CensoPy – Statistical computations for exponentially distributed lifetimes under hybrid censoring

Jakob I. Cramer, Erhard Cramer*

Institute of Statistics, RWTH Aachen University, D-52056 Aachen, Germany

This paper presents the Python package `CensoPy` for statistical computations of certain statistics (e.g., maximum likelihood estimators) for exponentially distributed lifetimes under hybrid censoring. Due to numerical instabilities caused by the complicated structure of the probability density function and cumulative distribution functions of the maximum likelihood estimator under hybrid censoring, the computation of confidence intervals is usually affected by numerical issues even for moderate sample sizes. In order to address these issues, we propose to use the mpmath library which allows to increase the number of digits to avoid these behavior. Of course, the price to pay is a higher computational time but, as is shown, it allows to use the pivoting method also for large sample sizes. Furthermore, we take advantage of both the modularization of hybrid censoring schemes and the B-spline representations to implement an efficient computation of the density functions and cumulative distribution functions, respectively. The code in its present form uses also the Cox-de Boor approach to compute B-spline basis functions efficiently and can be applied to both hybrid and progressive hybrid censored data.

Keywords: Likelihood inference, exponential distribution, hybrid censoring, generalized order statistics, progressive hybrid censoring, progressive censoring, modularization.

1 Introduction

Hybrid censoring of exponentially distributed lifetimes has been found great attention in the literature since the model has been introduced by Epstein (1954). Starting from a (complete) sample of order statistics $X_{1:n}, \ldots, X_{n:n}$ based on independent and identically distributed (iid) random variables, a random counter D_{HCS} determines the (observed) sample size under a particular hybrid censoring scheme, that is, the effectively observed sample is given by

$$X_{1:n}, \ldots, X_{D_{\mathsf{HCS}}:n} \tag{1.1}$$

with random sample size D_{HCS}. Thus, given $D_{\mathsf{HCS}} = d(> 0)$, the first d failures times $X_{1:n}, \ldots, X_{d:n}$ are observed. Note that D_{HCS} depends only on the design of the hybrid censoring scheme and is determined by the observed values of the sample. Of course, for some hybrid censoring schemes, it is possible that no failures are observed, that is, $D_{\mathsf{HCS}} = 0$. This happens when Type-I (or time) censoring with a threshold T is an essential part of the censoring scheme. However, in that case, it is known that all censored lifetimes must exceed T. Therefore, inference is usually carried out conditionally on the event $\{D_{\mathsf{HCS}} > 0\}$ for hybrid censoring schemes with $\Pr(D_{\mathsf{HCS}} = 0) > 0$. In is worth mentioning that the hybrid censoring schemes published so far can be categorized in two groups: hybrid censoring schemes with

*Corresponding author; Email: erhard.cramer@rwth-aachen.de

(i) bounded test duration but possibly sample size equal to zero, that is, $\Pr(D_{\mathsf{HCS}} = 0) > 0$, and

(ii) unbounded test duration but minimum guaranteed sample size, that is, $\Pr(D_{\mathsf{HCS}} > 0) = 1$.

A general account to construct hybrid censoring schemes has been proposed in van Bentum and Cramer (2025) who presented a method how hybrid censoring schemes can be obtained for given thresholds $T_1 < T_2 < \cdots < T_k$ and censoring numbers $m_1 > \cdots > m_k$. In Section 2, we will recall their construction and provide an improved algorithm that clearly explains the generation process. This insight will be very helpful in formulating an efficient algorithm for the computation of density functions and cumulative distribution functions of certain statistics and related quantities. Furthermore, we will address numerical issues that are usually present in recent implementations to compute these functions. Details are given below. To avoid these computational issues, we provide the Python package `CensoPy` which is available at https://censopy-738e78.pages.git.nrw/.

The study of the exact (conditional) distribution of the maximum likelihood estimator (MLE) of the expected value for exponentially distributed lifetimes has been addressed in many publications (see, e.g., Bartholomew (1963), Chen and Bhattacharyya (1988), Childs et al. (2003), Chandrasekar et al. (2004), Cramer et al. (2016), Górny and Cramer (2016)). Surveys of such results are provided in the recent monograph Balakrishnan et al. (2023) and the review Balakrishnan and Kundu (2013), respectively. In Górny and Cramer (2018b), a general account to the available hybrid censoring schemes has been proposed to study the distributions of certain statistics in a unifying way. This *modularization approach* has been found a useful tool to derive results for the MLEs. The result has been refined and further developed in Cramer (2025) who pointed out that the distributional assumptions on the data and the mechanism of the hybrid censoring scheme can be decoupled. The basic idea behind this approach is a decomposition of the joint distribution into some basic modules with an identical probabilistic structure. This paper revisits the approach proposed in Górny and Cramer (2018b), simplifies and refines it, so that it can be easily used for both analysing known hybrid censoring schemes and designing new ones. It is worth mentioning that similar issues arise in the computation of the MLE's distribution in further models based on censored exponentially distributed lifetimes. This includes models with step-stress testing (see, e.g., Balakrishnan et al. (2007)), competing risks (see, e.g., Kundu and Gupta (2007)), and stage life testing (see, e.g., Laumen and Cramer (2021), Uddin et al. (2025)). Further examples can be found in, e.g., Balakrishnan et al. (2023) and Kundu and Ganguly (2017) as well as references cited therein. The subsequently presented strategies can also be applied in these frameworks using appropriately adapted algorithms and approaches.

In this paper, we focus on the computation of density functions and cumulative distribution functions of certain statistics resulting in the statistical analysis of hybrid censored exponentially distributed independent lifetimes $X_1, \ldots, X_n$ where $X_i \sim \mathrm{Exp}(\vartheta)$, $\vartheta > 0$, that is, the corresponding density function is given by

$$f_\vartheta(t) = \begin{cases} \frac{1}{\vartheta} e^{-t/\vartheta}, & t > 0 \\ 0, & \text{otherwise} \end{cases}.$$

The scale parameter ϑ is supposed unknown, Θ denotes the correspondig sample space. Denoting by W_{HCS} the test duration of the hybrid censored life test,

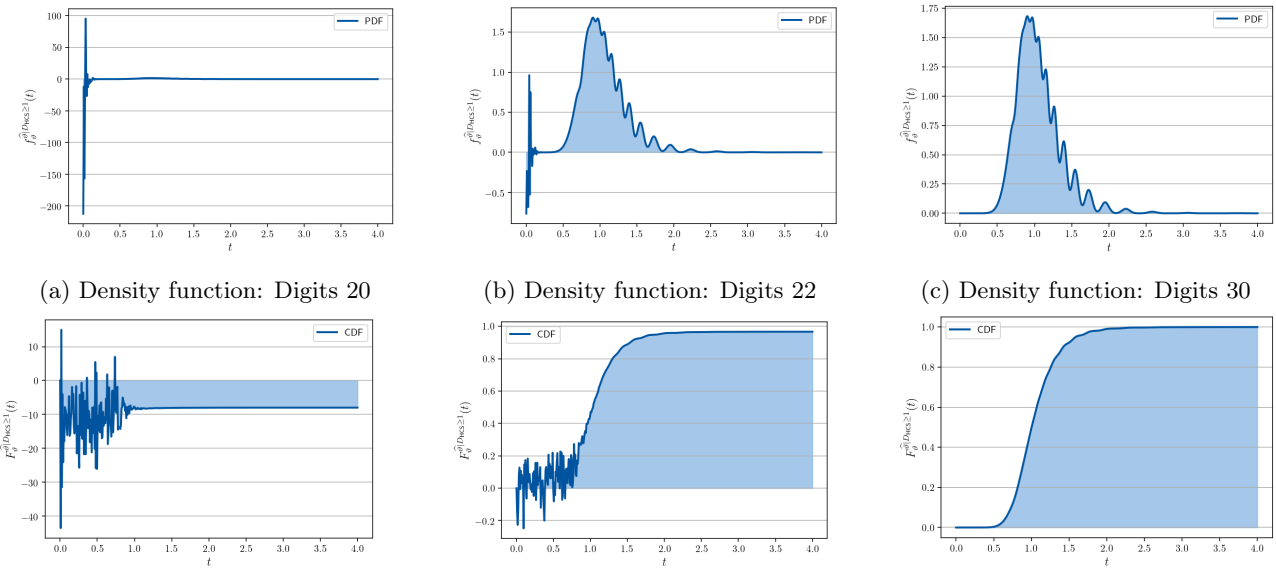
$$\mathtt{TTT}_{D_{\mathsf{HCS}}} = \sum_{j=1}^{D_{\mathsf{HCS}}} X_{j:n} + (n - D_{\mathsf{HCS}}) W_{\mathsf{HCS}}. \tag{1.2}$$

denotes the corresponding total time on test. Provided that $D_{\mathsf{HCS}} > 0$, it has been shown in Cramer (2025) that the MLE of ϑ is given by

$$\widehat{\vartheta} = \frac{\mathtt{TTT}_{D_{\mathsf{HCS}}}}{D_{\mathsf{HCS}}} = \frac{1}{D_{\mathsf{HCS}}} \left(\sum_{j=1}^{D_{\mathsf{HCS}}} X_{j:n} + (n - D_{\mathsf{HCS}}) W_{\mathsf{HCS}}. \right). \tag{1.3}$$

Many efforts have been made to derive the exact (conditional) distribution of the MLE, i.e., the derivation of the density function and cumulative distribution function. A major application of the cumulative distribution function is its use to determine exact (conditional) confidence intervals for ϑ using the method of pivoting the cumulative distribution function (see Casella and Berger (2002), Hahn et al. (2017), and Balakrishnan et al.

(2014)). A fundamental result to establish tractable representations of the cumulative distribution functions of the MLE (or the total time on test), has been initiated by Cramer and Balakrishnan (2013) who connected the representation of the corresponding density function to linear combinations of B-spline functions (for information on B-Spline functions, see de Boor (1972, 2001), Schumaker (2007)). Integrating the density functions, the cumulative distribution functions are immediate so that the resulting confidence intervals can principally be computed. However, it turned out that the resulting calculations are numerical instable (due to the incorporation of huge binomial coefficients, exponential functions with very small exponents, and alternating signs). As a result, the numerical computations took a long time and led to insufficient computational results even for moderate sample size n. The effect of the precision is illustrated in a series of plots of the density functions and cumulative distribution functions in Figure 1.1 for Type-I hybrid censoring. The situation will get worse the larger the sample size and the censoring numbers. Furthermore, the complexity of the considered hybrid censoring scheme has also a huge impact on the numerical stability of the computations. The presented plots have been generated by the Python package `CensoPy` proposed in this paper, by altering the valid digits in the computation as indicated by the captions of the plots. As the plots in Figure 1.1 show, the critical part of the computation occurs particularly for 'small' values of the argument (somewhat close to zero); in other words, computational problems therefore occur primarily at the left margin of the support. However, the same behavior is observed in alternative implementations using Python, R, etc. Clearly, these numerical issues have also a huge impact on the computation of the confidence intervals for the expected lifetime which are based on the cumulative distribution function of the MLE (see Section 3.3.2) so that there is need for a numerically stable code.



(a) Density function: Digits 20

(b) Density function: Digits 22

(c) Density function: Digits 30

(d) Cumulative distribution function: Digits 20

(e) Cumulative distribution function: Digits 22

(f) Cumulative distribution function: Digits 30

Figure 1.1: Effect of precision (in terms of digits) in plotting the density function and cumulative distribution function of a Type-I hybrid censoring scheme with $n = 40$, $T_1 = 0.5$, $m_1 = 20$, $\vartheta = 1$.

In this paper, we address these issues by coding the algorithm in Python using the packages *mpmath* (The mpmath development team (2026)) and *numpy* (Harris et al. (2020)). The code of the Python package `CensoPy` is available for download at GitLab (link is given above) where also a detailed documentation of its use is provided. The technical details of the computations and the statistical background are presented in Section 2 where we explain in general how hybrid censoring schemes are constructed and can be designed. The key factor that makes these computations possible is, in particular, the *mpmath* package, which allows computation for any number of digits. Although this does the job, it results, of course, in a reduction in speed. As mentioned above, an implementation in R yields also bad results due to limitations in precision of the values. The R package *Rmpfr* (see Maechler (2025)) can be applied to address this issue using R. However, if both a Python

and an R compiler are available on a computer, then the package `CensoPy` can also be easily used in R using the R package *reticulate* (see Kalinowski et al. (2026)). For easy use, a corresponding user file is also provided on GitLab. In this regard, our approach provides a flexible implementation that can be used in both communities.

