Reliability Engineering and System Safety 96 (2011) 425–431

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Utility based maintenance analysis using a Random Sign censoring model

J. Andrés Christen [a,*], Fabrizio Ruggeri [b], Enrique Villa [a]

[a] CIMAT, A.P. 402, Guanajuato, GTO 36000, Mexico
[b] CNR IMATI, Milano, Italy

ARTICLE INFO

Article history:
Received 7 January 2010
Received in revised form
29 October 2010
Accepted 6 November 2010
Available online 11 November 2010

Keywords:
Competing risks
Bayesian analysis
MCMC
Failure times
Optimal maintenance time

ABSTRACT

Industrial systems subject to failures are usually inspected when there are evident signs of an imminent failure. Maintenance is therefore performed at a random time, somehow dependent on the failure mechanism. A competing risk model, namely a Random Sign model, is considered to relate failure and maintenance times. We propose a novel Bayesian analysis of the model and apply it to actual data from a water pump in an oil refinery. The design of an optimal maintenance policy is then discussed under a formal decision theoretic approach, analyzing the goodness of the current maintenance policy and making decisions about the optimal maintenance time.

1. Introduction

Complex systems and components subject to failures are the object of preventive maintenance in order to avoid the more or less disruptive consequences of their failure. Examples range from nuclear power plants to heating systems and car tires. Both maintenance and failure have a cost, with the former being less expensive in general. At the same time, policies based on too early maintenance could be unacceptable since there should be an excessive use of new components well before failing of old ones and the continuity of services provided by system could not be guaranteed because of maintenance. Failure and maintenance times are naturally intertwined and there is a quest for optimal maintenance policies that act just before failures.

Different models have been proposed in the literature about optimal maintenance policies, going back, for example, to Barlow and Hunter [1]. Here we concentrate on a particular case, in which data are available as either failure or maintenance time and maintenance is performed when some warnings denote a possible incipient failure. Such situation arises quite naturally under a condition-based maintenance policy, which is addressed in many papers (e.g. [2]). Both maintenance and failure are modeled by random variables, as in the Random Sign model developed by Cooke [3], which will be considered in this paper. The model is a particular example of competing risks model, widely used in reliability and survival analysis. A review about competing risks is provided by Crowder [4].

The Random Sign model, illustrated in Section 2, considers the case of a component whose failure time is subject to a possible right censoring, due to maintenance; here censoring is supposed independent of the age X at which the component would expire but, given that the component is censored, the censoring time may depend on X.

The Random Sign model seems suitable for analyzing data about an oil refinery water pump, see Section 4. A limited number of data are collected and the actual maintenance policy is unknown. At the same time, company experts can provide opinions on the failure process of the pump and the maintenance policy followed so far. Maintenance and repair costs can be assessed as well, and combined with knowledge about the failure process to develop optimal maintenance policies using a decision theoretic approach. Therefore, a Bayesian analysis of the Random Sign model, novel in the literature, is presented in Section 3, along with a utility-based optimal maintenance policy; both are applied to the pump data in Section 4. Concluding remarks are presented in Section 5.

2. Random Sign model

Since Cox [5], lack of identifiability of marginal distributions of some competing risk models has been discussed in the literature; we will not further discuss this problem and refer the interested reader to, for example, Bunea [6]. Cooke [3] proposed the Random Sign censoring, which is probably the simplest model allowing for identifiability of marginal distributions. The model assumes that a component, which would fail at time X, could be subject to right

* Corresponding author.
E-mail addresses: jac@cimat.mx (J. Andrés Christen), fabrizio@mi.imati.cnr.it (F. Ruggeri), villadi@cimat.mx (E. Villa).

0951-8320/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2010.11.001

censoring. The event that the lifetime could be censored is independent of the age X but, once censoring occurs, then censoring time depends on X. A typical situation modeled by Random Sign censoring is about censoring occurring when some warning about failing age X is available. In particular, in our case study, censoring is due to maintenance and, therefore, an intervention occurs in advance to avoid failure once some warning denotes an incipient failure. We present the definition of a Random Sign censoring model (see, for example, [7, p. 359]).

