Estimators and designs for interval censored data

CHRISTINE H. MÜLLER*

June 26, 2017

1 Lifetimes with the same starting times

Let be $T_1, \ldots, T_N$ independent nonnegative random variables (life time variables), each with distribution given by the cumulative distribution function F_θ and the density f_θ, where θ is an unknown parameter. However, the realizations $t_1, \ldots, t_N$ of $T_1, \ldots, T_N$ are not observed. Only realizations n_i of

$$N_i := \sum_{n=1}^{N} \mathbb{1}_{(\tau_{i-1},\tau_i]}(T_n), \quad i = 1, \ldots, I+1,$$

are observed, where $0 = \tau_0 < \tau_1 < \ldots < \tau_I = \tau < \tau_{I+1} = \infty$ are given inspections times and $\mathbb{1}_A$ denotes the indicator function for the set A. In particular, we have $N = \sum_{i=1}^{I+1} n_i$.

The aim is to find estimators for θ and designs for choosing the inspections times $\tau_1 < \ldots < \tau_I$.

1.1 Exact maximum likelihood estimators

The likelihood function for a single observation n_i is given by

$$l_\theta(n_i) := \prod_{n=1}^{N} P_\theta\left(T_n \in (\tau_{i-1}, \tau_i]\right)^{\mathbb{1}_{(\tau_{i-1},\tau_i]}(t_n)} = \left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^{n_i}$$

for $i = 1, \ldots, I$ and

$$l_\theta(n_{I+1}) := \prod_{n=1}^{N} P_\theta\left(T_n \in (\tau_I, \infty)\right)^{\mathbb{1}_{(\tau_I,\infty)}(t_n)} = \left(1 - F_\theta(\tau_I)\right)^{n_{I+1}}$$

so that the common likelihood function of $n_* := (n_1, \ldots, n_{I+1})$ is given by

$$L_\theta(n_*) := \prod_{i=1}^{I} \left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^{n_i} \, \left(1 - F_\theta(\tau_I)\right)^{n_{I+1}}.$$

*Department of Statistics, TU Dortmund University, D-44227 Dortmund, Germany, cmueller@statistik.tu-dortmund.de

The loglikelihood function is then

$$\ln L_\theta(n_*) = \sum_{i=1}^{I} n_i \, \ln\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) + n_{I+1} \, \ln\left(1 - F_\theta(\tau_I)\right) \tag{1}$$

so that its derivative is given by

$$\frac{\partial}{\partial\theta} \ln L_\theta(n_*) = \sum_{i=1}^{I} n_i \, \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)}{F_\theta(\tau_i) - F_\theta(\tau_{i-1})} + n_{I+1} \, \frac{-\frac{\partial}{\partial\theta} F_\theta(\tau_I)}{1 - F_\theta(\tau_I)}. \tag{2}$$

For the special case $I = 1$, because $F_\theta(0) = 0$, this reduces to

$$\frac{\partial}{\partial\theta} \ln L_\theta((n_1, n_2)) = n_1 \, \frac{\frac{\partial}{\partial\theta} F_\theta(\tau_1)}{F_\theta(\tau_1)} + n_2 \, \frac{-\frac{\partial}{\partial\theta} F_\theta(\tau_1)}{1 - F_\theta(\tau_1)}. \tag{3}$$

In particualar, if T_n has an exponential distribution with unknown parameter λ then (3) becomes

$$\begin{aligned}
\frac{\partial}{\partial\lambda} \ln L_\lambda((n_1, n_2)) &= n_1 \, \frac{\frac{\partial}{\partial\lambda}\left(1 - \exp(-\lambda\,\tau_1)\right)}{1 - \exp(-\lambda\,\tau_1)} + n_2 \, \frac{-\frac{\partial}{\partial\lambda}\left(1 - \exp(-\lambda\,\tau_1)\right)}{\exp(-\lambda\,\tau_1)} \\
&= \tau_1 \, \exp(-\lambda\,\tau_1) \left(\frac{n_1}{1 - \exp(-\lambda\,\tau_1)} - \frac{n_2}{\exp(-\lambda\,\tau_1)} \right) \\
&= \tau_1 \left(\frac{n_1 \, \exp(-\lambda\,\tau_1)}{1 - \exp(-\lambda\,\tau_1)} - n_2 \right) = \tau_1 \left(\frac{n_1}{\frac{1}{\exp(-\lambda\,\tau_1)} - 1} - n_2 \right).
\end{aligned}$$

