

On the Statistical Modelling and Analysis of Repairable Systems

Bo Henry Lindqvist

Department of Mathematical Sciences, Norwegian University of
Science and Technology, N-7491, Norway
bo@math.ntnu.no

17th April 2006

Abstract

The present paper reviews basic modelling approaches for failure and maintenance data from repairable systems. These are systems which are repaired rather than replaced after a failure. In particular we consider imperfect repair models defined in terms of virtual age processes, and the trend-renewal process which extends the nonhomogeneous Poisson process and the renewal process. In the case where several systems of the same kind are observed, we show how observed covariates and unobserved heterogeneity can be included in the models. We also consider various approaches to trend testing, a topic which traditionally has received much attention in reliability theory. Modern reliability databases usually contain more information than just the failure times. For example, there may be information on the identity of failed components, the type of failure, the type of maintenance, etc., in addition to values of observed covariates. We introduce a framework where the observed events are modelled as marked point processes with marks labelling the types of events. Some recent literature in this direction is reviewed, where the aim is to extend the theory of repairable systems to a competing risks setting in time. Throughout the paper the emphasis is more on modelling than on statistical inference.

Key words and phrases: Repairable system, preventive maintenance, nonhomogeneous Poisson process, renewal process, marked point process, virtual age process, trend-renewal process, heterogeneity, trend, competing risks.

1 Introduction

According to a commonly used definition of a repairable system (Ascher and Feingold [5]), this is a system which, after failing to perform one or more of its functions satisfactorily, can be restored to fully satisfactory performance by any method other than replacement of the entire system. For the present paper and following recent literature on the subject, we suggest to extend this definition to include the possibility of additional maintenance actions which aim at servicing the system for better performance. We shall refer to this as preventive maintenance (PM), where one may further distinguish between condition based PM and planned PM. The former type of maintenance is due when the system exhibits inferior performance while the latter is performed at predetermined points in time. In this presentation we will concentrate on condition based PM, while the planned PM will be briefly touched in the concluding remarks.

Traditionally, the literature on repairable systems is concerned with modelling of the failure times, with point process theory being the main tool. The most commonly used models for the failure process of a repairable system are renewal processes (RP), including the homogeneous Poisson processes (HPP), and nonhomogeneous Poisson processes (NHPP). While such models often are sufficient for simple reliability studies, the need for more complex models has of course emerged.

There is currently a rapidly increasing literature concerning modelling and analysis of recurrent events, with a wide range of applications, including reliability analysis of repairable systems which is the present topic. In a recent review paper, Cook and Lawless [14] present several examples from medical studies where models and methods for recurrent events are appropriate. The review paper Peña [55] gives examples from both medical and reliability studies. The scope of our paper is biased towards reliability applications although most of the models considered have a wider applicability in the analysis of recurrent events.

We will in particular consider models which incorporate effects of different kinds of repair and maintenance, and with the possibility of handling several failure causes, for example.

In a review paper like this it is of course impossible to cover all models or methods which have been suggested in the literature. Our aim is rather to emphasize some important ideas, and in this respect there will be a clear bias towards work in the direction of our own interests and in work by ourselves and collaborators. Throughout the paper the emphasis will be more on modelling than on statistical inference. In addition we will try to give some historical perspectives on the theory and practice related to repairable systems, again not necessarily complete and possibly biased by our own views.

One of the first comprehensive treatments of statistical methods for recurrent events with reliability emphasis, is the talk by David R. Cox, read before the Royal Statistical Society in London in March 1955 and published in Cox [17]. Cox touched a large number of topics, most of them motivated from the clothing industry. Topics of particular importance for reliability applications were trend testing, testing whether a failure process is a Poisson process, autocorrelated time gaps, doubly stochastic Poisson processes, heterogeneity between systems, correlations between different types of events, mean repair times, availability of service, etc. Many results from the paper are contained in the subsequent book by Cox and Lewis [19], which still is a very useful and much cited source on the subject.

Another early contribution to the study of repairable systems is the heavily cited 1963 paper by Frank Proschan [57], “Theoretical explanation of observed decreasing failure rate”. This paper is particularly important since it led to the awareness that proper analysis of recurrent events is an important part of reliability theory. In particular it is one of the first treatments of heterogeneity in the theory of repairable systems.

What seems to be the first book devoted solely to repairable systems reliability was published by Ascher and Feingold [5] in 1984. For a long time this was the main reference

for repairable systems and it is still a major source. The subtitle of the book is “Modeling, inference, misconceptions and their causes”. The authors were complaining that reliability researchers and practitioners did not recognize the crucial difference between the statistical treatment of repairable systems and non-repairable components. They demonstrated by simple examples how conclusions from data may be very wrong if times between failures are treated as i.i.d. if there is a trend in them.

Data from repairable systems are usually given as ordered failure times T_1, T_2, \dots with data coming from a single system or from several systems of the same kind. The implicit assumption is usually that the system is repaired and put into new operation immediately after the failure. This restriction, disregarding repair times, is not serious if one is interested in modeling and estimation of the probability mechanisms behind failure occurrences. It is, moreover, justified if time scale is taken to be operation time, number of cycles, number of kilometers run, etc. We will impose this restriction in this paper, and we will therefore not cover important topics such as availability and unavailability of systems, where the standard tool is to use alternating renewal processes with operation periods alternating with repair periods (see e.g. Rausand and Høyland [59, Ch. 7]).

A common recipe for analysis of data from a repairable system is as follows. First, apply a test for trend in the inter-failure times $X_i \equiv T_i - T_{i-1}$. If no significant trend is found, then use an RP as a model, in which case the well established statistical tools for analysis of i.i.d. observations can be used. Otherwise, use an NHPP model, which handles trend through specification of an intensity function $\lambda(t)$. For example, a deteriorating system will then correspond to an increasing function $\lambda(t)$, while an improving system will correspond to a decreasing $\lambda(t)$. A homogeneous Poisson process, HPP(λ), corresponds to a constant intensity $\lambda(t) \equiv \lambda$ and is at the same time a renewal process with exponentially distributed inter-failure times.

An RP model is also called a perfect repair model, since the system is as good as new

after a failure. On the other hand, an NHPP model corresponds to what is called minimal repairs, meaning that the system after repair is only as good as it was immediately before the failure. Lindqvist, Elvebakk and Heggland [48] represent the problem of distinguishing between the two “extreme” kinds of repair as corresponding to the first “dimension” of a repairable system description in the form of a so called model cube (Figure 3). The second “dimension” is the appearance of trend or no trend in inter-failure times. This particular aspect of system behavior has traditionally received much attention in reliability theory and is resolved by considering trend tests. Finally, the third “dimension” corresponds to the existence of unobserved heterogeneity between systems. The last problem is of course relevant only when several systems of the same kind are observed. There is currently a large and increasing interest in the modelling of heterogeneity, usually known as “frailties” in the survival analysis literature. To some extent, heterogeneity may have been much overlooked in reliability studies, but there are important exceptions in the literature.

Several classes of models have in turn been suggested for cases not covered by the “extreme” models RP and NHPP. These include the so called imperfect repair models. The idea is that after a repair the “virtual” age of the unit is not necessarily reduced to 0, such as for a perfect repair, nor is the same as before the repair, such as for a minimal repair. Instead, the virtual age is reduced by a certain amount depending on the type of repair. We review the basic properties of such models, and we will see how the concept of virtual age can be generalized to more than one dimension.

Another class of alternatives to NHPP and RP models, which includes these models, is the so called Trend-Renewal Process (TRP). This model is a generalization of Berman’s modulated gamma process [9] and has been extensively studied in Lindqvist et al. [48]. In the present paper we will use TRP models and their extensions as our basic framework in order to illustrate some main issues on modelling and statistical analysis of data from repairable systems. The TRP is particularly suitable to illustrate the already mentioned

three “dimensions” of repairable systems.

Modern reliability databases usually contain more information than just the failure times. For example, there may be information on the times of preventive maintenance (PM), identity of failed component, type of failure, type of repair, cost of replacement, etc. Thus we shall more generally assume that observations from repairable systems are represented as marked point processes where the marks label the types of events. For example, the marks may be of two kinds, corresponding to the type of maintenance, repair or PM. We review some recent literature in this direction with the aim of extending theory of repairable systems to a competing risks setting.

In addition to information on types of events, the databases may contain covariates representing environmental conditions, measures of various forms of load and usage stress, etc. Such covariates could be constant, or are possibly varying with time. Regression models for repairable systems are useful for obtaining better understanding of the underlying failure and PM mechanisms, or for predicting the behavior of new items.

The outline of the paper is as follows. The basic notation and definitions used in this paper are given in Section 2, including the introduction of the marked point process setup. Section 3 reviews models for the case of failure data with a single type of events, with emphasis on virtual age models and trend-renewal processes. Section 4 is devoted to a discussion of unobserved heterogeneity in repairable systems data. The model cube for heterogeneous trend-renewal processes is considered in particular. In Section 5 we consider various approaches to trend testing, both for data coming from single systems or from several similar systems. The possible extension of virtual age models to the marked process case is considered in Section 6. This section is based on some recent papers on the subject. Some concluding remarks are given in Section 7, for example regarding topics not covered in the main text.

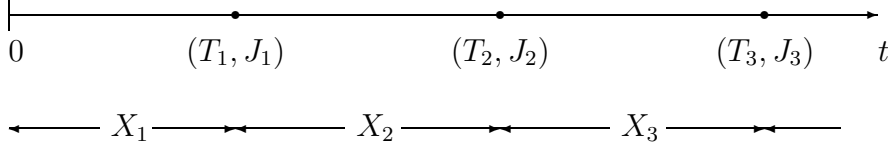


Figure 1: Event times (T_i) , event types (J_i) and sojourn times (X_i) of a maintained system.

2 Notation and Basic Definitions

We consider a repairable system where time usually runs from $t = 0$ and where events occur at ordered times T_1, T_2, \dots . Here time is not necessarily calendar time, but can in principle be any suitable measurement which is non-decreasing with calendar time, such as operation time, number of cycles, number of kilometers run, length of a crack etc. As already mentioned in the introduction, we shall disregard time durations of repair and maintenance, so we assume that the system is always restarted immediately after failure or a maintenance action. Types of events (type of maintenance, type of failure, etc.) are, when applicable, recorded as J_1, J_2, \dots with $J_i \in \mathcal{J}$ for some mark space \mathcal{J} which will depend on the current application. For simplicity we will here always assume that \mathcal{J} is a finite set. The observable process $(T_1, J_1), (T_2, J_2), \dots$ will be called the marked event process or occasionally the failure process. The inter-event, or inter-failure, times will be denoted X_1, X_2, \dots . Here $X_i = T_i - T_{i-1}$, $i = 1, 2, \dots$, where for convenience we define $T_0 \equiv 0$. Figure 1 illustrates the notation. We also make use of the counting process representation $N_j(t) = \text{number of events of type } j \text{ in } (0, t]$, which counts the number of events of type $j \in \mathcal{J}$, and $N(t) = \sum_{j \in \mathcal{J}} N_j(t)$, which counts the number of events irrespective of their types.

In order to describe probability models for repairable systems we use some notation from the theory of point processes. A key references is Andersen, Borgan, Gill and Keiding [4]. Let \mathcal{F}_{t-} dentote the history of the marked event process up to, but not including, time t . In

models without covariates we assume that \mathcal{F}_{t-} includes all information on event times and event types before time t . Formally, \mathcal{F}_{t-} is generated by the set $\{N_j(s) : 0 \leq s < t, j \in \mathcal{J}\}$.

