Nonparametric estimation of time trend for repairable systems data

Bo Henry Lindqvist
Dept of Mathematical Sciences
Norwegian University of Science and Technology
N-7491 Trondheim
Norway
bo@math.ntnu.no

Abstract

The trend-renewal-process (TRP) is defined to be a time-transformed renewal process, where the time transformation is given by a trend function $\lambda(\cdot)$ which is similar to the intensity of a nonhomogeneous Poisson process (NHPP). A nonparametric maximum likelihood estimator of the trend function of a TRP can be obtained much in the same manner as for the NHPP using kernel smoothing. But for a TRP one must consider the simultaneous estimation of the renewal distribution, which is here assumed to belong to a parametric class such as the Weibull-distribution. A weighted kernel estimator for $\lambda(\cdot)$ is derived using a general approach for kernel smoothing in counting processes.

1 Introduction

Failures of a repairable system are usually modeled by a stochastic point process in time. The most common models are the renewal process (RP), the homogeneous Poisson process (HPP), and the nonhomogeneous Poisson process (NHPP) (see Ascher and Feingold, 1984). As is well known, the RP model assumes what is called "perfect repair", indicating that after each failure, the system is renewed to its original condition. The NHPP model, on the other hand, assumes what is called "minimal repair". After each failure and following repair, the system is in the same state as it was just prior to that failure. This is often more plausible than the complete renewal assumption. Yet often the replaced part is not a minor part, or the repair may affect some other parts of the system, to the better or the worse. This could mean a small jump in the intensity, in either direction.

There is thus a need for models which allow the system to deteriorate (or improve) over time, yet still allow for the possibility that the system could have a drastic increase or decrease in its failure intensity just after a repair, because of damage done, or weak spots removed. Several models have been developed for this purpose.

A class of models of this kind is the trend-renewal process (TRP), which was defined and studied by Lindqvist, Elvebakk and Heggland (2003), see also Lindqvist (2006). This model contains the RP and NHPP as special cases, and in a simple manner the TRP fills some of the gap between the two extreme repair models. While Lindqvist et al. (2003) mostly considered parametric estimation for TRP's, not much has been done on nonparametric estimation in the TRP model. An exception is here the paper Heggland and Lindqvist (2007), where nonparametric estimation of time trend is done under the assumption that this trend is monotonic, leading to a problem of isotonic regression. The purpose of this note is to present an approach for nonparametric estimation without assuming a monotonicity property of the time trend, and using kernel smoothing as the basic estimation technique.

2 Definitions and preliminaries

Consider a repairable system, observed from time $t = 0$. Let $N(t)$ be the number of failures in $(0, t]$, let T_i be the time of the i'th failure, where we define $T_0 = 0$, and let X_i be the time between failure number $i - 1$ and failure number i, that is $X_i = T_i - T_{i-1}$. We assume that all repair times equal 0. This assumption is reasonable if the repair times are negligible compared to the times between failures, or if we let the time parameter be the operation time of the system. For a general treatment of repairable systems, see Ascher and Feingold (1984) or Meeker and Escobar (1998).

We next review the definitions of the RP and NHPP, and then we define the trend-renewal process which will be the main model used in this paper and which can be seen as a generalization of the first three models.

2.1 Models for repairable systems

2.1.1 The renewal process, RP(F):

The process $N(t)$ is an RP(F) if $X_1, X_2, \ldots$ are independent and identically distributed with cumulative distribution function (cdf) F, where we assume $F(0) = 0$. If F is the exponential distribution with parameter λ, then RP(F)=HPP(λ), the homogeneous Poisson process with intensity λ.

2.1.2 The nonhomogeneous Poisson process, NHPP($\lambda(\cdot)$):

Let $\lambda(t),\ t \geq 0$ be a nonnegative function, called the intensity of the process. The cumulative intensity function is then $\Lambda(t) = \int_0^t \lambda(u)du$. The process $N(t)$ is an NHPP($\lambda(\cdot)$) if the time-transformed process $\Lambda(T_1), \Lambda(T_2), \ldots$ is an HPP(1).

2.1.3 The trend-renewal process, TRP($F, \lambda(\cdot)$):

Let $\lambda(t)$ and let $\Lambda(t)$ be as for the NHPP. The process $N(t)$ is a TRP($F, \lambda(\cdot)$) if the time-transformed process $\Lambda(T_1), \Lambda(T_2), \ldots$ is an RP(F). The distribution F is called the renewal distribution, and $\lambda(\cdot)$ is called the trend function of the TRP.

It is easily seen that the TRP generalizes both the NHPP and the RP, since TRP$(1 - e^{-x}, \lambda(\cdot))$ = NHPP($\lambda(\cdot)$) and TRP$(F, 1)$ = RP(F).

