Imprecise system reliability using the survival signature

Citation for published version (APA):

Document license:
Unspecified

Document status and date:
Published: 26/01/2016

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 25. Dec. 2020
ABSTRACT: The survival signature has been introduced to simplify quantification of reliability of systems which consist of components of different types, with multiple components of at least one of these types. The survival signature generalizes the system signature, which has attracted much interest in the theoretical reliability literature but has limited practical value as it can only be used for systems with a single type of components. The key property for uncertainty quantification of the survival signature, in line with the signature, is full separation of aspects of the system structure and failure times of the system components. This is particularly useful for statistical inference on the system reliability based on component failure times. This paper provides a brief overview of the survival signature and its use for statistical inference for system reliability. We show the application of generalized Bayesian methods and nonparametric predictive inference, both these inference methods use imprecise probabilities to quantify uncertainty, where imprecision reflects the amount of information available. The paper ends with a discussion of related research challenges.

1 INTRODUCTION

In mathematical theory of reliability the main focus is on the functioning of a system given the functioning, or not, of its components and the structure of the system. The mathematical concept which is central to this theory is the structure function (Barlow & Proschan 1975). For a system with $m$ components, the state vector is $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in \{0, 1\}^m$, with $x_i = 1$ if the $i$th component functions and $x_i = 0$ if not. The labelling of the components is arbitrary but must be fixed to define $\mathbf{x}$. The structure function $\phi : \{0, 1\}^m \rightarrow \{0, 1\}$, defined for all possible $\mathbf{x}$, takes the value 1 if the system functions and 0 if the system does not function for state vector $\mathbf{x}$. Most practical systems are coherent, which means that $\phi(\mathbf{x})$ is not decreasing in any of the components of $\mathbf{x}$, so system functioning cannot be improved by worse performance of one or more of its components. The assumption of coherent systems is made throughout this paper and is convenient from the perspective of uncertainty quantification for system reliability. It is further logical to assume that $\phi(\mathbf{0}) = 0$ and $\phi(\mathbf{1}) = 1$, so the system fails if all its components fail and it functions if all its components function.

For larger systems, working with the full structure function may be complicated, and one may particularly only need a summary of the structure function in case the system has exchangeable components of one or more types. We use the term ‘exchangeable components’ to indicate that the failure times of the components in the system are exchangeable (De Finetti 1974). Recently, we introduced such a summary, called the survival signature, to facilitate reliability analyses for systems with multiple types of components (Coolen & Coolen-Maturi 2012). In
The survival signature is closely related to the system signature (Samaniego 2007), which is well-established and the topic of many research papers during the last decade. However, generalization of the signature to systems with multiple types of components is extremely complicated (as it involves ordering order statistics of different distributions), so much so that it cannot be applied to most practical systems. In addition to the possible use for such systems, where the benefit only occurs if there are multiple components of the same type, the survival signature is arguably also easier to interpret than the signature.

Consider a system with \( K \geq 1 \) types of components, with \( m_k \) components of type \( k \in \{1, \ldots, K\} \) and \( \sum_{k=1}^{K} m_k = m \). Assume that the random failure times of components of the same type are exchangeable (De Finetti 1974). Due to the arbitrary ordering of the components in the state vector, components of the same type can be grouped together, leading to a state vector that can be written as \( \mathbf{x} = (x_1, x_2, \ldots, x^K) \), with \( x_k = (x_1^k, x_2^k, \ldots, x_{m_k}^k) \) the sub-vector representing the states of the components of type \( k \).