Suppose then that a possibly time-dependent covariate vector $\mathbf{Z}(t)$ is observed for the system. In this case the covariate history $\{\mathbf{Z}(s) : 0 \leq s \leq t\}$ should be added to the history \mathcal{F}_{t-} for each $t > 0$. This will imply that just before any time t we have the complete information on the previous events, as well as the complete covariate history including the value of the covariate at time t . In the case of a time-constant covariate vector \mathbf{Z} , the information in \mathbf{Z} is added to each history \mathcal{F}_{t-} .

The conditional intensity of the process with respect to events of type $j \in \mathcal{J}$ is now defined as

$$\gamma_j(t) = \lim_{\Delta t \downarrow 0} \frac{Pr(\text{event of type } j \text{ in } [t, t + \Delta t) | \mathcal{F}_{t-})}{\Delta t}, \quad (1)$$

which we call the type-specific intensity for j . Thus, $\gamma_j(t)\Delta t$ is approximately the probability of an event of type j in the time interval $[t, t + \Delta t)$ given the history before time t . Further, we let $\gamma(t) = \sum_{j \in \mathcal{J}} \gamma_j(t)$ so that $\gamma(t)\Delta t$ is approximately the conditional probability of an event of any type in the time interval $[t, t + \Delta t)$. It has in this been tacitly assumed that the probability of more than one event in an interval $[t, t + \Delta t)$ is $o(\Delta t)$. Note that the $\gamma_j(\cdot)$ and hence the $\gamma(\cdot)$ may be functions of the covariate vector $\mathbf{Z}(\cdot)$ when appropriate. In typical applications, $\gamma_j(t)$ may depend on the covariate history only through the value $\mathbf{Z}(t)$ at time t . Further, it is common to assume that $\gamma_j(t) = \gamma_j^0(t)g(\mathbf{Z}(t))$, with $\gamma_j^0(t)$ depending only on the pure event history $\{N_j(s) : 0 \leq s < t, j \in \mathcal{J}\}$, and with $g(\cdot)$ being some parametric function of the covariate vector such as the exponential one, $g(\mathbf{z}) = \exp(\boldsymbol{\beta}'\mathbf{z})$, where $\boldsymbol{\beta}$ is a parameter vector.

For statistical inference we need an expression for the likelihood function. Suppose that a single system with a marked event process as described above is observed from time 0 to time τ , resulting in observations $(T_1, J_1), (T_2, J_2), \dots, (T_{N(\tau)}, J_{N(\tau)})$, in addition to the covariate vector $\mathbf{Z}(s)$ for $0 \leq s \leq \tau$ if applicable. The likelihood function is then given by

([4, Section II.7]),

$$L = \left\{ \prod_{i=1}^{N(\tau)} \gamma_{J_i}(T_i) \right\} \exp \left\{ - \int_0^\tau \gamma(u) du \right\}. \quad (2)$$

A rough verification of (2) can be given as follows. First, partition the interval $(0, \tau]$ into s equal pieces, each of length $h = \tau/s$. Assume that s is so large that at most one event can happen in an interval of length h . Then the conditional probability of an event of type j in the interval $[(k-1)h, kh)$, $(k = 1, \dots, s)$, given the history before $(k-1)h$, is roughly $\gamma_j(kh)h$, while the conditional probability of no event in this interval is roughly $1 - \gamma(kh)h$. The probability of a realization of the process from 0 to τ will therefore include a product of $N(\tau)$ terms of the type $\gamma_j(kh)h$, corresponding to the observed events, and which in the limit as $h \rightarrow 0$ (after dividing by the normalization $h^{N(\tau)}$) gives the product term on the right hand side of (2). The exponential part of (2) comes from taking the limit of the product of the terms $1 - \gamma(kh)h \approx \exp\{-\int_{(k-1)h}^{kh} \gamma(t)dt\}$ for the intervals containing no event, assuming continuity of $\gamma(\cdot)$.

The likelihood function (2) is valid under the assumption that τ is a stopping time, which means that its value depends stochastically only on the past history. This property holds for the standard censoring schemes used in practice and in particular when τ is independent of the event process. There is, however, an increasing awareness of the need to allow for dependent censoring in many applications (see for example Huang and Wang [33]).

In typical applications, data will be available for several similar systems, observed over possibly different time intervals. Under the assumption of stochastic independence and identical probability mechanisms for the systems, the total likelihood will be the product of expressions (2) computed for all systems. For both parametric and nonparametric models of this kind there is a well developed theory for estimation based on the martingale approach to point processes ([4] gives a comprehensive account). Relevant references for statistical

inference in reliability models are, among others, Ascher and Feingold [5], Rausand and Høyland [59], Crowder, Kimber, Smith and Sweeting [21] and Meeker and Escobar [52].

3 Models for Repairable Systems with a Single Type of Events

In the present section we assume that the observations are just the failure times T_1, T_2, \dots . Thus the mark space \mathcal{J} will be ignored.

A large number of models can be obtained in terms of a given hazard function $z(t)$, which we think of as being the hazard function of the time to first failure of a new system. The corresponding density and cumulative distribution functions are denoted, respectively, $f(t)$ and $F(t)$, so $z(t) = f(t)/(1 - F(t))$. The idea is to use the function $z(t)$ together with a specification of the repair strategy to define the conditional intensity function $\gamma(t)$ of the failure process. Models of this type are considered in Sections 3.1 and 3.2 below. The corresponding models may be extended to the case with observed covariates, although this will not be made explicit. As described in Section 2, the conditional intensities of the form $\gamma(t)$ as considered below, may be multiplied with a factor $g(\mathbf{Z}(t))$ defining the dependence of the covariate value at time t .

3.1 Perfect and Minimal Repair Models

Suppose first that after each failure, the system is repaired to a condition as good as new. In this case the failure process is modelled by a renewal process with inter-event time distribution F , denoted $\text{RP}(F)$. Clearly

$$\gamma(t) = z(t - T_{N(t-)})$$

where $t - T_{N(t-)}$ is the time since the last failure strictly before time t .

Suppose instead that after a failure, the system is repaired only to the state it had immediately before the failure, called a minimal repair. This means that the conditional

intensity of the failure process immediately after the failure is the same as it was immediately before the failure, and hence is exactly as it would be if no failure had ever occurred.

Thus we must have

$$\gamma(t) = z(t)$$

so that the process is an NHPP with intensity $z(t)$, denoted NHPP($z(\cdot)$). In practice a minimal repair usually corresponds to repairing or replacing only a minor part of the system.

3.2 Imperfect Repair Models and the Virtual Age of a System

A classical model, suggested by Brown and Proschan [13], assumes that at the time of each failure a perfect repair occurs with probability p and a minimal repair occurs with probability $1 - p$, independently of the previous failure history. This model is a simple example of what has been called imperfect repair models, and has later been generalized in several directions.

Kijima [34] suggested two imperfect repair models, both involving what is called the virtual age (or effective age) of the system. The idea is to distinguish between the system's age, which is the time elapsed since the system was new, usually at time $t = 0$, and the virtual age of the system which describes its present condition when compared to a new system. The virtual time is redefined at failures according to the type of repair performed, and runs along with the true time between repairs. More precisely, a system with virtual age $v \geq 0$ is assumed to behave exactly like a new system which has reached age v without having failed. The hazard rate of a system with virtual age v is thus $z_v(t) = z(v + t)$ for $t > 0$, where $z(\cdot)$ is the hazard rate of the time to first failure of the system.

It should be clear at this stage that models based on virtual ages make sense only if the underlying hazard functions $z(\cdot)$ are non-constant. In fact, if $z(\cdot)$ is constant, then a reduction of virtual age would not influence the rate of failures.

A variety of imperfect repair models can be obtained by specifying properties of the virtual age process in addition to the hazard function $z(t)$ of a new system. For this, suppose $v(i)$ is the virtual age of the system immediately after the i th event, $i = 1, 2, \dots$. The virtual age at time $t > 0$ is then defined by $A(t) = v(N(t-)) + t - T_{N(t-)}$, which is the sum of the virtual age after the last event before t and the time elapsed since the last event. The process $A(t)$, called the virtual age process by Last and Szekli [40], thus increases linearly between events and may jump only at events. It follows that

$$\gamma(t) = z_{v(N(t-))}(t - T_{N(t-)}) = z(A(t)) \quad (3)$$

assuming that $A(t)$ is included in \mathcal{F}_{t-} for all t . This means in turn that $v(i)$ is contained in \mathcal{F}_{T_i} for each t so that $v(i)$ depends on the history up to and including T_i . The likelihood then becomes

$$\begin{aligned} L &= \left\{ \prod_{i=1}^{N(\tau)} z(v(i-1) + X_i) \right\} \exp \left\{ - \sum_{i=1}^{N(\tau)} \int_0^{X_i} z(v(i-1) + u) du \right. \\ &\quad \left. - \int_0^{\tau - T_{N(\tau)}} z(v(N(\tau)) + u) du \right\}. \end{aligned}$$

This can be recognized as being the same as

$$\left\{ \prod_{i=1}^{N(\tau)} f_{v(i-1)}(X_i) \right\} \{1 - F_{v(N(\tau))}(\tau - T_{N(\tau)})\}$$

where $f_v(t) = f(v+t)/(1-F(v))$ and $F_v(\cdot)$ are the density and cumulative distribution function, respectively, of time to next failure for a system with virtual age v and hence with hazard rate $z_v(\cdot)$.

It is clear that the perfect repair and minimal repair models are the special cases where, respectively, $v(i) = 0$ and $v(i) = T_i$, $i = 1, 2, \dots$. In Kijima's [34] Model I, the virtual age $v(i)$ equals $\sum_{k=1}^i D_k X_k$, where D_1, D_2, \dots is a sequence of random variables on the interval $[0, 1]$ such that D_k is independent of \mathcal{F}_{T_k-} for each k . Note that \mathcal{F}_{T_k-} includes

the D_1, D_2, \dots, D_{k-1} so that in particular the D_k are independent. In Kijima's Model II the virtual age $v(i)$ is set to $\sum_{k=1}^i (\prod_{j=k}^i D_j) X_k$ with the same conditions for the D_k . This means that the virtual age after the i th failure equals D_i multiplied by the virtual age of the system just prior to the i th failure. The model by Brown and Proschan [13] is obtained when D_i is 1 with probability $1 - p$ and 0 with probability p for all i .

Dorado, Hollander and Sethuraman [22] studied nonparametric statistical inference in a model slightly more general than Kijima's models described above. Nonparametric statistical inference in the Brown-Proschan model was first studied by Whitaker and Samaniego [63] and later by Hollander, Presnell and Sethuraman [31].

Recall that for the above models, the D_i need to be observed for likelihood inference using (2) to be valid. This means in effect that the type of repair (minimal or perfect) must be reported for each repair action. In real applications, however, exact information on the type of repair is rarely available. The estimation problem in the case of unobserved D_i has been considered by, for example, Lim [45] (suggesting an EM-algorithm approach) and Langseth and Lindqvist [38, 39].

Doyen and Gaudoin [23] study classes of virtual age models based on deterministic reduction of virtual age due to repairs, and hence not requiring the observation of repair characteristics. The basic models of this type can be obtained simply by letting the D_i in Kijima's models above be replaced by parametric functions such as $1 - \rho$ where $0 < \rho < 1$ is an unknown parameter to be estimated.

