A Class of Tests for Trend in Time Censored Recurrent Event Data

Jan Terje Kvaløy
Department of Mathematics and Physics
University of Stavanger
N-4036 Stavanger, Norway
Email: jan.t.kvaloy@uis.no

Bo Henry Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7491 Trondheim, Norway
Email: bo.lindqvist@ntnu.no

April 5, 2019

Abstract

Statistical tests for trend in recurrent event data not following a Poisson process are generally constructed for event censored data. However, time censored data are more frequently encountered in practice. In this paper we contribute to filling an important gap in the literature on trend testing by presenting a class of statistical tests for trend in time censored recurrent event data, based on the null hypothesis of a renewal process. The class of tests is constructed by an adaption of a functional central limit theorem for renewal processes. By this approach a number of tests for time censored recurrent event data can be constructed, including among others a version of the classical Lewis-Robinson trend test and an Anderson-Darling type test. The latter test turns out to have attractive properties for general use by having good power properties against both monotonic and non-monotonic trends. Extensions to situations with several processes are considered. Properties of the tests are studied by simulations and some asymptotic calculations, and the approach is illustrated in data examples.

Keywords: Trend testing; Time truncation; Renewal process; Trend-renewal process; Brownian bridge.

1 Introduction

Analyzing recurrent event data is a challenge encountered in many fields, for instance engineering, medicine and economics to mention some. Generally, recurrent event data arise when the phenomenon studied can occur repeatedly. Some examples are the occurrence of a failure in a repairable system or the outbreak of a recurrent disease. One aspect of the data which typically is of interest is to examine whether there are any systematic alterations, i.e., trends, in the pattern of events. For example, does a repairable system have a tendency to fail more often as it gets older? Or is there any improvement in how often a recurrent disease occurs for a particular patient? Visual inspections of the data can be very useful and give important information on systematic tendencies in the data, but generally, in order to distinguish actual systematic alterations from random fluctuations, statistical methods are needed.

There is a rich literature on trend testing, see for instance the overviews in Cox and Lewis (1966), Ascher and Feingold (1984), Kvaløy and Lindqvist (1998), Lawless, Çiğşar and Cook (2012) and Cook and Lawless (2013). Trend tests are based on different assumptions for the data collection process and different definitions of trend. Many of the existing tests for trend are based on Poisson process theory and constructed for testing the null hypothesis of a homogeneous Poisson process (HPP), see for instance Cox and Lewis (1966), Ascher and Feingold (1984), Cohen and Sackrowitz (1993), Kvaløy and Lindqvist (1998), Lawless et al. (2012) and references therein. Such tests are, however, generally sensitive to departures from the Poisson process assumption. This fact was noted in the classical reference Lewis and Robinson (1974), who observed that the commonly used Laplace trend test often led to rejection of the null hypothesis of no trend, even in cases where a trend could not exist. More specifically, the authors observed that false rejections were particularly occurring in cases of overdispersion of the interevent times with respect to the exponential distribution. Their idea was to modify the Laplace test statistic to account for this overdispersion, which led to the test known under the name of Lewis-Robinson test, to be further considered later in this paper.

The immediate conclusion to draw from this seems to be that, unless the Poisson assumption can be verified, trend tests need to be based on more general null hypotheses than

the one of HPP. So how could one formalize a more useful null hypothesis? Lawless et al. (2012) concluded that there is no single definition which covers all cases that can naturally be thought of. Lewis and Robinson (1974) argued that a definition of no trend should state that the event process is stationary in some sense, possibly allowing some amount of serial correlation. On the other hand, because of analytical possibilities they found that the renewal process (RP) assumption would be the best choice for further investigations. Under this assumption they were able to repair the Laplace test and introduce the Lewis-Robinson test.

In this paper we shall consider trend tests assuming the null hypothesis of RP. In addition to the Lewis-Robinson test, there exist several trend tests in the literature based on this null hypothesis. We would like to mention first the nonparametric test by Mann (1945). Other tests are found in Ascher and Feingold (1984), Kvaløy and Lindqvist (2003), Viertävä and Vaurio (2009), Lawless et al. (2012) and references therein.

RP based tests for trend, including the classical Lewis-Robinson test are, however, usually constructed for *event censored* data, which means that the recurrent event process is censored when it has completed a fixed number of renewal events. On the other hand, *time censored* data, where the event process is censored after a predetermined observation period, are far more naturally occurring in practice. As pointed out by Lawless et al. (2012), there is still an unfortunate lack of available trend tests constructed for time censored data. Important issues when going from event censoring to time censoring are the fact that the number of events is now random, and that one needs to involve in a consistent manner the time interval from the last event to the censoring time. Lawless et al. (2012) argued that ignoring this interval may lead to considerable bias, see also the most interesting discussion of this and related issues in Aalen and Husebye (1991). The latter authors, furthermore, pointed out that it is far less critical to ignore an incomplete time at the start of the observation, which will not introduce bias although it might incur a certain loss of efficiency.

With the above as our motivation and point of departure, we demonstrate in this paper how a flexible class of trend tests for time censored data can be constructed under the RP null hypothesis. We thereby complement the above mentioned literature on trend tests for event censored data, in particular the paper by Lawless et al. (2012). Our construction is based on an adaption of Donsker's theorem (Donsker, 1952) to renewal processes following

the lines of Billingsley (1999). Among other tests, the class turns out to include a time censored version of the Lewis-Robinson test, an Anderson-Darling type test with power against both monotonic and non-monotonic trends and an extension of the Lewis-Robinson test with power against non-monotonic trend. After having studied tests for trend in single processes, we consider extensions to trend tests based on the joint observation of several processes.

The paper is organized as follows. In Section 2 we define the necessary notation and give some key results for renewal processes. The general construction of tests is presented in Section 3 and several specific tests are derived. Section 4 discusses extensions to cases where several similar processes are observed. Two other tests are described in Section 5 and are included in the simulation study presented in Section 6. Section 7 presents asymptotic power calculations for the main tests of the paper, considering a certain class of alternatives to no trend. Two case studies are considered in Section 8, while some concluding remarks are given in Section 9. The paper is ended by Appendices 1-3, providing details on the derivation of parameter estimators and a specific trend test, as well as a functional central limit theorem used in the power calculations.

2　The Basic Convergence Results for Renewal Processes

2.1　Setup and Notation

Consider a renewal process observed from time $t = 0$. The successive event times are denoted $T_1, T_2, \ldots$, and the corresponding interevent times, or gap times, are denoted $X_1, X_2, \ldots$ where $X_i = T_i - T_{i-1}$, $i = 1, 2, \ldots$ (with the convention $T_0 = 0$). The X_i are independent and identically distributed, with $\mathrm{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$, where it will be assumed throughout the paper that $\sigma^2 < \infty$.

We use the standard notation where $N(t)$ is the number of events in $(0, t]$ for all $t > 0$. For the theory of renewal processes we refer to, e.g., Ross (1983) and Gallager (2013).

2.2 A Functional Central Limit Theorem for Renewal Processes

The key result in our approach is a functional central limit theorem given in Billingsley (1999). With notation as above, define

$$V_{t,\mu,\sigma}(s) = \mu^{3/2}\frac{N(st) - st/\mu}{\sigma\sqrt{t}} \text{ for } 0 \le s \le 1,\ t > 0.$$

Then (Billingsley, 1999, thm. 14.6),

$$V_{t,\mu,\sigma} \Rightarrow W \text{ as } t \to \infty, \tag{1}$$

where $\Rightarrow$ denotes weak convergence and W is the Wiener measure (Billingsley, 1999, chap. 8).

Now define $W^0(s) = W(s) - sW(1)$ for $0 \le s \le 1$, so that W^0 is a Brownian bridge (Billingsley, 1999, chap. 8). It is straightforward to verify that (1) implies the following result which together with the succeeding corollary is the basis of our construction of trend tests.

Theorem 1 *Define*

$$V^0_{t,\mu,\sigma}(s) = V_{t,\mu,\sigma}(s) - sV_{t,\mu,\sigma}(1) = \mu^{3/2}\frac{N(st) - sN(t)}{\sigma\sqrt{t}} \text{ for } 0 \le s \le 1. \tag{2}$$

Then $V^0_{t,\mu,\sigma} \Rightarrow W^0$.

Let the coefficient of variation of the interevent times X_i be denoted $\gamma \equiv \sigma/\mu$. As will become clear, γ plays a special role in our construction of tests. First, define

$$\tilde{V}^0_{t,\gamma}(s) = \frac{1}{\gamma}\frac{N(st) - sN(t)}{\sqrt{N(t)}} \text{ for } 0 \le s \le 1. \tag{3}$$

Then Theorem 1 implies the following corollary:

Corollary 1 *With notation as above we have*

$$\tilde{V}^0_{t,\gamma} \Rightarrow W^0 \text{ as } t \to \infty.$$

Proof: We can write

$$\tilde{V}^0_{t,\gamma}(s) = \frac{\sqrt{1/\mu}}{\sqrt{N(t)/t}}V^0_{t,\mu,\sigma}(s).$$

5

From standard renewal process theory (Ross, 1983) it is well known that $N(t)/t \to 1/\mu$ a.s. The result then follows by use of Billingsley (1999, thm. 3.1), sometimes called 'the converging together lemma'. The argument, using the uniform norm, is as follows:

$$\sup_{0 \leq s \leq 1} |\tilde{V}_{t,\gamma}^0(s) - V_{t,\mu,\sigma}^0(s)| \leq \left| \frac{\sqrt{1/\mu}}{\sqrt{N(t)/t}} - 1 \right| \sup_{0 \leq s \leq 1} |V_{t,\mu,\sigma}^0(s)| \xrightarrow{p} 0,$$

where the convergence to 0 follows since the first factor tends to 0 a.s. and hence in probability, and the last factor converges in distribution to $\sup_{0 \leq s \leq 1} |W^0(s)|$ which has the Kolmogorov distribution (and will be considered below).

3   The Class of Tests for Trend

In the present section we consider event data from a single counting process $N(t)$ observed from time $t = 0$ until time censoring at the given time $\tau > 0$. With notation as in Section 2, we thus observe a random number $N(\tau)$ of events, at times $T_1, T_2, \ldots, T_{N(\tau)}$, and with fully observed interevent times $X_1, X_2, \ldots, X_{N(\tau)}$ and a censored interevent time $\tau - T_{N(\tau)}$.

From Theorem 1 and Corollary 1 it follows that, under the null hypothesis of RP, $V_{\tau,\mu,\sigma}^0$ and $\tilde{V}_{\tau,\gamma}^0$ will approximately be Brownian bridges. However, if there is a trend in the data, these processes are likely to deviate from a Brownian bridge. Tests for trend can therefore be based on measures of deviation from a Brownian bridge of the two asymptotically equivalent processes $V_{\tau,\mu,\sigma}^0$ and $\tilde{V}_{\tau,\gamma}^0$.