The survival signature for such a system, denoted by \( \Phi(l_1, \ldots, l_K) \), with \( l_k = 0, 1, \ldots, m_k \) for \( k = 1, \ldots, K \), is defined as the probability for the event that the system functions given that precisely \( l_k \) of its \( m_k \) components of type \( k \) function, for each \( k \in \{1, \ldots, K\} \) (Coolen & Coolen-Maturi 2012). There are \( \binom{m_k}{l_k} \) state vectors \( \mathbf{x}^k \) with \( \sum_{i=1}^{m_k} x_{i}^k = l_k \). Let \( S_{l_k}^k \) denote the set of these state vectors for components of type \( k \) and let \( S_{l_1, \ldots, l_K} \) denote the set of all state vectors for the whole system for which \( \sum_{i=1}^{m_k} x_{i}^k = l_k \), \( k = 1, \ldots, K \). We also introduce the notation \( l = (l_1, \ldots, l_K) \). Due to the exchangeability assumption for the failure times of the \( m_k \) components of type \( k \), all the state vectors \( \mathbf{x}^k \in S_{l_k}^k \) are equally likely to occur, hence (Coolen & Coolen-Maturi 2012)

\[
\Phi(l) = \left[ \prod_{k=1}^{K} \binom{m_k}{l_k} \right]^{-1} \times \sum_{\mathbf{x} \in S_{l_1, \ldots, l_K}} \phi(\mathbf{x})
\]

Let \( C_{i}^k \in \{0,1, \ldots, m_k\} \) denote the number of components of type \( k \) in the system that function at time \( t > 0 \). Then, for system failure time \( T_S \),

\[
P(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l) P(\bigcap_{k=1}^{K} \{C_{i}^k = l_k\})
\]

There are no restrictions on dependence of the failure times of components of different types, as the probability \( P(\bigcap_{k=1}^{K} \{C_{i}^k = l_k\}) \) can take any form of dependence into account, for example one can include common-cause failures quite straightforwardly into this approach (Coolen & Coolen-Maturi 2015). However, there is a substantial simplification if one assumes that the failure times of components of different types are independent, and even more so if one assumes that the failure times of components of type \( k \) are conditionally independent and identically distributed with CDF \( F_k(t) \). With these assumptions, we get

\[
P(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l) \times \prod_{k=1}^{K} \left( \frac{m_k}{l_k} \right) [F_k(t)]^{m_k-l_k}[1 - F_k(t)]^{l_k}
\]

The advantage of the survival signature, in line with this property of the signature for systems with a single type of components (Samaniego 2007), is that the information about the system structure is fully separated from the information about functioning of the components, which simplifies related statistical inference as well as considerations of optimal system design. In particular for study of system reliability over time, with the structure of the system, and hence the survival signature, not changing, this separation also enables relatively straightforward statistical inferences where even the use of imprecise probabilistic methods (Augustin, Coolen, de Cooman, & Troffaes 2014, Coolen & Utkin 2011) is quite straightforward. Such methods have the advantage that imprecision for the system survival function reflects the amount of information available. The next two sections briefly discuss such methods of statistical inference for the system failure time.

Inference (Coolen 2011), a frequentist statistical method which is based on relatively few assumptions, enabled through the use of imprecise probabilities, and which does not require the use of prior distributions. The paper ends with a brief discussion of research challenges, particularly with regard to upscaling the survival signature methodology for application to large-scale real-world systems and networks.

2 IMPRECISE BAYESIAN INFERENCE

The reliability of a system, for which the survival signature is available, is quite straightforwardly quantified through its survival function, as shown in the previous section. We briefly consider a scenario where we have test data that enable learning about the reliability of the components of different types in the system, where we assume independence of the failure times of components of different types. The numbers of components in the system, of each type, that are functioning at time \( t \), denoted by \( C_{i}^k \) for \( k = 1, \ldots, K \), are the random quantities of main interest. One attractive statistical method to learn about these random quantities from test data is provided by the Bayesian
framework of statistics, which can be applied with the assumption of a parametric distribution for the component failure times (Walter, Graham, & Coolen 2015) or in a nonparametric manner (Aslett, Coolen, & Wilson 2015). We briefly illustrate the latter approach.

