Reliability Engineering and System Safety 115 (2013) 10–18

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

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

Estimation of G-renewal process parameters as an ill-posed inverse problem

V. Krivtsov [a,]*, O. Yevkin [b]

[a] Ford Motor Company, 330 Town Center Drive, Dearborn, MI 48126, USA
[b] IHS, 155 Gordon Baker Road, Toronto, ON, Canada M2H 3N5

ARTICLE INFO

Article history:
Received 4 September 2012
Received in revised form
6 February 2013
Accepted 7 February 2013
Available online 19 February 2013

Keywords:
G-renewal process
Underlying distribution
Monte Carlo simulation
Inverse problems
Regularization

ABSTRACT

Statistical estimation of G-renewal process parameters is an important estimation problem, which has been considered by many authors. We view this problem from the standpoint of a mathematically ill-posed, inverse problem (the solution is not unique and/or is sensitive to statistical error) and propose a regularization approach specifically suited to the G-renewal process. Regardless of the estimation method, the respective objective function usually involves parameters of the underlying life-time distribution and *simultaneously* the restoration parameter. In this paper, we propose to regularize the problem by *decoupling* the estimation of the aforementioned parameters. Using a simulation study, we show that the resulting estimation/extrapolation accuracy of the proposed method is considerably higher than that of the existing methods.

1. Introduction

Parameter estimation of the G-renewal process is an important problem. In this paper, we considered this problem as well as the problem of the extrapolation accuracy of the G-renewal function based on restricted (in time) empirical data. This problem often arises in forecasting of warranty repairs/costs [14,15], maintenance optimization [12,13,18,19] and evaluation of the repair quality and/or effectiveness [17].

From the standpoint of mathematics, statistical estimation, i.e., evaluating model's parameters based on the data, can be viewed as an *inverse* problem. This is in contrast to a *forward* problem, which involves evaluating/predicting data points based on the (estimated) model parameters. Evans and Stark [5] draw some formal parallels between statistical estimation problems and mathematical inverse problems. For example, they point out that *identifiability* (distinct models yield distinct probability distributions for the observed data) is similar to *uniqueness* (the forward operator maps at most one model into the observed data). Further, *consistency* (model parameters can be estimated with arbitrary accuracy as the number of data points grow) is related to *stability of recovery* (small changes in the data produce small changes in the recovered model).

Most of the inverse problems are considered to be mathematically incorrect and/or ill-posed. A typical remedy in this case is the so-called *regularization*, i.e., the introduction of additional information in order to solve an ill-posed problem or to prevent model overfitting. Bayesian estimation, Ridge regression and Lasso regression are examples of regularization used in statistical science.

In this paper, we view the problem of estimating parameters of the G-renewal process as a mathematically ill-posed or incorrect (ill-conditioned) problem. In order to regularize this problem, we propose an approach, which is neither Bayesian, nor Ridge/Lasso regression related. It is based on separating the underlying distribution parameters from the GRP restoration factor in two consecutive steps. Using a simulation study, we show that the resulting extrapolation accuracy of the proposed method is considerably higher than that of the existing methods.

The paper is structured as follows. In Section II, we overview the G-renewal process and represent it as an inverse problem. In Section III, we discuss existing estimation methods of the G-renewal equation parameters and propose a regularization approach along with the respective estimation procedure. We use a simulation study to compare the accuracy of the proposed approach relative to the existing methods. In Section IV, we use a practical case study to show the efficiency of the proposed method.

2. G-renewal estimation as an inverse problem

Kijima and Sumita [6] introduced the *generalized renewal* (*G-renewal*) *process* using the notion of *virtual age*:

$V_n = qS_n,$

Abbreviations: CDF, cumulative distribution function; CIF, cumulative intensity function; GRP, generalized renewal (or G-renewal) process; HPP, homogeneous Poison process; IFR, increasing failure rate; LSQ, least squares estimation; MC, Monte Carlo; MLE, maximum likelihood estimation; MTTF, mean time to failure; NHPP, non-homogeneous Poison process; ORP, ordinary renewal process; PDF, probability density function; PFF, percent of first failures; RSS, residual sum of squares; SE, standard error; TTF, time to the first failure

* Corresponding author. Tel.: +1 313 323 8711.
E-mail address: vkrivtso@ford.com (V. Krivtsov).

0951-8320/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2013.02.005

Nomenclature

q	restoration (or repair effectiveness) factor
Θ	vector of parameters of the underlying life-time distribution
t	time
$W(t)$	G-renewal function denoting the expected cumulative number of events (failures)
$f(t)$	probability density function
$F(t)$	cumulative distribution function
λ,α	respectively, the scale and the shape parameters of Weibull distribution

where V_n and S_n is the system's age after and before the n-th repair, respectively, and q is the restoration (or repair effectiveness) factor. With $q=0$, the age of the system after the repair is "re-set" to zero, which corresponds to the ORP. With $q=1$, the system is restored to the "same-as-old" condition, which is the NHPP. The case of $0<q<1$ corresponds to the intermediate "better-than-old-but-worse-than-new" repair assumption. Finally, with $q>1$, the virtual age is $A_n>S_n$, so that the repair damages (ages) the system to a higher degree than it was just before the respective failure, which corresponds to the "worse-than-old" repair assumption.