Then we have

$$\begin{aligned}
&\frac{\partial}{\partial\lambda} \ln L_\lambda((n_1, n_2)) = 0 \Longleftrightarrow n_2 = \frac{n_1}{\frac{1}{\exp(-\lambda\,\tau_1)} - 1} \Longleftrightarrow \frac{n_1}{n_2} = \frac{1}{\exp(-\lambda\,\tau_1)} - 1 \\
\Longleftrightarrow \quad &\frac{n_1}{n_2} + 1 = \frac{1}{\exp(-\lambda\,\tau_1)} \Longleftrightarrow \frac{1}{\frac{n_1}{n_2} + 1} = exp(-\lambda\,\tau_1) \Longleftrightarrow \ln\left(\frac{1}{\frac{n_1}{n_2} + 1}\right) = -\lambda\,\tau_1
\end{aligned}$$

so that

$$\widehat{\lambda} = -\frac{1}{\tau_1} \, \ln\left(\frac{1}{\frac{n_1}{n_2} + 1}\right) = -\frac{1}{\tau_1} \, \ln\left(\frac{n_2}{n_1 + n_2}\right) \tag{4}$$

is a candidate for a maximum of $\ln L_\lambda((n_1, n_2))$. Since

$$\begin{aligned}
\frac{\partial^2}{\partial^2\lambda} \ln L_\lambda((n_1, n_2)) &= \frac{\partial}{\partial\lambda}\tau_1 \, \left(\frac{n_1}{\frac{1}{\exp(-\lambda\,\tau_1)} - 1} - n_2 \right) \\
&= -\tau_1^2 \, \frac{-n_1}{\left(\frac{1}{\exp(-\lambda\,\tau_1)} - 1\right)^2} \, \frac{-1}{\exp(-\lambda\,\tau_1)^2} \, \exp(-\lambda\,\tau_1) < 0,
\end{aligned}$$

we have indeed a maximum so that $\widehat{\lambda}$ given by (4) is the maximum likelihood estimator.

For $I > 2$ an explicit form of the maximum likelihood estimator cannot be derived. However, it can easily determined numerically by maximizing (1) or calculating the root of (1) if θ is one-dimensional as this is case for parameter $\theta = \lambda$ of the exponential distribution. However, for parameter dimensions higher than one, the calculation is more difficult but can be done for example with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method.

Alternatively, an approximate form of the likelihood function can be used.

1.2 Approximate maximum likelihood estimators

Using the mean value theorem, the derivative of the loglikelihood function given in (2) becomes for $\tau_i^* \in [\tau_{i-1}, \tau_i)$, $i = 1, \ldots, I$,

$$\frac{\partial}{\partial\theta} \ln L_\theta(n_*) = \sum_{i=1}^{I} n_i \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)}{F_\theta(\tau_i) - F_\theta(\tau_{i-1})} + n_{I+1} \frac{-\frac{\partial}{\partial\theta} F_\theta(\tau_I)}{1 - F_\theta(\tau_I)}$$
$$= \sum_{i=1}^{I} n_i \frac{\frac{\partial}{\partial\theta} f_\theta(\tau_i^*)(\tau_i - \tau_{i-1})}{f_\theta(\tau_i^*)(\tau_i - \tau_{i-1})} + n_{I+1} \frac{-\frac{\partial}{\partial\theta} F_\theta(\tau_I)}{1 - F_\theta(\tau_I)} = \sum_{i=1}^{I} n_i \frac{\frac{\partial}{\partial\theta} f_\theta(\tau_i^*)}{f_\theta(\tau_i^*)} + n_{I+1} \frac{-\frac{\partial}{\partial\theta} F_\theta(\tau_I)}{1 - F_\theta(\tau_I)}.$$

For the exponential distribution with unknown parameter λ, we get

$$\frac{\partial}{\partial\lambda} \ln L_\lambda(n_*) = \sum_{i=1}^{I} n_i \frac{\frac{\partial}{\partial\lambda}\lambda \exp(-\lambda\tau_i^*)}{\lambda \exp(-\lambda\tau_i^*)} + n_{I+1} \frac{-\frac{\partial}{\partial\lambda}\left(1 - \exp(-\lambda\tau_I)\right)}{\exp(-\lambda\tau_I)}$$
$$= \sum_{i=1}^{I} n_i \frac{\exp(-\lambda\tau_i^*)(1 - \lambda\tau_i^*)}{\lambda \exp(-\lambda\tau_i^*)} + n_{I+1} \frac{-\tau_I \exp(-\lambda\tau_I)}{\exp(-\lambda\tau_I)}$$
$$= \sum_{i=1}^{I} n_i \left(\frac{1}{\lambda} - \tau_i^*\right) - n_{I+1}\,\tau_I = \frac{1}{\lambda}(N - n_{I+1}) - \sum_{i=1}^{I} n_i\,\tau_i^* - n_{I+1}\,\tau_I = 0$$
$$\iff \widehat{\lambda} = \left(\frac{1}{N - n_{I+1}} \left(\sum_{i=1}^{I} n_i\,\tau_i^* + n_{I+1}\,\tau_I\right)\right)^{-1}. \tag{5}$$

Hence $\widehat{\lambda}$ is here the estimate which we would get if we would have observed n_i times τ_i^*, $i = 1, \ldots, I$, and have n_{I+1} censored observations censored at τ_{I+1}. For the special case $I = 1$ we obtain

$$\widehat{\lambda} = \left(\frac{1}{n_1}\,(n_1\,\tau_1^* + n_{I+1}\,\tau_I)\right)^{-1}. \tag{6}$$

Note the difference to the estimate given in (4).