Since the parameters μ, σ, γ are generally unknown, they must be estimated. It is clear that the results of Theorem 1 and Corollary 1 continue to hold under the RP assumption if μ, σ and γ are replaced by consistent estimators, $\hat{\mu}, \hat{\sigma}$ and $\hat{\gamma}$.

Below we first derive test statistics based on four different ways of measuring deviations from a Brownian bridge. This leads to test statistics of, respectively, Lewis-Robinson, Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling types. In addition we propose an extension of the Lewis-Robinson test which can be used to construct tests for non-monotonic trend. The test constructions are based on applications of Corollary 1. Finally we discuss how to estimate the parameters μ, σ and γ.

3.1 Lewis-Robinson Type Test

A classical measure of deviation from a Brownian bridge is the signed area under the path of the process. Using Corollary 1 this gives rise to the statistic $\int_0^1 \tilde{V}_{\tau,\hat{\gamma}}^0(s)ds$, which converges in distribution to $\int_0^1 W^0(s)ds$, which is normally distributed with expectation 0 and variance $1/12$.

In order to obtain the test statistic on the form that is most common for this test, we use instead the negative of the above suggested statistic, which will have the same limiting distribution. By scaling we obtain an asymptotically standard normally distributed test statistic given by

$$LR = -\sqrt{12} \int_0^1 \tilde{V}_{\tau,\hat{\gamma}}^0(s)ds = \frac{1}{\hat{\gamma}} \cdot \frac{\sqrt{12}}{\tau\sqrt{N(\tau)}} \left[\sum_{i=1}^{N(\tau)} T_i - \frac{N(\tau)}{2}\tau \right]. \tag{4}$$

If the factor $1/\hat{\gamma}$ is ignored, we actually get the well known Laplace test statistic for the null hypothesis of HPP for the time censored case, which can be derived from properties of Poisson-processes. The division by $\hat{\gamma}$ corresponds to the correction obtained by Lewis and Robinson (1974), who considered the event censored case.

The resulting test will primarily have power against deviations from an RP caused by monotonic trends. It is seen that positive (negative) values of the test statistic will correspond to an increasing (decreasing) trend. Although one-sided tests for increasing (decreasing) trend thus can be constructed, the typical application would be to construct a two-sided test which rejects when $|LR| > z_{\alpha/2}$, where $z_{\alpha/2}$ is the $\alpha/2$ quantile of the standard normal distribution and α is the level of the test.

3.2 Kolmogorov-Smirnov Type Test

Another classical measure of deviation from a Brownian bridge is the maximum deviation, giving rise to the statistic $\sup_{s\in[0,1]} |\tilde{V}_{\tau,\hat{\gamma}}^0(s)|$. By Corollary 1, this statistic converges in distribution to $\sup_{s\in[0,1]} |W^0(s)|$, which has the Kolmogorov distribution (Kolmogorov, 1933; Smirnov, 1948). A Kolmogorov-Smirnov type test for trend in the time censored case is

hence given by the test statistic

$$
\begin{aligned}
KS &= \sup_{s\in[0,1]} |\tilde{V}^0_{\tau,\hat{\gamma}}(s)| = \frac{1}{\hat{\gamma}} \frac{1}{\sqrt{N(\tau)}} \sup_{s\in[0,1]} |N(s\tau) - sN(\tau)| \\
&= \frac{1}{\hat{\gamma}} \frac{1}{\sqrt{N(\tau)}} \max_{i=1,\ldots,N(\tau)} \left\{ \max\left[\left| i - \frac{N(\tau)}{\tau}T_i \right|, \left| i - 1 - \frac{N(\tau)}{\tau}T_i \right| \right] \right\}.
\end{aligned}
\tag{5}
$$

The asymptotic cumulative distribution function and a corresponding table of critical values are given in Kolmogorov (1933), see also Billingsley (1999, p. 151). For $\alpha = 0.05$ the null hypothesis is rejected when $KS > 1.358$.

3.3 Cramér-von Mises Type Test

Using the Cramér-von Mises type measure we obtain

$$
CvM = \int_0^1 \tilde{V}^0_{\tau,\hat{\gamma}}(s)^2 ds \overset{d}{\to} \int_0^1 W^0(s)^2 ds,
$$

where the right hand side has the commonly known limit distribution of the Cramér-von Mises statistic (Anderson and Darling, 1952). Due to the squaring of $\tilde{V}^0_{\tau,\hat{\gamma}}(s)$ it is clear that a test which rejects the null hypothesis of RP for large values of CvM will have sensitivity against both monotonic and non-monotonic trends. Straightforward calculations give the statistic

$$
\begin{aligned}
CvM &= \frac{1}{\hat{\gamma}^2} \frac{1}{N(\tau)} \left\{ \sum_{i=0}^{N(\tau)-1} \left[i^2 \frac{X_{i+1}}{\tau} - iN(\tau)\frac{T_{i+1}^2 - T_i^2}{\tau} \right] \right. \\
&\qquad\qquad \left. + N(\tau)^2 \left[\frac{T_{N(\tau)}^2}{\tau^2} - \frac{T_{N(\tau)}}{\tau} + \frac{1}{3} \right] \right\}.
\end{aligned}
\tag{6}
$$

The asymptotic cumulative distribution function and a table of critical values are given in Anderson and Darling (1952). For $\alpha = 0.05$ the null hypothesis is rejected when $CvM > 0.461$.

3.4 Anderson-Darling Type Test

The Anderson-Darling type measure leads to

$$
AD = \int_0^1 \frac{V^0_{\tau,\hat{\gamma}}(s)^2}{s(1-s)} ds \overset{d}{\to} \int_0^1 \frac{W^0(s)^2}{s(1-s)} ds,
$$

8

which has the limit distribution of the Anderson-Darling statistic (Anderson and Darling, 1952; Anderson and Darling, 1954). As for the Cramér-von Mises type test it is clear that this test will have sensitivity against both monotonic and non-monotonic trends. The difference between the Cramér-von Mises and the Anderson-Darling statistics is that the latter puts more weight on the information at the beginning and the end of the observation interval. Straightforward but somewhat tedious calculations give that

$$
\begin{aligned}
AD \; = \; & \frac{1}{\hat{\gamma}^2} \frac{1}{N(\tau)} \left\{ \sum_{i=1}^{N(\tau)-1} \left[(N(\tau) - i)^2 \ln(\frac{\tau - T_i}{\tau - T_{i+1}}) + i^2 \ln(\frac{T_{i+1}}{T_i}) \right] \right. \\
& \left. + N(\tau)^2 \left[\ln(\frac{\tau}{\tau - T_1}) + \ln(\frac{\tau}{T_{N(\tau)}}) - 1 \right] \right\}.
\end{aligned} \tag{7}
$$

The asymptotic cumulative distribution function is given in Anderson and Darling (1952) and a table of critical values is given in Anderson and Darling (1954). For $\alpha = 0.05$ the null hypothesis is rejected when $AD > 2.492$.

3.5   The Extended Lewis-Robinson Test for Non-Monotonic Trend

Recall that the Lewis-Robinson type test for the time censored case was based on the integral $\int_0^1 \tilde{V}_{\tau,\hat{\gamma}}^0(s)ds$. This test is suited for alternatives of monotonic trend. Consider instead the expression

$$
\int_0^a \tilde{V}_{\tau,\hat{\gamma}}^0(s)ds - \int_a^1 \tilde{V}_{\tau,\hat{\gamma}}^0(s)ds, \tag{8}
$$

where $0 \leq a \leq 1$. It is seen that $a = 0$ in fact leads to the preferred test statistic (4) for the Lewis-Robinson test (of course, $a = 1$ gives the negative of the LR statistic (4)).

A test based on (8) will obviously have power to detect non-monotonic trends where the trend in $[0, a\tau]$ and $[a\tau, \tau]$ are in opposite directions. Clearly, (8) converges in distribution to $\int_0^a W^0(s)ds - \int_a^1 W^0(s)ds$, which is normally distributed with expectation 0 and variance $1/12 - a^2(1-a)^2$ (see Appendix 2). It follows from a calculation in Appendix 2 that (8), after a scaling to give an asymptotically standard normal distribution under the null hypothesis, can be written

$$
ELR = \frac{1}{\hat{\gamma}} \cdot \frac{1}{\tau\sqrt{N(\tau)}\sqrt{(1/12) - a^2(1-a)^2}} \left\{ \sum_{i=1}^{N(\tau)} |T_i - a\tau| - \left(\frac{1}{2} - a(1-a) \right) \tau N(\tau) \right\}. \tag{9}
$$

A disadvantage of the above test is that the value of a has to be given. This choice of a should be made prior to looking at the data. In some applications a reasonable choice could be made based on e.g. knowledge about operational or seasonal changes. Another possibility would of course be to allow an adaptive choice of a. This will, however, destroy the above distributional properties, and we will therefore not pursue this approach here. The impact of the choice of a will be illustrated in the simulation study. If no specific prior knowledge of a natural turning point for the trend is available, $a = 1/2$ is a natural default choice.

Viertävä and Vaurio (2009) suggested on an ad hoc basis, and for the event censored case, a test statistic similar to (9) with $a = 1/2$.

3.6   Parameter Estimation

If one assumes the null hypothesis of HPP, then $\gamma = 1$ is known, and hence no estimation is needed in the use of Corollary 1. If we more generally assume specific parametric models for the event process, then the parameters μ, σ, γ may be estimated by maximum likelihood methods since they are functions of the model parameters. In the case study of Section 8.1 we illustrate the parametric estimation by fitting Weibull RPs to the interevent times, taking into account also the censored time at the end of the observation. Since the Weibull distribution is a rather flexible distribution, the corresponding estimates of μ, σ and γ may be satisfactory also under the null hypothesis of RP when no parametric assumptions are made. But strictly, when fitting Weibull distributions under H_0, we test the null hypothesis that the events follow a Weibull RP.

When no distributional assumptions are made on the process, obvious choices for estimators of μ and σ are the sample mean $\hat{\mu}$ and sample standard deviation $\hat{\sigma}$ of the completely observed interevent times. These estimators are consistent as $\tau \to \infty$ (see Appendix 1), but have the disadvantage of not utilizing the censored times $\tau - T_{N(\tau)}$ at the end of the observation period. The corresponding estimator of γ is $\hat{\gamma} = \hat{\sigma}/\hat{\mu}$.