There is a large literature on reliability modeling using the virtual age process. For a review we refer to Pham and Wang [58] and Lindqvist [46]. Section 6 presents an attempt to define a multivariate virtual age process and corresponding repairable system models with several types of events.

3.3 Generalized Linear Model Types

Berman and Turner [10] considered estimation in parametric models with the conditional intensity being of the generalized linear model type

$$\gamma(t) = g \left\{ \sum_{i=0}^p \beta_i z_i(t) \right\}, \quad (4)$$

where g is a known monotonic continuous function, the $z_i(t)$ are known functions of t and the history \mathcal{F}_{t-} , and the β_i are unknown parameters. Note that the functions $z_i(t)$ may therefore be functions of the covariates if available. One aim of the paper was to demonstrate how to use software for generalized linear models to analyze repairable systems data.

The model (4) is closely related to the modulated renewal process introduced in Cox [18] for which Cox suggested a semiparametric approach using a partial likelihood.

The special case of (4) obtained when $g(y) = e^y$ was applied to repairable systems by Lawless and Thiagarajah [43]. In particular they considered the model

$$\gamma(t) = e^{\beta_0 + \beta_1 g_1(t) + \beta_2 g_2(t - T_{N(t-)})} \quad (5)$$

where g_1 and g_2 are known functions. This conditional intensity is a function of both the calendar time and the time since last failure. Note that $\beta_1 = 0$ gives an RP, $\beta_2 = 0$ gives an NHPP, while $\beta_1 = \beta_2 = 0$ gives an HPP.

3.4 The Trend-Renewal Process

A class of processes called inhomogeneous gamma processes was suggested by Berman [9]. Berman motivated the inhomogeneous gamma process by first considering the process T_1, T_2, \dots obtained by observing every κ th event of an NHPP, where κ is a positive integer. He then showed how to generalize to the case when κ is any positive number.

We present now a generalization of Berman's idea, called the trend-renewal process,

which was extensively studied in Lindqvist et al. [48]. We will use this process in particular to describe the “three dimensions” related to the properties of repairable systems.

The idea behind the trend-renewal process is to generalize the following well known property of the NHPP. First let the cumulative intensity function corresponding to an intensity $\lambda(\cdot)$ be defined by $\Lambda(t) = \int_0^t \lambda(u)du$. Then if T_1, T_2, \dots is an $\text{NHPP}(\lambda(\cdot))$, the time-transformed stochastic process $\Lambda(T_1), \Lambda(T_2), \dots$ is $\text{HPP}(1)$.

The trend-renewal process (TRP) is defined simply by allowing the above $\text{HPP}(1)$ to be any renewal process $\text{RP}(F)$. Thus, in addition to the intensity function $\lambda(t)$, for a TRP we need to specify a distribution function F of the inter-arrival times of this renewal process. Formally we can define the process $\text{TRP}(F, \lambda(\cdot))$ as follows:

Let $\lambda(t)$ be a nonnegative function defined for $t \geq 0$, and let $\Lambda(t) = \int_0^t \lambda(u)du$. The process T_1, T_2, \dots is called $\text{TRP}(F, \lambda(\cdot))$ if the transformed process $\Lambda(T_1), \Lambda(T_2), \dots$ is $\text{RP}(F)$, that is if the $\Lambda(T_i) - \Lambda(T_{i-1})$; $i = 1, 2, \dots$ are i.i.d. with distribution function F . The function $\lambda(\cdot)$ is called the trend function, while F is called the renewal distribution. In order to have uniqueness of the model it is usually assumed that F has expected value 1.

Figure 2 illustrates the definition. For an $\text{NHPP}(\lambda(\cdot))$, the $\text{RP}(F)$ would be $\text{HPP}(1)$. Thus $\text{TRP}(1 - e^{-x}, \lambda(\cdot)) = \text{NHPP}(\lambda(\cdot))$. Also, $\text{TRP}(F, 1) = \text{RP}(F)$, which shows that the TRP class includes both the RP and NHPP classes.

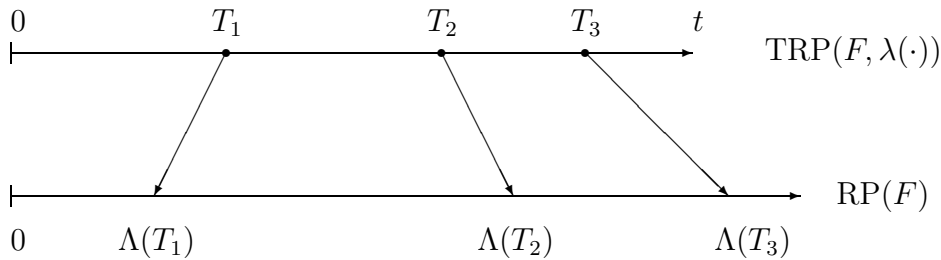


Figure 2: The defining property of the trend-renewal process.

As a motivation for the TRP model, suppose that failures of a particular system corre-

spond to replacement of a major part, for example the engine of a tractor (as in the data given by Barlow and Davis [6]), while the rest of the system is not maintained. Then if the rest of the system is not subjected to wear, a renewal process would be a plausible model for the observed failure process. In the presence of wear, on the other hand, an increased replacement frequency is to be expected. This is achieved in a TRP model by accelerating the internal time of the renewal process according to a time transformation $\Lambda(t) = \int_0^t \lambda(u)du$ which represents the cumulative wear. The TRP model thus has some similarities to accelerated failure time models.

It can be shown [48] that the conditional intensity function for the $\text{TRP}(F, \lambda(\cdot))$ is

$$\gamma(t) = z(\Lambda(t) - \Lambda(T_{N(t-)}))\lambda(t) \quad (6)$$

where $z(\cdot)$ is the hazard rate corresponding to F . This is a product of one factor, $\lambda(t)$, which depends on the age t of the system and one factor which depends on a transformed time from the last previous failure. However, time since last failure is measured on a scale depending on the current intensity of failures. This shows that the TRP class does not contain, nor is contained in, the classes of processes considered in the previous subsection.

Suppose now that a single system has been observed in $[0, \tau]$, with failures at $T_1, T_2, \dots, T_{N(\tau)}$. If a $\text{TRP}(F, \lambda(\cdot))$ is used as a model, then substitution of (6) into (2) gives the likelihood

$$L = \left\{ \prod_{i=1}^{N(\tau)} z[\Lambda(T_i) - \Lambda(T_{i-1})]\lambda(T_i) \right\} \exp\left\{ - \int_0^\tau z[\Lambda(u) - \Lambda(T_{N(u-)})]\lambda(u)du \right\}. \quad (7)$$

Equivalently, if f is the density function corresponding to F , we can write this likelihood as

$$L = \left\{ \prod_{i=1}^{N(\tau)} f[\Lambda(T_i) - \Lambda(T_{i-1})]\lambda(T_i) \right\} \{1 - F[\Lambda(\tau) - \Lambda(T_{N(\tau)})]\}. \quad (8)$$

The latter form of the likelihood follows directly from the definition of the TRP, since the conditional density of T_i given $T_1 = t_1, \dots, T_{i-1} = t_{i-1}$ is $f[\Lambda(t_i) - \Lambda(t_{i-1})]\lambda(t_i)$, and the probability of no failures in the time interval $(T_{N(\tau)}, \tau]$, given $T_1, \dots, T_{N(\tau)}$, is $1 - F[\Lambda(\tau) - \Lambda(T_{N(\tau)})]$.

A possible extension of the TRP to include covariates would be to multiply the trend function $\lambda(t)$ by a factor $g(\mathbf{Z}(t))$, for example of the form $\exp(\boldsymbol{\beta}'\mathbf{Z}(t))$ as suggested in Section 2. The $\lambda(t)$ would then play the role of a baseline trend function. This definition generalizes in a natural way the commonly used NHPP model with covariates, see for example Lawless [41].

4 Unobserved Heterogeneity in Repairable Systems

Analyses of reliability data often lead to an apparent decreasing failure rate which could be counterintuitive in view of wear and ageing effects. Proschan [57] pointed out that such observed decreasing rates could be caused by unobserved heterogeneity. Proschan presented failure data from 17 air conditioner systems on Boeing 720 airplanes. Applying Mann's [51] nonparametric trend test to each system and then combining to a global test statistic he argued that there is no significant trend in the failure times for each separate plane. He then showed by a similar test that "it seems safe to accept the exponential distribution as describing the failure interval, although to each plane may correspond a different failure rate". He demonstrated this last fact statistically by using a result from Barlow, Marshall and Proschan [7] which implies that a mixture of exponential distributions has a decreasing failure rate. More precisely, he used again a Mann type test which was obtained by pooling the inter-failure times from all the planes and which is sensitive to a decreasing failure rate. In this way he obtained a P-value of 0.007 for the null hypothesis of identical exponential distributions of the inter-failure times.

Heterogeneity in connection with Poisson-processes was in fact studied as early as 1920 by Greenwood and Yule [27], who used a compound Poisson distribution. Later, Maguire et al. [50], studying occurrences of industrial accidents, showed how Laplace transforms enter general expressions for resulting distributions of intervals and counts. Cox [17] considered the possibility of heterogeneity, which he called variance components,

between homogeneous Poisson-processes and listed several reasons for the interest in such models for repairable systems data.

It has similarly long been known in biostatistics that neglecting individual heterogeneity may lead to severe bias in estimates of lifetime distributions. The idea is that individuals or components have different “frailties”, and that those who are most “frail” will die or fail earlier than the others. This in turn leads to a decreasing population hazard, which has often been misinterpreted. Important references on heterogeneity in the biostatistics literature are Vaupel et al. [62], Hougaard [32] and Aalen [2]. It should be noted that heterogeneity is in general unidentifiable if being considered an individual quantity. For identifiability it is necessary that frailty is common to several individuals, for example in family studies in biostatistics, or if several events are observed for each individual, such as for the repairable systems considered in this paper and more generally for recurrent events data. The presence of heterogeneity is often apparent for data from repairable systems if there is a large variation in the number of events per system. However, it is not really possible to distinguish between heterogeneity and dependence of the intensity on past events for a single process. It is a fact, though, that ignorance of an existing heterogeneity may lead to sub-optimal or even wrong decisions.

4.1 Modelling of Heterogeneity for Repairable Systems

The common way of modelling heterogeneity is to include an unobservable multiplicative constant in the conditional intensity of the process, see for example Vaupel et al. [62]. For systems with a single type of event this is done by first replacing the conditional intensities $\gamma(t)$ in (1) by $a\gamma(t)$ where a is a random variable representing the “frailty” of the system and such that a is included in \mathcal{F}_{t-} for each t . Note that $\gamma(t)$ as described in Section 2 may well be a function of covariates. Now a can be viewed as being the effect of an unobserved covariate. Systems with a large value of a will have a larger failure proneness

than a system with a low value of a . Intuitively, the variation in the a between systems implies that the variation in observed number of failures among the systems is larger than would be expected if the failure processes were identically distributed. Now, since a is unobservable, one needs to take the expectation of the likelihood resulting from (2) with respect to the distribution of a in order to have a likelihood function for the observed data.