However, in the exact form, $\tau_1^*, \ldots, \tau_I^*$ depend on the unknown parameter λ so that the $\tau_1^*, \ldots, \tau_I^*$ in (5) and (6) must be approximated. This can be done by using

$$\tau_i^* \approx \tau_{i-1} \quad \text{or} \quad \tau_i^* \approx \tau_i \quad \text{or} \quad \tau_i^* \approx \frac{1}{2}(\tau_{i-1} + \tau_i) \quad \text{for} \quad i = 1, \ldots, I.$$

This leads to three alternatives to the exact maximum likelihood estimator given in Section 1.1.

1.3 Information matrix

The likelihood function for a single random variable T_n, although T_n is not observed, is

$$l_\theta(T_n) := \prod_{i=1}^{I+1} P_\theta\left(T_n \in (\tau_{i-1}, \tau_i]\right)^{1\!\!1_{(\tau_{i-1},\tau_i]}(T_n)}$$

so that

$$\ln l_\theta(T_n) = \sum_{i=1}^{I+1} \ln\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n)$$

and

$$\begin{aligned}\frac{\partial}{\partial\theta} \ln l_\theta(T_n) &= \sum_{i=1}^{I+1} \frac{\partial}{\partial\theta} \ln\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n) \\ &= \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)}{F_\theta(\tau_i) - F_\theta(\tau_{i-1})} 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n). \qquad (7)\end{aligned}$$

In particular, we have

$$\begin{aligned}E_\theta\left(\frac{\partial}{\partial\theta} \ln l_\theta(T_n)\right) &= \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)}{F_\theta(\tau_i) - F_\theta(\tau_{i-1})} \left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) \\ &= \frac{\partial}{\partial\theta} \sum_{i=1}^{I+1} \left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) = \frac{\partial}{\partial\theta}\left(F_\theta(\infty) - F_\theta(0)\right) = \frac{\partial}{\partial\theta} 1 = 0\end{aligned}$$

so that a maximum likelihood estimator based on $\ln l_\theta(T_n)$ is asymptotically unbiased. Because of

$$\frac{\partial}{\partial\theta} \ln l_\theta(T_n) \frac{\partial}{\partial\theta} \ln l_\theta(T_n)^\top = \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)}{F_\theta(\tau_i) - F_\theta(\tau_{i-1})} \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^\top}{F_\theta(\tau_i) - F_\theta(\tau_{i-1})} 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n),$$

the information matrix is given as

$$\begin{aligned}I_\theta(\tau_1, \ldots, \tau_I) &:= E_\theta\left(\frac{\partial}{\partial\theta} \ln l_\theta(T_n) \frac{\partial}{\partial\theta} \ln l_\theta(T_n)^\top\right) \\ &= \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) \frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^\top}{\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^2} \left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right).\end{aligned}$$

1.4 Approximate information matrix

The mean value theorem, provides again the following form of the information matrix:

$$I_\theta(T_n) := \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right) \frac{\partial}{\partial\theta}\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^\top}{\left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)^2} \left(F_\theta(\tau_i) - F_\theta(\tau_{i-1})\right)$$

$$
\begin{aligned}
&= \sum_{i=1}^{I} \frac{\frac{\partial}{\partial\theta} f_\theta(\tau_i^*) \frac{\partial}{\partial\theta} f_\theta(\tau_i^*)^\top}{\left(f_\theta(\tau_i^*)\right)^2} f_\theta(\tau_i^*)\,(\tau_i - \tau_{i-1}) + \frac{\frac{\partial}{\partial\theta} F_\theta(\tau_I) \frac{\partial}{\partial\theta} F_\theta(\tau_I)^\top}{(1 - F_\theta(\tau_I))} \\
&= \sum_{i=1}^{I} \frac{\frac{\partial}{\partial\theta} f_\theta(\tau_i^*) \frac{\partial}{\partial\theta} f_\theta(\tau_i^*)^\top}{f_\theta(\tau_i^*)}\,(\tau_i - \tau_{i-1}) + \frac{\frac{\partial}{\partial\theta} F_\theta(\tau_I) \frac{\partial}{\partial\theta} F_\theta(\tau_I)^\top}{(1 - F_\theta(\tau_I))}.
\end{aligned}
$$

1.5 Optimal inspection time in the case of one inspection time

If $I = 1$ then

$$
\ln l_\theta(T_n) = \ln F_\theta(\tau_1) 1\!\!1_{(0,\tau_1]}(T_n) + \ln(1 - F_\theta(\tau_1))\, 1\!\!1_{(\tau_1,\infty)}(T_n).
$$