Alternative estimators which involve the censored time $\tau - T_{N(\tau)}$ may be derived from standard renewal process theory. Again we refer to Appendix 1 for justification of the

following estimators,

$$\tilde{\mu} = \frac{\tau}{N(\tau)}, \quad \tilde{\sigma}^2 = \frac{1}{N(\tau)} \left\{ \sum_{i=1}^{N(\tau)} X_i^2 + (\tau - T_{N(\tau)})^2 \right\} - \tilde{\mu}^2, \quad \tilde{\gamma} = \tilde{\sigma}/\tilde{\mu}. \tag{10}$$

Another variance estimator (see Appendix 1 for its verification) is

$$\sigma^{*2} = \frac{1}{2(N(\tau) - 1)} \sum_{i=1}^{N(\tau)-1} (X_{i+1} - X_i)^2. \tag{11}$$

The potential advantage of this estimator is that it tends to be smaller than $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ under alternatives with positive dependence between subsequent interevent times. This makes the estimated γ become smaller, which leads to larger (absolute) values of the test statistics and hence higher rejection probability under alternatives of monotonic trend, see for example Viertävä and Vaurio (2009). We will, however, in our simulation and data examples use $\hat{\sigma}$ or $\tilde{\sigma}$ and not σ^*, due to apparent less satisfactory significance level properties, as experienced in simulations.

4 Tests for Trend in Multiple Processes

Suppose now that $m > 1$ similar processes are observed. Under the assumption that the processes are stochastically independent it may be of interest to test the null hypothesis that they all have no trend. One possible formulation of the null hypothesis is to let H_0 state that all the m processes are independent RPs, but that they are not necessarily identically distributed. A stronger null hypothesis would be to state that the m processes are independent RPs with the same distribution of the interevent times. We will below stick to the former interpretation. The latter interpretation may lead to false conclusions of trend. A discussion and illustration of these issues in a Poisson process setting is found in Kvaløy and Lindqvist (1998).

Cook and Lawless (2013) considered a type of modulated renewal model and showed that heterogeneities between processes, when not accounted for, may lead to an apparent trend in the gap time distribution. That is, even if each individual process has no trend, not accounting for the heterogeneity creates an apparent trend. While Cook and Lawless (2013) account for the heterogeneity by introducing a random effect, our approach can be

considered a fixed effects type approach where we make individual parameter estimates in each process.

Construction of the test statistics is based on the following fact, which we state as a lemma:

Lemma 1 *Let $W_1^0, W_2^0, \ldots, W_m^0$ be independent Brownian bridges and let $a_1, a_2, \ldots, a_m$ be real numbers with $\sum_{j=1}^m a_j^2 = 1$. Then*

$$W^0 = \sum_{j=1}^m a_j W_j^0$$

is a Brownian bridge.

Proof: By linearity it is clear that W^0 is a Gaussian process with expectation 0. The result follows by a straightforward calculation of the covariance function.

Let τ_j, μ_j, σ_j and γ_j be, respectively, the censoring time, mean, standard deviation and coefficient of variation corresponding to process j, $j = 1, \ldots, m$. Let further $A_1, \ldots, A_m$ be random variables where A_j depends on the data from process j only, and assume that $A_j \xrightarrow{p} a_j$, $j = 1, \ldots, m$, where the a_j are constants with $\sum_{j=1}^m a_j^2 = 1$. Then from Lemma 1, Corollary 1 and the already cited 'converging together lemma' it follows that

$$\sum_{j=1}^m A_j \tilde{V}_{\tau_j, \gamma_j}^0(s) = \sum_{j=1}^m A_j \frac{1}{\gamma_j} \frac{N_j(s\tau_j) - sN_j(\tau_j)}{\sqrt{N_j(\tau_j)}} \Rightarrow W^0 \text{ as } \tau_j \to \infty, \ j = 1, \ldots, m \qquad (12)$$

under the null hypothesis that the m processes are independent RPs.

4.1 Lewis-Robinson Type Test for m Processes

By the same arguments as in Section 3.1, and with the assumption on the weights given above, the following statistic will be asymptotically standard normally distributed under H_0,

$$LR^m = -\sqrt{12} \int_0^1 \sum_{j=1}^m A_j \tilde{V}_{\tau_j, \gamma_j}^0(s) ds = \sum_{j=1}^m A_j \frac{1}{\hat{\gamma}_j} \cdot \frac{\sqrt{12}}{\sqrt{N_j(\tau_j)}} \left[\sum_{i=1}^{N_j(\tau_j)} \frac{T_{ij}}{\tau_j} - \frac{N_j(\tau_j)}{2} \right]. \qquad (13)$$

Here T_{ij} denotes the time until failure number i in process j, $i = 1, \ldots, N_j(\tau_j)$, $j = 1, \ldots, m$.

Different choices of the weights A_j will lead to different tests. For instance, $A_j = 1/\sqrt{m}$, $j = 1, \ldots, m$ will mean equal weighting of the information from each process. This might, however, not be an optimal choice in cases where there is a large variation in the number of events per process, or where the coefficient of variation of interevent times differs much between processes.

We will show that the choice

$$A_j \propto \sqrt{N_j(\tau_j)/\hat{\gamma}_j^2} \tag{14}$$

is optimal in terms of power against certain alternatives, using results from Section 7.1.2. These weights are, moreover, reasonable in that they give higher weight to systems with many failures, and with lower coefficient of variation of interevent times.

Suppose now that the τ_j tend to infinity in such a manner that, for a τ tending to infinity, $\tau_j/\tau \to c_j$ for positive constants c_j, $j = 1, \ldots, m$. Since the $N_j(\tau_j)/\tau_j \to 1/\mu_j$ a.s. and assuming $\hat{\gamma}_j \overset{p}{\to} \gamma_j$, we have

$$A_j = \frac{\sqrt{N_j(\tau_j)/\hat{\gamma}_j^2}}{\sqrt{\sum_{k=1}^m N_k(\tau_k)/\hat{\gamma}_k^2}} = \frac{\sqrt{\tau_j/\tau}\sqrt{N_j(\tau_j)/\tau_j\hat{\gamma}_j^2}}{\sqrt{\sum_{k=1}^m (\tau_k/\tau)N_k(\tau_k)/(\tau_k\hat{\gamma}_k^2)}} \overset{p}{\to} \frac{\sqrt{c_j}/(\gamma_j\sqrt{\mu_j})}{\sqrt{\sum_{k=1}^m c_k/(\gamma_k^2\mu_k)}} \equiv a_j. \tag{15}$$

Clearly, $\sum_{j=1}^m a_j^2 = 1$, so the statistic (13) will converge to a standard normal distribution under the null hypothesis H_0. Moreover, the $a_j \propto \sqrt{c_j}/(\gamma_j\sqrt{\mu_j})$, which by Section 7.1.2 are optimal choices for alternatives with power law type trend, t^b, as $b \to 1$.

Inserting the weights A_j from (15) and rearranging we can write the test statistic (13) as

$$LR^m = \frac{\sqrt{12}}{\sqrt{\sum_{k=1}^m N_k(\tau_k)/\hat{\gamma}_k^2}} \sum_{j=1}^m \frac{1}{\hat{\gamma}_j^2} \left[\sum_{i=1}^{N_j(\tau_j)} \frac{T_{ij}}{\tau_j} - \frac{N_j(\tau_j)}{2} \right]. \tag{16}$$

4.2 Further Tests for m Processes

For the other tests considered in Section 3 it is in principle possible to replace the $\tilde{V}^0_{\tau,\gamma}(s)$ by $\sum_{j=1}^m A_j \tilde{V}^0_{\tau_j,\gamma_j}(s)$, using (12), and applying the same operations as for the case $m = 1$. This corresponds to what we did for the Lewis-Robinson test in the previous subsection, but there things are easy due to the linearity of integrals. This also applies to the extended Lewis-Robinson test. For the remaining tests it is not straightforward, however, to derive explicit expressions for the test statistics, and it is neither clear what would be the best

weights to use. The problem associated with the Cramér-von Mises and Anderson-Darling tests are of course that the integrand is a square, while for the Kolmogorov-Smirnov test the various processes are mixed together before taking the absolute value, making tractable expressions impossible.

Another possibility for these last mentioned tests would therefore be to use (weighted) sums of the individual test statistics to define the new test statistics. Such an approach requires, on the other hand, the distributions of sums or linear combinations of the limiting distributions for the single process cases. These may be determined by simulations or, for larger m, by normal approximations. Note also that Scholz and Stephens (1987) have considered the distribution of sums of independent Anderson-Darling statistics.

For such linear combinations there are no obvious choices for the weights given to each process. A reasonable choice under the assumption of the same interevent distribution in all processes would be to let the weights be proportional to τ_j. Otherwise, it may be tempting to use weights similar to (14), hence taking into account the the number of observed events as well as the coefficient of variation of the interevent times. A problem would then of course be that these weights are random, making exact simulation of the distribution under the null hypothesis impossible.

In practice we have found that the normal approximation works fairly well for the Cramér-von Mises test, but less well for the Anderson-Darling test due to the very skew distribution of the Anderson-Darling statistic.

5   Other Tests

Below we present a couple of other tests for trend based on robust null hypotheses. Although these tests are constructed for somewhat different conditions than our tests we will include these tests for comparison in simulations and real data examples.

5.1   Linear rank test

A general presentation of linear rank tests can be found in Kalbfleisch and Prentice (2002, chap. 7.2). We shall here consider the specific linear rank test with exponential ordered scores introduced by Cox and Lewis (1966, chap. 3.4) and used in Lawless et al. (2012).

This test is constructed for event censored processes, and we let n_j denote the prespecified number of failures in process j. Let X_{ij} be interarrival time number i in process j and let r_{ij} be the rank of X_{ij} in process j. Then the exponential ordered scores for process j are

$$e_{ij} = \frac{1}{n_j} + \cdots + \frac{1}{n_j - r_{ij} + 1}, \qquad i = 1, \ldots, n_j.$$

The test is based on the statistic

$$V_j = \sum_{i=1}^{n_j} e_{ij}(i - (n_j + 1)/2)$$

which under the null hypothesis has mean zero and variance

$$\mathrm{Var}(V_j) = \frac{\sum_{i=1}^{n_j}(i - (n_j + 1)/2)}{\sum_{i=1}^{n_j}(e_{ij} - \bar{e}_j)/(n_j - 1)},$$

where $\bar{e}_j = \sum_{i=1}^{n_j} e_{ij}/n_j$. A combined test based on m processes can then be based on

$$R^m = \frac{\sum_{j=1}^{m} V_j}{\sqrt{\sum_{j=1}^{m} \mathrm{Var}(V_j)}}$$

which under the null hypothesis is asymptotically standard normal both when $m \to \infty$ and when all $n_j \to \infty$ and m is fixed (Lawless et al., 2012).

Although this test is constructed for event censored processes, Lawless et al. (2012) showed that it is still suitable in a time censoring situation, conditioning on $N_j(\tau_j) = n_j$. However, the test is not using the information in the time interval between $T_{N_j(\tau_j)j}$ and τ_j, which is not satisfactory (Aalen and Husebye, 1991).

This version of the linear rank test will only have power against monotonic trend. Other choices of scores e_{ij} or weights in the V_j-statistic can lead to tests with other properties.