The G-renewal process gained its *practical* popularity only after methods for estimating its parameters had become available. The nonlinear LSQ estimation of the G-renewal process was first offered by Kaminskiy and Krivtsov [1]. The maximum likelihood procedures were subsequently discussed by Yañez, et. al [2] and Mettas and Zhao [3]. The estimation of the G-renewal restoration factor was addressed in detail by Kahle and Love [4].

Mathematically, estimation of the G-renewal process amounts to solving the following G-renewal equation with respect to its parameters, q and θ:

$$W(t) = \int_0^t \left(g(\tau|0)+\int_0^\tau w(x)g(\tau-x|x)dx \right) d\tau, \quad (1)$$

where

$$g(t|x)=\frac{f(t+qx,\theta)}{1-F(qx,\theta)}, \quad t,x\geq 0; \quad w(x)=\frac{dW(x)}{dx};$$

and $F(t)$ and $f(t)$ are the *cumulative distribution function* (CDF) and *probability density function* (PDF) of the underlying failure time distribution (note that $g(t|0)=f(t)$); q is the restoration factor, and θ is the vector of parameters of the underlying life-time distribution.

It can be shown that this inverse problem is ill-posed or incorrect. According to Hadamard [7], a *well-posed problem* is such for which: (a) a solution exist, (b) the solution is unique, (c) the solution depends continuously on the data in some reasonable topology. A *correct problem* has almost the same definition except for (c) the solution must be stable (meaning that small statistical errors in the data should not much influence the solution). The problem becomes *incorrect*, if at least one condition in the definition of a *correct problem* is violated.

Consider a G-renewal process with arbitrary and infinitely increasing cumulative intensity function, $W(t)$, corresponding to some underlying failure-time distribution function, $F(t)$, and the restoration factor $q\neq 1$. In this setting, one can derive another solution corresponding to $q=1$: $F(t)=1-e^{-W(t)}$, simply because the cumulative intensity function, $W(t)$, of the NHPP (i.e., G-renewal process with $q=1$) is formally equal to the cumulative hazard, $H(t)$, of the respective underlying failure-time distribution [9]. This is to say that for a G-renewal process with *any* value of $q\neq 1$, one can *always* find a solution in the above form of $F(t)$ and $q=1$. Hence, the solution is not unique in general case, and, therefore, the respective (inverse) problem is ill-posed. In other words, any CIF can be *theoretically* modeled by the NHPP, but it is not a unique presentation of the solution and, in a *practical* setting, may be not a correct reflection of the underlying physical process (for example, if we know apriori that the system is "better than old" after the restoration). With the proposed regularization approach, we can expect a more accurate estimation of GRP parameters (including the restoration factor, q), which, in turn, provides a better extrapolation accuracy of the CIF.

Even the *ordinary* renewal process ($q=0$) can be shown as an ill-posed (or ill-conditioned) reverse problem. Recall that the ordinary renewal equation involves a convolution integral (which, of course, is also present in Eq. 1):

$$W(t) = F(t)+\int_0^t F(t-\tau)\, dW(\tau) \quad (2)$$

Hence, the respective inverse problem can be considered as a *deconvolution* problem, which is typically ill-conditioned. Its solution is not stable with respect to the calculation error and/or empirical noise [10,11]. Moreover, as time tends to infinity, $W(t)$ becomes linear, and infinite number of underlying $F(t)$ can produce the same (linear) behavior of $W(t)$. Conversely, if time tends to zero, the CIF does not depend on restoration factor and approximates the underlying CDF.

The inverse (estimation) problem becomes even more complicated for the G-renewal process, because, in addition to the parameters of the underlying life-time distribution, one has to also deal with the restoration parameter. When discussing the estimation of the G-renewal process with the underlying Weibull distribution in [8], we noticed that two "competing" vectors of significantly different GRP parameters yielded practically indistinguishable values of the CIF at various time cross-sections.

As an illustration, let us choose the class of Weibull distribution functions,$F(t)=1-\exp(-\lambda t)^\alpha$, with scale parameter, λ, and shape parameter, α as the underlying failure-time distribution of the G-renewal process. Now, consider Fig. 1, where we show three CIF's simulated (under $n=10^7$ trials) with three sets of the underlying parameters: Case 1: $\{\lambda_1=1.0,\ \alpha_1=2.0,\ q_1=0\}$, Case 2: $\{\lambda_2=0.949,\ \alpha_2=1.675,\ q_2=0\}$, Case 3: $\{\lambda_3=0.949,\ \alpha_3=1.675,\ q_3=0.1\}$. Cases 2 and 3 can be considered as empirical data fluctuating close to the exact solution (Case 1).

Note that with as much as 20% difference in the Weibull shape parameter between Cases 1 and 2, the maximum difference in the respective values of the CIF is around 4% (at $t=1.214$). Moreover, even though all cases presented in Fig. 1 can be considered as good approximations (interpolations) in interval,$0\leq W(t)\leq 1$, they have significantly different GRP parameters and, as a consequence, yield significantly different extrapolations of the G-renewal function, as shown in Fig. 2. This indicates that even in the class of Weibull distribution functions, the inverse problem is ill-conditioned.

There are many developed general methods for regularization of inverse problems [10], however all of them require significant amount of additional calculations, because typically another optimization parameter (e.g., a Lagrange multiplier) is introduced in the problem. These methods are efficient in the case when corresponding forward problem is easy to solve. In our case, the forward problem is described by integral Eq. (1), which does not have a closed form solution.