In particular, for the exponential distribution we get

$$
\ln l_\lambda(T_n) = \ln(1 - \exp(-\lambda\tau_1))\, 1\!\!1_{(0,\tau_1]}(T_n) - \lambda\tau_1 1\!\!1_{(\tau_1,\infty)}(T_n)
$$

and

$$
\frac{\partial}{\partial\lambda} \ln l_\lambda(T_n) = \frac{\tau_1 \exp(-\lambda\tau_1)}{1 - \exp(-\lambda\tau_1)} 1\!\!1_{(0,\tau_1]}(T_n) - \tau_1 1\!\!1_{(\tau_1,\infty)}(T_n)
$$

so that

$$
\begin{aligned}
I_\lambda(\tau_1) &= \left(\frac{\tau_1 \exp(-\lambda\tau_1)}{1 - \exp(-\lambda\tau_1)}\right)^2 (1 - \exp(-\lambda\tau_1)) + \tau_1^2 \exp(-\lambda\tau_1) \\
&= \frac{\tau_1^2 \exp(-\lambda\tau_1)^2}{1 - \exp(-\lambda\tau_1)} + \tau_1^2 \exp(-\lambda\tau_1) = \frac{\tau_1^2 \exp(-\lambda\tau_1)}{1 - \exp(-\lambda\tau_1)} (\exp(-\lambda\tau_1) + 1 - \exp(-\lambda\tau_1)) \\
&= \frac{\tau_1^2 \exp(-\lambda\tau_1)}{1 - \exp(-\lambda\tau_1)} = \frac{1}{\lambda^2} \frac{(\lambda\tau_1)^2 \exp(-\lambda\tau_1)}{1 - \exp(-\lambda\tau_1)}. \qquad (8)
\end{aligned}
$$

The aim is now to find a τ_1 which maximizes (8). Therefore, we have to find the maximum of

$$
f(x) := \frac{x^2 \exp(-x)}{1 - \exp(-x)} = \frac{x^2}{\exp(x) - 1}.
$$

Since

$$
\begin{aligned}
f'(x) &= \frac{2x}{\exp(x) - 1} - \frac{x^2 \exp(x)}{(\exp(x) - 1)^2} = \frac{2x \exp(x) - 2x + x^2 \exp(x)}{(\exp(x) - 1)^2} \\
&= \frac{x}{\exp(x) - 1} \left(2 - \frac{x \exp(x)}{\exp(x) - 1}\right) = \frac{x}{\exp(x) - 1} \left(2 - \frac{x}{1 - \exp(-x)}\right),
\end{aligned}
$$

an extremum x_* of f is given by

$$
2 = \frac{x_*}{1 - \exp(-x_*)}
$$

which can be determined numerically. Then $\tau_1 = \tau(\lambda) = \frac{1}{\lambda} x_*$ maximizes (8). In particular, it depends on the unknown λ so that it is only a locally optimal design and an initial estimate for λ is needed.

The equation

$$2 = \frac{\lambda \tau(\lambda)}{1 - \exp(-\lambda \tau(\lambda))}$$

yields

$$1 - \exp(-\lambda \tau(\lambda)) = \frac{\lambda \tau(\lambda)}{2} \text{ and } \exp(-\lambda \tau(\lambda)) = 1 - \frac{\lambda \tau(\lambda)}{2}$$

so that the information matrix at the local optimal design is

$$\begin{aligned} I_\lambda(\tau(\lambda)) &= \frac{1}{\lambda^2} \frac{(\lambda \tau(\lambda))^2 \exp(-\lambda \tau(\lambda))}{1 - \exp(-\lambda \tau(\lambda))} \\ &= \frac{1}{\lambda^2} \frac{2(\lambda \tau(\lambda))^2 \left(1 - \frac{\lambda \tau(\lambda)}{2}\right)}{\lambda \tau(\lambda)} = \frac{1}{\lambda^2} \left(\lambda \tau(\lambda) \left(2 - \lambda \tau(\lambda)\right)\right) = \frac{2}{\lambda} \tau(\lambda) - \tau(\lambda)^2. \end{aligned}$$

To derive a maximin-efficient inspection time τ_*, note at first that

$$\lim_{\lambda \to \infty} I_\lambda(\tau_1) = \lim_{\lambda \to \infty} \frac{1}{\lambda^2} \frac{(\lambda \tau_1)^2 \exp(-\lambda \tau_1)}{1 - \exp(-\lambda \tau_1)} = 0,$$

i.e. the information becomes arbitrary small at any inspection point τ_1 if λ becomes very large, which means that the underlying observation T_n is concentrated near zero. Hence we have to restrict λ by an upper bound U.