5.2   Generalized Laplace Test

For situations with several processes Lawless et al. (2012) presented a score based test which they call the generalized Laplace test since it is similar in form to the Laplace test for Poisson process data. This is a test for time censored data, but under the slightly different null hypothesis that all the m processes have constant rate functions, and with asymptotics as $m \to \infty$. Let $U_j = \sum_{i=1}^{N_j(\tau_j)} T_{ij} - N_j(\tau_j)\tau_j/2$ for $j = 1, \ldots, m$. The test statistic is

$$GL = \frac{\sum_{j=1}^{m} U_j}{\sqrt{\sum_{j=1}^{m} U_j^2}}, \qquad (17)$$

15

which under the null hypothesis is asymptotically standard normal as $m \to \infty$. Like the Laplace and the Lewis-Robinson test this test will only have power against monotonic trend.

6 Simulation Study

We have done various simulations to study and compare the properties of the tests. When we report results for single processes we do not include the Cramér-von Mises test as this test had less power than the Anderson-Darling test, while for several processes we do not include the Anderson-Darling test as the Cramér-von Mises test had better level properties in this case as discussed in Section 4.2. For the extended Lewis-Robinson test we set $a = 1/2$ in (9) as default, but we also explore the impact of other choices in some simulations. We report this test for non-monotonic trend only, as it has inferior power against monotonic trends. For comparison, the linear rank test is included in single process simulations, while the generalized Laplace test is considered in simulations of several processes.

In the reported simulations we estimated rejection probabilities by simulating 100 000 data sets for each choice of model and parameter values, and recorded the relative number of rejections of each test. The standard errors of the simulated rejection probabilities are then ≤ 0.0016. The nominal significance level was set to $\alpha = 0.05$. All simulations were done in R.

To simulate data with trend, we used the trend-renewal processes (TRP) (Lindqvist, Elvebakk and Heggland, 2003) which in short is defined as follows: Let $\lambda(t)$ be a non-negative function defined for $t \geq 0$ and let $\Lambda(t) = \int_0^t \lambda(u)du$. Then the process $T_1, T_2, \ldots$ is a TRP with *trend function* $\lambda(t)$ and *renewal distribution* F, if $\Lambda(T_1), \Lambda(T_2), \ldots$ is an RP with interevent times having the distribution F.

The RP, the non-homogeneous Poisson process (NHPP) and the HPP are all special cases of the TRP. For example, if the trend function is constant, then the TRP is an RP, while if the distribution F is the unit exponential distribution, then the TRP is an NHPP with intensity function $\lambda(t)$. The trend in a TRP is hence governed by the trend function $\lambda(t)$, and by letting the distribution F be any positive-valued distribution, we are left with a large class of processes with trend. In our simulations we will use parameterizations of the TRP where the renewal distribution F is a Weibull-distribution and the trend function is either

16

of so called power law or bathtub type, see Section 6.2 below.

6.1 One Process - Level Properties

First the level properties of the tests were studied by generating data sets from Weibull RPs with shape parameters respectively 0.75 and 1.5, corresponding respectively to a process which is overdispersed and a process which is underdispersed relative to an HPP. In Figure 1 the simulated level of the tests for systems with the expected number of events ranging from 10 to 60 are reported.
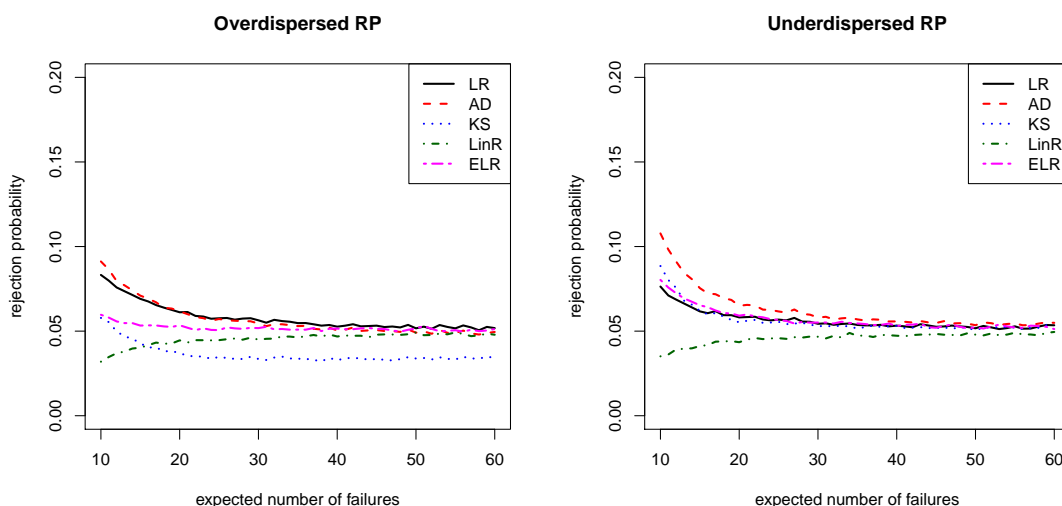

Figure 1: Simulated level properties as a function of expected number of events, with data generated from Weibull RPs with shape parameters respectively 0.75 (overdispersed RP) and 1.5 (underdispersed RP). A nominal significance level of 0.05 was used in all tests. Abbreviations: LR = Lewis-Robinson, KS = Kolmogorov-Smirnov, AD = Anderson-Darling, LinR = Linear rank test, ELR = Extended Lewis-Robinson test.

The tests mostly have adequate level properties, but being based on asymptotic distributions the achieved levels tend to deviate a bit from the nominal level for small sample sizes. In these cases most of the tests are a bit non-conservative, except the linear rank which is a bit too conservative for small samples and the Kolmogorov-Smirnov test which is too conservative in the overdispersed case.

6.2 One Process - Power Properties

Data sets with a monotonic trend were generated by simulating data from TRPs with the renewal distribution being Weibull and the trend function $\lambda(t)$ being of the power law form $\lambda(t) = bt^{b-1}$. The rejection probability as a function of b was simulated, where $b < 1$ corresponds to a decreasing trend, $b = 1$ corresponds to no trend and $b > 1$ corresponds to an increasing trend. Two different values of the shape parameter β of the Weibull renewal

Figure 2: Simulated power properties as a function of trend parameter b, with data simulated from TRPs with trend function bt^{b-1} and Weibull renewal distribution with shape parameters respectively $\beta = 0.75$ (overdispersed TRP) and $\beta = 1.5$ (underdispersed TRP). A nominal significance level of 0.05 was used in all tests. The censoring time was adjusted such that the expected number of failures was 30. Abbreviations: LR = Lewis-Robinson, KS = Kolmogorov-Smirnov, AD = Anderson-Darling, LinR = Linear rank test.

distribution were considered, $\beta = 0.75$ and $\beta = 1.5$. The censoring times were adjusted such that the expected number of failures was 30. The results are displayed in Figure 2. We see in this figure that the Anderson-Darling test is the most powerful test against decreasing trend, but is a bit less powerful than the Lewis-Robinson test for increasing trend. The Kolmogorov-Smirnov and linear rank tests have substantially lower power than the other tests.

Data sets with a bathtub trend were generated by simulating data from TRPs with trend function $\lambda(t)$ on the form displayed in Figure 3. Here e is the average of $\lambda(t)$ over $[0, \tau]$.
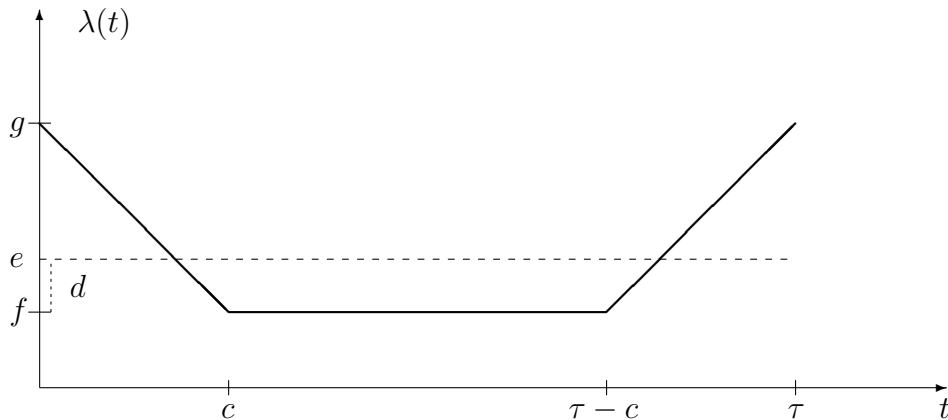

Figure 3: Bathtub-shaped trend function.

The degree of bathtub shape can be expressed by the parameter d, with $d = 0$ corresponding to a horizontal line (no trend).

The rejection probability as a function of d was simulated with c and τ in each case set to values such that the expected number of failures in each phase (decreasing, no, increasing trend) were equal to 20. The shape parameter of the Weibull renewal distribution was set to respectively $\beta = 0.75$ and $\beta = 1.5$. The results are displayed in Figure 4.

We see in Figure 4 that the extended Lewis-Robinson test and the Anderson-Darling test have the ability to detect this non-monotonic trend, while the other tests have no power in such cases. Not surprisingly, the trend is easier to detect in the underdispersed case. The extended Lewis-Robinson test with $a = 1/2$ (9) is by its construction particularly well suited for picking up non-monotonic trends which are symmetric around the mid-point of the observation interval, $\tau/2$, as we have in this case. We see that the performance of this test drops when a less optimal choice of a is made. The Anderson-Darling test is in many cases better than the extended Lewis-Robinson test when $a = 1/3$.