Throughout, we illustrate the power of the proposed approach by a new hybrid censoring scheme called generalized unified Type-IV hybrid censoring scheme that came to our mind when writing down the simple construction procedure presented in Section 2 (see Figure 2.1). The general results on the decoupling and decomposition are presented for general hybrid censoring schemes, too. In Section 3, the computation of density functions, cumulative distribution functions, and confidence intervals for hybrid censoring schemes are presented where the data is given by order statistics of an iid sample. The computation is based on explicit sum expressions valid for these B-spline basis functions. In Section 3.6, the easy use of the package is illustrated by an application to four data sets available in the literature. Section 4 is devoted to the extension to progressively Type-II censored data (for details, see Balakrishnan and Cramer (2014), Balakrishnan et al. (2023)). In order to implement the computation efficiently, we make use of the Cox-de Boor recurrence relation for B-spline basis functions (see de Boor (2001), p. 89 and eq. (4.4) in this paper).

2 Design of general hybrid censoring schemes and decomposition of distributions

In order to define general hybrid censoring schemes for $k \in \mathbb{N}$ and sample size $n \in \mathbb{N}$, we consider as in van Bentum and Cramer (2025) sequences of thresholds and censoring numbers given by

$$\begin{aligned} &\text{increasing thresholds} && -\infty = T_0 < T_1 < T_2 < \cdots < T_{k-1} < T_k = \infty \\ &\text{decreasing censoring numbers} && \infty \geq n \geq m_1 > m_2 > \cdots > m_{k-1} > m_k \geq 0. \end{aligned}$$

Thus, we have always the same number of censoring numbers and thresholds. It has already been pointed out in the literature (see Balakrishnan et al. (2023)), that the hybrid censoring schemes can be categorized in two classes:

C-I : hybrid censoring schemes with bounded test duration, but not guaranteeing measurements,

C-II: hybrid censoring schemes with unbounded test duration, but ensuring a minimum sample size.

This fact becomes very clear when one considers the following structure of the hybrid censoring schemes and the illustration in Table 2.1 where the parameterizations of those hybrid censoring schemes already discussed in the literature have been summarized. The table is organized such that it becomes evident how new hybrid censoring schemes are constructed based on previous ones. First, the censoring schemes are reported in pairs, which which differ only by the fact that a maximum number of allowed observations m_1 replaces the sample size n. Clearly, the censoring number $m_1 \leq n$ induces a Type-II censoring of the data, that is, instead of the complete ordered data $X_{1:n} \leq \cdots \leq X_{n:n}$ the Type-II censored sample $X_{1:n} \leq \cdots \leq X_{m_1:n}$ is subject to hybrid censoring. In particular, $m_1 = n$ corresponds to the base hybrid censoring schemes of these pairs. As an example, Type-II hybrid censoring can be interpreted as generalized Type-I hybrid censoring with the (maximal) choice $m_1 = n$ of the scheme parameter m_1.

As illustrated in Table 2.1, the construction of the hybrid censoring schemes results from an alternating application of Type-I and Type-II censoring mechanisms. We start with a Type-II censored sample (**C-II**) (where, as mentioned above, a complete sample is subsumed by choosing $m_1 = n$). In particular, the following construction principle applies.

(i) For a given k and a pair of **C-II**-hybrid censoring schemes, a Type-I censoring is applied to this data, that is, the test duration is bounded by introducing a threshold $T_k \in [T_{k-1}, \infty)$. The constructed pair of **C-I**-hybrid censoring schemes with k parameters is thus obtained by replacing the 'bound' ∞ by the threshold T_k (see, e.g., blue down arrow in Table 2.1 and orange colored cells).

(ii) For a given k and a pair of **C-I**-hybrid censoring schemes, the resulting pair of **C-II**-hybrid censoring schemes with $k+1$ parameters is obtained by introducing the censoring number $m_{k+1} > 0$ as a required minimum sample size (see, e.g., red up arrow in Table 2.1 and yellow colored cells). It thus overrules the Type-I censoring, by enforcing a minimum number m_{k+1} of measurements. Implicitly, this number is zero in the **C-I**-hybrid censoring schemes.

Table 2.1: Hybrid censoring schemes with types and parameters.

	Hybrid censoring scheme		censoring numbers m_i				thresholds T_j			
		k	1	2	3	4	1	2	3	4
C-II-HCS	complete sample	1	n				∞			
	Type-II	1	m_1				∞			
	Type-II hybrid[b]	2	n	m_2			T_1	∞		
	gen. Type-I hybrid[c]	2	m_1	m_2			T_1	∞		
	uni. Type-II hybrid[e]	3	n	m_2	m_3		T_1	T_2	∞	
	uni. Type-IV hybrid[f]	3	m_1	m_2	m_3		T_1	T_2	∞	
	gen. uni. Type-II hybrid[g]	4	n	m_2	m_3	m_4	T_1	T_2	T_3	∞
	gen. uni. Type-IV hybrid[new]	4	m_1	m_2	m_3	m_4	T_1	T_2	T_3	∞
C-I-HCS	Type-I	1	n				T_1			
	Type-I hybrid[a]	1	m_1				T_1			
	gen. Type-II hybrid[c]	2	n	m_2			T_1	T_2		
	uni. Type-I hybrid[d]	2	m_1	m_2			T_1	T_2		
	uni. Type-III hybrid[f]	3	n	m_2	m_3		T_1	T_2	T_3	
	gen. uni. Type-I hybrid[new]	3	m_1	m_2	m_3		T_1	T_2	T_3	

Notes: Hybrid censoring schemes introduced by [a]Epstein (1954), [b]Childs et al. (2003), [c]Chandrasekar et al. (2004), [d]Huang and Yang (2010), [e]Balakrishnan et al. (2008), [f]Park and Balakrishnan (2012), [g]Górny and Cramer (2018a)

This procedure shows how the hybrid censoring schemes are related and how new ones can be constructed by alternating the preceding steps. As indicated in the last line of Table 2.1, the illustration shows that one hybrid censoring scheme has been overlooked so far. In accordance with the other hybrid censoring schemes, we call it *generalized unified Type-I hybrid censoring scheme*. For completeness, we have also introduced a new hybrid censoring scheme called *generalized unified Type-IV hybrid censoring* which results from the censoring scheme *generalized unified Type-II hybrid* by introducing a Type-II censoring.

Remark 2.1. *It should be noted that odd type numbers typically correspond to a* ***C-I****-hybrid censoring scheme (like Type-I censoring) whereas even ones correspond to* ***C-II****-hybrid censoring schemes (like Type-II censoring). For the generalized Type-I and Type-II hybrid censoring schemes proposed by Chandrasekar et al. (2004), this is unfortunately not true as the authors have reversed the names of the types (see Table 2.1).*

For each threshold T_i, define the random counter $D_i = D(T_i)$ by

$$D_i = D(T_i) = \sum_{j=1}^{n} \mathbb{1}_{(-\infty, T_i]}(X_{j:n}), \quad i \in \{1, \ldots, k\}. \tag{2.1}$$

Note that, for $T_k = \infty$, we get $D_k = n$.

Remark 2.2. *Here, it should be noted that, for consistency, we have adapted the decision trees of the hybrid censoring schemes with $m_1 = n$ to those with $m_1 < n$ (see Figures 2.1a and 2.1b for Type-II hybrid censoring). In Balakrishnan et al. (2023) (or Cramer (2025)), the module type* ***A*** *is not present in the decision tree for these hybrid censoring schemes since it can be merged with the module type* ***O****. This is caused by the identity of the events $\{D_1 \geq n\} = \{D_1 = n\}$ since n is the maximum value that can be attained by D_1. However, for a consistent notation, we preserve it in this presentation and consider the decision trees valid in the case m_1 and choose $m_1 = n$.*

Table 2.2: Hybrid censoring schemes with decompositions induced by the random counters. Parameters are as in Table 2.1.

		module type						
		$\mathbf{A}_1$	$\mathbf{O}_1$	$\mathbf{AB}_{1,2}$	$\mathbf{O}_2$	$\mathbf{AB}_{2,3}$	$\mathbf{O}_3$	$\mathbf{B}_3$
		$\{D_1 \geq m\}$	$\{D_1 = d\}$	$\{D_1 < m, D_2 \geq m\}$	$\{D_2 = d\}$	$\{D_2 < m, D_3 \geq m\}$	$\{D_3 = d\}$	$\{D_3 < m\}$
	sample size ℓ of $X_{1:n}, \ldots, X_{\ell:n}$	m	d	m	d	m	d	m
C-II	complete	n						
	Type-II	m_1						
C-I	Type-I	n	$n-1, \ldots, 0$					
	Type-I hybrid	m_1	$m_1-1, \ldots, 0$					
C-II	Type-II hybrid	n	$n-1, \ldots, m_2$	${m_2}^*$				
	generalized Type-I hybrid	m_1	$m_1-1, \ldots, m_2$	${m_2}^*$				
C-I	generalized Type-II hybrid	n	$n-1, \ldots, m_2$	m_2	$m_2-1, \ldots, 0$			
	unified Type-I hybrid	m_1	$m_1-1, \ldots, m_2$	m_2	$m_2-1, \ldots, 0$			
C-II	unified Type-II hybrid	n	$n-1, \ldots, m_2$	m_2	$m_2-1, \ldots, m_3$	${m_3}^{**}$		
	unified Type-IV hybrid	m_1	$m_1-1, \ldots, m_2$	m_2	$m_2-1, \ldots, m_3$	${m_3}^{**}$		
C-I	unified Type-III hybrid	n	$n-1, \ldots, m_2$	m_2	$m_2-1, \ldots, m_3$	m_3	$m_3-1, \ldots, 0$	
	generalized unified Type-I hybrid	m_1	$m_1-1, \ldots, m_2$	m_2	$m_2-1, \ldots, m_3$	m_3	$m_3-1, \ldots, 0$	
C-II	generalized unified Typ-I hybrid	n	$n-1, \ldots, m_2$	m_2	$m_2-1, \ldots, m_3$	m_3	$m_3-1, \ldots, m_4$	m_4
	generalized unified Typ-III hybrid	m_1	$m_1-1, \ldots, m_2$	m_2	$m_2-1, \ldots, m_3$	m_3	$m_3-1, \ldots, m_4$	m_4

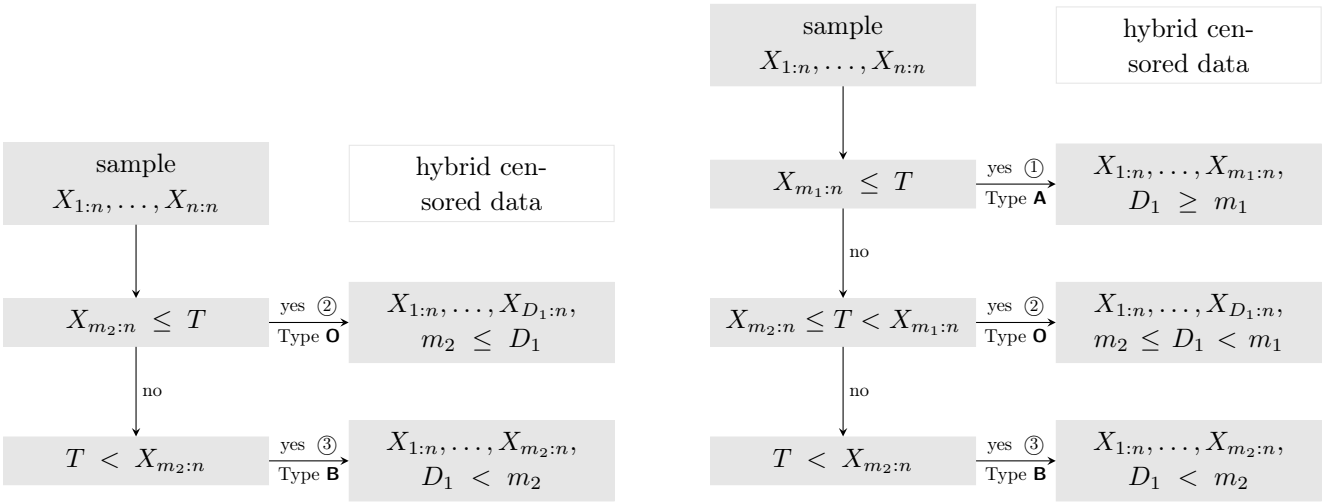
Notes: ${}^*T_2 = \infty$ *so that* $D_2 = n$ *and* $\{D_1 < m_2, D_2 \geq m_2\} = \{D_1 < m_2\}$ *Hence, module type* $\mathbf{AB}_{1,2}$ *becomes* $\mathbf{B}_1$. ${}^{**}T_3 = \infty$ *so that the same argument applies as in the case with* $T_2 = \infty$.