In the marked point process formulation of Section 2 we may more generally assume that there are different frailty variables for each event type $j \in \mathcal{J}$. More precisely, we assume that there is a random vector $\mathbf{a} = (a_j, j \in \mathcal{J})$ such that the type-specific intensities for given \mathbf{a} are $a_j \gamma_j(t)$, respectively, where $\gamma_j(t)$ corresponds to the type specific conditional intensity defined in Section 2. The resulting likelihood including heterogeneity is thus

$$L = E_{\mathbf{a}} \left[\left(\prod_{i=1}^{N(\tau)} a_{J_i} \gamma_{J_i}(T_i) \right) \exp \left\{ - \sum_{j \in \mathcal{J}} a_j \int_0^\tau \gamma_j(u) du \right\} \right] \quad (9)$$

where the expected value is taken with respect to the joint distribution of \mathbf{a} . Multivariate frailty distributions are considered by, for example, Hougaard [32] and Aalen [1].

In the case of several independent systems, it is assumed that the \mathbf{a} corresponding to each system are i.i.d. from the given joint distribution. The total likelihood is then the product of factors (9), one for each system. Note that for identifiability it may be necessary to introduce a normalization of \mathbf{a} , for example assuming that $E(\|\mathbf{a}\|) = 1$. This is because otherwise a scale factor may be moved from a_j to $\gamma_j(\cdot)$ or vice versa without changing the value of (9). Alternatively one may let the a_j act as free random scale parameters in the model if the $\gamma_j(\cdot)$ themselves do not include scale parameters.

For the special case of single type of event one obtains the following simplification of the likelihood function in (9),

$$L = E_a \left[a^{N(\tau)} \left(\prod_{i=1}^{N(\tau)} \gamma(T_i) \right) \exp \left\{ -a \int_0^\tau \gamma(u) du \right\} \right], \quad (10)$$

where the expectation is with respect to the distribution of the random variable a and

where for normalization one will usually assume $E(a) = 1$.

The expression (10) suggests that a gamma-distribution for a is mathematically convenient, since a closed form expression of the likelihood is obtained. More generally, for the version (9), a multivariate gamma-distribution for \mathbf{a} leads to a simplified expression (see for example [1, 32] regarding multivariate gamma-distributions).

Consider now the likelihood (10), and suppose that a is gamma-distributed with $E(a) = 1$, $Var(a) = \delta$. Then a straightforward computation gives

$$\begin{aligned} L &= \left\{ \prod_{i=1}^{N(\tau)} \gamma(T_i) \right\} \frac{\Gamma(N(\tau) + 1/\delta)}{\delta^{1/\delta} \Gamma(1/\delta) [1/\delta + \int_0^\tau \gamma(u) du]^{N(\tau)+1/\delta}} \\ &= \left\{ \prod_{i=1}^{N(\tau)} \gamma(T_i) \right\} \frac{[\delta(N(\tau) - 1) + 1][\delta(N(\tau) - 2) + 1] \cdots 1}{[\delta \int_0^\tau \gamma(u) du + 1]^{N(\tau)+1/\delta}} \end{aligned} \quad (11)$$

where we have used the fact that $\Gamma(r + 1) = r\Gamma(r)$. Recall that $\gamma(T_i)$ may well include covariates. This likelihood expression is applicable, for example, together with the virtual age model (3) and the generalized linear model types (4) and (5). It is also the likelihood function for NHPPs with heterogeneity and possibly covariates, as studied in Lawless [41], and results in the likelihood of the so called compound power law model studied by Engelhardt and Bain [26].

We remark that (11) converges to (2) (assuming a single type of event) as $\delta \rightarrow 0$.

4.2 Heterogeneity in the TRP Model, the HTRP Model

Lindqvist et al [48] introduced heterogeneity into the TRP model by including an unobservable random multiplicative constant a in the trend function $\lambda(t)$, thus considering the conditional model $\text{TRP}(F, a\lambda(\cdot))$ with a renewal distribution F not depending on a . This definition is consistent with the regression version of TRP as suggested at the end of Section 3.4. Now the a replaces the function $g(\mathbf{Z}(t))$ used there. Note that in practice one might need to include both the frailty a and a covariate factor $g(\mathbf{Z}(t))$. In order to simplify

the discussion, we will, however, not consider covariates in our presentation below. Considering (6) it is seen, however, that the conditional intensity function given a is no longer of the simple multiplicative form $a\gamma(t)$ which was assumed in the previous subsection. This is because the $\Lambda(\cdot)$ in (6) is also multiplied by a .

Instead of the expression (10), the relevant likelihood from one system becomes, using (7),

$$L = E_a \left[\left\{ \prod_{i=1}^{N(\tau)} z[a(\Lambda(T_i) - \Lambda(T_{i-1}))]a\lambda(T_i) \right\} \exp \left\{ -a \int_0^\tau z[a(\Lambda(u) - \Lambda(T_{N(u-)}))] \lambda(u) du \right\} \right] \quad (12)$$

or, using (8),

$$L = E_a \left\{ \prod_{i=1}^{N(\tau)} f[a(\Lambda(T_i) - \Lambda(T_{i-1}))]a\lambda(T_i) \right\} \{1 - F[a(\Lambda(\tau) - \Lambda(T_{N(\tau)}))]\}. \quad (13)$$

Here f and z are respectively, as before, the density and hazard function of the distribution F .

The expressions (12) and (13) appear to be less tractable than the expression (10). Lindqvist et al. [48] obtain, however, a rather simple expression for the likelihood in the case of an inhomogeneous gamma process with gamma distributed heterogeneity factor a . In this case the last factor of (13) disappears, and letting F be the gamma-distribution with unit expectation and variance γ , while a is gamma-distributed with unit expectation and variance δ , one obtains

$$\begin{aligned} L &= \left\{ \prod_{i=1}^{N(\tau)} (\Lambda(T_i) - \Lambda(T_{i-1}))^{1/\gamma-1} \lambda(T_i) \right\} \\ &\quad \times \frac{\Gamma(N(\tau)/\gamma + 1/\delta)}{\gamma^{N(\tau)/\gamma} [\Gamma(1/\gamma)]^{N(\tau)} \delta^{1/\delta} \Gamma(1/\delta) [1/\delta + (1/\gamma)\Lambda(T_{N(\tau)})]^{N(\tau)/\gamma + 1/\delta}}. \end{aligned}$$

Note that for $\gamma = 1$ this is of the same form as in (11).

More generally, we use the notation $\text{HTRP}(F, \lambda(\cdot), H)$ for the model with likelihood (12) or, equivalently, (13). The 'H' in 'HTRP' here stands for 'heterogeneity', and the H

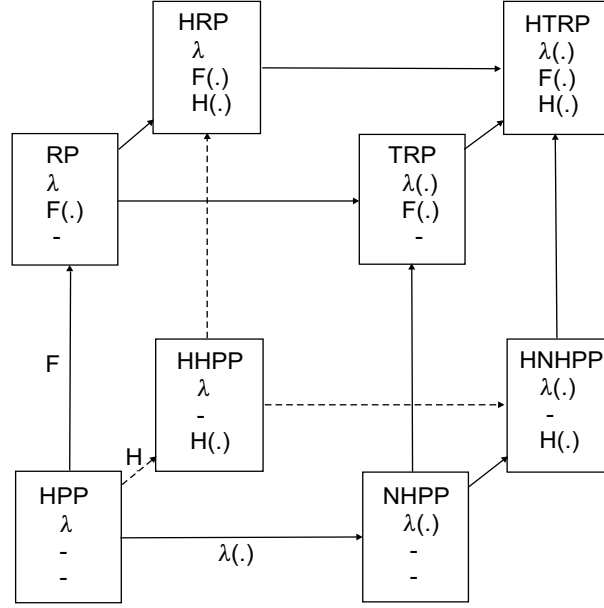


Figure 3: The model cube illustrating the $\text{HTRP}(F, \lambda(\cdot), H)$ and the submodels obtained by restricting one or more of $F, \lambda(\cdot), H$ to their basic versions, respectively F being standard exponential (using notation ‘-’ in the figure), $\lambda(t) \equiv \lambda$ being constant in time, and H being the distribution deterministic at 1 (‘-’ in the figure).

which is added to $(F, \lambda(\cdot))$ in the notation, is the distribution of the variable a , which could be any positive distribution with expected value 1.

4.3 The Three Dimensions of a Repairable System Description: The Model Cube and the log-likelihood Cube

A useful feature of the HTRP model is that several models for repairable systems can be represented as sub-models. With the notation HPP, NHPP, RP and TRP used as before, and with an H in front meaning the model which includes heterogeneity, Figure 3 shows how the HTRP and the seven sub-models can be represented in a cube. Each vertex of the cube represents a model, and the lines connecting them correspond to changing one of the three “coordinates” $(F, \lambda(\cdot), H)$ in the HTRP-notation. Going to the right corresponds to introducing a time trend, going upwards corresponds to entering a non-Poisson (renewal) case, and going backwards (inwards) corresponds to introducing heterogeneity.

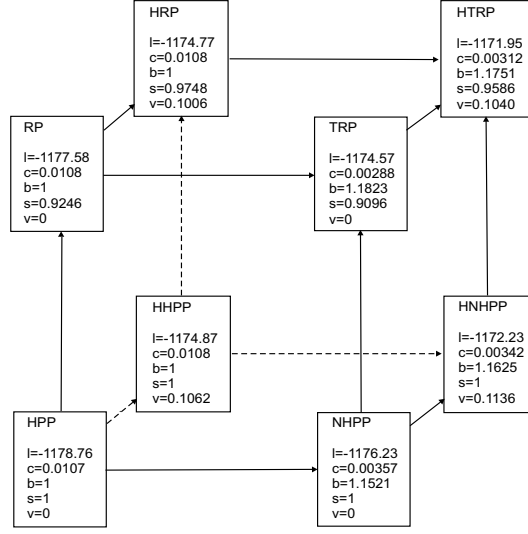


Figure 4: The log-likelihood cube for the data of Proschan [57] concerning failures of airconditioner systems on airplanes, fitted with a parametric $\text{HTRP}(F, \lambda(\cdot), H)$ model and its sub-models. Here F is a Weibull-distribution with expected value 1 and shape parameter s , $\lambda(t) = cbt^{b-1}$ is a power function of t , and H is a gamma-distribution with expected value 1 and variance v . The maximum value of the log likelihood is denoted l .

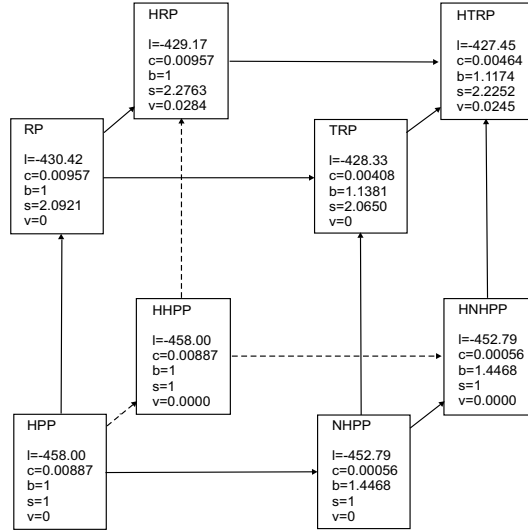


Figure 5: The log-likelihood cube for the data of Aalen and Husebye [3] concerning migratory motor complex periods, fitted with a parametric $\text{HTRP}(F, \lambda(\cdot), H)$ model and its sub-models. Here F is a Weibull-distribution with expected value 1 and shape parameter s , $\lambda(t) = cbt^{b-1}$ is a power function of t , and H is a gamma-distribution with expected value 1 and variance v . The maximum value of the log likelihood is denoted l .