2.2 The likelihood function for the TRP model

The conditional intensity function of a point process (Andersen et al., 1993) is defined by

$$\gamma(t) = \lim_{\Delta t \to 0} \frac{\mathrm{P}(\text{failure in } [t, t + \Delta t)|\mathcal{F}_{t-})}{\Delta t},$$

where $\mathcal{F}_{t-}$ is the history of the process $N(t)$ up to, but not including time t. The conditional intensity function will, in general, be stochastic.

For a TRP($F, \lambda(\cdot)$) the conditional intensity function is given by

$$\gamma(t) = z(\Lambda(t) - \Lambda(T_{N(t-)}))\lambda(t). \tag{1}$$

Consider now a point process $N(t)$, observed from time $t = 0$ to time $t = \tau$, with corresponding failure times $T_1, T_2, \ldots, T_{N(\tau)}$ and conditional intensity function $\gamma(t)$. The likelihood function of the process is then given by (Andersen et al., 1993)

$$L = \{\prod_{i=1}^{N(\tau)} \gamma(T_i)\} \exp\{-\int_0^\tau \gamma(u)du\}. \tag{2}$$

The likelihood function of a TRP is obtained by substituting (1) into (2), giving

$$\begin{aligned} L &= \{\prod_{i=1}^{N(\tau)} z(\Lambda(T_i) - \Lambda(T_{i-1}))\lambda(T_i)\} \exp\left\{-\sum_{i=1}^{N(\tau)} \int_{T_{i-1}}^{T_i} z(\Lambda(u) - \Lambda(T_{i-1}))\lambda(u)du\right\} \\ &\quad \cdot \exp\left\{-\int_{T_{N(\tau)}}^{\tau} z(\Lambda(u) - \Lambda(T_{N(\tau)}))\lambda(u)du\right\}. \end{aligned}$$

By making the substitution $v = \Lambda(u) - \Lambda(T_{i-1})$ and taking the log we get the log likelihood function

$$\begin{aligned} l = \ln L &= \sum_{i=1}^{N(\tau)} \{\ln(z(\Lambda(T_i) - \Lambda(T_{i-1}))) + \ln(\lambda(T_i)) - \int_0^{\Lambda(T_i)-\Lambda(T_{i-1})} z(v)dv\} \\ &\quad - \int_0^{\Lambda(\tau)-\Lambda(T_{N(\tau)})} z(v)dv. \end{aligned} \tag{3}$$

3 Nonparametric estimation of $\lambda(\cdot)$

While Heggland and Lindqvist (2007) considered the case where $\lambda(t)$ is monotone but otherwise is completely unspecified, in the present note we assume that $\lambda(t)$ is completely nonparametric; not necessarily monotone. The renewal distribution is, on the other hand, assumed to be given on parametric form with hazard rate $z(t;\theta)$, as in Heggland and Lindqvist (2007). Due to an identifiability issue in the defintition of the TRP, we assume that θ does not include a scale parameter. For convenience we also assume that there is a parameter value $\theta^{(0)}$ such that $z(t;\theta^{(0)}) \equiv 1$.

The estimation in the next subsection will be done using an iterative scheme switching between maximizing the log likelihood (3) for estimation of θ, and estimating the trend function $\lambda(t)$ using in an ad hoc way the kernel smoothing method for counting process intensities due to Ramlau-Hansen (1983).

3.1 The kernel smoothing algorithm

Let $K(t)$ be a positive density function and let h until further be a fixed window size. We consider the following algorithm for estimation of $\lambda(\cdot)$ and θ.

Step 1 Let $\theta = \theta^{(0)}$. Then $\{N(t)\}$ is an NHPP with intensity $\lambda(t)$, and it is well known that this can be estimated by a standard kernel estimator (Ramlau-Hansen, 1983):

$$\lambda^{(1)}(t) = \frac{1}{h} \sum_{i=1}^{N(\tau)} K\left(\frac{t - T_i}{h}\right). \tag{4}$$

Let also the cumulative time trend $\Lambda^{(1)}(t) = \int_0^t \lambda^{(1)}(s)ds$ be given.

Step 2 Substitute $\lambda^{(1)}(\cdot)$, $\Lambda^{(1)}(\cdot)$ and $z(\cdot;\theta)$ in (3) and maximize the expression with respect to θ to find the maximum at $\theta = \theta^{(1)}$.