The subscripts at the module types show the particular counters/thresholds involved in the definition of the module type.

Table 2.3: Hybrid censoring schemes with sample sizes and test duration per module type. Parameters are as in Table 2.1.

		module type													
		A_1		O_1		$\mathsf{AB}_{1,2}$		O_2		$\mathsf{AB}_{2,3}$		O_3		B_3	
		ℓ	w	ℓ	w	ℓ	w	ℓ	w	ℓ	w	ℓ	w	ℓ	w
C-II	complete	n	$X_{n:n}$												
	Type-II	m_1	$X_{m_1:n}$												
C-I	Type-I	n	$X_{n:n}$	d_1	T_1										
	Type-I hybrid	m_1	$X_{m_1:n}$	d_1	T_1										
C-II	Type-II hybrid	n	$X_{n:n}$	d_1	T_1	m_2	$X_{m_2:n}$								
	generalized Type-I hybrid	m_1	$X_{m_1:n}$	d_1	T_1	m_2	$X_{m_2:n}$								
C-I	generalized Type-II hybrid	n	$X_{n:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2						
	unified Type-I hybrid	m_1	$X_{m_1:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2						
C-II	unified Type-II hybrid	n	$X_{n:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2	m_3	$X_{m_3:n}$				
	unified Type-IV hybrid	m_1	$X_{m_1:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2	m_3	$X_{m_3:n}$				
C-I	unified Type-III hybrid	n	$X_{n:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2	m_3	$X_{m_3:n}$	d_3	T_3		
	generalized unified Type-I hybrid	m_1	$X_{m_1:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2	m_3	$X_{m_3:n}$	d_3	T_3		
C-II	generalized unified Typ-I hybrid	n	$X_{n:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2	m_3	$X_{m_3:n}$	d_3	T_3	m_4	$X_{m_4:n}$
	generalized unified Typ-III hybrid	m_1	$X_{m_1:n}$	d_1	T_1	m_2	$X_{m_2:n}$	d_2	T_2	m_3	$X_{m_3:n}$	d_3	T_3	m_4	$X_{m_4:n}$

Notes: *The subscripts at the module types show the particular counters/thresholds involved in the definition of the module type.*



(a) Condensed decision tree as in Balakrishnan et al. (2023) with merged types **A**/**O**.

(b) Decision tree used in this paper for consistency reasons with generalized Type-I hybrid censoring (choose here $m_1 = n$).

Figure 2.1: Variants of decision tree for Type-II hybrid censoring.

Based on the decomposition of the hybrid censoring scheme as given in Table 2.2, we can obtain various quantities, like the joint distribution of the data, the (conditional) density function or cumulative distribution function of the total time on test or the maximum likelihood estimator. As an example, we present the results for both the Type-I- and Type-II-hybrid censoring scheme.

From Cramer (2025), we have the following module types induced by the counters involved in the hybrid censoring scheme ($i \in \{1, \ldots, k\}, j \in \{2, \ldots, k\}$):

$$\mathbf{O}_i\colon\ D_i = d, \qquad \mathbf{A}_i\colon\ D_i \geq m, \qquad \mathbf{B}_i\colon\ D_i < m, \qquad \mathbf{AB}_{j,j+1}\colon\ D_{j-1} < m_j, D_j \geq m_j$$

with appropriately chosen values for d, m, m_j. To make it easier to distinguish between the module types, subscripts have been introduced for the module types so that the relevant counters/thresholds can be easily identified. It has been shown in Cramer (2025) that the random counters $D_1, D_2, \ldots, D_k$ can be used to decompose hybrid censored data into the above mentioned module types. In particular, given the random sample size D_{HCS} generated by a particular hybrid censoring scheme, the counter decomposition of a **C-I**-hybrid censoring scheme is given by

$$\{D_{\mathsf{HCS}} \geq 1\} = \{D_1 \geq m_1\} \cup \bigcup_{j=1}^{k} \bigcup_{d=m_{j+1}}^{m_j - 1} \{D_j = d\} \cup \bigcup_{j=2}^{k} \{D_{j-1} < m_j, D_j \geq m_j\} \tag{2.2}$$

(where $m_{k+1} = 1$ for convenience) and that of a **C-II**-hybrid censoring scheme is given by

$$\{D_{\mathsf{HCS}} \geq 0\} = \{D_1 \geq m_1\} \cup \bigcup_{j=1}^{k-1} \bigcup_{d=m_{j+1}}^{m_j - 1} \{D_j = d\} \cup \bigcup_{j=2}^{k-1} \{D_{j-1} < m_j, D_j \geq m_j\} \cup \{D_{k-1} < m_k\}, \tag{2.3}$$

respectively.

Remark 2.3. *The sample size D_{HCS} and the test duration W_{HCS} under a certain hybrid censoring scheme can be written in terms of minima $\wedge$ and maxima $\vee$ of m_i, D_i and $X_{m_i:n}, T_i$, respectively, as follows:*

C-I : $D_{HCS} = \bigvee_{i=1}^{k} (m_i \wedge D_i)$, $W_{HCS} = \bigvee_{i=1}^{k} (X_{m_i:n} \wedge T_i)$.

Therefore, $0 \leq D_{HCS} \leq m_1$ *and* $X_{m_1:n} \wedge T_1 \leq W_{HCS} \leq T_k$.

C-II: $D_{HCS} = m_k \vee \bigvee_{i=1}^{k-1} (m_i \wedge D_{k-i})$, $W_{HCS} = X_{m_k:n} \vee \bigvee_{i=1}^{k-1} (X_{m_i:n} \wedge T_{k-i})$.

Therefore, $m_k \leq D_{HCS} \leq m_1 \wedge D_{k-1}$ *and* $X_{m_k:n} \leq W_{HCS}$.

Note that for some hybrid censoring schemes $m_1 = n$ *has to be chosen (see Table 2.3).*

The decomposition of the random counter D_{HCS} in (2.2) and (2.3), respectively, can now be used to establish a representation for the distribution of relevant statistics based on the hybrid censored data. Consider a statistic $Y = Y(X_{1:n}, \ldots, X_{D_{\mathsf{HCS}}:n})$ based on the hybrid censored sample $X_{1:n}, \ldots, X_{D_{\mathsf{HCS}}:n}$ and denote by $Y_{\ell,\mathbf{M}}$ the resulting expression of the statistic under the corresponding module type $\mathbf{M}$. Then, e.g., we get for the cumulative distribution function of Y and a **C-II**-hybrid censoring scheme (see also (2.3))

$$\begin{aligned}(2.4)\qquad \Pr(Y \leq t) &= \Pr(Y_{m_1,\mathbf{A}_1} \leq t, D_1 \geq m_1) + \sum_{j=1}^{k-1} \sum_{d=m_{j+1}}^{m_j-1} \Pr(Y_{d,\mathbf{O}_j} \leq t, D_j = d) \\ &\quad + \sum_{j=2}^{k-1} \Pr(Y_{m_j,\mathbf{AB}_{j-1,j}} \leq t, D_{j-1} < m_j, D_j \geq m_j) + \Pr(Y_{m_k,\mathbf{B}_{k-1}} \leq t, D_{k-1} < m_k).\end{aligned}$$

A similar decomposition results for **C-II**-hybrid censoring scheme considering the event $\{D_{\mathsf{HCS}} \geq 1\}$ which ensures that at least one measurement has been observed.

In the present discussion, we consider the (conditional) distribution of the total time on test statistic (see, e.g., Cramer (2025), Theorem 4.4) in (1.2) and the corresponding MLE $\widehat{\vartheta}$ in (1.3) for hybrid censored exponentially distributed lifetimes where the test duration $W_{D_{\mathsf{HCS}}}$ is given in Table 2.3 (see also Cramer (2025), Table 2, for a representation in terms of Boolean functions). Using these expressions, we get for the total time on test $Y_{\ell,\mathbf{M}} = \mathtt{TTT}_{\ell,\mathbf{M}}$ under a hybrid censoring scheme the representations

$$\mathtt{TTT}_{\ell,\mathbf{M}} = \begin{cases} \mathtt{TTT}_{d_i} = \sum_{j=1}^{d_i} X_{j:n} + (n-d_i)T_i, & \mathbf{M} = \mathbf{O}_i, i = 1, \ldots, k \\ \mathtt{TTT}_{m_1} = \sum_{j=1}^{m_1} X_{j:n} + (n-m_1)X_{m_1:n}, & \mathbf{M} = \mathbf{A}_1 \\ \mathtt{TTT}_{m_i} = \sum_{j=1}^{m_i} X_{j:n} + (n-m_i)X_{m_i:n}, & \mathbf{M} = \mathbf{AB}_{i-1,i}, i = 2, \ldots, k-1 \\ \mathtt{TTT}_{m_k} = \sum_{j=1}^{m_k} X_{j:n} + (n-m_k)X_{m_k:n}, & \mathbf{M} = \mathbf{B}_k \end{cases}.$$

This illustrates that finally we have only two different forms of the total time on test. Furthermore, we can write

$$\Pr_\vartheta(\mathtt{TTT}_{m_i} \leq t, D_{i-1} < m_i, D_i \geq m_i) = F_\vartheta^{\mathtt{TTT}_{m_i}, D_i \geq m_i}(t) - F_\vartheta^{\mathtt{TTT}_{m_i}, D_{i-1} \geq m_i}(t)$$

since $D_i \leq D_{i+1}$ so that module type $\mathbf{AB}_{i-1,i}$ can be evaluated using module type $\mathbf{A}_i$, $i = 2, \ldots, k$, with different thresholds. Finally, we have

$$\Pr_\vartheta(\mathtt{TTT}_{m_k} \leq t, D_k < m_k) = F_\vartheta^{\mathtt{TTT}_{m_k}}(t) - F_\vartheta^{\mathtt{TTT}_{m_k}, D_k \geq m_k}(t),$$

where, for exponentially distributed lifetimes, $\mathtt{TTT}_{m_k}$ has a gamma distribution with parameters m_k and ϑ. These identities show that we only need the following cumulative distribution functions

$$(2.5)\qquad F_\vartheta^{\mathtt{TTT}_d, D=d}, \quad F_\vartheta^{\mathtt{TTT}_m, D \geq m}$$

for certain values of d, m and thresholds T to compute the necessary cumulative distribution functions in the decompositions (2.2) and (2.3).

As pointed out in Cramer (2025), the law of total probability can be applied to establish the desired quantities from the decompositions (2.2) and (2.3). Thus, in view of (2.4), we get for a general **C-I**-hybrid censoring scheme

$$\begin{aligned}\Pr_\vartheta(\mathtt{TTT}_{D_{\mathsf{HCS}}} \leq t, D_{\mathsf{HCS}} \geq 1) &= \Pr_\vartheta(\mathtt{TTT}_{m_1} \leq t, D_1 \geq m_1) \\ &\quad + \sum_{j=1}^{k} \sum_{d=m_{j+1}}^{m_j-1} \Pr_\vartheta(\mathtt{TTT}_d \leq t, D_j = d) + \sum_{j=2}^{k} \Pr_\vartheta(\mathtt{TTT}_{m_j} \leq t, D_{j-1} < m_j, D_j \geq m_j)\end{aligned}$$

$$= F_\vartheta^{\mathtt{TTT}_{m_1}, D_1 \geq m_1}(t) + \sum_{j=1}^{k} \sum_{d=m_{j+1}}^{m_j - 1} F_\vartheta^{\mathtt{TTT}_d, D_j = d}(t) + \sum_{j=2}^{k} \left[F_\vartheta^{\mathtt{TTT}_{m_j}, D_j \geq m_j}(t) - F_\vartheta^{\mathtt{TTT}_{m_j}, D_{j-1} \geq m_j}(t) \right]$$

$$= \sum_{j=1}^{k} \sum_{d=m_{j+1}}^{m_j - 1} F_\vartheta^{\mathtt{TTT}_d, D_j = d}(t) + \sum_{j=1}^{k} F_\vartheta^{\mathtt{TTT}_{m_j}, D_j \geq m_j}(t) - \sum_{j=2}^{k} F_\vartheta^{\mathtt{TTT}_{m_j}, D_{j-1} \geq m_j}(t). \tag{2.6}$$

By analogy, we find the representation

$$\begin{aligned} \Pr_\vartheta(\mathtt{TTT}_{D_{\mathsf{HCS}}} \leq t) &= \Pr_\vartheta(\mathtt{TTT}_{m_1} \leq t, D_1 \geq m_1) \\ &\quad + \sum_{j=1}^{k} \sum_{d=m_{j+1}}^{m_j - 1} \Pr_\vartheta(\mathtt{TTT}_d \leq t, D_j = d) + \sum_{j=2}^{k-1} \Pr_\vartheta(\mathtt{TTT}_{m_j} \leq t, D_{j-1} < m_j, D_j \geq m_j) \\ &\quad + \Pr_\vartheta(\mathtt{TTT}_{m_k} \leq t, D_{k-1} < m_k) \\ &= F_{\Gamma(m_k, \vartheta)}(t) + \sum_{j=1}^{k-1} \sum_{d=m_{j+1}}^{m_j - 1} F_\vartheta^{\mathtt{TTT}_d, D_j = d}(t) + \sum_{j=1}^{k-1} F_\vartheta^{\mathtt{TTT}_{m_j}, D_j \geq m_j}(t) - \sum_{j=2}^{k} F_\vartheta^{\mathtt{TTT}_{m_j}, D_{j-1} \geq m_j}(t) \end{aligned} \tag{2.7}$$

for a general **C-II**-hybrid censoring scheme.