In analyzing data by parametric HTRP models we may use the cube to facilitate the presentation of maximum log-likelihood values and parameter estimates for the different models in a convenient, visual manner which may guide in model choice (see Lindqvist et al. [48]). Figure 4 and Figure 5 show maximum likelihood values computed from the data of Proschan [57] and Aalen and Husebye [3], respectively. The latter data set is taken from a medical study and is included here in order to demonstrate results for data which are clearly non-Poisson distributed.

For the Proschan data we conclude that the renewal distribution can be taken to be exponential, leaving us with the bottom face of the cube. Further, when comparing the front face to the back face there is clear reason to conclude that there is heterogeneity between the systems, with $Var(a)$ being estimated to approximately 0.11. The conclusions so far are thus in accordance with the conclusions of Proschan [57]. However, a comparison of the left and right faces of the cube reveals a slight time trend. In fact, twice the log-likelihood difference from HHPP to HNHPP amounts to 5.24, giving a p-value of 0.022 assuming a chi-square distribution with one degree of freedom of the corresponding likelihood ratio test statistic. The power parameter b of the trend function is, furthermore, estimated to 1.16.

The most obvious conclusion for the Aalen and Husebye [3] data is that the renewal distribution is not exponential, implying that the upper face of the cube applies. Further, the differences in log likelihood obtained by introducing heterogeneity are seen to be small enough to conclude no significant heterogeneity. However, as for the Proschan data, there seems to be a slight time trend. Here, twice the log-likelihood difference from RP to TRP amounts to 4.18, giving a p-value of 0.041, while the power parameter b is estimated to 1.14 for the TRP model. Note the large difference in log likelihood value between, for example, the TRP and NHPP models, for example. As shown by the parameter estimates (Figure 4), the NHPP estimates seem to compensate the large estimated shape parameter

for the renewal distribution of the TRP, by increasing the power parameter b of the trend function (from 1.14 to 1.45). It is also seen that for the Poisson models (bottom face) there is no gain in introducing heterogeneity. In fact, the maximum likelihood estimates of the heterogeneity variance v are given by the border value 0. This is so since the profile likelihood of v can be shown to be a decreasing function of $v > 0$ near 0 (see [48] for a further discussion of this effect).

5 Trend Testing

In applications involving repairable systems, the main aim is often to detect trends in the pattern of failures occurring over time. These may be revealed as monotonic trends in the inter-failure times, either corresponding to an improving system or a deteriorating system. Various types of non-monotonic trends may also be present, for example a cyclic trend or a bathtub shaped trend.

5.1 Graphical methods

A simple but informative way of checking for a possible trend in the pattern of failures is to study plots like Figure 6, plotting cumulative failure number versus failure time for a single system. The present data are failures of the airconditioner system of airplane 7913 of the Proschan [57] data. A convex plot would be indicative of a deteriorating system, while a concave plot would indicate an improving system. In Figure 6 there seems to be no significant deviation from a straight line, however, thus indicating no trend in inter-failure times.

5.1.1 Nelson-Aalen plot

The plot of Figure 6 is a special case of the Nelson-Aalen plot to be described next. Assume that m systems are observed, with the individual failure processes being independent and

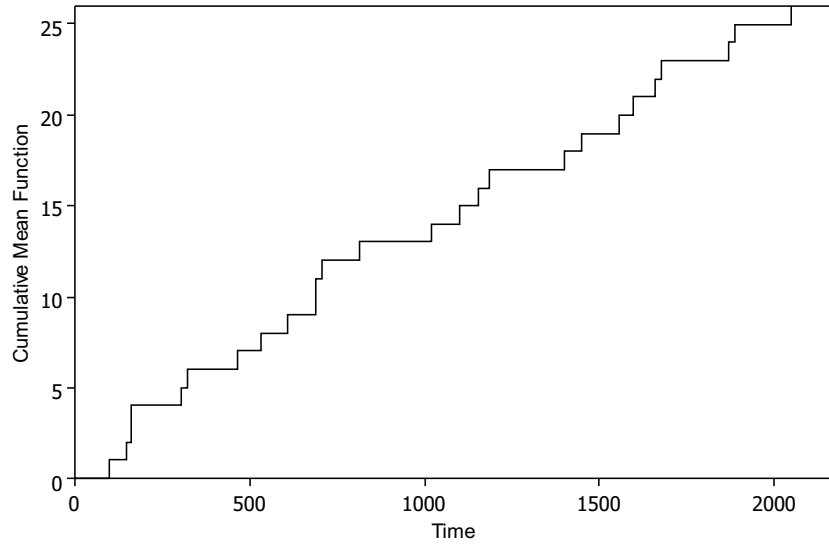


Figure 6: Plot of cumulative number of failures, $N(t)$, for airconditioner failures of plane 7913 in the Proschan [57] data.

identically distributed. Suppose further that the i th process is observed on the time interval $(0, \tau_i]$ and let $y(t)$ denote the number of processes under observation at time t . Note that $y(t)$ is a function of the τ_i and not of the failure times. Let T_k denote the k th arrival time in the superposed process, that is T_i is a failure time in one of the processes and $0 < T_1 \leq T_2 \leq \dots \leq T_N \leq \tau$ where $\tau = \max\{\tau_i : i = 1, \dots, m\}$. Define the cumulative mean function of a single process to be $M(t) = E(N(t))$. The Nelson-Aalen estimator of $M(t)$ is given by

$$\hat{M}(t) = \sum_{T_k \leq t} \frac{1}{y(T_k)}$$

where the sum is taken over all failure times T_k before or at time t . Figure 7 shows the plot of $\hat{M}(t)$ for the data on times of valve-seat replacements in a fleet of $m = 41$ diesel engines, taken from Nelson [53]. The plot indicates that the replacement frequency is fairly constant up to 550 days, and then increases as revealed from the convex shape of the curve at the right end.

The plot as defined here is studied, for example, in Nelson [53] and Lawless and Nadeau

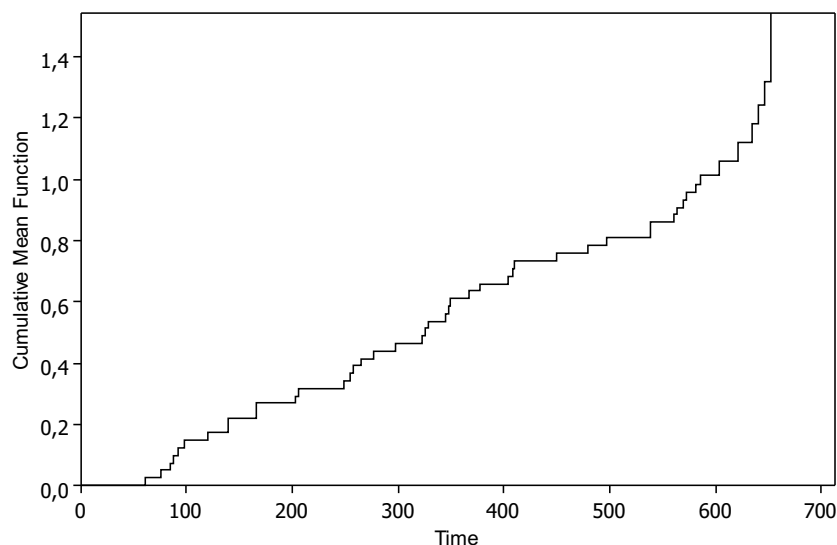


Figure 7: Nelson-Aalen plot of the estimated cumulative mean function $\hat{M}(t)$ for the valve-seat replacement data as given by Nelson [53].

[42]. These papers also derive robust non-parametric estimates of the variance of $\hat{M}(t)$, valid under any distributional properties of the individual processes $N(t)$.

5.1.2 TTT plot

Consider the special case of the above where the m processes are independent NHPPs with a common intensity function $\lambda(t)$. The superposed process is now an NHPP with intensity function $\phi(t) = \lambda(t)y(t)$, and hence (see Section 3.4) the process $\int_0^{T_1} \phi(u)du, \int_0^{T_2} \phi(u)du, \dots$ is HPP(1) on $(0, \tau)$. Define the total time on test, TTT, at time t by

$$r(t) = \int_0^t y(u)du.$$

Barlow and Davis [6] introduced the TTT plot for repairable systems data as a plot of the points

$$\left(\frac{i}{N}, \frac{r(T_i)}{r(\tau)} \right), \quad i = 1, \dots, N.$$

The idea is that if $\lambda(t)$ is a constant, so that the processes are HPP, then the $r(T_i)/r(\tau)$, $i = 1, \dots, N$, form an HPP(1) on $[0, 1]$. In this case the TTT plot is by its definition expected to

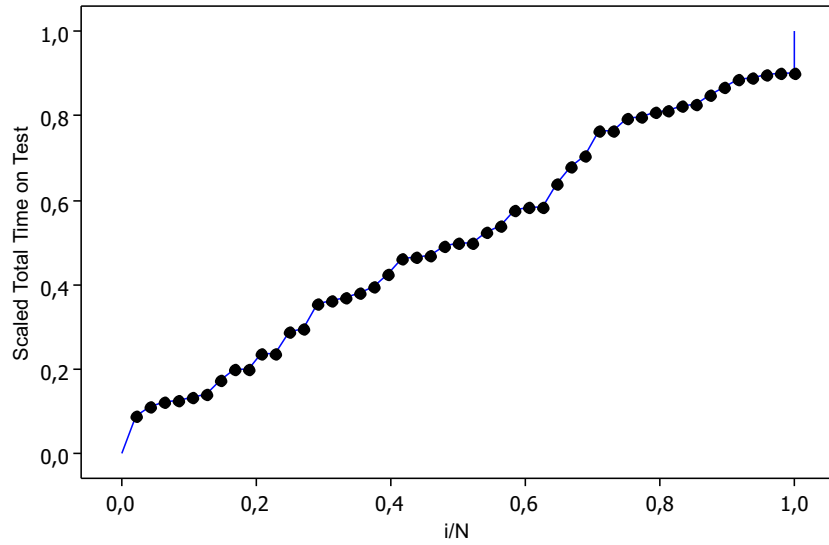


Figure 8: TTT plot of valve seat data as given in Nelson [53].

be located near the main diagonal of the unit square. Under the alternatives of decreasing, increasing and bathtub shaped intensity $\lambda(t)$, on the other hand, the TTT plots appear to be, respectively, convex, concave and S-shaped. Figure 8 shows the TTT plot of the valve-seat replacement data of Nelson [53]. The plot appears to be fairly straight, but with a slightly concave shape near the end corresponding to the increasing intensity here as revealed by the Nelson-Aalen plot in Figure 7.

5.2 Statistical trend tests

Statistical trend tests for repairable systems data were extensively discussed by Ascher and Feingold [5, Ch. 5B]. A trend test is a statistical test for the null hypothesis that the failure process is stationary, in some sense to be made precise, versus alternatives depending on the kind of trend one would like to detect. Here we restrict attention to the null hypothesis that the process is an HPP or more generally an RP. However, as will be discussed below, some care should be taken when determining the relevant null hypothesis.

The null hypothesis of HPP is the most common and often the most useful one in

reliability applications. The corresponding null property, under the name of “randomness”, was studied in several papers in the 1950s, and various tests for randomness in time were devised. Here randomness pertained to the property that counts in given time intervals are Poisson-distributed. Maguire et al. [50], however, discussed the advantages of using inter-event times rather than counts in order to test for changes with time of the occurrence rate of events. Cox [17] stated eight different kinds of possible alternatives to “randomness”, one of them being trend in the sense that the conditional intensity is a smooth function of time.