Definition 1. Given a random variable X, consider $Y=X-W\delta$, where W is a random variable with $0<W<X$ and $\delta=\{-1,1\}$ is also a random variable independent of X. The variable $Z\equiv[\min(X,Y), 1(Y<X)]$, with $1(\cdot)$ denoting the set function, is called a Random Sign censoring of X by Y.

Bunea et al. [7, p. 359] present an alternative definition using the indicator function $1(X<Y)$; our definition above should lead to no confusion.

3. Bayesian analysis

We consider the Random Sign censoring model given in Definition 1 and denote maintenance and failure times by Y and X, respectively. Unlike the more standard notation used in Definition 1, we take $\delta=2\varepsilon-1$, so that $\varepsilon=\{0,1\}$ (failure and maintenance, respectively). Then ε and X are independent and X and Y are related by

$$Y=X-(2\varepsilon-1)W, \tag{1}$$

$0<W<X$. Let $T=\min(X,Y)$ and note that $Z=(T,\varepsilon)$. It follows that

$$f(T=t,\varepsilon=0)=f(X=t,\varepsilon=0)=f(X=t)f(\varepsilon=0), \tag{2}$$

given the assumed independence of ε and X. In this case W takes any value in (0, t). Here we use f to denote any density; from its argument, it is evident which random variable is related to.

Conversely, censoring implies

$$\begin{aligned} f(T=t,\varepsilon=1)&=f(Y=t,\varepsilon=1)=f(Y=t|\varepsilon=1)f(\varepsilon=1)\\ &=f(\varepsilon=1)\int f(X=t+w,W=w)\,dw\\ &=f(\varepsilon=1)\int f(X=t+w|W=w)f(W=w)\,dw, \end{aligned} \tag{3}$$

since $X=Y+W$, with $Y=t$ and $W>0$. Distributions on ε and X are to be specified for both cases, whereas the distribution on W has to be specified only for the censoring case. In this case W has the meaning of the "time elapsed between the censoring (maintenance) time and the failure time which would have been observed if no censoring had been performed". Now we present a particular (parametric) proposal for these distributions.

3.1. Models for failure and censoring times

We consider a Gamma model $\mathcal{G}(\alpha,\lambda)$ for the failure time X, with $\alpha,\lambda>0$. This is a standard failure time model with sufficient flexibility, and adequate when light tails are expected. We now turn to define the distribution for $W|X=x$. This distribution should have support in $[0,x]$. Moreover, since we expect W, the elapsed time between maintenance and (unobserved) failure, to be definitely closer to 0 rather than to its largest value $X=x$, we should have that this conditional distribution is decreasing. A reasonable assumption would be

$$f(W=w|X=x)=\frac{\alpha-1}{x}\left(\frac{w}{x}\right)^{\alpha-2}I_{(0,x)}^{(w)}, \tag{4}$$

a truncated power law. We choose $1<\alpha\le 2$, so that (4) is non-increasing. The shape of this distribution is not dependent on x and is solely governed by the parameter α.

Under such assumptions on the distributions of X and $W|X$, it can be easily shown that, marginally, $W\sim\mathcal{G}(\alpha-1,\lambda)$, $Y|\varepsilon=1\sim\mathcal{E}(\lambda)$ and $f(X=x|W=w)=\lambda\exp\{-\lambda(x-w)\}I_{(w,\infty)}(x)$ ($\mathcal{E}$ is the exponential distribution). As a special case, $\alpha=2$ implies $X\sim\mathcal{G}(2,\lambda)$, $W|X=x\sim\mathcal{U}(0,x)$ and $W\sim\mathcal{E}(\lambda)$. Regarding the censoring mechanism, we take $\varepsilon\sim\mathcal{B}er(\theta)$. This model, arising from the Gamma model for failure times (restricting $\alpha\in(1,2]$) and the truncated power law model for $W|X$, leads to a setting where all marginals are well defined, having known distributions, and is governed by the parameters (θ,λ,α), providing reasonable flexibility for typical applications.