6.3 Several Processes

When considering several processes, the number of processes is one of the important factors for the behavior of the tests. We show here some simulations which illustrate power and
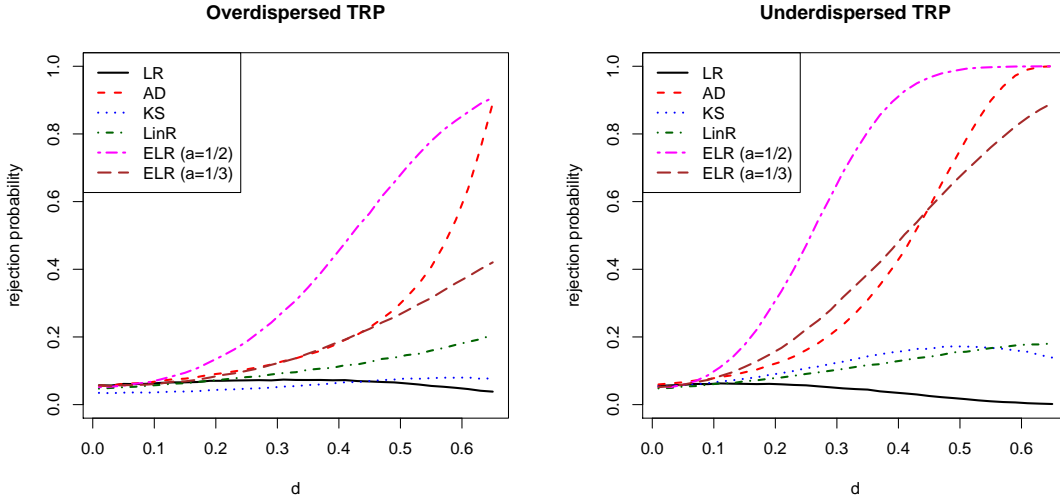
Figure 4: Simulated power properties as a function of trend parameter d, with data simulated from TRPs with bathtub trend function (Figure 3), Weibull renewal distribution with shape parameter $\beta = 0.75$ (overdispersed TRP) and $\beta = 1.5$ (underdispersed TRP), and expected number of failures in each phase equal to 20. A nominal significance level of 0.05 was used in all tests. Abbreviations: LR = Lewis-Robinson, KS = Kolmogorov-Smirnov, AD = Anderson-Darling, LinR = Linear rank test, ELR = Extended Lewis-Robinson test. For ELR results for two choices of a, $a = 1/2$ and $a = 1/3$, are reported.

level properties for the test with different numbers of processes. In this setting with several processes the generalized Laplace test also applies. For the Lewis-Robinson test the weights in (15) were used, and for the Cramér-von Mises weights proportional to τ_j for each process and a normal distribution approximation as discussed in Section 4.2 were used.

Figures 5 and 6 show power properties for cases with respectively 5 and 25 processes and with censoring time chosen such that the expected number of events in each process is 20. Simulations with other expected number of failures showed similar behavior, just with lower or higher power depending on whether the expected number of failures was lower or higher. These simulations show that the Lewis-Robinson type test has the best power properties in these monotonic trend cases. For the generalized Laplace test which has asymptotics as $m \to \infty$ we observe that the achieved significance level using the asymptotic critical value is far too low in the case with 5 processes. For a more fair comparison we have added a version

of the test which achieves the significance level 0.05, obtained by adjusting the critical value empirically. In the case with 25 processes the Lewis-Robinson and the Laplace tests are very similar.
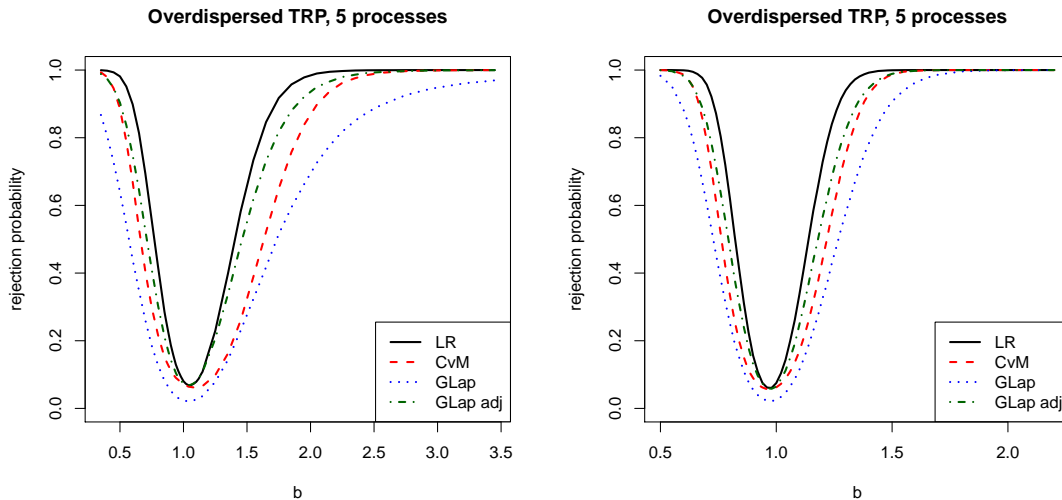

Figure 5: Simulated power properties as a function of trend parameter b, with data simulated from 5 TRPs with trend function bt^{b-1} and Weibull renewal distribution with shape parameters respectively $\beta = 0.75$ (overdispersed TRP) and $\beta = 1.5$ (underdispersed TRP). A nominal significance level of 0.05 was used in all tests. The censoring time was adjusted such that the expected number of failures in each process was 20. Abbreviations: LR = Lewis-Robinson, CvM = Cramér-von Mises, GLap = Generalized Laplace, GLap adj = Generalized Laplace test with adjusted critical value.

7 Asymptotic properties of trend tests under alternatives

In this section we shall complement the simulation study with some asymptotic considerations of the tests studied in this paper. As in the simulation study, we shall assume that data come from trend-renewal processes (TRP) with a power law trend function, $\lambda(t) = bt^{b-1}$, and hence a cumulative trend function $\Lambda(t) = t^b$, where $b > 0$.
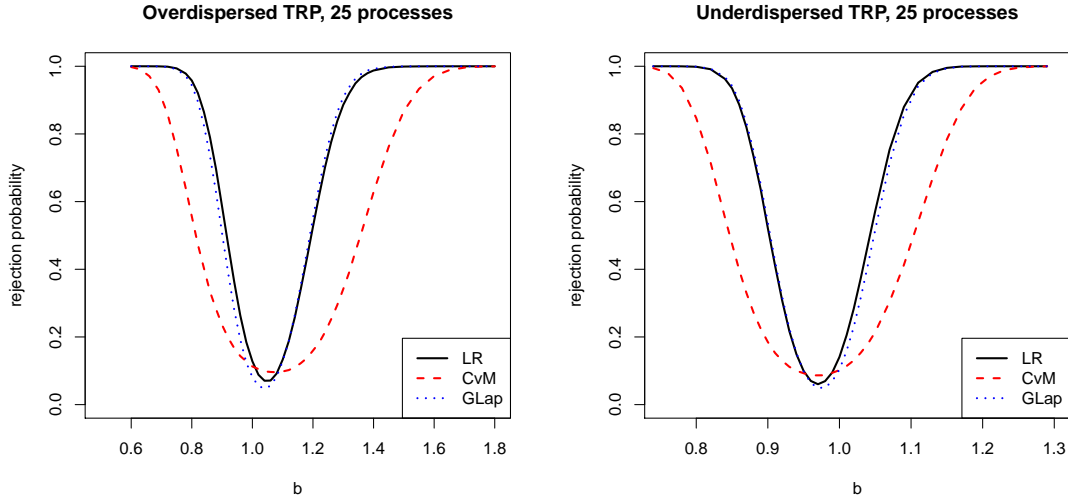
Figure 6: Simulated power properties as a function of trend parameter b, with data simulated from 25 TRPs with trend function bt^{b-1} and Weibull renewal distribution with shape parameters respectively $\beta = 0.75$ (overdispersed TRP) and $\beta = 1.5$ (underdispersed TRP). A nominal significance level of 0.05 was used in all tests. The censoring time was adjusted such that the expected number of failures in each process was 20. Abbreviations: LR = Lewis-Robinson, CvM = Cramér-von Mises, GLap = Generalized Laplace Test.

Suppose first that $T_1, T_2, \ldots$ follow a TRP with general trend function $\lambda(t)$ and renewal distribution F, as defined in Chapter 6. Let $N(t) = \#\{\text{events in } (0, t]\}$ for the TRP. By definition of the TRP, $\Lambda(T_1), \Lambda(T_2), \ldots$ is an RP with interevent distribution F. Let the corresponding counting process be $N^*(u) = \#\{\text{events in } (0, u]\}$. Then

$$N^*(u) = N(\Lambda^{-1}(u)). \tag{18}$$

In Appendix 3 a functional central limit theorem for TRPs is proved, and used to show that for TRPs with the power law trend function as given above, we have

$$\tilde{V}_{t,\gamma}^0(s) - \frac{s^b - s}{\gamma}\sqrt{N(t)} \Rightarrow W^0(s^b) \quad \text{as } t \to \infty, \tag{19}$$

where the left and right hand sides are seen as random functions on $[0, 1]$ and W^0 is the Brownian bridge.

We shall in the following also make use of the weak convergence result

$$W^0(s^b) \Rightarrow W^0(s) \quad \text{as } b \to 1,$$

which can be proven by noting that all finite dimensional distributions converge, and that the family of $W^0(s^b)$ is tight (see, e.g., Billingsley, 1999, Chapter 2).

7.1 The Lewis-Robinson test

7.1.1 The Lewis-Robinson test for single processes

From the definition of the Lewis-Robinson statistic (4) and integration from 0 to 1 in (19), it follows that for $\tau \to \infty$ we have

$$LR - \frac{\sqrt{N(\tau)}}{\hat{\gamma}} \frac{\sqrt{3}(b-1)}{b+1} \xrightarrow{d} -\sqrt{12} \int_0^1 W^0(s^b)ds. \tag{20}$$

Note now that $N(\tau)$ tends to infinity as $\tau \to \infty$, which is a consequence of $N^*(u)$ tending to infinity as $u \to \infty$ for the RP underlying the TRP. Since the right hand side of (20) is a finite (normally distributed) random variable, this shows that LR tends to $+\infty$ or $-\infty$ according to whether $b > 1$ or $b < 1$ when these are fixed values and $\tau \to \infty$. This shows that the LR test is consistent for these kinds of alternatives.

Next, suppose $b \to 1$ as $\tau \to \infty$ in such a way that

$$b = 1 + \frac{\theta}{\sqrt{\tau}} \tag{21}$$

for a given $\theta \in (-\infty, +\infty)$. Then

$$-\sqrt{12} \int_0^1 W^0(s^b)ds \xrightarrow{d} -\sqrt{12} \int_0^1 W^0(s)ds \sim N(0,1),$$

while

$$\frac{\sqrt{N(\tau)}}{\hat{\gamma}} \frac{\sqrt{3}(b-1)}{b+1} = \sqrt{\frac{N(\tau)}{\tau}} \frac{1}{\hat{\gamma}} \frac{\sqrt{3}\,\theta}{2+\theta/\sqrt{\tau}} \xrightarrow{p} \frac{1}{\gamma}\sqrt{\frac{1}{\mu}} \frac{\sqrt{3}}{2}\theta. \tag{22}$$

Here we assume that $\hat{\gamma}$ is a consistent estimator of γ, the coefficient of variation of the underlying renewal distribution. We also use the fact that $N(\tau)/\tau \to 1/\mu$ a.s. as $\tau \to \infty$, where μ is the expected value corresponding to the renewal distribution of the TRP. The argument is as follows, with notation as before,

$$\frac{N(\tau)}{\tau} = \frac{N^*(\tau^b)}{\tau^b} \cdot \frac{\tau^b}{\tau}.$$

Here the first factor tends a.s. to $1/\mu$, since $N^*(\cdot)$ is a renewal process. In the second factor, we use (21) to get

$$\lim_{\tau \to \infty} \tau^{b-1} = \lim_{\tau \to \infty} \tau^{\theta/\sqrt{\tau}} = \lim_{\tau \to \infty} e^{\theta \tau^{-1/2} \ln \tau} = e^0 = 1.$$

23

It follows from the above that the LR-statistic converges in distribution as

$$LR \xrightarrow{d} N\left(\frac{1}{\gamma}\sqrt{\frac{1}{\mu}}\frac{\sqrt{3}}{2}\theta, 1\right).\tag{23}$$

The asymptotic power of an α-level two-sided LR-test, considered as a function of θ defined in (21), is hence

$$\Phi\left(-z_{\alpha/2} + \frac{1}{\gamma}\sqrt{\frac{1}{\mu}}\frac{\sqrt{3}}{2}\theta\right) + \Phi\left(-z_{\alpha/2} - \frac{1}{\gamma}\sqrt{\frac{1}{\mu}}\frac{\sqrt{3}}{2}\theta\right).$$

Note that the expression to the left (right) is dominant if $\theta > 0$ ($\theta < 0$). It is thus seen that the power increases with decreasing γ and μ. Clearly, a low value of μ increases the number of events, while a low γ implies less variation in the interevent times.