5.2.1 Tests of the Null Hypothesis of HPP

Single process. Suppose first that the null hypothesis is “the process is an HPP”, with alternative being an NHPP with monotone intensity. Two classical trend tests for this case are the Laplace test and the Military Handbook test (see e.g. Ascher and Feingold [5, p. 79]). To see how they are obtained, consider a single system observed on $[0, \tau]$. If the failure process is an HPP, then given $N(\tau) = n$, the failure times T_1, T_2, \dots, T_n are distributed as the ordering of n i.i.d. uniform random variables on $[0, \tau]$. Equivalently, the T_i/τ ($i = 1, \dots, n$) are distributed as ordered i.i.d. uniforms on $[0, 1]$ conditionally given $N(\tau) = n$. From this we can in principle obtain trend-tests from any test for detecting deviations from a uniform sample. The Laplace test statistic is simply a normalization of $\sum_{i=1}^n T_i$, while the Military Handbook test is similarly a normalization of $\sum_{i=1}^n \log T_i$. The Laplace test and the Military Handbook test are optimal tests against the alternatives of NHPPs with, respectively, log linear intensity and power intensity functions (Ascher and Feingold [5, p. 79]).

Several processes. As in Section 5.1.2, assume that m independent NHPPs with a common intensity function $\lambda(t)$ are observed, where the i th process is observed on the time interval $(0, \tau_i]$. Recall that, under the null hypothesis that $\lambda(t)$ is a constant, the $r(T_i)/r(\tau)$, $i =$

$1, \dots, N$, form an HPP(1) on $[0, 1]$. Kvaløy and Lindqvist [35] suggested from this that formal trend tests could be defined by substituting the $r(T_i)/r(\tau)$ into the Laplace and Military Handbook test statistics. While these tests are powerful against monotone alternatives, the authors suggested to use a test statistic based on the Anderson-Darling statistic as a general test with power against several kinds of trend.

For many applications, the null hypothesis need to be weakened to state that each process is an HPP, but that intensities may differ from system to system. For example, in the data of Proschan [57], one may be interested in a simultaneous trend test for the systems, allowing there to be heterogeneities between them. Kvaløy and Lindqvist [35] suggested tests for this case, called combined tests. A precise setting for these tests has recently been defined by Kvist, Andersen and Kessing [37], who consider a model where the conditional intensity function for a particular system is given by $ag(\mathbf{Z})\lambda(t)$ where a is an unobservable frailty variable as considered in Section 4.1, \mathbf{Z} is a fixed-time covariate vector observed for each system, while $\lambda(t)$ is a baseline intensity function. Suppose that such a process is observed on the time interval $[0, \tau]$ with events at times $T_1, T_2, \dots, T_{N(\tau)}$. Then, conditional on $(a, \mathbf{Z}, \tau, N(\tau))$, the $T_1/\tau, T_2/\tau, \dots, T_{N(\tau)}/\tau$ are distributed as $N(\tau)$ ordered standard uniform variables on $[0, 1]$. We are hence back to the setting of Section 5.2.1. In practice one observes m independent processes of this kind, with a common $\lambda(t)$, with the a being i.i.d. unobservable random variables and the \mathbf{Z} being observed covariate vectors for each system. The above mentioned combined tests by Kvaløy and Lindqvist [35] can thus be used. Kvist et al. [37] apply the Laplace type test of this kind on data from the Danish register on psychiatric hospital admissions.

5.2.2 Tests of the Null Hypothesis of RP

The Laplace-test and the Military Handbook test are tests for the null hypothesis that the data come from HPPs. Thus rejection of the null hypothesis means merely that the process

is not an HPP. It could still, however, be an RP and thus still have "no trend" according to our definition above. Lawless and Thiagaraja [43] and Elvebakk [25] concluded from simulations that the Laplace and Military Handbook tests in fact may be seriously misleading when used to detect trend departures from general renewal processes. Similarly, Lewis and Robinson [44] noted that these tests are not able to discriminate properly between trends in the data and the appearance of sequences of very long intervals.

In order to test the null hypothesis of RP, Lewis and Robinson [44] suggested to modify the Laplace test by dividing the test statistic by an estimate of the coefficient of variation of the inter-failure times under the null hypothesis of an RP. This test, called the Lewis-Robinson test, is thus a simple modification of the Laplace test. Another classical trend test for the null hypothesis of RP is the rank test developed by Mann [51] and used in Proschan [57] (see Section 4).

Kvaløy and Lindqvist [36] presented a general class of tests for renewal process versus both monotonic and nonmonotonic trend for which the Lewis-Robinson and a useful Anderson-Darling type test are special cases.

Elvebakk [25] suggested to arrive at new trend tests by resampling test statistics for the Poisson case under the null hypothesis of RP. The general conclusion of Elvebakk [25] is that resampling trend tests are usually favorable to the Lewis-Robinson test. A comparison with the Laplace test and the Military Handbook test, shows that the corresponding resampling tests have lower power than their original versions for NHPP alternatives, but that this difference is small for large sample sizes. This fact, together with the fact that the Laplace and Military Handbook tests are invalid in non-Poisson cases, lead to a recommendation of the resampling tests as an overall conclusion.

5.2.3 Tests of the Null Hypothesis of Stationary Inter-failure Times

Lewis and Robinson [44] presented a test in order to distinguish between a general stationary sequence of inter-failure times X_i , and a monotonic trend in inter-failure times. Elvebakk [25] extended the resampling trend testing approach described in the previous subsection, in order to cover the case when “no trend” corresponds to stationary inter-failure times. The idea is to resample data under this new null hypothesis assumption. Elvebakk did this both by a parametric approach assuming an autoregressive model, and by employing a block bootstrap technique adapted from Hall [28]. Simulations indicated rather satisfactory performance of the method.

5.2.4 Trend Tests Obtained as Likelihood Ratio Tests

In parametric models which include separate parameters for trend, trend tests may be performed as likelihood ratio tests involving these parameters. An example is to test the null hypothesis $\beta_1 = 0$ in (5) which was suggested in [43]. Trend tests can also be obtained in models of the form $\text{HTRP}(F, \lambda(\cdot), H)$ by testing the null hypothesis that $\lambda(\cdot) \equiv \lambda$ using likelihood ratio tests. Note that this leads to tests of the null hypothesis that the processes are all renewal processes, but possibly heterogeneous.

A nonparametric likelihood ratio test for the null hypothesis of an HPP versus the alternative of an NHPP with monotone intensity $\lambda(\cdot)$ was derived by Boswell [11]. A generalization to the null hypothesis of RP can be obtained using the nonparametric monotone estimator of $\lambda(\cdot)$ in the TRP model derived by Heggland and Lindqvist [29].

6 Repairable Systems with Several Types of Events

In this section we consider the general marked event process as described in Section 2. The purpose is to show how new classes of maintenance and repair models can be obtained by generalizing the approach of the imperfect repair models for single type events considered

in Section 3.2. In order to simplify the presentation we shall not allow covariates or heterogeneity in the models considered here.

As in Section 3.2, we consider first a non-repairable unit. Assume that this unit may fail due to one of several causes, or may be stopped for PM before it fails, in which case failure is prevented.

We can formally think of this as having a system or component with, say, n components, denoted $\{C_1, C_2, \dots, C_n\}$, where a unique failing component can be identified at failures of the system, and where PM if present is represented by one of these components in order to simplify notation. Let W_j be the potential failure time due to failure of component C_j , $j = 1, 2, \dots, n$. What is observed is the failure time $T = \min(W_1, \dots, W_n)$ and the identity of the failing component, say $J = j$ if the component C_j fails. This determines a competing risks situation with n competing risks and with the observed outcome (T, J) (Crowder [20, Ch. 7.1]). The joint distribution of (T, J) is thus identifiable, as are the so called type-specific hazards defined by

$$h_j(t) = \lim_{\Delta t \downarrow 0} \frac{\Pr(t < T \leq t + \Delta t, J = j \mid T > t)}{\Delta t}. \quad (14)$$

However, neither the joint nor the marginal distributions of the individual potential failure times W_1, \dots, W_n are identifiable in general. This follows from the so-called Cox-Tsiatis impasse, see Crowder [20]. On the other hand, in reliability applications, these marginal and joint distributions are indeed of interest, for example in connection with maintenance optimization. An example is given in the next paragraph.

Consider the setup of Cooke [15, 16] involving a competing risks situation with a potential component failure at some time W_1 and an action of preventive maintenance to be performed at time W_2 . Thus $n = 2$, while C_1 corresponds to failure of the component ($J = 1$), and C_2 ($J = 2$) corresponds to the action of PM. Knowing the marginal distribution of W_1 would be particularly important since it is the basic failure time distribution

of the component. However, as noted above, the marginal distributions of W_1 and W_2 are not identifiable unless specific assumptions are made on the dependence between W_1 and W_2 . The most common assumption of this kind is that W_1 and W_2 are independent, in which case identifiability follows ([61, 20]). However, this assumption is unreasonable in the present application, since the maintenance crew is likely to have some information regarding the component's state during operation. This insight is used to perform maintenance in order to avoid component failures. Thus we are in practice faced with a situation of dependent competing risks between W_1 and W_2 , and hence identifiability of marginal distributions require additional assumptions.

Lindqvist, Støve and Langseth [49] suggested a model called the repair alert model for describing the joint behavior of the failure time W_1 and time W_2 of PM. This model is a special case of random signs censoring, Cooke [15, 16], under which the marginal distribution of W_1 is always identifiable. Recall that W_2 is said to be a random signs censoring of W_1 if the event $\{W_2 < W_1\}$ is stochastically independent of W_1 , that is if the event of having a PM before failure is not influenced by the time W_1 at which the component fails or would have failed without PM. The idea is that the component emits some kind of signal before failure, and that this signal is discovered with a probability which does not depend on the age of the component. The repair alert model extends this idea by introducing a so called repair alert function which describes the “alertness” of the maintenance crew as a function of time.

Another possibility to obtain identifiability would be to use the result of Zheng and Klein [64], who proved identifiability of marginal distributions when the dependence is given by a known copula.

Return now to the general case. Suppose that the system is repaired after failure and then put into operation, then may fail again, and so on. This leads to a marked event

process as in Section 2 with marks in $\mathcal{J} = \{1, 2, \dots, n\}$, so that the mark at each event time is the number of the failing component.

The properties of this process depend on the repair strategy. Several classes of interesting models can be described in terms of a generalization of the virtual age concept introduced in Section 3.2, as discussed in the next subsection.

6.1 Virtual Age Models with Several Types of Events

Recall from Section 3.2 that the class of virtual age models generalize the perfect repair and minimal repair models, and that the approach more generally leads to a large class of models. The main inputs are a hazard function $z(\cdot)$, which is thought of as the hazard function of a new unit, and a virtual age process which is a stochastic process which depends on the actual repair actions performed.

Several generalizations of the standard imperfect repair models are found in the literature. Shaked and Shantikumar [60] suggested a multi-component imperfect repair model with components having dependent life-lengths. Langseth and Lindqvist [38] suggested a model, involving imperfect maintenance and repair in the case of several components and several failure causes. In a recent paper, Doyen and Gaudoin [24] develop the ideas further by presenting a general point process framework for modeling of imperfect repair by a competing risks situation between failure and PM. Bedford and Lindqvist [8] considered a series system of n repairable components where only the failing component is repaired at failures.

Inspired by the mentioned approaches, we suggest in this section a generalization of the imperfect repair models to the case where there are more than one type of events, and where the virtual age process is multidimensional.