Moreover, note that the efficiency is given by

$$\begin{aligned} \frac{I_\lambda(\tau_1)}{I_\lambda(\tau(\lambda))} &= \frac{1}{\lambda^2} \frac{(\lambda \tau_1)^2 \exp(-\lambda \tau_1)}{1 - \exp(-\lambda \tau_1)} \frac{1}{\frac{2}{\lambda} \tau(\lambda) - \tau(\lambda)^2} \\ &= \frac{1}{\lambda^2} \frac{(\lambda \tau_1)^2 \exp(-\lambda \tau_1)}{1 - \exp(-\lambda \tau_1)} \frac{1}{\frac{1}{\lambda^2} (2\, x_* - x_*^2)} = \frac{(\lambda \tau_1)^2 \exp(-\lambda \tau_1)}{1 - \exp(-\lambda \tau_1)} \frac{1}{2\, x_* - x_*^2} \\ &= \frac{(\lambda \tau_1)^2}{\exp(\lambda \tau_1) - 1} \frac{1}{2\, x_* - x_*^2} \end{aligned}$$

so that

$$\lim_{\lambda \to \infty} \frac{I_\lambda(\tau_1)}{I_\lambda(\tau(\lambda))} = 0, \quad \lim_{\lambda \to 0} \frac{I_\lambda(\tau_1)}{I_\lambda(\tau(\lambda))} = 0.$$

This means that we have to restrict λ also by a lower bound L. Since f given by $f(x) := \frac{x^2}{\exp(x) - 1}$ is a unimodal function, a maximin-efficient inspection time τ_* is defined by

$$\begin{aligned} \tau_* &:= \arg\max_{\tau_1 > 0} \min_{\lambda \in [L,U]} \frac{I_\lambda(\tau_1)}{I_\lambda(\tau(\lambda))} = \arg\max_{\tau_1 > 0} \min_{\lambda \in [L,U]} \frac{(\lambda \tau_1)^2}{\exp(\lambda \tau_1) - 1} \frac{1}{2\, x_* - x_*^2} \\ &= \arg\max_{\tau_1 > 0} \min \left\{ \frac{(L \tau_1)^2}{\exp(L \tau_1) - 1}, \frac{(U \tau_1)^2}{\exp(U \tau_1) - 1} \right\} \frac{1}{2\, x_* - x_*^2}. \end{aligned}$$

1.6 Optimal inspection times in the case of a given number of several inspections

The assumption is here that the number I of inspections are given and that an optimal choice of the inspections times $\tau_1, \ldots, \tau_I$ shall be determined. Again we assume that T_n are exponential distributed with unknown parameter λ. Then (7) becomes

$$\frac{\partial}{\partial \lambda} \ln l_\lambda(T_n) = \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial \lambda}\left(F_\lambda(\tau_i) - F_\lambda(\tau_{i-1})\right)}{F_\lambda(\tau_i) - F_\lambda(\tau_{i-1})} 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n)$$
$$= \sum_{i=1}^{I+1} \frac{\frac{\partial}{\partial \lambda}\left(e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}\right)}{e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}} 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n) = \sum_{i=1}^{I+1} \frac{\tau_i e^{-\lambda\tau_i} - \tau_{i-1} e^{-\lambda\tau_{i-1}}}{e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}} 1\!\!1_{(\tau_{i-1},\tau_i]}(T_n)$$

so that the information matrix is

$$\begin{aligned} I_\lambda(\tau_1, \ldots, \tau_I) &= E_\lambda\left(\left(\frac{\partial}{\partial \lambda} \ln l_\lambda(T_n)\right)^2\right) \\ &= \sum_{i=1}^{I+1} \left(\frac{\tau_i e^{-\lambda\tau_i} - \tau_{i-1} e^{-\lambda\tau_{i-1}}}{e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}}\right)^2 \left(e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}\right) \\ &= \sum_{i=1}^{I} \frac{\left(\tau_i e^{-\lambda\tau_i} - \tau_{i-1} e^{-\lambda\tau_{i-1}}\right)^2}{e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}} + \tau_I^2 e^{-\lambda\tau_I} \\ &= \frac{1}{\lambda^2} \sum_{i=1}^{I} \frac{\left(\lambda\tau_i e^{-\lambda\tau_i} - \lambda\tau_{i-1} e^{-\lambda\tau_{i-1}}\right)^2}{e^{-\lambda\tau_{i-1}} - e^{-\lambda\tau_i}} + (\lambda\tau_I)^2 e^{-\lambda\tau_I}. \end{aligned} \tag{9}$$

Hence to find $\tau_1^*, \ldots, \tau_I^*$ so that $I_\lambda(\tau_1, \ldots, \tau_I)$ is maximized, we have to find $x_1^*, \ldots, x_I^*$ which maximizes

$$f(x_1, \ldots, x_I) := \sum_{i=1}^{I} \frac{\left(x_i e^{-x_i} - x_{i-1} e^{-x_{i-1}}\right)^2}{e^{-x_{i-1}} - e^{-x_i}} + x_I^2 e^{-x_I},$$

which can be determined numerically. Then (9) is maximized by $\tau_1^* = \frac{x_1^*}{\lambda}, \ldots, \tau_1^* = \frac{x_1^*}{\lambda}$.
If the aim is to find the interval length Δ of equidistant inspections times $\tau_1 = \Delta, \tau_2 = 2\Delta, \ldots, \tau_I = I\Delta$, then (9) becomes

$$I_\lambda(\Delta) = \frac{1}{\lambda^2} \sum_{i=1}^{I} \frac{\left(\lambda i\Delta e^{-\lambda i\Delta} - \lambda(i-1)\Delta e^{-\lambda(i-1)\Delta}\right)^2}{e^{-\lambda(i-1)\Delta} - e^{-\lambda i\Delta}} + (\lambda I\Delta)^2 e^{-\lambda I\Delta}.$$