7.1.2 The Lewis-Robinson test with several processes

Another application of (23) is to derive optimal weights for the LR test in the case of several processes. The following lemma is probably known, and is easy to prove by a geometric argument.

Lemma 2 *Let $X_j \sim N(\delta_j, 1)$ for $j = 1, \ldots, m$ be independent. Let*

$$W = \sum_{j=1}^{m} a_j X_j, \text{ where } \sum_{j=1}^{m} a_j^2 = 1.$$

Then $W \sim N(\sum_{j=1}^{m} a_j\delta_j, 1)$, where the expected value $\sum_{j=1}^{m} a_j\delta_j$ is maximized by choosing

$$a_j = \frac{\delta_j}{\sqrt{\sum_{k=1}^{m} \delta_k^2}}.$$

Suppose now that we have m systems, each of which is governed by a TRP with cumulative trend function $\Lambda_j(t) = t^{b_j}$ for some $b_j > 0$. Assume further that the jth system is observed on the time interval τ_j, and has renewal distribution F_j with expected value μ_j and coefficient of variation equal to γ_j. We now want to combine LR-test statistics for the m systems as $\sum_{j=1}^{m} a_j LR_j$. Here LR_j is the LR statistic for the jth system and a_j are weights satisfying $\sum_{j=1}^{m} a_j^2 = 1$. In view of (23) and Lemma 2, in the limit as the τ_j tend to infinity, we should choose the a_j such that

$$a_j \propto \frac{\theta_j}{\gamma_j}\sqrt{\frac{1}{\mu_j}}$$

24

in order to optimize the asymptotic power at a certain alternative corresponding to $(\theta_1, \ldots, \theta_m)$, with θ_j defined through $b_j = 1 + \theta_j/\sqrt{\tau_j}$.

Suppose that under the alternative, the $b_j = b$ are the same for all systems, so that $\theta_j = (b-1)\sqrt{\tau_j}$ for each $j = 1, \ldots, m$. Suppose further that the τ_j tend to infinity such that $\tau_j = c_j \tau$ for constants c_j and $\tau \to \infty$. Then clearly $\theta_j \propto \sqrt{c_j}$ for $j = 1, \ldots, m$ and the optimal weights are hence $a_j \propto \sqrt{c_j}/(\gamma_j \sqrt{\mu_j})$.

7.2   Kolmogorov-Smirnov test

The test statistic is

$$KS = \sup_{s \in [0,1]} |\tilde{V}^0_{\tau,\hat{\gamma}}(s)|. \tag{24}$$

Consider now (19). If $b < 1$ ($b > 1$) is fixed, then $(s^b - s)\sqrt{N(\tau)}/\hat{\gamma}$ tends to $+\infty$ ($-\infty$) as $\tau \to \infty$. Since the supremum of the absolute value of the right hand side of (19) has the (finite) Kolmogorov distribution, it follows that the Kolmogorov-Smirnov test is consistent for the considered alternatives.

Now assume that b tends to 1 as $\tau \to \infty$ in the way given by (21). Using that

$$\lim_{b \to 1} \frac{s^b - s}{(b-1)s \ln s} = 1 \tag{25}$$

we obtain

$$\frac{s^b - s}{\hat{\gamma}} \sqrt{N(t)} = \frac{s^b - s}{(b-1)s \ln s} \frac{\theta s \ln s}{\hat{\gamma}} \sqrt{\frac{N(t)}{\tau}} \xrightarrow{p} \frac{s \ln s}{\gamma} \theta \sqrt{\frac{1}{\mu}}.$$

as $\tau \to \infty$. Thus by (19) we have that

$$\tilde{V}^0_{\tau,\hat{\gamma}}(s) \Rightarrow W^0(s) + \frac{s \ln s}{\gamma} \theta \sqrt{\frac{1}{\mu}}$$

and hence by (24) the Kolmogorov-Smirnov test statistic has asymptotically the same distribution as

$$\sup_{s \in [0,1]} \left| W^0(s) + \frac{s \ln s}{\gamma} \theta \sqrt{\frac{1}{\mu}} \right|,$$

where W^0 is a Brownian bridge.

7.3 The Cramér-von Mises test

This test is based on the test statistic $CvM = \int_0^1 (\tilde{V}^0_{\tau,\hat{\gamma}}(s))^2 ds$. Squaring each side of (19) and integrating from 0 to 1 we get

$$CvM - \frac{2}{\hat{\gamma}}\sqrt{N(\tau)} \int_0^1 (s^b - s)\tilde{V}^0_{\tau,\hat{\gamma}}(s)ds + \frac{N(\tau)}{\hat{\gamma}^2} \int_0^1 (s^b - s)^2 ds \xrightarrow{d} \int_0^1 (W^0(s^b))^2 ds \qquad (26)$$

as $\tau \to \infty$. Considering (19) we can write the expression on the left of the $\xrightarrow{d}$ in (26) as

$$CvM - \frac{2}{\hat{\gamma}}\sqrt{N(\tau)} \int_0^1 (s^b - s)\left(\tilde{V}^0_{\tau,\hat{\gamma}}(s) - \frac{s^b - s}{\hat{\gamma}}\sqrt{N(\tau)}\right) ds - \frac{N(\tau)}{\hat{\gamma}^2} \int_0^1 (s^b - s)^2 ds. \qquad (27)$$

Let now τ tend to ∞ in (27). For fixed $b \neq 1$, the rightmost term is given as $N(\tau)$ multiplied by a positive constant. It hence converges to ∞ a.s. as $\tau \to \infty$. By (19), the middle term of (27) equals $\sqrt{N(\tau)}$ times a random variable that converges to a normally distributed random variable. Thus the last term on the left hand side of (27) dominates, making the test statistic CvM converge to infinity as $\tau \to \infty$. This shows that the Cramér-von Mises test is consistent for the given type of alternatives.

Assume next that b tends to 1 as $\tau \to \infty$ in the manner given by (21). The right hand side of (26) then converges to the ordinary Cramér-von Mises distribution. For the last term of (27) we can find the limit as follows:

$$\frac{N(\tau)}{\hat{\gamma}^2} \frac{2(b-1)^2}{(2b+1)(b+2)} = \frac{1}{\hat{\gamma}^2} \frac{N(\tau)}{\tau} \frac{2(b-1)^2\tau}{(2b+1)(b+2)} \xrightarrow{p} \frac{1}{\gamma^2} \frac{1}{\mu} \frac{2}{9}\theta^2$$

which is a constant.

It remains to consider the middle term of (27). Using (25) we get

$$\begin{aligned}
\frac{2}{\hat{\gamma}}\sqrt{N(\tau)} \int_0^1 (s^b - s)\tilde{V}^0_{\tau,\hat{\gamma}}(s))ds &= \frac{2}{\hat{\gamma}}\sqrt{\frac{N(\tau)}{\tau}} \int_0^1 \sqrt{\tau}(b-1)s\ln s\, W^0(s^b)ds \\
&= \frac{2}{\hat{\gamma}}\sqrt{\frac{N(\tau)}{\tau}}\theta \int_0^1 s\ln s\, W^0(s^b)ds \\
&\xrightarrow{d} \frac{2}{\gamma}\sqrt{\frac{1}{\mu}}\theta \int_0^1 s\ln s\, W^0(s)ds
\end{aligned}$$

It finally follows from (26) and (27) that the statistic CvM under (19) converges in distribution to the random variable

$$\int_0^1 (W^0(s))^2 ds + \frac{2}{\gamma}\sqrt{\frac{1}{\mu}}\theta \int_0^1 s\ln s\, W^0(s)ds + \frac{1}{\gamma^2} \frac{1}{\mu} \frac{2}{9}\theta^2,$$

26

as $\tau \to \infty$, where W^0 is a Brownian bridge.

Since $\mathrm{E}[\int_0^1 s \ln s W^0(s) ds] \doteq 0$, it is seen that the expected value of the CvM statistic under the aternative equals the expected value under the null hypothesis plus the positive term

$$\frac{1}{\gamma^2} \frac{1}{\mu} \frac{2}{9} \theta^2.$$

7.4 The Anderson-Darling test

The test statistic and the derivation of its asymptotic properties are similar to the ones for the Cramer-von Mises test. We therefore omit a further treatment of this test.

8 Case Studies

8.1 Load-Haul-Dump Machine Data (Kumar et al., 1989)

Kumar, Klefsjö and Granholm (1989) reported failure data for a load-haul-dump machine operating in a Swedish mine. For the purpose of this example we considered the data to be time censored at $\tau = 2000$ hours. The recorded failure times of the machine up to this time are reported in Table 1, and a plot of the observed process $N(t)$ for $0 \le t \le 2000$ is given in the left panel of Figure 7. The plot seems to indicate a non-monotonic trend, apparently in the form of a bathtub trend.

For illustration we also show, in the right panel of Figure 7, a plot of the function $\tilde{V}_{\tau,1}^0(s)$ for $0 \le s \le 1$. This is the transformed and tied down version of $N(t)$, and should, if the null hypothesis holds, be close to a Brownian bridge. However, this plot too indicates a non-monotonic trend with an upward deviation in the first part and a downward deviation in the second part.

Table 1: Load-haul-dump data. Failure times in hours. The data are time censored at 2000 hours.

16	39	71	95	98	110	114	226	294	344	555	599
757	822	963	1077	1167	1202	1257	1317	1345	1372	1402	1536
1625	1643	1675	1726	1736	1772	1796	1799	1814	1868	1894	1970

Table 2: Load-haul-dump data. Parameter estimates using methods of Section 3.6
.