We let the first part of a virtual age model for n components be given by a vector process $A(t) = (A_1(t), \dots, A_n(t))$ containing the virtual ages of the n components at time t .

The crucial assumption is that $A(t) = (A_1(t), \dots, A_n(t)) \in \mathcal{F}_{t-}$, which means that the component ages are functions of the history up to time t .

As for the case with $n = 1$ in Section 3.2, it is assumed that the $A_j(t)$ increase linearly with time between events, and may jump only at event times. We define $v_j(i)$ to be the virtual age of component j immediately after the i th event. The virtual age process for component j is therefore defined by

$$A_j(t) = v_j(N(t-)) + t - T_{N(t-)}.$$

The second part of a virtual age model consists of the hazard function $z(\cdot)$ in the case $n = 1$. For general n we replace this by functions $\nu_j(v_1, \dots, v_n)$ for $v_1, v_2, \dots, v_n \geq 0$, such that the conditional intensity of type j events, given the history \mathcal{F}_{t-} , is

$$\gamma_j(t) = \nu_j(A_1(t), \dots, A_n(t)).$$

Thus $\nu_j(v_1, \dots, v_n)$ is the intensity of an event of type j when the component ages are v_1, \dots, v_n , respectively. The conditional intensity thus depends on the history only through the virtual ages of the components.

The family $\{\nu_j(v_1, \dots, v_n) : v_1, v_2, \dots, v_n \geq 0\}$ describes the failure mechanisms of the components and the dependence between them in terms of the ages of all the components. The statistical inference problem therefore consists of estimating these functions from field data. The case $n = 1$ has already been discussed in Section 3.2, but we shall see that identifiability problems can occur when $n > 1$.

6.2 Repair Models and their Virtual Age Processes

Most of the virtual age processes considered for the case $n = 1$ can be generalized to the present case of several event types. There are often, however, several ways of doing this. Some examples are given below. Additional examples would include, for example, generalization of Kijima's [34] models which may be plausible in applications.

6.2.1 Perfect repair of complete system

Suppose that all the components are repaired to as good as new at each failure of the system. In this case we have $v_j(i) = 0$ for all j and i , and hence $A_j(t) = t - T_{N(t-)}$ for all j . It follows that we can only identify the "diagonal" values $\nu_j(t, \dots, t)$ of the functions ν_j . As noted in Section 6.3, these are given by the type-specific hazards defined in (14) for the non-repairable competing risks case. This is not surprising in view of the fact that the present case of perfect repair essentially corresponds to observation of i.i.d. realizations of the non-repairable competing risks situation.

6.2.2 Minimal repair of complete system

In the given setting a minimal repair will mean that following an event, the process is restarted in the same state as was experienced immediately before the event. In mathematical terms, this implies that $v_j(i) = T_i$ for all i, j and hence that $A_j(t) = t$ for all j . Note that the complete set of functions ν_j is again not identifiable. Moreover, for a single component it is well known that minimal repair results in a failure time process which is an NHPP. In the present case of several components which are minimally repaired, it follows similarly that the failure processes of the individual components are independent NHPPs with the intensity for component j given by $\nu_j(t, \dots, t)$, which as already noted equals the type-specific hazard (14).

6.2.3 A partial repair model

Bedford and Lindqvist [8] suggested a partial repair model for the n component case. The virtual age process is defined by letting $A_j(t) =$ time since last event of type j . Equivalently, the process could be defined by

$$v_j(i) = \begin{cases} 0 & \text{if } J_i = j \\ v_j(i-1) + X_i & \text{if } J_i \neq j. \end{cases}$$

Thus, the age of the failing component is reset to 0 at failures, whereas the ages of the other components are unchanged. The authors considered a single realization of the process, with the main result being that under reasonable conditions pertaining to ergodicity, the functions $\nu_j(v_1, \dots, v_n)$ are identifiable. The intuitive idea of their proof is that the ages v_1, \dots, v_n will mix in such a manner that the complete set of $\nu_j(v_1, \dots, v_n)$ can be identified.

6.2.4 Age reduction models

Doyen and Gaudoin [24] consider a single component or system and two types of events, $C_1 = \text{failure}$, $C_2 = \text{PM}$. In their basic model the virtual ages of the two types of events are equal, $A_1(t) = A_2(t) = A(t)$. They indicate, however, that this restriction is not necessary. Various choices of virtual age processes are considered. In particular they consider age reduction models generalizing the ones described at the end of Section 3.2. More precisely, assume that there are given age reduction factors $0 < \rho_1, \rho_2 < 1$ for the two types of events. The virtual age immediately after the i th repair is then

$$v(i) = (1 - \rho_{J_i})(v(i-1) + X_i)$$

which means that the virtual age immediately before the i th failure, $v(i-1) + X_i$, is reduced due to repair by the factor $1 - \rho_{J_i}$. Alternatively, if only the additional age X_i is reduced by the repair, it is assumed that $v(i) = v(i-1) + (1 - \rho_{J_i})X_i$.

6.3 Modelling of the Intensity Functions ν_j

In principle the functions $\nu_j(v_1, \dots, v_n)$ could be any functions of the component ages. Bedford and Lindqvist [8] motivated these functions by writing, for $j = 1, \dots, n$,

$$\nu_j(v_1, \dots, v_n) = \lambda_j(v_j) + \lambda_{j*}(v_1, \dots, v_n) \tag{15}$$

with the convention that $\lambda_{j*}(v_1, \dots, v_n) = 0$ when all the component ages except the j th is 0, in order to have uniqueness. The $\lambda_j(v_j)$ is then thought of as the intensity of component

j when working alone or together with only new components, while $\lambda_{j*}(v_1, \dots, v_n)$ is the additional failure intensity imposed on component j caused by the other components when they are not all new. Note that any functions of v_1, \dots, v_n can be represented this way, by allowing the λ_{j*} to be negative as well as positive.

Langseth and Lindqvist [38] and Doyen and Gaudoin [24] extended the competing risk situation between failure and PM, as described in the beginning of Section 6, and suggested how to define functions ν_j . The main ideas of these approaches can be described for general n as follows. Starting from a state where the component ages are, respectively, v_1, \dots, v_n , let the time to next event be governed by the competing risk situation between the random variables W_1^*, \dots, W_n^* with distribution equal to the conditional distribution of $W_1 - v_1, \dots, W_n - v_n$ given $W_1 > v_1, \dots, W_n > v_n$, where the W_i are defined in the non-repairable case as described in the beginning of Section 6. It is then rather straightforward to show that this implies

$$\nu_j(v_1, \dots, v_n) = \frac{-\partial_j R(v_1, \dots, v_n)}{R(v_1, \dots, v_n)} \quad (16)$$

where $R(v_1, \dots, v_n) = P(W_1 > v_1, \dots, W_n > v_n)$ is the joint survival function of the W_i , and ∂_j means the partial derivative with respect to the j th entry in R . Note that this generalizes the usual hazard rate in the case $n = 1$ considered in Section 3.2. Further, we have $\nu_j(t, t, \dots, t) = h_j(t)$ where the latter is the type specific hazard rate given in (14).

A final remark on the suggested construction of the functions ν_j is due. It was demonstrated by Bedford and Lindqvist [8] that, even in the case with $n = 2$, it is not always possible to derive the functions $\nu_j(v_1, \dots, v_n)$ from a single joint survival distribution as in (16). A simple counter example is given in [8]. Thus for generality one should stick to completely general representations like (15).

7 Concluding Remarks

In the present paper we have reviewed some main approaches for the analysis of data from repairable systems. To a large extent the emphasis has been on describing the underlying principles and structures of common models. Essential features of such models correspond to the three "dimensions" of the model-cube in Figure 3; renewal-property, time trend and heterogeneity. The presentation has less emphasis on statistical inference than on modelling. However, it has been an intention to show how likelihood functions are obtained for the different models. It is also indicated how covariates can be included in the models and the corresponding likelihood functions. While the derived likelihood functions can be used in a rather straightforward manner in parametric statistical inference, there turn out to be several challenging problems connected to nonparametric estimation in some of the models.

Two main types of models with rather simple and transparent basic structures have been considered. These are the virtual age type models and the TRP type models. The former type combines two basic ingredients, a hazard rate $z(\cdot)$ of a new system together with a particular repair strategy which governs the virtual age process $A(t)$. The renewal dimension is taken care of by the virtual age process, while trend is determined by the distribution of a new system. For the TRP($F, \lambda(\cdot)$), the renewal dimension corresponds to the renewal distribution F , while the trend is explicitly given by the trend function $\lambda(\cdot)$. For both types of processes, heterogeneity can be included by multiplicative factors working on the intensities. A noticeable difference between the two types of models is that the virtual age type model usually requires that the virtual age process is observable. Such observations may, however, often be lacking in real data.

Many processes show some degree of clustering of failures. This may be due to various causes, see for example Cox [17]. Several models have been suggested in the literature, a

classical one being the Neyman and Scott [54] model. As pointed out by a referee, even the TRP model can pick up the clustering effect by allowing the renewal distribution to be a mixture with a substantial amount of probability near zero.

Peña [55] reviews a class of models suggested in Peña and Hollander [56]. This is a virtual age model which includes the possibility of heterogeneity between systems, time-dependent covariates, and for which in addition the conditional intensities may depend on the number of previous events. This last feature adds an interesting flexibility to the model. In particular it enables modelling of certain load sharing processes and software failure processes.

Certain systems, for example alarm systems, are tested only at fixed times which are usually periodic. If the system is found in a failed state, then it is repaired or replaced. Thus repair is not done at the same time as the failure, and the situation is not covered by the methods considered in the paper. A simple model of this situation was suggested by Hokstad and Frøvig [30] and further studied and extended by Lindqvist and Amundrustad [47]. Consider a system which starts the operation at time $t = 0$ and is tested at time epochs $\tau, 2\tau, 3\tau, \dots$. When time is running between testing epochs, the state of the system is modelled by an absorbing Markov chain. Having thus defined the probabilistic behavior of the system state between testing, one needs to add to the model a specification of the repair policy. In Lindqvist and Amundrustad [47] this is modelled in the form of a transition matrix on the state space of the Markov chain, which defines the possible changes of state and their probabilities following the repair actions.

In a given study there is usually a choice between several types of models. It is thus important to have tools for model checking and goodness-of-fit procedures. For model checking in parametric estimation of the HTRP-model we refer to Lindqvist et al. [48], who use a type of Cox-Snell residuals together with plots using the TTT technique. The general underlying idea, which in principle can be used with all estimation methods considered in this

paper, is that the process of integrated conditional intensities, $\int_0^{T_1} \gamma(t)dt, \int_0^{T_2} \gamma(t)dt, \dots$, is an HPP(1) (Brémaud [12]). In turn this gives rise to computable residual processes when estimates are inserted for parameters and distributions. The use of these processes in model checking is demonstrated for three different data sets in [48]. Typically, one would check (i) the distribution of the residuals with respect to departures from the unit exponential distribution, (ii) the possible presence of time trends in residuals within each system. Trend tests for the null hypothesis of HPP can be used in this part of the check.