Assume that there are \(m_k\) components of type \(k\) in the system, and we are interested in the probability distribution of \(C^k_t\). Suppose that \(n_k\) components of the same type \(k\) were tested, these are not the components that are in the system but their failure times are assumed to be exchangeable with those in the system. We assume that for all tested components the failure time has been observed, let \(s^k_t\) denote the number of these components that still functioned at time \(t\). A convenient and basic model for \(C^k_t\) is the Binomial distribution, where the probability of ‘success’, that is a component still to be functioning at time \(t\), can, in the Bayesian framework, be conveniently modelled as a random quantity with a Beta prior distribution. Different to the standard parameterization for the Beta distribution, we define a Beta prior distribution as a pseudocount of components and the expected number of these components that still functioned at time \(t\). The updating results in a similar distribution of \(C^k_t\) and \(y^k_{k,t}\) still function at time \(t\) (Walter 2013). Doing this leads to straightforward updating, using the test information consisting of observations of \(n_k\) components of which \(s^k_t\) were still functioning at time \(t\). The updating results in a similar Beta distribution as the prior, but now with parameter values \(n^k_{k,t} = n^k_{k,t} + n_k\) and \(y^k_{k,t}\) as interpretations a pseudocount of components and the expected value of the success probability, respectively. Hence, these parameters can be interpreted in the sense that the prior distribution represents beliefs reflecting the same information as would result from observing \(n^k_{k,t}\) components of which \(s^k_t\) still function at time \(t\) (Walter 2013). This model enables conflict between prior beliefs and data to be shown through increased imprecision, that is difference between upper and lower probabilities for an event of interest (Walter 2013). We illustrate the use of this model, together with the survival signature, for a small system in Example 1, without attention to such prior-data conflict, further details on this will be presented elsewhere (Walter, Aslett, & Coolen 2016).

Example 1

As a small example, consider the system with three types of components presented in Figure 1. The survival signature of this system is given in Table 1, where all cases with \(l_3 = 0\) have been omitted as the system cannot function if the component of Type 3 does not function, hence \(\Phi(l_1, l_2, 0) = 0\) for all \((l_1, l_2)\).

![Figure 1: System with 3 types of components](image)

Table 1: Survival signature for the system in Figure 1 for cases with \(l_3 = 1\).

<table>
<thead>
<tr>
<th>(l_1)</th>
<th>(l_2)</th>
<th>(\Phi(l_1, l_2, 1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

For component types 1 and 2, we consider a near-noninformative set of prior survival functions. For components of type 3, we consider an informative set of prior survival functions as given in Table 2. This set could result from eliciting prior survival probabilities at times \(t = 0, 1, 2, 3, 4, 5\) only, and using those values to deduce such prior probabilities for all other values of \(t\) without further assumptions. These prior assumptions, together with sets of posterior survival functions, are illustrated in Figure 3 (presented at the end of the paper); test data for components of type 1 and 2 are taken as \{2.2, 2.4, 2.6, 2.8\} and such prior sets, with only an interval of possible values specified for each parameter, provides much flexibility for modelling prior beliefs and indeterminacy, together with interesting ways in which the corresponding sets of posterior (predictive) distributions and related inferences can vary. Most noticeably, this model enables conflict between prior beliefs and data to be shown through increased imprecision, that is difference between upper and lower probabilities for an event of interest (Walter 2013). We illustrate the use of this model, together with the survival signature, for a small system in Example 1, without attention to such prior-data conflict, further details on this will be presented elsewhere (Walter, Aslett, & Coolen 2016).
Table 2: Lower and upper prior functioning probability bounds for component type 3 in the system of Figure 1.