Maximization of $I_\lambda(\Delta)$ with respect to Δ is equivalent to the maximization of

$$f(x) := \sum_{i=1}^{I} \frac{\left(i\,x e^{-i\,x} - (i-1)\,x e^{-(i-1)\,x}\right)^2}{e^{-(i-1)x} - e^{-i\,x}} + (I\,x)^2 e^{-I\,x}.$$

Hence the maximum Δ_* of $I_\lambda(\Delta)$ is given by $\Delta_* = \frac{x_*}{\lambda}$ where f has a maximum at x_*. The time horizon is here $\tau^* = I\Delta_*$. Here Δ_* depends on I so that we may also write $\Delta_*(I)$.

If the time horizon should not exceed a given value τ then I_* might me chosen so that

$$I_* := \max\{I \in \mathbb{N};\ I\,\Delta_*(I) \leq \tau\}.$$

2 Lifetimes with different starting times

2.1 Known individual death and birth dates

Let be $D_1, \ldots, D_N$ independent nonnegative random variables (death dates), but here with different starting times (birth dates) $B_1, \ldots, B_N$ so that $D_1 - B_1, \ldots, D_N - B_N$ are identically distributed nonnegative random variables, each with distribution given by the cumulative distribution function F_θ and the density f_θ, where θ is again the unknown parameter. However neither the realizations $d_1, \ldots, d_N$ of $D_1, \ldots, D_N$ nor the realizations $b_1, \ldots, b_N$ of $B_1, \ldots, B_N$ are observed. We observe only whether d_n or b_n is lying in $(\tau_{i-1}, \tau_i]$ or not. This means that the realizations $\tilde{d}_n$ and $\tilde{b}_n$ of the discretized random variables

$$\tilde{D}_n := \sum_{i=1}^{I+1} \tau_i \mathbb{1}_{(\tau_{i-1},\tau_i]}(D_n) \quad \text{and} \quad \tilde{B}_n := \sum_{i=1}^{I+1} \tau_i \mathbb{1}_{(\tau_{i-1},\tau_i]}(B_n)$$

are observed for $n = 1, \ldots, N$. Hence

$$n_i := \sum_{n=1}^{N} \mathbb{1}_{(\tau_{i-1},\tau_i]}(\tilde{d}_n - \tilde{b}_n)$$

can be observed and is a realization of

$$N_i := \sum_{n=1}^{N} \mathbb{1}_{(\tau_{i-1},\tau_i]}(\tilde{D}_n - \tilde{B}_n).$$

Since

$$P_\theta(\tilde{D}_n - \tilde{B}_n \in (\tau_{i-1}, \tau_i]) \approx P_\theta(D_n - B_n \in (\tau_{i-1}, \tau_i]) = F_\theta(\tau_i) - F_\theta(\tau_{i-1}),$$

we can treat $\tilde{D}_n - \tilde{B}_n$ as the T_n of Section 1 so that all results of Section 1 can be used for $\tilde{D}_n - \tilde{B}_n$.

2.2 Unknown individual death and birth dates

Sometimes it could happen that only

$$n_i := \text{ number of deaths in } (\tau_{i-1}, \tau_i]$$

and

$$k_i := \text{ number of births in } (\tau_{i-1}, \tau_i]$$

is reported for $i = 1, \dots, I$, so that individual death and birth dates are not available. Then additionally

$$m_i = N - n_i + k_i, \;\; m_0 = N,$$

are available where N is the starting number of individuals. We assume here that shortly after a death a birth happens. This is a realistic assumption in repair systems where a failed component is replaced by a new component. If the replacement would happen at once then $m_i = N$ for all $i = 1, \dots, I$. However, we allow here that the replacement might happen with some delay so that m_i can be different from N. Additionally we assume that the lifetime D of an individual has exponential distribution with unknown parameter λ. Then the hazard rate is given by

$$\begin{aligned} h_i &:= P(D \in (\tau_{i-1}, \tau_i] | D > \tau_{i-1}) = \frac{P(D \in (\tau_{i-1}, \tau_i])}{P(D > \tau_{i-1})} \\ &= \frac{\int_{\tau_{i-1}}^{\tau_i} \lambda e^{-\lambda t} dt}{\int_{\tau_{i-1}}^{\infty} \lambda e^{-\lambda t} dt} = \frac{e^{-\lambda \tau_{i-1}} - e^{-\lambda \tau_i}}{e^{-\lambda \tau_{i-1}}} = 1 - e^{-\lambda(\tau_i - \tau_{i-1})} = 1 - e^{-\lambda \Delta_i} \end{aligned}$$

with $\Delta_i := \tau_i - \tau_{i-1}$. If $\Delta_i = \Delta$ for all $i = 1, \dots, I$, i.e. the inspections times are equidistant, then

$$h := h_1 = \dots h_I,$$

i.e. the hazard rates are equal. The hazard rate h can be estimated by

$$\widehat{h} := \frac{1}{I} \sum_{i=1}^{I} \frac{n_i}{m_{i-1}}$$

since n_i is the number of observed events (deaths) and m_{i-1} is the number of individuals at risk (see Tutz and Schmid 2016, p.17).