Remark 2.4. *For **C-I**-hybrid censoring schemes, the life test can end without observing a failure, that is, the probability $Pr_\vartheta(D_{\mathsf{HCS}} = 0)$ is positive for any $\vartheta > 0$. In such a situation, the distribution of statistics of interest will typically be obtained conditionally on the event $\{D_{\mathsf{HCS}} > 0\}$, that is, we assume that at least one failure has been observed. Therefore, the conditional cumulative distribution function*

$$Pr_\vartheta(\mathtt{TTT}_{D_{\mathsf{HCS}}} \leq t \mid D_{\mathsf{HCS}} > 0) = \frac{Pr_\vartheta(\mathtt{TTT}_{D_{\mathsf{HCS}}} \leq t, D_{\mathsf{HCS}} \geq 1)}{1 - Pr_\vartheta(D_{\mathsf{HCS}} = 0)}$$

is considered which can be easily obtained from (2.6). Furthermore, note that $\{D_{\mathsf{HCS}} > 0\} = \{D_k > 0\}$ so that we have

$$Pr_\vartheta(D_{\mathsf{HCS}} = 0) = Pr_\vartheta(D_k = 0) = Pr_\vartheta(X_{1:n} > T_k) = e^{-nT_k/\vartheta}.$$

Remark 2.5. *The preceding expressions have been derived assuming exponentially distributed lifetimes and the total time on test. However, the decompositions hold for any statistic Y depending on the hybrid censored data and any baseline cumulative distribution function F. However, for exponentially distributed lifetimes and the total time on test (or the MLE of ϑ) explicit expressions are available for the density functions and cumulative distribution functions, respectively (see Section 3. Therefore, the results can also be applied for other distributions when expressions for the specific module types can be established (or at least approximated).*

3 Computations with Python package `CensoPy`

In its present form, the Python package `CensoPy` offers the following features:

(i) Implementation of general hybrid censoring schemes as shown in Table 2.1 for order statistics (`HCS`: hybrid censoring schemes) and progressively Type-II censored order statistics (`PCS`: progressive hybrid censoring schemes)

(ii) Computation and plotting of density functions and cumulative distribution functions for general (progressive) hybrid censoring schemes.

(iii) Computation of exact (conditional) confidence intervals for both given data and simulated data.

(iv) Simulation of (progressive) hybrid censored data for exponential distributions.

(v) Plots of (conditional) cumulative distribution functions as a function of the parameter.

(vi) Plots of bivariate (conditional) density functions and cumulative distribution functions as a function of the parameter and the argument.

A detailed documentation is available on GitLab. In the following, the theoretical background used in the computations is presented and the some examples of its use are shown.

3.1 Computation and plotting of (conditional) density functions and cumulative distribution functions of total time on test statistic using B-spline basis functions

First, we consider the distribution of the total time on test statistic. The distribution of the MLE is then easily obtained from these expressions.

Expressions for density functions of the total time on test under hybrid censoring in terms of B-spline basis functions have been derived first by Cramer and Balakrishnan (2013). In this case, the following expressions for B-spline basis functions are used

$$B_0(s \mid a, b) = \frac{1}{b-a}\mathbb{1}_{[a,b)}(s), \quad a < b, s \in \mathbb{R} \tag{3.1}$$

For $d \in \{1, \ldots, n\}$, we have

$$B_{d-1}(s \mid (n-d)T, \ldots, nT) = \frac{1}{T^d(d-1)!}\sum_{i=0}^{d}(-1)^i\binom{d}{i}\big[(n-i)T - s\big]_+^{d-1}, \tag{3.2}$$

$$\begin{aligned} B_{d-1}(s \mid 0, (n-d+1)T, \ldots, nT) = {} & \frac{1}{T^d(d-1)!}\sum_{i=0}^{d-1}(-1)^i\binom{d}{i}\frac{d-i}{n-i}\big[(n-i)T - s\big]_+^{d-1} \\ & + (-1)^d\frac{d(n-d)!}{n!T^d}[-s]_+^{d-1}, \end{aligned} \tag{3.3}$$

where $[z]_+ = \max(z, 0)$ and $0^0 = 1$. The B-spline basis functions in (3.1)-(3.3) are normalized, that is, they can be interpreted as density functions (of linear combinations of uniform order statistics; see Ignatov and Kaishev (1989), Lemma 2.2, Agarwal et al. (2002), Theorem 3.1) so that $\int_{c_0}^{c_d} B_{d-1}(s \mid c_0, \ldots, c_d)dt = 1$ where $c_0 < \cdots < c_d$. Note that, for $s \geq 0$, the representation in (3.3) simplifies to

$$B_{d-1}(s \mid 0, (n-d+1)T, \ldots, nT) = \frac{1}{T^d(d-1)!}\sum_{i=0}^{d-1}(-1)^i\binom{d}{i}\frac{d-i}{n-i}\big[(n-i)T - s\big]_+^{d-1}$$

since the last term equals zero. Using the B-spline functions defined in (3.1)–(3.3), the density functions of the total time on test can be written as in Theorem 3.1.

Theorem 3.1 (Górny and Cramer (2018b), Cramer (2025))**.** *Let $1 \leq d \leq m-1$ and $D = D(T)$ as in (2.1). The density function $f^{\mathit{TTT}_d, D=d}$ is given by*

$$f_\vartheta^{\mathit{TTT}_d, D=d}(s) = \binom{n}{d}\frac{T^d}{\vartheta^d}B_{d-1}(s \mid (n-d)T, \ldots, nT)\mathsf{e}^{-s/\vartheta}, \quad 0 \leq s \leq nT. \tag{3.4}$$

Furthermore, the density function $f^{\mathit{TTT}_m, D \geq m}$ is given by

$$f_\vartheta^{\mathit{TTT}_m, D \geq m}(s) = \binom{n}{m}\frac{T^m}{\vartheta^m}B_{m-1}(s \mid 0, (n-m+1)T, \ldots, nT)\mathsf{e}^{-s/\vartheta}, \quad 0 \leq s \leq nT.$$

The density function of TTT_m and $D < m$ is given by

$$f_\vartheta^{\mathit{TTT}_m, D<m}(s) = \frac{s^{m-1}\mathsf{e}^{-s/\vartheta}}{(m-1)!\vartheta^m} - f_\vartheta^{\mathit{TTT}_m, D \geq m}(s), \quad s \geq 0.$$

The corresponding cumulative distribution functions are now directly obtained by integration. In order to find suitable expressions for the cumulative distribution functions, we derive the corresponding anti-derivative of the density functions. We use the representations in (3.1) and (3.3) to calculate the integral $\int_0^x \big[M - s\big]_+^{d-1}\mathsf{e}^{-s/\vartheta}\,ds$ for $M = (n-i)T$. Thus, let $x \geq 0$ and $i \in \{0, \ldots, n-1\}$. For $d \in \{1, \ldots, n\}$, we have

$$\begin{aligned} \int_0^x \Big[M - s\Big]_+^{d-1}\mathsf{e}^{-s/\vartheta}\,ds &= \int_0^{\min(x,M)} \Big(M - s\Big)^{d-1}\mathsf{e}^{-s/\vartheta}\,ds \\ &= -\mathsf{e}^{-M/\vartheta}\int_M^{M-\min(x,M)} y^{d-1}\mathsf{e}^{y/\vartheta}\,dy = \mathsf{e}^{-M/\vartheta}\int_{M-\min(x,M)}^{M} y^{d-1}\mathsf{e}^{y/\vartheta}\,dy \end{aligned} \tag{3.5}$$

$$= \mathsf{e}^{-M/\vartheta} \int_{M+\max(-x,-M)}^{M} y^{d-1}\mathsf{e}^{y/\vartheta}\,dy = \mathsf{e}^{-M/\vartheta} \int_{[M-x]_+}^{M} y^{d-1}\mathsf{e}^{y/\vartheta}\,dy$$
$$= \vartheta^d \mathsf{e}^{-M/\vartheta} \int_{[M-x]_+/\vartheta}^{M/\vartheta} y^{d-1}\mathsf{e}^{y}\,dy = \vartheta^d \mathsf{e}^{-M/\vartheta}\Big(H_d\big(M/\vartheta\big) - H_d\big(M-x\big]_+/\vartheta\big)\Big)$$

where $H_d(t) = \mathsf{e}^t(d-1)!\sum_{k=0}^{d-1}(-1)^{d-1-k}\frac{t^k}{k!}$, $t \geq 0$ (see Gradshteyn and Ryzhik (2007), 2.321.2).

Combining the B-spline representations and (3.5), we arrive at the following explicit sum expressions for the cumulative distribution functions.

Theorem 3.2. *Let $1 \leq d \leq m-1$ and $D = D(T)$ as in (2.1). The cumulative distribution function $F_\vartheta^{\mathit{TTT}_d, D=d}$ is given by*

$$F_\vartheta^{\mathit{TTT}_d, D=d}(s) = \binom{n}{d}\frac{T^d}{\vartheta^d}\int_0^s B_{d-1}(t \mid (n-d)T, \ldots, nT)\mathsf{e}^{-t/\vartheta}dt$$

$$= \frac{1}{(d-1)!}\binom{n}{d}\sum_{i=0}^{d-1}(-1)^i\binom{d}{i}\Big(H_d\big((n-i)T/\vartheta\big) - H_d\big((n-i)T-s\big]_+/\vartheta\big)\Big)\mathsf{e}^{-(n-i)T/\vartheta}, \quad 0 \leq s \leq nT. \tag{3.6}$$

Furthermore, the cumulative distribution function $F_\vartheta^{\mathit{TTT}_m, D\geq m}$ is given by

$$F_\vartheta^{\mathit{TTT}_m, D\geq m}(s) = \binom{n}{m}\frac{T^m}{\vartheta^m}\int_0^s B_{m-1}(t \mid 0, (n-m+1)T, \ldots, nT)\mathsf{e}^{-t/\vartheta}dt$$

$$= \frac{1}{(m-1)!}\binom{n}{m}\sum_{i=0}^{m-1}(-1)^i\binom{m}{i}\frac{m-i}{n-i}\Big(H_m\big((n-i)T/\vartheta\big) - H_m\big((n-i)T-s\big]_+/\vartheta\big)\Big)\mathsf{e}^{-(n-i)T/\vartheta}, \tag{3.7}$$
$$0 \leq s \leq nT.$$

The cumulative distribution function of $F_\vartheta^{\mathit{TTT}_m, D<m}$ is given by

$$F_\vartheta^{\mathit{TTT}_m, D<m}(s) = \gamma_{m,\vartheta}(s) - F_\vartheta^{\mathit{TTT}_m, D\geq m}(s), \quad s \geq 0, \tag{3.8}$$

where $\gamma_{m,\vartheta}(\cdot)$ denotes the cumulative distribution function of a gamma distribution with shape parameter m and scale parameter ϑ.