$$\begin{array}{c|cccc}
 t & (0, 1) & (1, 2) & (2, 3) & (3, 4) & (4, 5) \\
\hline
\tilde{P}_{3,t}^{(0)} & 0.625 & 0.375 & 0.250 & 0.125 & 0.010 \\
\tilde{P}_{3,t}^{(0)} & 0.999 & 0.875 & 0.500 & 0.375 & 0.250 \\
\end{array}$$
component type \( k \), the function \( \bar{D} \) ensures that maximum possible probability, corresponding to NPI for Bernoulli data (Coolen 1998), is assigned to the event \( C^k_t = 0 \), so \( \bar{D}(C^k_t = 0) = \bar{P}(C^k_t = 0) \). Then, \( \bar{D}(C^k_t = 1) \) is defined by putting the maximum possible remaining probability mass, from the total probability mass available for the event \( C^k_t \leq 1 \), to the event \( C^k_t = 1 \). This is achieved by \( \bar{D}(C^k_t = 1) = \bar{P}(C^k_t \leq 1) - \bar{P}(C^k_t = 0) \). This argument is continued, by assigning for increasing \( l_k \) the maximum possible remaining probability mass \( \bar{D}(C^k_t = l_k) \). As the survival signature is increasing in \( l_k \) for coherent systems, as assumed in this paper, and the resulting \( \bar{D} \) is a precise probability distribution, the right-hand side of the inequality above is indeed a lower bound and it is the maximum possible lower bound. As such, it is the NPI lower probability for the event \( T_S > t \), giving the NPI lower survival function for the system failure time (for \( t > 0 \))

\[
P(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l) \prod_{k=1}^{K} \bar{D}(C^k_t = l_k)
\]

The corresponding NPI upper survival function for \( T_S \) is similarly derived, using the upper bound

\[
P(T_S > t) \leq \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l) \prod_{k=1}^{K} D(C^k_t = l_k)
\]

where

\[
D(C^k_t = l_k) = P(C^k_t \leq l_k) - P(C^k_t \leq l_k - 1) =
\]

\[
\begin{pmatrix}
  n_k + m_k \\
  n_k
\end{pmatrix}^{-1} \begin{pmatrix}
  s^k_t + l_k \\
  s^k_t
\end{pmatrix} \times
\begin{pmatrix}
  n_k - s^k_t + m_k - l_k - 1 \\
  n_k - s^k_t
\end{pmatrix}
\]

In this expression, \( P \) denotes the NPI lower probability for Bernoulli data (Coolen 1998). This construction ensures that minimum possible weight is given to small values of \( C^k_t \), resulting in the NPI upper survival function for the system failure time

\[
\bar{P}(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l) \prod_{k=1}^{K} D(C^k_t = l_k)
\]

We illustrate this NPI method for system reliability using the survival signature in Example 2 (Coolen, Coolen-Maturi, & Al-nefaiee 2014).

Example 2.
Consider the system with \( K = 2 \) types of components as presented in Figure 2. The survival signature for this system is presented in Table 3. It is easily verified by checking all possible combinations of the specific components of each type which function or not.

To illustrate NPI for the system survival time, suppose that \( n_1 = 2 \) components exchangeable with those of type 1 and \( n_2 = 2 \) components exchangeable with those of type 2 were tested. First suppose that failure times \( t^2_4 < t^1_4 < t^2_3 < t^1_3 \) were observed, with \( t_j^k \) the \( j \)-th ordered failure time of a component of type \( k \). The resulting NPI lower and upper survival functions for the system failure time \( T_S \) are specified in Table 4, together with the results for the case with the test failure times ordered as \( t^1_1 < t^2_2 < t^1_3 < t^2_3 \).

For the ordering \( t^2_1 < t^1_1 < t^2_2 < t^2_3 \), in the first interval in Table 4 we have not yet seen a failure in the test data, so the NPI upper probability that the system will function is equal to one, which is logical as we base the inferences on the data with few additional assumptions. In the second interval, one failure of type 2 has occurred but we do not have any evidence from the data against the possibility that a component of type 1 will certainly function at times in this interval, so the NPI upper survival function remains one. In the fourth interval, both type 2 components have failed but only one component of type 1 has failed. In this interval, to consider the lower survival function the system is effectively reduced to a series system consisting of three components of type 1, with one ‘success’ and one ‘failure’ as data, denoted by (2, 1). As such a series system only functions if all three components function, the NPI lower survival function within this fourth interval is equal to