Let be n_i a realizations of N_i and m_{i-1} a realizations of M_{i-1} for $i = 1, \dots, I$. Given $M_{i-1} = m_{i-1}$ we have

$$N_i = \sum_{m=1}^{m_{i-1}} \mathbb{1}_{(\tau_{i-1}, \tau_i]}(D_m) \sim Bin(m_{i-1}, h)$$

where $D_1, \dots, D_{m_{i-1}}$ are the variables under risk in $(\tau_{i-1}, \tau_i]$ and $Bin(m_{i-1}, h)$ denotes the binomial distribution with parameters m_{i-1} and h. Then we have $\mathrm{E}_\lambda(N_i | M_{i-1} = m_{i-1}) = m_{i-1} h$ as conditional expectation and $\mathrm{var}_\lambda(N_i | M_{i-1} = m_{i-1}) = m_{i-1} h(1 - h)$ as conditional variance which implies

$$\begin{aligned} &\mathrm{E}_\lambda(\widehat{h} | M_0 = m_0, \dots, M_{I-1} = m_{I-1}) \\ &= \mathrm{E}_\lambda \left(\frac{1}{I} \sum_{i=1}^{I} \frac{N_i}{M_{i-1}} \Big| M_0 = m_0, \dots, M_{I-1} = m_{I-1} \right) = \frac{1}{I} \sum_{i=1}^{I} \mathrm{E}_\lambda \left(\frac{N_i}{M_{i-1}} \Big| M_{i-1} = m_{i-1} \right) = h. \end{aligned}$$

Since we assume that shortly after a death a birth happens we can assume that $N_1, \dots, N_I$ are almost stochastically independent so that

$\mathrm{var}_\lambda(\widehat{h} | M_0 = m_0, \dots, M_{I-1} = m_{I-1})$

$$
\begin{aligned}
&= \operatorname{var}_\lambda\left(\frac{1}{I}\sum_{i=1}^{I}\frac{N_i}{M_{i-1}}\Big|M_0=m_0,\ldots,M_{I-1}=m_{I-1}\right) \approx \frac{1}{I^2}\sum_{i=1}^{I}\operatorname{var}_\lambda\left(\frac{N_i}{M_{i-1}}\Big|M_{i-1}=m_{i-1}\right)\\
&= \frac{1}{I^2}\sum_{i=1}^{I}\frac{1}{m_{i-1}^2}\operatorname{var}_\lambda\left(N_i\Big|M_{i-1}=m_{i-1}\right) = \frac{1}{I^2}\sum_{i=1}^{I}\frac{h(1-h)}{m_{i-1}}.
\end{aligned}
$$

Usually, the interest does not ly in h but in λ which is a nonlinear aspect of h, namely we have

$$
h = 1-e^{-\lambda\Delta} \Longleftrightarrow e^{-\lambda\Delta} = 1-h \Longleftrightarrow -\lambda\Delta = \ln(1-h) \Longleftrightarrow \lambda = -\frac{\ln(1-h)}{\Delta}
$$

so that

$$
\widehat{\lambda} := -\frac{\ln(1-\widehat{h})}{\Delta}.
$$

Taylor expansion leads to the approximation

$$
\widehat{\lambda} := -\frac{\ln(1-\widehat{h})}{\Delta} \approx -\frac{\ln(1-h)}{\Delta} + \frac{1}{\Delta}\frac{1}{1-h}(\widehat{h}-h)
$$

implying

$$
\begin{aligned}
&\mathrm{E}_\lambda(\widehat{\lambda}|M_0=m_0,\ldots,M_{I-1}=m_{I-1})\\
\approx\ & -\frac{\ln(1-h)}{\Delta} + \frac{1}{\Delta}\frac{1}{1-h}(\mathrm{E}_\lambda(\widehat{h}|M_0=m_0,\ldots,M_{I-1}=m_{I-1})-h) = -\frac{\ln(1-h)}{\Delta} = \lambda
\end{aligned}
$$

and

$$
\begin{aligned}
&\operatorname{var}_\lambda(\widehat{\lambda}|M_0=m_0,\ldots,M_{I-1}=m_{I-1})\\
\approx\ & \frac{1}{\Delta^2}\frac{1}{(1-h)^2}\operatorname{var}_\lambda(\widehat{h}|M_0=m_0,\ldots,M_{I-1}=m_{I-1})\\
\approx\ & \frac{1}{\Delta^2}\frac{1}{(1-h)^2}\frac{1}{I^2}\sum_{i=1}^{I}\frac{h(1-h)}{m_{i-1}} = \frac{1}{\Delta^2}\frac{h}{1-h}\frac{1}{I^2}\sum_{i=1}^{I}\frac{1}{m_{i-1}}.
\end{aligned}
$$