Step 3 Now use the estimated $\Lambda^{(1)}(\cdot)$ and $\theta^{(1)}$ as if they were known, to get from (1) an expression for the conditional intensity of $\{N(t)\}$ given by

$$z\left(\Lambda^{(1)}(t) - \Lambda^{(1)}(T_{N(t-)}); \theta^{(1)}\right) \; \lambda(t).$$

This is of the form of Aalen's multiplicative intensity model (Andersen et al., 1993), which assumes that the intensity of the counting process $\{N(t)\}$ is $\lambda(t)Y(t)$, where $\{Y(t)\}$ is an observable stochastic process, adapted to the process $\{N(t)\}$ and predictable. In our case,

$$Y(t) = z\left(\Lambda^{(1)}(t) - \Lambda^{(1)}(T_{N(t-)}); \theta^{(1)}\right)$$

which clearly satisfies these criterions. From equation (3.12) in Ramlau-Hansen (1983) we get from this a new kernel estimate for $\lambda(t)$ given by

$$\lambda^{(2)}(t) = \frac{1}{h} \sum_{i=1}^{N(\tau)} K\left(\frac{t - T_i}{h}\right) \cdot \frac{1}{z\left(\Lambda^{(1)}(T_i) - \Lambda^{(1)}(T_{i-1}); \theta^{(1)}\right)} \tag{5}$$

which is seen to be a weighted version of the kernel estimator $\lambda^{(1)}(t)$ in (4).

Step 4 Now return to the step 2 above with $\lambda^{(1)}(\cdot)$ and $\Lambda^{(1)}(\cdot)$ replaced by $\lambda^{(2)}(\cdot)$, $\Lambda^{(2)}(\cdot)$, to find $\theta^{(2)}$; then use these in step 3 and continue with steps 2 and 3 until convergence. Here $\Lambda^{(k)}(\cdot)$ is defined in the obvious way by integrating $\lambda^{(k)}(\cdot)$.

3.1.1 Example: The Weibull renewal distribution

By the non-uniqueness problem of the TRP we need to use only one parameter, letting

$$z(t;b) = bt^{b-1},$$

where $\beta = 1$ corresponds to an NHPP.

In this case we have the (inverse) weights

$$z(\Lambda(T_i) - \Lambda(T_{i-1})) = b(\Lambda(T_i) - \Lambda(T_{i-1}))^{b-1}.$$

3.2 Maximum likelihood kernel estimation

Jones and Henderson (2008) study kernel density estimation using variable weights or variable locations of the kernels, and then maximizing the nonparametric likelihood with respect to these weights or locations. Since our estimator (5) in effect is a weighted kernel estimator, it might be worthwhile to search for better weights than the ones suggested by (5). It is then tempting to try to maximize the log likelihood (3) simultaneously with respect to θ and the weights $w = (w_i; i = 1, 2, \ldots, N(\tau))$, which sum to $N(\tau)$, using

$$\lambda(t; w) = \frac{1}{h} \sum_{i=1}^{N(\tau)} K\left(\frac{t - T_i}{h}\right) w_i.$$

According to Jones and Henderson (2008) it may be even better to, instead of using variable weights, change the location of the ith term of the kernel estimator to be an unknown value m_i instead of the observed point T_i. In this case one may maximize the log likelihood (3) with respect to θ and the location points $m = (m_i; i = 1, 2, \ldots, N(\tau))$, using

$$\lambda(t; m) = \frac{1}{h} \sum_{i=1}^{N(\tau)} K\left(\frac{t - m_i}{h}\right).$$

Note then that each T_i has its own location m_i, while the weights are now all equal to 1.

It should finally be noted that Jones and Henderson (2008) maximize a likelihood corresponding to assuming $z(t) \equiv 1$ in (3). Thus the task of maximizing (3) for a general parametric $z(t; \theta)$ will obviously be a harder problem than theirs.

3.3 Choice of window size h

In the above we have assumed throughout a fixed h, tacitly assuming that the value of h should be chosen by some additional procedure. This is in fact also the approach of Jones and henderson, who advocate the use of the ordinary CV criterion. For the method of subsection 3.1 we suggest that an optimal value of h is chosen at the first step (NHPP case) and that this value is kept throughout the iterative procedure.

References

[1] Andersen, P., Borgan, O., Gill, R., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. New York: Springer.

[2] Ascher, H. and Feingold, H. (1984). *Repairable Systems Reliability: Modeling, Inference, Misconceptions and their Causes.* New York: Marcel Dekker.

[3] Heggland, K. and Lindqvist, B. H. (2007). A non-parametric monotone maximum likelihood estimator of time trend for repairable systems data. *Reliability Engineering and System Safety 92*, 575–584.

[4] Jones, M. C. and Henderson, D. A. (2008). Maximum likelihood kernel density estimation: on the potential of convolution sieves. Technical Report 08/09, The Open University, Milton Keynes, UK.

[5] Lindqvist, B. H. (2006). On the statistical modelling and analysis of repairable systems. *Statistical Science 21*, 532-551.

[6] Lindqvist, B. H., Elvebakk, G. and Heggland, K. (2003). The trend-renewal process for statistical analysis of repairable systems. *Technometrics 45,* 31-44. Reliability.

[7] Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data.* New York: Wiley.

[8] Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *The Annals of Statistics 11,* 453-466.