Remark 3.3. *Since m is a positive integer, the cumulative distribution function of a gamma distribution has an explicit finite sum representation, that is, it corresponds to the cumulative distribution function of an Erlang distribution given by*

$$\gamma_{m,\vartheta}(s) = 1 - \sum_{j=0}^{m-1}\frac{1}{j!}\left(\frac{s}{\vartheta}\right)^j e^{-s/\vartheta}, \quad s \geq 0.$$

Using the representation in Theorem 3.1 and 3.2, the (conditional) density function and cumulative distribution function of the total time on test can be directly obtained from the decomposition approach as given in Section 2 (see eq. (2.6)).

3.2 Application to the MLE under hybrid censoring

By analogy with the density function and cumulative distribution function of the total time on test, we also establish expressions for the density function and cumulative distribution function of the MLE $\widehat{\vartheta}$. In particular, we need the cumulative distribution functions (for $t \geq 0$)

$$F_\vartheta^{\widehat{\vartheta}, D=d}(t) = F_\vartheta^{\mathrm{TTT}_d, D=d}(dt), \quad F_\vartheta^{\widehat{\vartheta}, D\geq m}(t) = F_\vartheta^{\mathrm{TTT}_m, D\geq m}(mt), \quad t \in \mathbb{R},$$

and density functions

$$f_\vartheta^{\widehat{\vartheta}, D=d}(t) = d \cdot f_\vartheta^{\mathrm{TTT}_d, D=d}(dt), \quad f_\vartheta^{\widehat{\vartheta}, D\geq m}(t) = m \cdot f_\vartheta^{\mathrm{TTT}_m, D\geq m}(mt), \quad t \in \mathbb{R},$$

which are then appropriately replaced in (2.2) and (2.3), respectively. Again, the decomposition approach yields the required expressions for the density function and the cumulative distribution function of the MLE. The plots provided in Figure 1.1 are generated using the results of Sections 3.1 and 3.2.

3.3 Application to compute exact (conditional) confidence intervals

Finally, the representations of the (conditional) cumulative distribution function in Theorem 3.2 are used to compute exact (conditional) confidence intervals via the method of pivoting the cumulative distribution function as described in Casella and Berger (2002), Hahn et al. (2017), and Balakrishnan et al. (2014). The application of the pivoting method requires two conditions. First, the stochastic monotonicity of the distributions in the parameter ϑ must be ensured (see, e.g., Balakrishnan and Iliopoulos (2009), van Bentum and Cramer (2019), and, for a more general approach van Bentum (2024), Theorem 4.11). As this is not necessary for the computation of the confidence limits themselves, we do not provide further details here and refer to the references mentioned above. In case of a **C-II**-hybrid censoring scheme, the second condition for a two-sided confidence interval is given by the existence of solutions of the two equations

$$F_{\vartheta}^{\widehat{\vartheta}}(\widehat{\vartheta}_{\mathrm{obs}}) = \frac{\alpha}{2}, \quad F_{\vartheta}^{\widehat{\vartheta}}(\widehat{\vartheta}_{\mathrm{obs}}) = 1 - \frac{\alpha}{2} \tag{3.9}$$

for a given desired nominal level $\alpha \in (0,1)$ where $\widehat{\vartheta}_{\mathrm{obs}}$ denotes the maximum likelihood estimate obtained from the data. For one-sided confidence intervals only one of these equations has to be taken into account (with $\frac{\alpha}{2}$ replaced by α). For a **C-I**-hybrid censoring scheme, the conditional cumulative distribution function will be considered (see Section 3.3.2).

Therefore, one must ensure that $F_{\vartheta}^{\widehat{\vartheta}}(\widehat{\vartheta}_{\mathrm{obs}})$ as a function of ϑ attains any value in $(0,1)$. Since the function is continuous in ϑ, it is sufficient to consider the limits at the boundaries of the parameter space $\Theta = (0,\infty)$. Such considerations are presented in Balakrishnan et al. (2014) and Iliopoulos (2015) for some simple hybrid censoring schemes. Using the representations provided in Theorem 3.2, we will establish the existence (and possibly non-existence) for any general hybrid censoring scheme so that the results available so far are extended to these general scenarios. An extension to general progressive hybrid censoring schemes is obtained by analogy (see comments in Section 4).

3.3.1 C-II-hybrid censoring schemes

First, consider an arbitrary **C-II**-hybrid censoring scheme. From Theorem 3.2, it follows that, for $s > 0$,

$$\lim_{\vartheta\to\infty} F_{\vartheta}^{\widehat{\vartheta},D=d}(s) = \lim_{\vartheta\to\infty} F_{\vartheta}^{\widehat{\vartheta},D\geq m}(s) = 0, \quad \text{for } d,m \geq 1,$$

since we conclude $0 \leq \int_0^s B_{d-1}(t \mid (n-d)T,\ldots,nT)\mathsf{e}^{-t/\vartheta}dt \leq \int_0^\infty B_{d-1}(t \mid (n-d)T,\ldots,nT)dt = 1$ from the density property of the B-spline basis functions. On the other hand, we get from the representations in (3.6) and (3.7) that

$$\lim_{\vartheta\to 0+} F_{\vartheta}^{\widehat{\vartheta},D=d}(s) = \lim_{\vartheta\to 0+} F_{\vartheta}^{\widehat{\vartheta},D\geq m}(s) = 0, \quad d,m \geq 1.$$

Finally, (3.8) yields for $s > 0$

$$\lim_{\vartheta\to 0+} F_{\vartheta}^{\widehat{\vartheta},D<m}(s) = \lim_{\vartheta\to 0+} \gamma_{m,\vartheta}(s) = \lim_{\vartheta\to 0+} \gamma_{m,1}(s/\vartheta) = 1, \quad m \geq 1.$$

Hence, for $t > 0$, we obtain

$$\lim_{\vartheta\to 0+} F_{\vartheta}^{\widehat{\vartheta}}(t) = 1, \quad \lim_{\vartheta\to\infty} F_{\vartheta}^{\widehat{\vartheta}}(t) = 0$$

so that the confidence bounds exist uniquely for any $t = \widehat{\vartheta}_{\mathrm{obs}}$ under a **C-II**-hybrid censoring scheme.

3.3.2 C-I-hybrid censoring schemes

For a **C-I**-hybrid censoring scheme, the situation is slightly different since we have to ensure the observation of at least one measurement. In this case, we consider the conditional cumulative distribution function by analogy with the equations in (3.9), that is, we have to solve the equations

$$F_{\vartheta}^{\widehat{\vartheta}\mid D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\mathrm{obs}}) = \frac{\alpha}{2}, \quad F_{\vartheta}^{\widehat{\vartheta}\mid D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\mathrm{obs}}) = 1 - \frac{\alpha}{2}$$

for $\vartheta \in \Theta = (0, \infty)$. As pointed out in Balakrishnan et al. (2014), it is possible that these equations will not yield any solutions. Therefore, it is important to know in the computational process when this issue occurs. As above, we consider the limits of the function $F_{\vartheta}^{\widehat{\vartheta}|D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\text{obs}})$ for $\vartheta \to 0+$ and $\vartheta \to \infty$, respectively, since the function is continuous in ϑ. The resulting ratio of probabilities has the probability $\Pr_{\vartheta}(D_{\mathsf{HCS}} \geq 1) = 1 - e^{-nT_k/\vartheta}$ as denominator where T_k is the largest threshold involved in the **C-I**-hybrid censoring scheme. For $\vartheta \to 0+$, this probability tends to 1 so that

$$B_u = \lim_{\vartheta \to 0+} F_{\vartheta}^{\widehat{\vartheta}|D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\text{obs}}) = 1$$

as above. However, for $\vartheta \to \infty$, the probability goes to 0 so that both nominator and denominator tend to 0. Therefore, We need to look at this case more closely. First, note that $\{D_{\mathsf{HCS}} = 1\} = \{D_k = 1\}$ holds for a **C-I**-hybrid censoring scheme, meaning that $X_{1:n} < T_k \leq X_{2:n}$ (see also Table 2.2). Then, applying the limit result

$$\vartheta^d (1 - e^{-nT_k/\vartheta}) \xrightarrow{\vartheta \to \infty} \begin{cases} nT_k, & D_{\mathsf{HCS}} = d = 1 \\ \infty, & D_{\mathsf{HCS}} = d \geq 2 \end{cases},$$

we conclude that, for given $s > 0$,

$$\begin{aligned} B_l = B_l(s) = \lim_{\vartheta \to \infty} F_{\vartheta}^{\widehat{\vartheta}|D_{\mathsf{HCS}}\geq 1}(s) = \lim_{\vartheta \to \infty} \frac{F_{\vartheta}^{\mathsf{TTT}_1, D_k = 1}(s)}{1 - \mathsf{e}^{-nT_k/\vartheta}} = \int_0^s B_0(t \mid (n-1)T_k, nT_k)\, dt \\ = \begin{cases} 0, & s \leq (n-1)T_k \\ \frac{s-(n-1)T_k}{T_k}, & (n-1)T_k \leq s \leq nT_k \\ 1, & nT_k \leq s \end{cases}. \end{aligned} \tag{3.10}$$

Therefore, for a solution $\vartheta > 0$ of the equation $F_{\vartheta}^{\widehat{\vartheta}|D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\text{obs}}) = \beta$ to exist for a given $\beta \in (0, 1)$, the condition

$$\beta \in [B_l, B_u] = [B_l, 1]$$

must necessarily hold. Thus, $B_l(\widehat{\vartheta}_{\text{obs}}) = \beta_{\min}$ denotes the minimal value such that a solution of the equation exist. The situation is depicted in Figure 3.1.

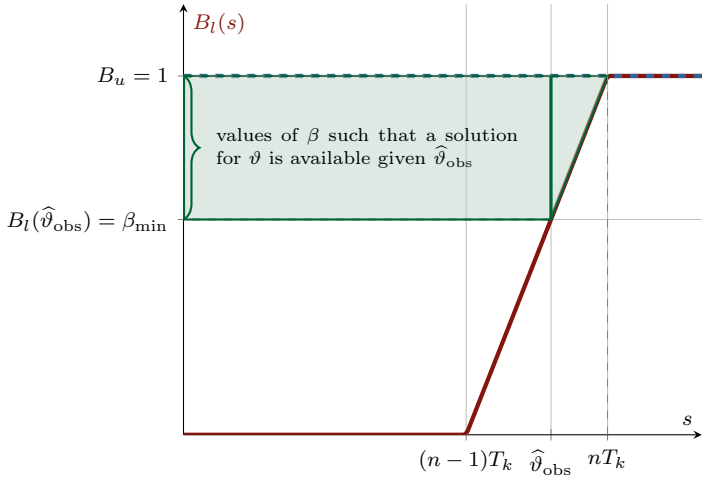


Figure 3.1: Plot of limit function $B_l(\cdot)$ for a **C-I**-hybrid censoring scheme where $\beta_{\min}$ denotes the minimal value of β such that a solution of the equation exists (provided $\widehat{\vartheta}_{\text{obs}}$ has been observed).

Now, we can see from (3.10) that for $\widehat{\vartheta}_{\text{obs}} \leq (n-1)T_k$, the confidence limits exist for a given $\alpha \in (0, 1)$ in any case. Note that this condition is satisfied iff $D_{\mathsf{HCS}} \geq 2$. In particular, this result extends that one for Type-I hybrid censoring presented in Balakrishnan et al. (2014), Theorem 2. On the other hand, for any given $\alpha \in (0, 1)$, the equation does never have a solution when $nT_k \leq \widehat{\vartheta}_{\text{obs}}$. In that case, a workaround is provided in

Balakrishnan et al. (2014). If $(n-1)T_k \leq \widehat{\vartheta}_{\text{obs}} \leq nT_k$ then the equation $F_{\vartheta}^{\widehat{\vartheta}|D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\text{obs}}) = \beta$ with $\beta \in (0,1)$ does not have a solution when

$$\beta < \frac{\widehat{\vartheta}_{\text{obs}} - (n-1)T_k}{T_k} \iff \widehat{\vartheta}_{\text{obs}} > (n-1+\beta)T_k \tag{3.11}$$

(see also Balakrishnan et al. (2014), Corollary 1, for Type-I hybrid censoring). This condition has also been implemented in the Python code so that the existence of solutions is verified before the computation of the confidence bounds is started.

3.4 Use in simulations

The package includes the option to simulate hybrid censored data from exponentially distributed lifetimes so that it can be easily used for simulation studies. Furthermore, histograms and kernel estimates of the distribution of the maximum likelihood estimator can be generated. The kernel estimates are computed using the package scikit-learn (Pedregosa et al. (2011)).