3.2. Prior choice

We consider the parameter vector $\eta=(\theta,\lambda,\alpha)$ and we take independent priors $\theta\sim\mathcal{B}e(\alpha_1,\beta_1)$, $\lambda\sim\mathcal{G}(\alpha_2,\beta_2)$ and $\alpha\sim\mathcal{U}(1,2)$.

The choice of the hyperparameters is a complex aspect in Bayesian analysis and it is important to explore some features of the involved random variables, especially the observable ones, to specify their values. It is worth observing that $X\sim\mathcal{G}(\alpha,\lambda)$ and $Y\sim\mathcal{E}(\lambda)$ imply that $E(X)=\alpha/\lambda$ and $E(Y)=1/\lambda$, so that $\lambda=1/E(Y)$ and $\alpha=E(X)/E(Y)$. The choice of the hyperparameters α_2 and β_2 could be performed noting that

$$E(X)=E[E(X|\alpha,\lambda)]=E(\alpha)E\left(\frac{1}{\lambda}\right)=E(\alpha)\frac{\beta_2}{\alpha_2-1}. \tag{5}$$

Similarly, using $Var(X)=E[Var(X|\alpha,\lambda)]+Var[E(X|\alpha,\lambda)]$, it can be proved that

$$C_X^2=\frac{1}{\alpha_2-2}\{(\alpha_2-1)(\mu_\alpha^{-1}+C_\alpha^2)+1\}, \tag{6}$$

where C_X and C_α are the coefficients of variation (standard deviation over the mean) of X and α, and $\mu_\alpha=E(\alpha)$. To ensure the existence of $E(X)$ and C_X^2 we need $\alpha_2>2$; furthermore, (5) and (6) imply $\alpha_2=1+(C_X^2+1)/(C_X^2-(\mu_\alpha^{-1}+C_\alpha^2))$ and $\beta_2=\mu_\alpha^{-1}E(X)(\alpha_2-1)$. Both α_2 and β_2 are positive if

$$C_X>\sqrt{\mu_\alpha^{-1}+C_\alpha^2}.$$

The last condition is a property of our model, and for the uniform prior assumed for α entails to $\mu_\alpha=3/2$, $C_\alpha^2=1/27$ and thus $C_X>0.8388$; the *a priori* coefficient of variation for X cannot be lower than 0.8388. If we elicit an *a priori* expected value m_X and a standard error s_X for X (e.g. arising from the manufacturer's specifications for continuous operation of the system), we take $C_X=s_X/m_X$ to obtain α_2 and β_2 as above.

Regarding the choice of parameters for the Beta prior on θ, they could be chosen specifying some quantiles or by observing that $E(\theta|\alpha_1,\beta_1)=\alpha_1/(\alpha_1+\beta_1)$ and $Var(\theta|\alpha_1,\beta_1)=\alpha_1\beta_1/((\alpha_1+\beta_1)^2(\alpha_1+\beta_1+1))$. When scarce prior information is available about θ, a default non-informative prior may be stated as $\alpha_1=\beta_1=0.5$ (see [8]).

3.3. Likelihood

Suppose data are given by $(\underline{t},\underline{\varepsilon})=\{(t_i,\varepsilon_i)\}_{i=1}^n$; therefore the likelihood would be

$$f(\underline{t},\underline{\varepsilon}|\eta)=\prod_{i=1}^n f(t_i,\varepsilon_i|\eta), \tag{7}$$

where $f(t_i,\varepsilon_i|\eta)$ would be given by (2) and (3), depending on $\varepsilon=0$ or 1, respectively. The integrals involved in (3) are an evident drawback of the approach using the full likelihood (7). Therefore, we