Estimators	μ	σ	γ
Sample estimators - not including censored time	54.72	48.61	0.888
Sample estimators - including censored time	55.56	47.23	0.850
Parametric: Weibull - including censored time	55.46	47.22	0.851

For estimation of the coefficient of variation under the null hypothesis, we estimated the parameters μ, σ, γ using methods considered in Section 3.6. The results are given in Table 2. It is seen that the estimates which use the censored time are very close, while the one that disregard this time gives a slightly higher estimated coefficient of variation. This might be a coincidence, however, and will not be generally valid.

In order to calculate the LR-test statistic (4), we first calculated the Laplace test statistic, and then divided by the estimated coefficient of variation, to get $0.605/0.888 = 0.681$ using the estimates in the first row of Table 2. This gave the p-value 0.50 for a two-sided test. We also calculated the estimator σ^* of (11), which gave the result 42.77, which is lower than the estimates of σ in Table 2, and would give an estimated coefficient of variation of $42.77/54.72 = 0.782$ and a test statistic of $0.605/0.782 = 0.774$ and a p-value of 0.44. This illustrates the effect of using σ^*, as estimator of σ, as discussed in Section 3.6, namely to possibly give a lower estimated coefficient of variation, and in turn a lower calculated p-value.

Two-sided p-values for all tests are reported in Table 3. In the extended Lewis-Robinson test we used $a = 1/2$, and it is interesting to see that this test detected a significant trend in the data while the tests for monotonic trend had fairly high p-values. The example thus illustrates the need for trend tests with power against non-monotonic trend.

8.2 Hydraulic System of Load-Haul-Dump Machines (Kumar and Klefsjö, 1992)

Kumar and Klefsjö (1992) reported and studied failure data for the hydraulic system of six load-haul-dump machines operating in a Swedish mine. These data were also studied by Lawless et al. (2012) and for the purpose of this example we will follow their analysis and
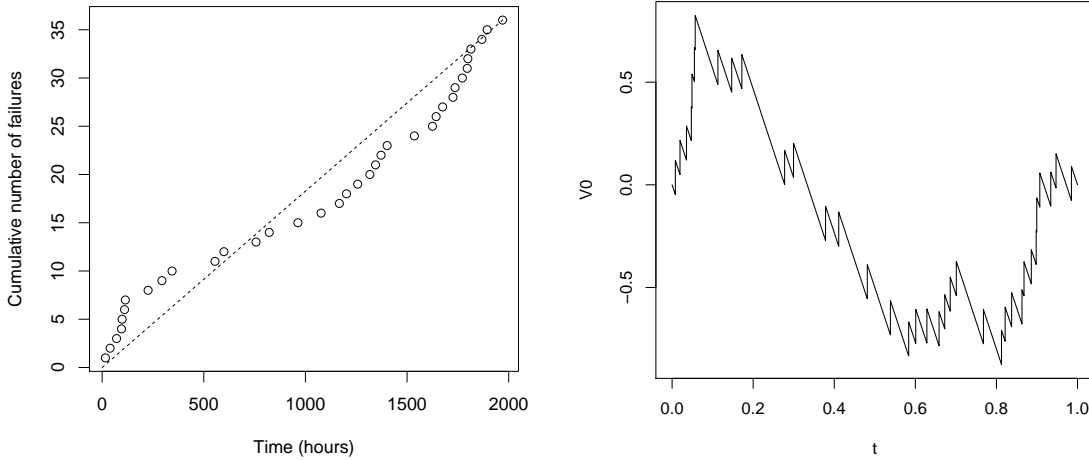
Figure 7: Load-haul-dump data. Left: Plot of $N(t)$ for $0 \leq t \leq 2000$. Right: Plot of the process $\tilde{V}^0_{\tau,1}(s)$ for $0 \leq s \leq 1$.

Table 3: Load-haul-dump data. The table reports p-values for trend tests applied to the load-haul-dump machine data in Table 1. Abbreviations: LR = Lewis-Robinson, KS = Kolmogorov-Smirnov, CvM = Cramér-von Mises, AD = Anderson-Darling, ELR = Extended Lewis-Robinson, LinR = Linear rank. In ELR $a = 1/2$ was used.

LR	KS	CvM	AD	ELR	LinR
0.50	0.29	0.13	0.086	0.011	0.76

assume that for each machine the data are time censored at the last recorded event time.

Figure 8 shows the Nelson-Aalen estimate of the common mean function $E[N(t)]$ for the machines, see Nelson (1988) and Lawless and Nadeau (1995) for the motivation and validity of the plot. As shown by Lawless and Nadeau (1995), the Nelson-Aalen estimator is unbiased and consistent for $E[N(t)]$ under fairly general conditions. Here we present the plot as an illustration of a possible trend in the data.

The p-values obtained for different tests are reported in Table 4. We have here estimated γ individually in each process. For the ordinary and extended Lewis-Robinson tests we used the weights in (15). For the Cramér-von Mises and Anderson-Darling tests we used
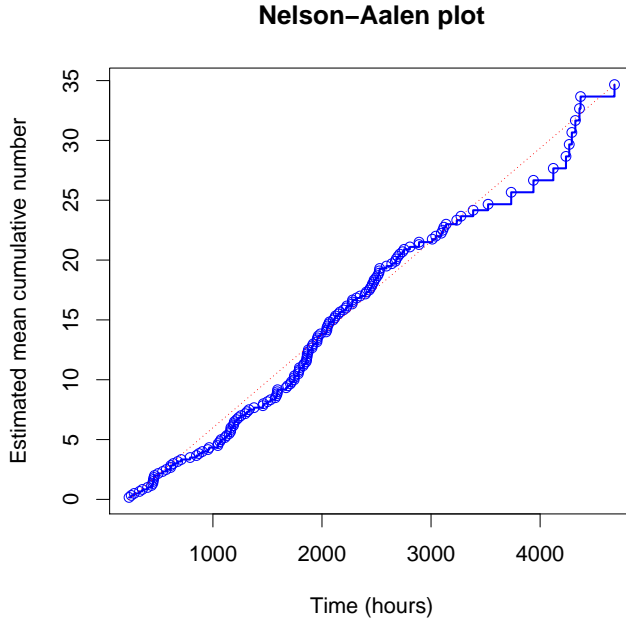
29

Nelson–Aalen plot

Figure 8: Hydraulic system of load-haul-dump machines data. Plot of the the cumulative number of failures over time for the hydraulic system of six machines.

weights proportional to τ_j for each process and simulated the null distribution as discussed in Section 4.2. For the extended Lewis-Robinson test, recall that for a given value of $a \in (0, 1)$, it is assumed that a process, observed on the time interval $[0, \tau]$, changes the direction of a trend at the point $a\tau$. Since we do not have any prior information about a natural turning point for a trend in the present case, we have chosen to let the turning point for all processes be at time $t_0 = 4743/2 = 2371.5$, where $4743 = \max_k \tau_k$. Thus, for the jth system with end of observation at τ_j, we choose the corresponding value a_j of a such that $a_j \tau_j = t_0$.

The two Lewis-Robinson tests identify a significant trend, and the other tests are borderline significant. It should be noted that for the generalized Laplace test the small number of processes ($m = 6$) might imply that the asymptotic normal distribution is not very accurate in the tail in this case as observed in the simulation study. Lawless et al. (2012) report a permutation p-value of 0.037 for this test.

30

Table 4: Hydraulic system of load-haul-dump machines data. The table reports p-values for trend tests applied to the hydraulic system of load-haul-dump machines data in Figure 8. Abbreviations: LR = Lewis-Robinson, CvM = Cramér-von Mises, AD = Anderson-Darling, ELR = Extended Lewis-Robinson, GL = Generalized Laplace.

LR	CvM	AD	ELR	GL
0.019	0.064	0.076	0.003	0.062

9 Conclusion

We have presented a novel class of tests for trend in time censored recurrent event processes, based on the general null hypothesis of an RP. This class includes, among other tests, new versions of the Lewis-Robinson test and the Anderson-Darling test, extending these tests to time censored processes. For the single process case, the Anderson-Darling test turns out to have attractive properties when used as a test for general alternatives, both monotonic and non-monotonic trends. If power against monotonic trends is of main interest, the Lewis-Robinson type test is on the other hand a safe choice, both for single and multiple processes.

The derived test statistics are based on asymptotic results for renewal processes. The calculated critical values are hence only approximate when used in small and medium sized samples. The simulation study shows, however, satisfactory performance of the tests, with some exceptions in cases with very small samples. In such cases an alternative procedure would be to simulate the null distribution of the test statistic by a permutation approach, permuting the order of the completely observed interevent times. Lawless et al. (2012) showed that this is a valid approach even for time censored processes, and we have confirmed this in simulations not reported here.

It is clear that the basic result of Corollary 1 in principle may give rise to a very large class of tests. We have in Section 3 considered four tests based on standard goodness-of-fit statistics, and as an example of the variety of other possible tests we added and studied in some detail a non-standard test, which led to a further extension of the Lewis-Robinson test.

An interesting fact of the constructed test statistics is that they may be viewed as test statistics for the case of Poisson processes, with null hypothesis corresponding to HPP, that are adjusted according to the coefficient of variation of the observed interevent times. This

is exactly the way Lewis and Robinson (1974) obtained their test statistic for the event censored case, starting from the Laplace test.

R code for the tests is available at `https://github.com/jtkgithub/trendtests`.

Acknowledgments

Appendix 1

Consistent Estimator of μ

It is clear from the strong law of large numbers for renewal processes (see, e.g., Ross (1983)) that, as $\tau \to \infty$

$$\hat{\mu} = \frac{\sum_{i=1}^{N(\tau)} X_i}{N(\tau)} \to \mu \text{ a.s.}$$

since $N(\tau) \to \infty$. Note that by standard renewal process theory we have as $\tau \to \infty$

$$\frac{N(\tau)}{\tau} \to \frac{1}{\mu} \text{ a.s.}$$

Thus another consistent estimator of μ is given by $\tilde{\mu} = \tau/N(\tau)$. Note that we can write $\hat{\mu} = T_{N(\tau)}/N(\tau)$, so we have $\tilde{\mu} > \hat{\mu}$.

Consistent Estimator of σ^2

By the strong law of large numbers we have as $\tau \to \infty$

$$\frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} (X_i - \mu)^2 \to \sigma^2 \text{ a.s.}$$

Writing

$$\frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} (X_i - \mu)^2 = \frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} (X_i - \hat{\mu})^2 + (\hat{\mu} - \mu)^2$$

it follows from Slutsky's theorem that

$$\hat{\sigma}^2 = \frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} (X_i - \hat{\mu})^2$$

is a consistent estimator of σ^2.

A disadvantage of the estimator $\hat{\sigma}$, as with $\hat{\mu}$, is that they do not take into account the censored time $\tau - T_{N(\tau)}$. Gallager (2013, chap. 5) shows that

$$\lim_{\tau \to \infty} \frac{\int_0^\tau (t - T_{N(t)})dt}{\tau} = \frac{\mathrm{E}(X^2)}{2\mu} \quad \text{a.s.} \tag{28}$$

Here the left hand side is the long run average length of time since the last previous event, and the result says that this equals $(\mu/2)(1 + \gamma^2)$ where γ is the coefficient of variation of the distribution of X.