3.5 Implementation in Python code

In order to perform the computation with the Python package `CensoPy`, the hybrid censoring scheme under study has to be specified first as

C-I : `HCS=C_I_HCS(n, [T1,...,Tk], [m1,...,mk], Digits, Method)`

C-II: `HCS=C_II_HCS(n, [T1,...,Tk-1,'inf'], [m1,...,mk], Digits, Method)`

where `HCS` is an object of one of the classes with parameters

- `n` (integer): sample size
- [`T1`, ..., `Tk`] (list of integers): thresholds (increasingly ordered)
- [`m1`, ..., `mk`] (list of integers): censoring numbers (decreasingly ordered)
- `Digits` (integer): Precision (number of decimals used in the computation)
- `Method` (`analytic`, `recursive`): Method used to evaluate the B-spline basis functions. The option `recursive` must be used for progressive hybrid censoring (see Sections 3 and 4).

The special form of the input variables for a given hybrid censoring scheme is provided in Table 3.1. Then, the objective function has to be specified as `HCS.function(t, theta)` where

$$\texttt{function} \in \{\texttt{pdf_TTT, cdf_TTT, pdf_MLE, cdf_MLE, calc_MLE, monte_carlo}\}$$

with variables `t` (float) and `theta` (float; true distribution parameter value). The `return`-value (float) is always the computed value of the chosen function given `t`, `theta`. For each function, we provide a plot as output. Of course, adaptions can be carried out by the user. Moreover, one can compute bounds of an exact (conditional) confidence interval by providing a (complete) data set as an array. Then, for a specified hybrid censoring scheme, the hybrid censored data is obtained and used to compute the total time on test statistic. For a given nominal level $\alpha \in (0,1)$, a two-sided exact (conditional) confidence interval is computed according to the pivotal method described in Section 3.3.2.

For easy conduction of simulation studies based on hybrid censored data, one can simulate the values of N hybrid censored data sets. For given $\vartheta > 0$, N independent runs to generate data sets

$$X_{1:n}^{(i)}, \ldots, X_{n:n}^{(i)}, \quad 1 \leq i \leq N,$$

of size n from an $\text{Exp}(\vartheta)$-distribution are performed. For each data set, the MLE $\widehat{\vartheta}^{(i)}$ is computed according to the hybrid censoring scheme under study (if $d_{\mathsf{HCS}} = 0$, the MLE does not exist for the corresponding data which,

thus, is excluded from the computation). The proportion of excluded data sets as well as the mean and standard error of the maximum likelihood estimates are reported. Furthermore, a plot of the histogram (with bin width b) of the estimates $\widehat{\vartheta}^{(i)}$, $1 \leq i \leq N$, is generated. Alternatively, as a smoother option, a kernel estimate (with bin width b) using a Gaussian kernel function can be chosen. This can also be plotted in combination with the corresponding theoretical density function. Finally, the empirical coverage probabilities of the corresponding confidence intervals can be reported.

Remark 3.4. *Currently, the package* `CensoPy` *focuses on exponentially distributed random variables. The main application here is the computation of (conditional) density functions and cumulative distribution functions in the case of the exponential distribution, in order to determine, e.g., expected values and exact (conditional) confidence intervals. However, for the simulation, it is very easy to adapt it to alternative baseline distributions. The generator of the random variables has to be chosen appropriately or an appropriate transformation has to be applied. For instance, applying the power transformation $x \mapsto x^{\beta}$ with $\beta > 0$ to the exponential data, one gets Weibull distributed lifetime data.*

Furthermore, the censoring mechanism can be used for any given sample in order to determine the corresponding hybrid censored sample as it operates only on the ordered data. Therefore, the censoring procedure can also be applied to record/minimal repair data (see Berzborn and Cramer (2023)), sequential order statistics (see Burkschat et al. (2016)), and generalized order statistics (see Cramer (2025)).

Table 3.1: Implementation of hybrid censoring schemes and input used in the command.

			sample size	censoring numbers	thresholds
	Hybrid censoring scheme	k	n	$[m_1, \ldots, m_k]$	$[T_1, \ldots, T_k]$
C-II-HCS	complete sample	1	`n`	`[n]`	`['inf']`
	Type-II	1	`n`	`[m1]`	`['inf']`
	Type-II hybrid	2	`n`	`[n,m2]`	`[T1,'inf']`
	gen. Type-I hybrid	2	`n`	`[m1,m2]`	`[T1,'inf']`
	uni. Type-II hybrid	3	`n`	`[n,m2,m3]`	`[T1,T2,'inf']`
	uni. Type-IV hybrid	3	`n`	`[m1,m2,m3]`	`[T1,T2,'inf']`
	gen. uni. Type-II hybrid	4	`n`	`[n,m2,m3,m4]`	`[T1,T2,T3,'inf']`
	gen. uni. Type-IV hybrid	4	`n`	`[m1,m2,m3,m4]`	`[T1,T2,T3,'inf']`
C-I-HCS	Type-I	1	`n`	`[n]`	`[T1]`
	Type-I hybrid	1	`n`	`[m1]`	`[T1]`
	gen. Type-II hybrid	2	`n`	`[n,m2]`	`[T1,T2]`
	uni. Type-I hybrid	2	`n`	`[m1,m2]`	`[T1,T2]`
	uni. Type-III hybrid	3	`n`	`[n,m2,m3]`	`[T1,T2,T3]`
	gen. uni. Type-I hybrid	3	`n`	`[m1,m2,m3]`	`[T1,T2,T3]`

Notes: For **C-II**-hybrid censoring schemes, the infinite bound $T_k =$ `inf` has to be written as `'inf'` in the parametrization. The dimension k of the arrays is automatically determined from the input.

3.6 Application of `CensoPy` to data examples

To illustrate the usage of the package, we consider four data sets taken from the literature. Its use in plotting density functions and cumulative distribution functions has already been shown in Figure 1.1.

Example 3.5. *Sundberg (2001) analyzed the data*

4	5	8	11	20	29	35	40	66	70

with sample size $n = 10$ originally taken from Kalbfleisch (1985). He compared various approaches to construct confidence intervals under Type-I censoring. He considered thresholds $T = 50$ and $T = 75$ with a nominal level $\alpha = 5\%$.

For $T = 50$, the confidence interval is given by $[17.164, 70.231]$ ($\widehat{\vartheta}_{obs} = 31.5$). For $T = 75$, it is obtained as $[16.467, 60.919]$ ($\widehat{\vartheta}_{obs} = 28.8$). Note that both computed confidence intervals are comparable to the approximate confidence intervals C4-C6 presented in Sundberg (2001, Table 1) (which were evaluated as the best solution there by means of simulations).

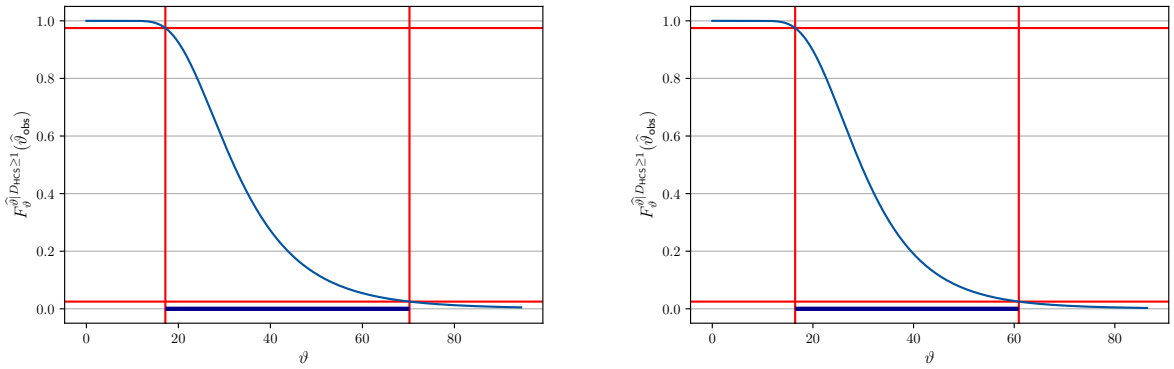


Figure 3.2: Plots of $\vartheta \mapsto F_{\vartheta}^{\widehat{\vartheta}|D_{\mathsf{HCS}}\geq 1}(\widehat{\vartheta}_{\mathrm{obs}})$ from Example 3.5 with confidence bounds (vertical red lines) and resulting confidence interval (dark blue horizontal line) [left: $T = 50$; right: $T = 75$].

Example 3.6. *Gupta and Kundu (1998) considered a Type-I hybrid censored data with original sample size $n = 20$, $T_1 = 150$, and $m_1 = 13$ leading to the data*

3	19	23	26	37	38	41	45	58	84	90	109	138

They computed various confidence intervals which are close to that one obtained in our case, that is, $[79.928, 242.124]$ ($\widehat{\vartheta}_{obs} = 129.0$). The same data has been discussed in Algarni et al. (2022) under Type-I censoring who also presented alternative confidence intervals. As a result, we get $[83.248, 244.776]$ ($\widehat{\vartheta}_{obs} = 135.461538$) which is also in accordance with their suggestions.

The results are simply obtained using the command

```
C_I_HCS(n, [T1], [m1], Digits, "analytic")
```

with $T_1 = 150$ as well as $m_1 = n$ (Type-I censoring) and $m_1 = 13$ (Type-I hybrid censoring), respectively. For illustration, a template file is provided on GitLab.

Example 3.7. *Bartholomew (1963) studied the Type-I censored data*

3	19	23	26	27	37	38	41	45	58	84	90	99	109	138

with threshold $T = 150$ and initial sample size $n = 20$ resulting in $D = 15$ measurements. As a result, we get the 95%-confidence interval $[66.955, 184.545]$ and the 99%-confidence interval $[58.673, 224.158]$, respectively ($\widehat{\vartheta}_{obs} = 105.8$). Both are close to the corresponding approximate confidence intervals presented in Bartholomew (1963, Table 4). The result for the 95%-confidence interval can also be found in Chandrasekar et al. (2004, Table 3). The confidence intervals for generalized Type-II hybrid censoring with $T_1 = 50, T_2 = 100$ and $m_1 = n, m_2 \in \{7, 13, 15, 20\}$ can easily reproduced using `CensoPy` by the command

```
C_I_HCS(n, [T1,T2], [m1,m2], Digits, "analytic")
```

Note that, by design, this scheme is a **C-I***-hybrid censoring scheme. The term originally introduced is therefore misleading. The same applies to the subsequently discussed generalized Type-II hybrid censoring scheme, which represents a* **C-I***-hybrid censoring scheme.*

Bartholomew (1963)'s data has also been studied to analyze Type-II hybrid censored data in Childs et al. (2003, Illustration 3) and generalized Type-I hybrid censored data in Chandrasekar et al. (2004). The thresholds have to be specified as $T_1 = 50, T_2 = \infty$ *whereas* $m_1 = n$ *equals the sample size for Type-II hybrid censoring (see Table 3.1). The commands is given by*

```
C_II_HCS(n, [T1,T2], [m1,m2], Digits, "analytic")
```

with appropriate choices of the parameters, that is,

(i) Type-II hybrid censoring: `T1=50, T2='inf', m1=n, m2=7`.

In Childs et al. (2003), the hybrid censoring schemes with $m_2 \in \{7, 15\}$ *have been discussed.*

*(ii) Generalized Type-I hybrid censoring (***C-II***-hybrid censoring scheme):* `T1=50, T2='inf', m1=15, m2=12`.

In Chandrasekar et al. (2004, Table 2), the pairs $(m_1, m_2) \in \{(15, 12), (11, 5), (7, 4), (20, 7)\}$ *were considered for a nominal level* $\alpha \in \{0.01, 0.05\}$.

Please note that, due to Type-I censoring, the data are not actually suitable for this purpose. This is because the observations above 150 *are not available. However, as the parameters of the hybrid censoring scheme were chosen wisely with regard to the available measurement data, the Type-I censoring of the original data does not come into play and all quantities can be determined.*

Example 3.8. *As demonstrated in the final example, the package is also effective for large data sets. Nevertheless, a higher level of computational accuracy is required. Prajapat et al. (2023) considered data which has been obtained from the data in Table 3.2 reported in Birnbaum and Saunders (1969) via a Cox-Box-transformation* $z \mapsto \left(\frac{z-65}{76.322}\right)^{3.2571}$ *so that it fits the exponential distribution well.*

70	90	96	97	99	100	103	104	104	105	107	108	108	108	109
109	112	112	113	114	114	114	116	119	120	120	120	121	121	123
124	124	124	124	124	128	128	129	129	130	130	130	131	131	131
131	131	132	132	132	133	134	134	134	134	134	136	136	137	138
138	138	139	139	141	141	142	142	142	142	142	142	144	144	145
146	148	148	149	151	151	152	155	156	157	157	157	157	158	159
162	163	163	164	166	166	168	170	174	196	212				

Table 3.2: Data taken from Birnbaum and Saunders (1969).

