similar to the proof of Theorem 2.1, but this time the score functions are given by

$$\frac{\partial}{\partial p_j}\ell\left(p_1,\ldots,1-\sum_{k=1}^{m-1}p_k\right)$$

$$=\sum_{i=1}^{n}\frac{\psi(E_i,S_{L,i},S_{R,i},t_j)-\psi(E_i,S_{L,i},S_{R,i},t_m)}{\sum_{k=1}^{m}\psi(E_i,S_{L,i},S_{R,i},t)\,p_k}, \qquad j=1,\ldots,m-1, \tag{6.2}$$

where ψ is defined by (2.5).

A key part of the treatment of the doubly censored case is the rewrite of the log likelihood for one observation, using the integration by parts

$$\int_{t\in(s_L,s_R]}\{F(t)-F(t-E)\}\,dt=\int\psi(e,s_L,s_E,t)\,dF(t),$$

where ψ is defined by (2.5). □

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] CHERNOFF, H. (1964). Estimation of the mode. *Ann. Inst. Statist. Math.* **16** 31–41. MR0172382 (30 ##2601)

[4] GROENEBOOM, P. (1987). Asymptotics for incomplete censored observations. Report 87-18, Mathematical Institute, University of Amsterdam.

[5] GROENEBOOM, P. (2020). Incubation Time. https://github.com/pietg/incubationtime.

[6] GROENEBOOM, P. (2021). Estimation of the incubation time distribution for COVID-19. *Stat. Neerl.* **75** 161–179. MR4245907

[7] GROENEBOOM, P. (2023). Nonparametric estimation of the incubation time distribution. https://arxiv.org/abs/2205.04399.

[8] GROENEBOOM, P., JONGBLOED, G. and WELLNER, J. A. (2008). The support reduction algorithm for computing non-parametric function estimates in mixture models. *Scand. J. Statist.* **35** 385–399. MR2446726 (2009m:62115)

[9] GROENEBOOM, P., JONGBLOED, G. and WITTE, B. I. (2010). Maximum smoothed likelihood estimation and smoothed maximum likelihood estimation in the current status model. *Ann. Statist.* **38** 352–387.

[10] GROENEBOOM, P. and WELLNER, J. A. (2001). Computing Chernoff's distribution. *J. Comput. Graph. Statist.* **10** 388–400. MR1939706

[11] GRÜGER, J. (1986). Nichtparametrische Analyse sporadisch beobachtbarer Krankheitsverlaufsdaten, Ph.D. dissertation, Technische Universität, Dortmund, Germany.

[12] JACOBSEN, M. (1982). *Statistical analysis of counting processes*. *Lecture Notes in Statistics* **12**. Springer-Verlag, New York-Berlin. MR676128

[13] LAUER, S. A. (2020). ncov_incubation. https://github.com/HopkinsIDD/ncov_incubation.

[14] LAUER, S. A., GRANTZ, K. H., BI, Q., JONES, F. K., ZHENG, Q., MEREDITH, H. R., AZMAN, A. S., REICH, N. G. and LESSLER, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine* **172** 577–582.

[15] MAATHUIS, M. H. and HUDGENS, M. G. (2011). Nonparametric inference for competing risks current status data with continuous, discrete or grouped observation times. *Biometrika* **98** 325–340. MR2806431

[16] MEYER, M. C. (2013). A simple new algorithm for quadratic programming with applications in statistics. *Comm. Statist. Simulation Comput.* **42** 1126–1139. MR3039672

[17] PETO, R. (1973). Experimental survival curves for interval-censored data. *J.R. Statist. Soc. Series C* **22** 86–91.

[18] 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. MR2750164

[19] ROCKAFELLAR, R. T. (1970). *Convex analysis*. *Princeton Mathematical Series, No. 28*. Princeton University Press, Princeton, NJ. MR274683

[20] TURNBULL, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. Roy. Statist. Soc. Ser. B* **38** 290–295. MR652727