\[
\bar{S}_{T_S}(t) = \frac{1}{2} \times \frac{1}{4} \times \frac{1}{3}
\]
hence this second component will also function with the system, so with combined data the probability for the third component to function, conceptually easy, for example one can straightforwardly generalize the survival signature to multi-state systems such that it again summarizes the structure function in a manner that is sufficient for a range of uncertainty quantifications for the system reliability. The survival signature can also be used with a generalization of the system structure function where the latter is a probability instead of a binary function, or even an imprecise probability. This enables uncertainty of system functioning for given states of its components to be taken into account, which may be convenient, for example, to take uncertain demands or environments for the system into consideration. In this paper, we only considered test data with observed failure times for all tested components. If test data also contain right-censored observations, this can also be dealt with, both in the imprecise Bayesian and NPI approaches (Walter, Graham, & Coolen 2015, Coolen & Yan 2004, Maturi 2010) (more information about NPI is available from www.npi-statistics.com). This generalization is further relevant as, instead of assuming availability of test data, it allows us to take process data for the actual components in a system into account while this system is operating, hence enabling inference on the remaining time until system failure.

Upscaling the survival signature to large real-world systems and networks, consisting of thousands of components, is a major challenge. However, even for such systems the fact that one only needs to derive the survival signature once for a system is an advantage, and also the monotonicity of the survival signature for coherent systems is very useful if one can only derive it partially. For small to medium-sized systems and networks, the survival signature is particularly easy to compute using the ReliabilityTheory R package (Aslett 2016b), available from www.louisaslett.com. Using this package it is straightforward to express your system in terms of an undirected graphical structure, after which a single call to the function computeSystemSurvivalSignature suffices. The function will compute all of the cut sets of the system and perform the combinatorial analysis, returning a table which contains the survival signature just as in Table 2 and 3. For example, computation of the survival signature for the system in Figure 1 is achieved with 3 simple commands.

```r
s <- graph.formula(s-1-2-3-t, s-1-4-5-t, 2:4-6:3:5)
setCompTypes(s, list("T1"=c(2,4,3,5), "T2"=6, "T3"=1))
computeSystemSurvivalSignature(s)
```

Full instructions and some worked examples are available within the package. There are numerous other functions in the package, enabling computation of: the legacy system signature (Samaniego 2007); the
continuous-time Markov chain representation of repairable systems; as well as numerous inference algorithms for Bayesian inference on the system signature using only system-level data (Aslett 2013).

The survival signature enables some interesting applications which would otherwise be intractably difficult. For example, often a system designer may consider the design (structure) of their system to be a trade secret and so be unwilling to release it to component manufacturers, while at the same time component manufacturers are frequently unwilling to release anything more than summary figures for components, e.g. mean-time-between-failures. These two opposing goals lead to a situation in which it would seem unrealistic to achieve a full probabilistic reliability assessment and to honour the privacy requirements of all parties. However, recent work (Aslett 2016a) makes use of the survival signature to allow cryptographically secure evaluation of the system reliability function, where the functional form resulting from the survival signature decomposition in Equation (1) is crucial to enabling encrypted computation using so-called homomorphic encryption schemes (Aslett, Esperança, & Holmes 2015). The equivalent decomposition in terms of the structure function leads to difficulties in encrypted computation, so that this application may be intractable without use of the survival signature.

ACKNOWLEDGEMENTS

The authors wish to thank Professor Radim Bris for his kind invitation to present this work at the ICAMER 2016 conference. Louis Aslett was supported by the i-like project (EPSRC grant reference number EP/K014463/1). Gero Walter was supported by the DINALOG project ‘Coordinated Advanced Maintenance and Logistics Planning for the Process Industries’ (CAMPI).

REFERENCES


Figure 3: Prior and posterior sets of survival functions for the system in Figure 1 and its three component types. The component failure times, that form the test data, are denoted with tick marks near the time axis.