For the (transformed) complete data, we get the exact 95%*-confidence interval* $[0.8304136, 1.227729]$ *(*$\widehat{\vartheta}_{obs} = 1.000005379$*). In Table 3.3, we report some computational results for various hybrid censoring schemes. Note that Prajapat et al. (2023) generated three random subsamples of sizes* $16, 20, 25$ *each in order to illustrate results mor multiple samples under Type-I hybrid censoring. Thus, the confidence intervals can not directly compared with those provided in this paper.*

For a Type-I censored data with $T = 120$*, we obtain the* 95%*-confidence interval Due to the large value of* $n = 101$*, the computation has been performed with a digit number of* `Digits=300` *to avoid numerical instabilities. The approximation C6 given in Sundberg (2001) yields the confidence interval* $[0.837977, 1.241342]$*. Note that this approximation coincides with the proposal of Fairbanks et al. (1982).*

4 Extension to progressive to censoring

The above discussion is based on hybrid censored order statistics data which mainly yields a simplified expression for the B-spline basis functions given in (3.1)–(3.3). However, as pointed out in, e.g., Balakrishnan et al. (2023)

HCS	m_i	T_i	T_i^{o*}	p^{**}	MLE $\widehat{\vartheta}_{\text{obs}}$	lower bound	upper bound
complete	101			1.00	1.000	0.830	1.228
Type-I	101	0.3440	120	0.24	1.115	0.779	1.664
	101	1	141.3	0.65	0.986	0.780	1.273
	101	2	159.4	0.89	0.956	0.783	1.189
Type-I hybrid	48	1	141.3	0.65	1.049	0.806	1.423
	60	1	141.3	0.65	1.000	0.789	1.311
	75	1	141.3	0.65	0.986	0.780	1.273
	48	2	159.4	0.89	1.049	0.806	1.423
	60	2	159.4	0.89	1.000	0.789	1.311
	75	2	159.4	0.89	0.933	0.753	1.186

Notes: $^{*}T_i^o$ corresponds to the threshold in the original data. $^{**}p$ denotes the proportion of failure times which remain after Type-I censoring.

Table 3.3: MLEs and confidence intervals for the Box-Cox-transformed data given in Table 3.2.

the density functions of the total time on test and MLE can also be written in terms of B-spline basis functions by adopting the knot sequence and the normalizing constant appropriately. Given a progressively Type-II censored sample $X^{\mathscr{R}}_{1:n:\gamma_1},\dots,X^{\mathscr{R}}_{n:n:\gamma_1}$ with censoring plan $\mathscr{R}=(R_1,\dots,R_n)$ with $R_i\in\mathbb{N}_0$ and baseline sample size $\sum_{j=1}^n R_i+n=\gamma_1$, we define

$$\gamma_i=\sum_{j=i}^{n}(R_j+1),\quad 1\le i\le n,$$

as number of items on test before the i-th failure in the progressively Type-II censored life test (see Balakrishnan and Cramer (2014), p. 7). Note that n denotes the observed sample size. Then, $\gamma_1>\cdots>\gamma_n\ge 1$ and the density function of the corresponding total time on test

$$\mathtt{TTT}^*_{D_{\mathsf{HCS}}}=\sum_{j=1}^{D_{\mathsf{HCS}}}(R_j+1)X^{\mathscr{R}}_{j:n:\gamma_1}+\gamma_{D_{\mathsf{HCS}}+1}W_{\mathsf{HCS}} \tag{4.1}$$

can be decomposed as in the order statistics' model. As has been pointed out in Cramer (2025) (see also Górny and Cramer (2018b)), the hybrid censoring scheme operates in the same way on these ordered data as it does for order statistics. In particular, by analogy with the representations in Theorem 3.1, the joint density function of $\mathtt{TTT}^*_D$ in (4.1) and D that can be written as

$$f_\vartheta^{\mathtt{TTT}^*_d,D=d}(t)=\frac{T^d\prod_{j=1}^d\gamma_j}{d!\vartheta^d}B_{d-1}(t\mid\gamma_{d+1}T,\dots,\gamma_1T)\,\mathsf{e}^{-t/\vartheta},\quad 0\le t, \tag{4.2}$$

with $d\in\{1,\dots,n\}$, $\gamma_{n+1}\equiv 0$, and random counters

$$D=D(T)=\sum_{j=1}^{n}\mathbb{1}_{(-\infty,T]}\big(X^{\mathscr{R}}_{j:n:\gamma_1}\big)$$

defined via the progressively Type-II censored sample $X^{\mathscr{R}}_{1:n:\gamma_1},\dots,X^{\mathscr{R}}_{n:n:\gamma_1}$. Thus, we have the same representation in terms of a B-spline basis function with the knot sequence $\gamma_{d+1}T,\dots,\gamma_1T$. Since order statistics can be seen as a submodel of progressive Type-II censoring with the model parameters $\gamma_i=\gamma_1-i+1$, $1\le i\le n$, and $\gamma_1=n$ (see Balakrishnan and Cramer (2014)), this also includes the preceding representations as a special case. Moreover, we have

$$f_\vartheta^{\mathtt{TTT}^*_n,D\ge n}(t)=\frac{T^n\prod_{j=1}^n\gamma_j}{n!\vartheta^n}B_{n-1}(t\mid 0,\gamma_nT,\dots,\gamma_1T)\mathsf{e}^{-t/\vartheta},\quad 0\le t\le\gamma_1T$$

(see Cramer and Balakrishnan (2013)) and

$$f_{\vartheta}^{\mathrm{TTT}_n^*, D<n}(t) = f_{\Gamma(\vartheta,n)}(t) - \frac{T^n \prod_{j=1}^n \gamma_j}{\gamma_1! \vartheta^n} B_{n-1}(t \mid 0, \gamma_n T, \ldots, \gamma_1 T) \mathsf{e}^{-t/\vartheta}, \quad t \geq 0.$$

Note that, by defining $\gamma_{n+1} = 0$, the knot sequence in the above B-spline function reads $\gamma_{n+1}T, \ldots, \gamma_1 T$ so that it is consistent with the notation in (4.2) for $d = n$.

Due to the different knot sequence, we can not use the explicit expressions for the B-spline basis functions given in (3.1)–(3.3) to compute density functions and cumulative distribution functions. Therefore, we implemented Cox-de Boor's recurrence relation for B-spline basis functions in the Python code. The algorithm considers a knot sequence $0 = \gamma_{n+1}T < \gamma_n T < \cdots < \gamma_1 T$. Since this recurrence relation is usually presented in terms of the non-normalized B-spline basis functions, we have to adapt the B-spline basis functions appropriately. Since the notation of the B-spline basis functions is not unique in the literature, we illustrate in the following how our computation fits into the Cox-de Boor procedure by using the notation used in de Boor (2001).

According to de Boor (2001), the (non-normalized) B-splines can be defined for a strictly increasing knot sequence $\cdots < t_1 < \cdots < t_m < \ldots$ by

(i) $B[t_j, t_{j+1}](\cdot) = \mathbb{1}_{[t_j, t_{j+1}]}$ for some j and $d = 1$, and by

(ii) by the recurrence relation

$$B[t_j, \ldots, t_{j+d}](t) = \frac{t - t_j}{t_{j+d-1} - t_j} B[t_j, \ldots, t_{j+d-1}](t) + \frac{t_{j+d} - t}{t_{j+d} - t_{j+1}} B[t_{j+1}, \ldots, t_{j+d}](t), \quad t \in \mathbb{R},$$

for $d \geq 2$.

We shall now bring our notation into line with de Boor's one. In our notation, we start with the B-spline of order 0

$$B_0(t \mid \gamma_2 T, \gamma_1 T) = \begin{cases} \frac{1}{(\gamma_1 - \gamma_2)T}, & \gamma_2 T \leq t < \gamma_1 T \\ 0, & \text{otherwise} \end{cases} = \frac{1}{(\gamma_1 - \gamma_2)T} B[t_j, t_{j+1}](t). \tag{4.3}$$

Using the relation

$$B_{d-1}(t \mid \gamma_{j+d}T, \ldots, \gamma_j T) = \frac{d}{\gamma_j T - \gamma_{j+d} T} B[\gamma_{j+d}T, \ldots, \gamma_j T](t),$$

the corresponding recurrence relation is given by, for $d > 1$,

$$\begin{aligned} &B_{d-1}(t \mid \gamma_{j+d}T, \ldots, \gamma_j T) \\ &\quad = \frac{d}{d-1} \Big(\frac{t - t_j}{t_{j+d} - t_j} B_{d-2}(t \mid \gamma_{j+d-1}T, \ldots, \gamma_j T) + \frac{t_{j+d} - t}{t_{j+d} - t_j} B_{d-2}(t \mid \gamma_{j+d}T, \ldots, \gamma_{j+1} T) \Big), \quad t \in \mathbb{R}. \end{aligned} \tag{4.4}$$

This relation is equal to that presented in Schumaker (2007, Theorem 4.15) except for the factor $\frac{d}{d-1}$ which is due to the fact that we standardize the B-spline basis functions so that their integral equals one (see also Dahmen and Micchelli (1986)). Thus, they can be interpreted as density functions. Note that $B_{d-1}(t \mid \gamma_{d+1}T, \ldots, \gamma_1 T) = 0$ if $t \notin [\gamma_{d+1}T, \gamma_1 T]$.

As a result, we can now compute the density functions recursively. In order to compute also the cumulative distribution functions, we need expressions for integrals of the following type (using de Boor's notation)

$$\int_{t_j}^{x} B[t_j, \ldots, t_{j+d}](t) e^{-t/\vartheta} \, dt, \quad t_j \leq x.$$

First, from de Boor et al. (1976, p. 126), we get the identity

$$\frac{\partial}{\partial t} B[t_j, \ldots, t_{j+d}](t) = \frac{d-1}{t_{j+d-1} - t_j} B[t_j, \ldots, t_{j+d-1}](t) - \frac{d-1}{t_{j+d} - t_{j+1}} B[t_{j+1}, \ldots, t_{j+d}](t)$$

for the derivative of a B-spline basis function. Then, by partial integration, we get for $t_j < x$ and $d \geq 2$ the recurrence relation

$$\begin{aligned}
\mathcal{I}_{j,d} = \int_{t_j}^{x} B[t_j,\ldots,t_{j+d}](t)e^{-t/\vartheta}\,dt &= -\vartheta B[t_j,\ldots,t_{j+d}](t)e^{-t/\vartheta}\Big|_{t_j}^{x} + \vartheta\int_{t_j}^{x} e^{-t/\vartheta}\frac{\partial}{\partial t}B[t_j,\ldots,t_{j+d}](t)\,dt \\
&= \vartheta B[t_j,\ldots,t_{j+d}](t_j)e^{-t_j/\vartheta} - \vartheta B[t_j,\ldots,t_{j+d}](x)e^{-x/\vartheta} \\
&\quad + \vartheta\int_{t_j}^{x} e^{-t/\vartheta}\frac{\partial}{\partial t}B[t_j,\ldots,t_{j+d}](t)\,dt \\
&= \vartheta B[t_j,\ldots,t_{j+d}](t_j)e^{-t_j/\vartheta} - \vartheta B[t_j,\ldots,t_{j+d}](x)e^{-x/\vartheta} \\
&\quad + \vartheta\frac{d-1}{t_{j+d-1}-t_j}\mathcal{I}_{j,d-1} - \vartheta\frac{d-1}{t_{j+d}-t_{j+1}}\mathcal{I}_{j+1,d-1}.
\end{aligned}$$

For $d = 1$, we obtain the initial integral

$$\mathcal{I}_{j,1} = \int_{t_j}^{x} B[t_j,\ldots,t_{j+1}](t)e^{-t/\vartheta}\,dt = \int_{t_j}^{\min\{x,t_{j+1}\}} e^{-t/\vartheta}\,dt = \vartheta\Big(e^{-t_j/\vartheta} - e^{-\min\{x,t_{j+1}\}/\vartheta}\Big).$$

As a result, we get also a simple recurrence relation which involves both the B-spline basis functions and the integrals $\mathcal{I}_{j,d}$.