We use (28) in the following way. A straightforward calculation shows that

$$\int_0^\tau (t - T_{N(t)})dt = \frac{1}{2} \left\{ \sum_{i=1}^{N(\tau)} X_i^2 + (\tau - T_{N(\tau)})^2 \right\},$$

which after substitution in (28), noting that $\tau/N(\tau) \to \mu$, gives the following consistent estimator for σ^2,

$$\tilde{\sigma}^2 = \frac{1}{N(\tau)} \left\{ \sum_{i=1}^{N(\tau)} X_i^2 + (\tau - T_{N(\tau)})^2 \right\} - \tilde{\mu}^2.$$

An alternative variance estimator, σ^{*2}, was presented in Section 3.6, see equation (11). To prove consistency of σ^{*2} under the null hypothesis of RP, we can consider separately the sum over odd i and even i and use the strong law of large numbers on each of the two resulting sums, which are now sums of i.i.d. variables.

Appendix 2

The Extended Lewis-Robinson Test

The test statistic (9) is obtained as follows. Note first that we can write $N(t) = \sum_{i=1}^{N(\tau)} I_{[T_i,\tau)}(t)$, where $I_A(t)$ is the indicator function of the set A. From (3) it follows that we can consider

$$
\begin{aligned}
\tau \int_0^a (N(s\tau) - sN(\tau))ds &= \int_0^{\tau a} \left(N(t) - \frac{N(\tau)}{\tau} t \right) dt \\
&= \left(\sum_{i=1}^{N(\tau)} \int_0^{a\tau} I_{[T_i,\tau)}(t)dt - \int_0^{a\tau} \frac{N(\tau)}{\tau} t \, dt \right) dt \\
&= \sum_{i=1}^{N(\tau)} (a\tau - \min\{T_i, a\tau\}) - \frac{1}{2}a^2\tau N(\tau) \qquad (29)
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\tau \int_a^1 (N(s\tau) - sN(\tau))ds &= \int_{a\tau}^{\tau} \left(N(t) - \frac{N(\tau)}{\tau} t \right) dt \\
&= \left(\sum_{i=1}^{N(\tau)} \int_{a\tau}^{\tau} I_{[T_i,\tau)}(t)dt - \int_{a\tau}^{\tau} \frac{N(\tau)}{\tau} t \, dt \right) dt \\
&= \sum_{i=1}^{N(\tau)} (\tau - \max\{T_i, a\tau\}) - \frac{1}{2}(1 - a^2)\tau N(\tau). \qquad (30)
\end{aligned}
$$

Subtracting (30) from (29), we get

$$
\sum_{i=1}^{N(\tau)} [(a\tau - \min\{T_i, a\tau\}) - (\tau - \max\{T_i, a\tau\})] + \left(\frac{1}{2} - a^2 \right) \tau N(\tau)
$$

$$
= \sum_{i=1}^{N(\tau)} |T_i - a\tau| - \left(\frac{1}{2} - a(1-a) \right) \tau N(\tau)
$$

The statistic (9) is hence obtained from (3).

We finally prove that $\int_0^a W^0(s)ds - \int_a^1 W^0(s)ds$ is normal with mean 0 and variance $(1/12) - a^2(1-a)^2$. For this we use the fact that, for a Gaussian process $G(t)$ with mean function $E(G(t)) = m(t)$ and covariance function $Cov(G(s), G(t)) = k(s,t)$, we have

$$
\int_a^b G(t)dt \sim N(\int_a^b m(t)dt, \int_a^b \int_a^b k(s,t)dsdt).
$$

The covariance function of the Brownian bridge is $k(s,t) = \min(s,t) - st$. Hence $U = \int_0^a W^0(s)ds$ is normal with mean 0 and variance

$$\int_0^a \int_0^a (\min(s,t) - st)dsdt = \frac{a^3(4-3a)}{12}. \tag{31}$$

Similarly, $V = \int_a^1 W^0(s)ds$ is normal with mean 0 and variance

$$\int_a^1 \int_a^1 (\min(s,t) - st)dsdt = (1/12) - \frac{a^2(6 - 8a + 3a^2)}{12}. \tag{32}$$

Now we want $\mathrm{Var}(U-V) = \mathrm{Var}(U) + \mathrm{Var}(V) - 2\mathrm{Cov}(U,V)$, and thus we seemingly need also $\mathrm{Cov}(U,V)$. We use, however, the following trick. Since $U + V = \int_0^1 W^0(s)ds$ is known to have variance $1/12$, we can solve for $\mathrm{Cov}(U,V)$ in the equation $\mathrm{Var}(U+V) = \mathrm{Var}(U) + \mathrm{Var}(V) + 2\mathrm{Cov}(U,V)$ to get

$$\mathrm{Var}(U-V) = 2(\mathrm{Var}(U) + \mathrm{Var}(V)) - \mathrm{Var}(U+V) = \frac{1}{12} - a^2(1-a)^2$$

where we used (31-32) and the fact that $\mathrm{Var}(U+V) = 1/12$.

Appendix 3

A Functional Central Limit Theorem for Trend-Renewal Processes

Consider a trend-renewal process (TRP) with trend function $\lambda(t)$ and renewal distribution F. As in Section 7, let $N(t) = \#\{\text{events in } (0,t]\}$ for the TRP and $N^*(u) = \#\{\text{events in } (0,u]\}$ for the RP corresponding to F. Then, see (18), $N^*(u) = N(\Lambda^{-1}(u))$.

By Corollary 1,

$$\frac{1}{\gamma} \cdot \frac{N^*(su) - sN^*(u)}{\sqrt{N^*(u)}} \Rightarrow W^0(s) \text{ as } u \to \infty, \tag{33}$$

where the left and right hand sides are seen as random functions of $s \in [0,1]$, and γ is the coefficient of variation of F.

Applying (18) to (33) and substituting $\Lambda^{-1}(u) = t$, we have

$$\frac{1}{\gamma} \cdot \frac{N(\Lambda^{-1}(s\Lambda(t))) - sN(t)}{\sqrt{N(t)}} \Rightarrow W^0(s) \text{ as } t \to \infty. \tag{34}$$

This can be viewed as a functional central limit theorem for the TRP, thus generalizing Corollary 1 to TRPs.

Application to Power Law TRP

Consider now the special case where the cumulative trend function of the TRP is $\Lambda(t) = t^b$ for $b > 0$. Then $\Lambda^{-1}(u) = u^{1/b}$ and (34) becomes

$$\frac{1}{\gamma} \cdot \frac{N(s^{1/b}t) - sN(t)}{\sqrt{N(t)}} \Rightarrow W^0(s) \text{ as } t \to \infty,$$

or, by replacing s by s^b,

$$\frac{1}{\gamma} \cdot \frac{N(st) - s^b N(t)}{\sqrt{N(t)}} \Rightarrow W^0(s^b) \text{ as } t \to \infty, \tag{35}$$

The right hand side is then a time-transformed Brownian bridge and is still a Gaussian process.

The key result for the power law TRP is the following, which is a consequence of (35):

$$\tilde{V}^0_{t,\gamma}(s) - \frac{s^b - s}{\gamma}\sqrt{N(t)} \Rightarrow W_0(s^b) \quad \text{as } t \to \infty,$$

valid for any TRP with cumulative trend function $\Lambda(t) = t^b$ and renewal distribution with coefficient of variation γ, and any fixed $b > 0$.

References

Aalen, O. and Husebye, E. (1991). Statistical analysis of repeated events forming renewal processes, *Statistics in Medicine* **10**: 1227–1240.

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain goodness of fit criteria based on stochastic processes, *Annals of Mathematical Statistics* **23**: 193–212.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit, *Journal of the American Statistical Association* **49**: 765–769.

Ascher, H. and Feingold, H. (1984). *Repairable Systems Reliability. Modeling, Inference, Misconceptions and Their Causes*, Marcel Dekker, Inc., New York.

Billingsley, P. (1999). *Convergence of Probability Measures.*, Wiley Series in Probability and Statistics.

Cohen, A. and Sackrowitz, H. B. (1993). Evaluating tests for increasing intensity of a Poisson process, *Technometrics* **35**: 446–448.

Cook, R. J. and Lawless, J. F. (2013). Concepts and tests for trend in recurrent event processes, *Journal of The Iranian Statistical Society* **12**(1): 35–69.

Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events*, Methuen, London.

Donsker, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov–Smirnov theorems, *Annals of Mathematical Statistics* **23**: 277–281.

Gallager, R. G. (2013). *Stochastic Processes: Theory for Applications*, Cambridge University Press.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edn, Wiley, New York.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione, *Giornale dell' Istituto Italiano degli Attuari* **4**: 83–91.

Kumar, U. and Klefsjö, B. (1992). Reliability analysis of hydraulic systems of lhd machines using the power law process model, *Reliability Engineering and System Safety* **35**: 217–224.

Kumar, U., Klefsjö, B. and Granholm, S. (1989). Reliability investigation for a fleet of load haul dump machines in a Swedish mine, *Reliability Engineering and System Safety* **24**: 341–361.

Kvaløy, J. T. and Lindqvist, B. (2003). A class of tests for renewal process versus monontonic and nonmonotonic trend in repairable systems data, *in* B. Lindqvist and K. Doksum (eds), *Mathematical and Statistical Methods in Reliability*, Series on Quality, Reliability and Engineering Statistics, Vol. 7, World Scientific Publishing, Singapore, pp. 401–414.

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.

Lawless, J. F. and Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent events, *Technometrics* **37**(2): 158–168.

Lawless, J., Çiğşar, C. and Cook, R. (2012). Testing for monotone trend in recurrent event processes, *Technometrics* **54**: 147–158.

Lewis, P. A. W. and Robinson, D. W. (1974). Testing for a monotone trend in a modulated renewal process, *in* F. Proschan and R. J. Serfling (eds), *Reliability and Biometry*, SIAM, Philadelphia, pp. 163–182.

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

Mann, H. B. (1945). Nonparametric tests against trend, *Econometrica* **13**: 245–259.

Nelson, W. (1988). Graphical analysis of system repair data, *Journal of Quality Technology* **20**(1).

Ross, S. M. (1983). *Stochastic Processes*, John Wiley, New York.

Scholz, F. and Stephens, M. (1987). *K*-sample anderson-darling tests, *Journal of the American Statistical Association* **82**: 918–924.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions, *Annals of Mathematical Statistics* **19**: 279–281.

Viertävä, J. and Vaurio, J. K. (2009). Testing statistical significance of trends in learning, ageing and safety indicator, *Reliability Engineering and System Safety* **94**: 1128–1132.