Acknowledgements

The author is grateful to Drs Sallie Keller-McNulty and Alyson Wilson for the invitation to contribute to this special issue of Statistical Science. The article was prepared while the author was working in an international research group studying Statistical Analysis of Complex Event History Data at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2005/2006. The author would like to thank for the pleasant hospitality extended to him. Discussions with other members of the group, in particular Odd Aalen, Per Kragh Andersen and Ørnulf Borgan, have been very helpful and are greatly acknowledged. Thanks are also due to the author's former PhD students Georg Elvebakk, Jan Terje Kvaløy and Helge Langseth for long time collaborations and important contributions to the theory presented here. The author finally thanks two anonymous referees for careful reading of the manuscript and for their important and constructive comments which led to a much improved article.

References

- [1] AALEN, O. O. (1987). Mixing distributions on a Markov chain. *Scand. J. Statist.* **14** 281–289.
- [2] AALEN, O. O. (1988). Heterogeneity in survival analysis. *Statistics in Medicine* **7** 1121–1137.
- [3] AALEN, O. O. AND HUSEBYE, E. (1991). Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine* **10** 1227–1240.
- [4] ANDERSEN, P., BORGAN, O., GILL, R., KEIDING, N. (1992). *Statistical Models Based on Counting Processes*. Springer, New York.
- [5] ASCHER, H. AND FEINGOLD, H. (1984). *Repairable Systems – Modeling, inference, misconceptions and their causes*. Marcel Dekker, New York.
- [6] BARLOW, R. E., AND DAVIS, B. (1977). Analysis of time between failures of repairable components. In *Nuclear Systems Reliability Engineering and Risk Assessment*, eds. J. B. Fussel and G. R. Burdick, SIAM Philadelphia, pp. 543–561.
- [7] BARLOW, R. E., MARSHALL, A. W. AND PROSCHAN, F. (1963). Properties of probability distributions with monotone hazard rate. *Annals of Mathematical Statistics* **34** 375–389.
- [8] BEDFORD, T. AND LINDQVIST, B. H. (2004). The identifiability problem for repairable systems subject to competing risks. *Advances in Applied Probability* **36** 774–790.
- [9] BERMAN, M. (1981). Inhomogeneous and modulated gamma processes. *Biometrika* **68** 143–152.

- [10] BERMAN, M. AND TURNER, T. R. (1992). Approximate point process likelihoods with GLIM. *Applied Statistics* **41** 31-38.
- [11] BOSWELL, M. T. (1966). Estimating and testing trend in a stochastic process of Poisson type. *Annals of Mathematical Statistics* **37** 1564-1573.
- [12] BRÉMAUD, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York.
- [13] BROWN, M. AND PROSCHAN, F. (1983). Imperfect repair. *Journal of Applied Probability* **20** 851-859.
- [14] COOK, R. J. AND LAWLESS, J. F. (2002). Analysis of repeated events. *Statistical Methods in Medical Research* **11** 141-166.
- [15] COOKE, R. M. (1993). The total time on test statistics and age-dependent censoring. *Statistics and Probability Letters* **18** 307-312.
- [16] COOKE, R. M. (1996). The design of reliability databases, Part I and II. *Reliability Engineering and System Safety* **51** 137-146 and 209-223.
- [17] COX, D. R. (1955). Some statistical methods connected with series of events (with discussion). *Journal of the Royal Statistical Society, Series B* **17** 129-164.
- [18] COX, D. R. (1972). The statistical analysis of dependencies in point processes. In *Stochastic Point Processes*, Ed. P.A. Lewis, pp. 55-66. Wiley, New York.
- [19] COX, D. R. AND LEWIS, P. W. (1966). *The Statistical Analysis of Series of Events*. Methuen, London.
- [20] CROWDER, M. J. (2001). *Classical competing risks*. Chapman & Hall/CRC, Boca Raton.

- [21] CROWDER, M. J., KIMBER, A. C., SMITH, R. L. AND SWEETING, T. J. (1991). *Statistical Analysis of Reliability Data*. Chapman & Hall, Great Britain.
- [22] DORADO, C., HOLLANDER, M. AND SETHURAMAN, J. (1997). Nonparametric estimation for a general repair model. *Annals of Statistics* **25** 1140-1160.
- [23] DOYEN, L. AND GAUDOIN, O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering and System Safety* **84** 45-56.
- [24] DOYEN, L. AND GAUDOIN, O. (2005). Imperfect maintenance in a generalized competing risk framework. Submitted manuscript.
- [25] ELVEBAKK, G. (1999). Analysis of Repairable Systems Data: Statistical Inference for a Class of Models Involving Renewals, Heterogeneity and Time Trends. Ph.D. dissertation. Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway.
- [26] ENGELHARDT, M. AND BAIN, L. J. (1987). Statistical analysis of a compound power-law model for repairable systems. *IEEE Transactions on Reliability* **R-36** 392-396.
- [27] GREENWOOD, M. AND YULE, G. U. (1920). An enquiry into the nature of frequency distributions representative of multiple happenings with particular reference of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society* **83** 255–279.
- [28] HALL, P. (1985). Resampling a coverage pattern. *Stochastic Processes and their Applications* **20** 231-246.
- [29] HEGGLAND, K. AND LINDQVIST, B. H. (2006). A nonparametric monotone maximum likelihood estimator of time trend for repairable systems data. Submitted to

Reliability Engineering and System Safety, Special Issue on Engineering Applications of Stochastic Point Processes.

- [30] HOKSTAD, P. AND FRØVIG, A. T. (1996). The modelling of degraded and critical failures for components with dormant failures. *Reliability Engineering and System Safety* **51** 189-199.
- [31] HOLLANDER, M., PRESNELL, B. AND SETHURAMAN, J. (1992). Nonparametric methods for imperfect repair models. *Annals of Statistics* **20** 879–896.
- [32] HOUGAARD, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika* **71** 75–83.
- [33] Huang, C-Y. and Wang, M-C. (2004). Joint modelling and estimation for recurrent event processes and failure time data. *Journal of American Statistical Association* **99** 1153–1165.
- [34] KIJIMA, M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability* **26** 89–102.
- [35] KVALØY, J. T. AND LINDQVIST, B. H. (1998). TTT-based tests for trend in repairable systems data. *Reliability Engineering and System Safety* **60** 13-28.
- [36] KVALØY, J. T. AND LINDQVIST, B. H. (2003). A class of tests for renewal process versus monotonic and nonmonotonic trend in repairable system data. In: Mathematical and Statistical Methods in Reliability. In *Mathematical and Statistical Methods in Reliability*. (Lindqvist, B.H., Doksum, K.A., eds.), Series on Quality, Reliability and Engineering Statistics, Vol. 7, World Scientific Publishing, Singapore, pp 401-414.

- [37] KVIST, K., ANDERSEN, P. K. AND KESSING, L. V. (2006). Repeated events and total time on test. Research Report No. 7/2006. Institute of Public Health, University of Copenhagen.
- [38] LANGSETH, H. AND LINDQVIST, B. H. (2003). A maintenance model for components exposed to several failure mechanisms and imperfect repair. In *Mathematical and Statistical Methods in Reliability*. (Lindqvist, B.H., Doksum, K.A., eds.), Series on Quality, Reliability and Engineering Statistics, Vol. 7, World Scientific Publishing, Singapore, pp 415-430.
- [39] LANGSETH, H. AND LINDQVIST, B. H. (2006). Competing risks for repairable systems: A data study. *Journal of Statistical Planning and Inference* **136** 1687–1700.
- [40] LAST, G. AND SZEKLI, R. (1998). Stochastic comparison of repairable systems by coupling. *Journal of Applied Probability* **35** 348–70.
- [41] LAWLESS, J. F. (1987). Regression methods for Poisson process data. *Journal of American Statistical Association* **82** 808-815.
- [42] LAWLESS, J. F. AND NADEAU, C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics* **37** 158-168.
- [43] LAWLESS, J. F. AND THIAGARAJAH, K. (1996). A point-process model incorporating renewals and time trends, with application to repairable systems. *Technometrics* **38** 131-138.
- [44] LEWIS, P. A. W. AND ROBINSON, D. W. (1974). Testing for a monotone trend in a modulated renewal process. In *Reliability and Biometry. Statistical Analysis of Lifetimes*, eds. Proschan, F. and Serfling, R. J., SIAM Philadelphia, pp. 163–182.

- [45] LIM, T. J. (1998). Estimating system reliability with fully masked data under Brown-Proschan imperfect repair model. *Reliability Engineering and System Safety* **59** 277-289.
- [46] LINDQVIST, B. (1999). Repairable systems with general repair. In *Safety and Reliability. Proceedings of the European Conference on Safety and Reliability, ESREL '99, Munich, 13-17 September 1999*, eds. Schueller, G. and Kafka, P. Balkema, pp. 43-48.
- [47] LINDQVIST, B. H. AND AMUNDRUSTAD, H. (1998). Markov models for periodically tested components. In: *Safety and Reliability. Proceedings of the European Conference on Safety and Reliability - ESREL '98*, ed. S. Lydersen, G. K. Hansen & H. A. Sandtorv, Rotterdam: A. A. Balkema, pp 191-197.
- [48] LINDQVIST, B. H., ELVEBAKK, G. AND HEGGLAND, K. (2003). The trend-renewal process for statistical analysis of repairable systems. *Technometrics* **45** 31-44.
- [49] LINDQVIST, B. H., STØVE, B. AND LANGSETH, H. (2006). Modelling of dependence between critical failure and preventive maintenance: The repair alert model. *Journal of Statistical Planning and Inference* **136** 1701-1717.
- [50] MAGUIRE, B. A., PEARSON, E. S. AND WYNN, A. H. A. (1952). The time intervals between industrial accidents. *Biometrika* **39** 168-180.
- [51] MANN, H. B. (1945). Nonparametric tests against trend. *Econometrica* **13** 245-259.
- [52] MEEKER, W. Q. AND ESCOBAR, L. A. (1998). *Statistical methods for reliability data*. Wiley, New York.
- [53] NELSON, W. (1995). Confidence limits for recurrence data - applied to cost or number of product repair. *Technometrics* **37** 147-157.

- [54] NEYMAN, J. AND SCOTT, E. L. (1958). A statistical approach to problems of cosmology. *Journal of the Royal Statistical Society, Series B* **20** 1-43.
- [55] PEÑA, E. A. (2006). Dynamic modelling and statistical analysis of event times. Submitted to *Statistical Science*.
- [56] PEÑA, E. A. AND HOLLANDER, M. (2004). Models for recurrent events in reliability and survival analysis. In *Mathematical Reliability: An Expository Perspective*, eds. R. Soyer, T. Mazzuchi and N. Singpurwalla. Kluwer Academic Publishers.
- [57] PROSCHAN, F. (1963). Theoretical explanation of observed decreasing failure rates. *Technometrics* **5** 375-383.
- [58] PHAM, H. AND WANG, H. (1996). Imperfect maintenance. *European Journal of Operations Research* **94**, 425-428.
- [59] RAUSAND, M. AND HØYLAND, A. (2004). *System reliability theory: Models, statistical methods, and applications. 2nd ed.* Wiley-Interscience, Hoboken, N.J.
- [60] SHAKED, M. AND SHANTIKUMAR, J. G. (1986). Multivariate imperfect repair. *Operations Research* **34** 437-448.
- [61] TSIATIS, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences, USA* **72** 20-22.
- [62] VAUPEL, J.W., MANTON, K.G. AND STALLARD, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16** 439-454.
- [63] WHITAKER, L. R. AND SAMANIEGO, F. J. (1989). Estimating the reliability of systems subject to imperfect repair. *Journal of American Statistical Association* **84** 301-309.

- [64] ZHENG, M. AND KLEIN, J.P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, **82**, 127-138.