Then we get

$$
\mathrm{E}_\lambda(\widehat{\lambda}) \approx \mathrm{E}_\lambda(\mathrm{E}_\lambda(\widehat{\lambda}|M_1,\ldots,M_I)) = \lambda.
$$

We can also assume that $M_0,\ldots,M_{I-1}$ are identically distributed with $\mathrm{E}_\lambda\left(\frac{1}{M_i}\right) = \mu$. Since we assume that shortly after a death a birth happens, we obtain the approximation $\mathrm{E}_\lambda\left(\frac{1}{M_i}\right) = \mu \approx \frac{1}{N}$ where N is the starting number of individuals. In particular $\mathrm{E}_\lambda\left(\frac{1}{M_i}\right)$ is independent of the interval length Δ. With these assumptions, we get

$$
\begin{aligned}
&\operatorname{var}_\lambda(\widehat{\lambda})\\
&= \mathrm{E}_\lambda\left(\operatorname{var}_\lambda(\widehat{\lambda}|M_0,\ldots,M_{I-1})\right) + \operatorname{var}_\lambda\left(\mathrm{E}_\lambda(\widehat{\lambda}|M_0,\ldots,M_{I-1})\right) \approx \mathrm{E}_\lambda\left(\operatorname{var}_\lambda(\widehat{\lambda}|M_0,\ldots,M_{I-1})\right)\\
&\approx \mathrm{E}_\lambda\left(\frac{1}{\Delta^2}\frac{h}{1-h}\frac{1}{I^2}\sum_{i=1}^{I}\frac{1}{M_{i-1}}\right) \approx \frac{1}{\Delta^2}\frac{h}{1-h}\frac{1}{I}\frac{1}{N} = \frac{1}{\Delta^2}\frac{1}{I}\frac{1}{N}\frac{1-e^{-\lambda\Delta}}{e^{-\lambda\Delta}}\\
&= \frac{1}{\Delta^2}\frac{1}{I}\frac{1}{N}\left(e^{\lambda\Delta}-1\right) = \frac{\lambda^2}{(\lambda\Delta)^2}\frac{1}{I}\frac{1}{N}\left(e^{\lambda\Delta}-1\right). \qquad (10)
\end{aligned}
$$

2.3 Optimal length of inspection intervals for unknown individual death and birth dates

2.3.1 Optimal length of inspection intervals with unrestricted time horizon

The optimal length Δ_* of inspection intervals has to minimize (10). Finding a minimum of (10) is equivalent to find a maximum of

$$f(x) := \frac{x^2}{\exp(x) - 1}$$

which was already considered in Section 1.5. Again the optimal length Δ_* is given by $\Delta_* = \frac{x_*}{\lambda}$ where f has a maximum at x_*. The time horizon is here $\tau^* = I\Delta_*$.

2.3.2 Optimal length of inspection intervals with given time horizon

If $\tau = \tau_I$ is a given time horizon then $\Delta = \frac{\tau_I}{I} = \frac{\tau}{I} \Leftrightarrow I = \frac{\tau}{\Delta}$ and (10) becomes

$$\mathrm{var}_\lambda(\widehat{\lambda}) = \frac{1}{\Delta^2}\frac{\Delta}{\tau}\frac{1}{N}\left(e^{\lambda\Delta} - 1\right) = \frac{1}{\Delta}\frac{1}{\tau}\frac{1}{N}\left(e^{\lambda\Delta} - 1\right) = \frac{\lambda}{\tau}\frac{1}{N}\frac{1}{\lambda\Delta}\left(e^{\lambda\Delta} - 1\right). \tag{11}$$

Hence to find a Δ_* which minimizes $\mathrm{var}_\lambda(\widehat{\lambda})$, we have to find the minimum of $\frac{1}{x}\left(e^x - 1\right)$ or the maximum of

$$f(x) := \frac{x}{\exp(x) - 1}.$$

Since

$$\begin{aligned} f'(x) &= \frac{1}{\exp(x) - 1} - \frac{x\exp(x)}{(\exp(x) - 1)^2} \\ &= \frac{1}{\exp(x) - 1}\left(1 - \frac{x\exp(x)}{\exp(x) - 1}\right) = \frac{1}{\exp(x) - 1}\left(1 - \frac{x}{1 - \exp(-x)}\right) < 0, \end{aligned}$$

for $x \geq 0$ if and only if

$$g(x) := 1 - \exp(-x) - x < 0$$

and $g(0) = 0$ with

$$g'(x) = \exp(-x) - 1 < 0 \text{ for all } x \geq 0,$$

the function f is monotone decreasing from 1 to 0 on $[0, \infty)$. Hence its maximum is at $x_* = 0$ which means that Δ should be chosen as small as possible.

References

[1] Tutz, G. and Schmid, M. (2016). *Modeling Discrete Time-to-Event Data*. Springer. New York.