Finally, it should be mentioned that, concerning the existence of confidence limits as in Section 3.3.2, we get the same condition on the observation $\widehat{\vartheta}_{\text{obs}}$ as given in (3.11). Therefore, the computation of the confidence intervals can also been performed by the given procedure.

Example 4.1. *As an example of the application of* `CensoPy` *for progressive hybrid censoring, we consider the situation discussed in Childs et al. (2008) who introduced progressive Type-I and Type-II hybrid censoring. They studied the data given in Table 4.1.*

progressively Type-II censored sample													
x_i	0.19	0.78	0.96	1.31	2.78	4.85	6.50	7.35					
R_i	0	0	3	0	3	0	0	5					
extended progressively Type-II censored sample													
x_i	0.19	0.78	0.96	1.31	2.78	4.85	6.50	7.35	8.27	12.06	32.52	33.91	72.89
R_i	0	0	3	0	3	0	0	0	0	0	0	0	0

Table 4.1: Data taken from Viveros and Balakrishnan (1994) and corresponding extended sample which includes the right censored measurements (see also Balakrishnan and Cramer (2014, p. 9)).

For progressive Type-I hybrid censoring, the command

```
C_I_PCS(R, [T1], [m1], Digits, "recursive")
```

with $\mathscr{R} = (0,0,3,0,3,0,0,5)$, $T_1 = 6$, $n = 8$ (length of censoring plan $\mathscr{R}$) reproduces exactly the values presented in Childs et al. (2008, Table 23.1). Note that the option `recursive` *is mandatory for progressive hybrid censoring.*

For progressive Type-II hybrid censoring, the situations must be discussed in more detail as the understanding of the baseline sample is different in the literature. This is due to the fact that the final censoring number corresponds to right censoring of the data. under progressive hybrid censoring, one can either interpret the progressively Type-II sample as a complete sample of ordered values or as right censored sample. In the last case, one can add the right censored values (if available). This supplemented progressively Type-II censored sample is called extended progressively Type-II censored sample. However, it can be considered as a progressively Type-II censored sample with an adapted censoring plan. In particular, the corresponding censoring plan is obtained by replacing R_n by 0 and adding R_n zeros to the censoring plan. Its length is therefore given by $n + R_n$ the final censoring number R_n (for details of this operation, we refer to Cramer et al. (2016)). Using

```
C_II_PCS(R, [T1,'inf'], [m1,m2], Digits, "recursive")
```

we obtain the desired results with $T_1 = 6$, $T_2 = \infty$, $m_2 \in \{6, 8\}$, *and*

(i) standard sample: $m_1 = n = 8$ *with censoring plan* $\mathscr{R} = (0, 0, 3, 0, 3, 0, 0, 5)$

(ii) extended sample: $m_1 = n + R_n = 13$ *with censoring plan* $\mathscr{R} = (0, 0, 3, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0)$

(see Table 4.2). The results obtained for the extended sample coincide with those presented in Childs et al. (2008, Table 23.1).

$\mathscr{R}$	m_2	MLE $\widehat{\vartheta}_{\text{obs}}$	lower bound	upper bound
$(0, 0, 3, 0, 3, 0, 0, 5)$	6	10.682	5.449	26.091
	8^*	9.086	5.040	21.046
$(0, 0, 3, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0)$	6	10.682	5.448	26.084
	8	9.086	5.027	20.935

Notes: * corresponds to the full progressively Type-II censored sample. Thus, the command `C_II_PCS(R, ['inf'], [m1], Digits, "recursive")` must be used.

Table 4.2: MLEs and confidence intervals for progressive Type-II hybrid censored data given in Table 4.1.

Code availability

The code and a documentation of the Python package `CensoPy` is available at GitLab

https://censopy-738e78.pages.git.nrw/.

The repository also provides template files to generate the examples, plots, and an R-file to include the package in R. In order to use it in R, a Python compiler must be available on the computer.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Agarwal, G. G., Dalpatadu, R. J., and Singh, A. K. (2002). Linear functions of uniform order statistics and B-splines. *Comm. Statist. Theory Methods*, 31:181–192.

Algarni, A., Almarashi, A. M., Kundu, D., Abd-Elmougod, G. A., and Abdel-Khalek, S. (2022). Comparison of different confidence intervals under Type-I censoring scheme. *Journal of Mathematics*, 2022(1).

Balakrishnan, N. and Cramer, E. (2014). *The Art of Progressive Censoring. Applications to Reliability and Quality*. Birkhäuser, New York.

Balakrishnan, N., Cramer, E., and Iliopoulos, G. (2014). On the method of pivoting the cdf for exact confidence intervals with illustration for exponential mean under life-test with time constraints. *Statist. Probab. Lett.*, 89:124–130.

Balakrishnan, N., Cramer, E., and Kundu, D. (2023). *Hybrid Censoring Know-How: Designs and Implementations*. Academic Press.

Balakrishnan, N. and Iliopoulos, G. (2009). Stochastic monotonicity of the MLE of exponential mean under different censoring schemes. *Ann. Inst. Statist. Math.*, 61:753–772.

Balakrishnan, N. and Kundu, D. (2013). Hybrid censoring: Models, inferential results and applications (with discussions). *Comput. Statist. Data Anal.*, 57:166–209.

Balakrishnan, N., Kundu, D., Ng, H. K. T., and Kannan, N. (2007). Point and interval estimation for a simple step-stress model with Type-II censoring. *Journal of Quality Technology*, 39(1):35–47.

Balakrishnan, N., Rasouli, A., and Farsipour, N. (2008). Exact likelihood inference based on an unified hybrid censored sample from the exponential distribution. *J. Stat. Comput. Simul.*, 78(5):475–488.

Bartholomew, D. J. (1963). The sampling distribution of an estimate arising in life testing. *Technometrics*, 5(3):pp. 361–374.

Berzborn, M. and Cramer, E. (2023). Inference for Type-I and Type-II hybrid censored minimal repair and record data. *Journal of Statistical Theory and Practice*, 17(53).

Birnbaum, Z. and Saunders, S. (1969). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, 6(2):328–347.

Burkschat, M., Cramer, E., and Górny, J. (2016). Type-I censored sequential k-out-of-n systems. *Appl. Math. Model.*, 40(19–20):8156–8174.

Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Duxbury Press, Boston, 2 edition.

Chandrasekar, B., Childs, A., and Balakrishnan, N. (2004). Exact likelihood inference for the exponential distribution under generalized Type-I and Type-II hybrid censoring. *Naval Res. Logist.*, 51(7):994–1004.

Chen, S.-M. and Bhattacharyya, G. K. (1988). Exact confidence bounds for an exponential parameter under hybrid censoring. *Comm. Statist. Theory Methods*, 16(8):2429–2442.

Childs, A., Chandrasekar, B., and Balakrishnan, N. (2008). Exact likelihood inference for an exponential parameter under progressive hybrid censoring schemes. In Vonta, F., Nikulin, M., Limnios, N., and Huber-Carol, C., editors, *Statistical Models and Methods for Biomedical and Technical Systems*, pages 323–334. Birkhäuser, Boston.

Childs, A., Chandrasekar, B., Balakrishnan, N., and Kundu, D. (2003). Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution. *Ann. Inst. Statist. Math.*, 55(2):319–330.

Cramer, E. (2025). Structure of hybrid censoring schemes and its implications. *Metrika*, 88:393–417.

Cramer, E. and Balakrishnan, N. (2013). On some exact distributional results based on Type-I progressively hybrid censored data from exponential distributions. *Statist. Methodology*, 10:128–150.

Cramer, E., Burkschat, M., and Górny, J. (2016). On the exact distribution of the MLEs based on Type-II progressively hybrid censored data from exponential distributions. *J. Stat. Comp. Simul.*, 86:2036–2052.

Dahmen, W. and Micchelli, C. A. (1986). Statistical encounters with B-splines. In *Function Estimates (Arcata, Calif., 1985)*, pages 17–48. Amer. Math. Soc., Providence, RI.

de Boor, C. (1972). On calculating with B-splines. *Journal of Approximation Theory*, 6(1):50–62.

de Boor, C., Lyche, T., and Schumaker, L. L. (1976). *On Calculating with B-Splines II. Integration*, pages 123–146. Birkhäuser, Basel.

de Boor, C. C. (2001). *A Practical Guide to Splines*. Springer, New York. Revised edition.

Epstein, B. (1954). Truncated life tests in the exponential case. *Ann. Math. Stat.*, 25:555–564.

Fairbanks, K., Madsen, R., and Dykstra, R. (1982). A confidence interval for an exponential parameter from a hybrid life test. *Journal of the American Statistical Association*, 77(377):137–140.

Górny, J. and Cramer, E. (2016). Exact likelihood inference for exponential distributions under generalized progressive hybrid censoring schemes. *Statistical Methodology*, 29:70–94.

Górny, J. and Cramer, E. (2018a). Exact inference for a new flexible hybrid censoring scheme. *Journal of the Indian Society for Probability and Statistics*, 19(1):169–199.

Górny, J. and Cramer, E. (2018b). Modularization of hybrid censoring schemes and its application to unified progressive hybrid censoring. *Metrika*, 81:173–210.

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products*. Academic Press, Boston, 7 edition.

Gupta, R. D. and Kundu, D. (1998). Hybrid censoring schemes with exponential failure distribution. *Comm. Statist. Theory Methods*, 27(12):3065–3083.

Hahn, G. J., Meeker, W. Q., and Escobar, L. A. (2017). *Statistical Intervals: A Guide for Practitioners*. John Wiley & Sons, New York, 2 edition.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A.,

del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

Huang, W.-T. and Yang, K.-C. (2010). A new hybrid censoring scheme and some of its properties. *Tamsui Oxford Journal of Mathematical Sciences*, 26(4):355–367.

Ignatov, Z. G. and Kaishev, V. K. (1989). A probabilistic interpretation of multivariate B-splines and some applications. *Serdica*, 15:91–99.

Iliopoulos, G. (2015). On exact confidence intervals in a competing risks model with generalized hybrid type-i censored exponential data. *J. Stat. Comp. Simul.*, 85:2953–2961.

Kalbfleisch, J. G. (1985). *Probability and Statistical Inference*, volume 2. Springer, New York [u.a.], 2. ed. edition.

Kalinowski, T., Ushey, K., Allaire, J., and Tang, Y. (2026). *reticulate: Interface to 'Python'*. R package version 1.47.0.

Kundu, D. and Ganguly, A. (2017). *Analysis of Step-Stress Models*. Academic Press Inc., London, UK.

Kundu, D. and Gupta, R. D. (2007). Analysis of hybrid life-tests in presence of competing risks. *Metrika*, 65(2):159–170.

Laumen, B. and Cramer, E. (2021). k-step stage life testing. *Stat. Neerlandica*, 75(2):203–233.

Maechler, M. (2025). *Rmpfr: Interface R to MPFR - Multiple Precision Floating-Point Reliable*. R package version 1.1-2.

Park, S. and Balakrishnan, N. (2012). A very flexible hybrid censoring scheme and its Fisher information. *J. Stat. Comp. Simul.*, 82(1):41–50.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Prajapat, K., Mondal, S., Mitra, S., and Kundu, D. (2023). Meta-analysis of exponential lifetime data from type-i hybrid censored samples. *Communications in Statistics - Theory and Methods*, 53(11):3973–3991.

Schumaker, L. (2007). *Spline Functions: Basic Theory*. Cambridge University Press, Cambridge, NY, 3 edition.

Sundberg, R. (2001). Comparison of confidence procedures for type I censored exponential lifetimes. *Lifetime Data Analysis*, 7(4):393–413.

The mpmath development team (2026). *mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 1.4.0)*. http://mpmath.org/.

Uddin, A., Cramer, E., and Sultana, F. (2025). Exact inference for the parameter of the exponential distribution based on type-i censored stage life testing data. *Journal of Statistical Computation and Simulation*, pages 1–38.

van Bentum, T. (2024). *Point processes and hybrid censored data*. PhD thesis, RWTH Aachen University, Aachen.

van Bentum, T. and Cramer, E. (2019). Stochastic monotonicity of MLEs of the mean for exponentially distributed lifetimes under sequential hybrid censoring. *Statist. Probab. Letters*, 148:1–8.

van Bentum, T. and Cramer, E. (2025). Design of generalized hybrid censoring schemes with an application to likelihood prediction. *REVSTAT – Statistical Journal*, 23:645–677.

Viveros, R. and Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data. *Technometrics*, 36:84–91.