Extending the Use of Some Traditional Trend Tests for Repairable Systems Data by Resampling Techniques

Georg Elvebakk

Abstract

Trend tests for repairable systems data are often based on the Poisson process assumption. As a consequence these tests may lack robustness against non-trend deviations from the homogeneous Poisson process, like renewal processes. A method to extend the null hypothesis of some of these tests, making them more robust, is by using resampling techniques. Compared to other approches, this method seems to work well for small samples, a typical feature of data from repairable systems. Different resampling schemes are presented and a simulation study is performed to explore the properties of the resampled tests.

1 INTRODUCTION

When analysing failure time data from repairable systems it is important to decide if there is any trend in the times between failures. A trend is here defined as a tendency to get either shorter or longer interfailure times (monotonic trends), or that the times may follow any other non-stationary model (non-monotonic trends). Means to detect such trends include graphical techniques, statistical models and statistical tests for trends. This paper concentrates on tests for trends, called trend tests, which is a formal way to decide whether there is a significant trend in the interfailure times.

Trend tests can be categorized by their null and alternative hypotheses. Ascher & Feingold (1984) consider three different null hypotheses: a) homogeneous Poisson process (HPP), b) renewal process (RP) and c) general stationary sequence (GSS). The possible alternative hypotheses are split into two: I) Monotonic trend (MT), and II) Non-monotonic trend (NT). Many of the standard tests in the reliability field are tests for HPP against MT. A problem with these tests is a lack of robustness versus other stationary processes like renewal processes (Lindqvist, Kjønstad & Meland, 1994).

In the category HPP versus monotonic trend we have two well-known tests: the Laplace test and the so-called military handbook test. A test of HPP versus non-monotonic trend is the so-called Anderson-Darling test (Kvaløy & Lindqvist, 1998) (Lewis, 1963). Of course, when using these tests, even if the null hypothesis is rejected it does not necessarily indicate a trend, just that the process is not an HPP. An RP or any other GSS may cause rejection, even if there is no trend present.

Therefore another category to consider is tests of RP versus monotonic trend. In this category there exist in particular several rank tests. One of them, the so-called Mann test (Mann, 1945), has received some attention in the reliability field and is supposed to be a fairly good all-round test (Ascher & Feingold, 1984). It should be noted, however, that rank tests are optimal against an alternative hypothesis of independent interfailure times from the same distribution but shifted in location (Aiyar, Guillier & Albers, 1979). Poisson processes (among others) do not share this property. One possibility to construct tests based on numerical values rather than ranks is to modify tests for the HPP null hypothesis. This approach was taken by Lewis & Robinson (1974) who modified the Laplace test to incorporate renewal processes or general stationary sequences in the null hypothesis. We pay for this extension of the null hypothesis by losing some power for nonhomogeneous Poisson process (NHPP) alternatives compared to the Laplace test. Another alternative is likelihood based tests in models incorporating both renewal processes and trend. Such tests have been constructed by Lawless & Thiagarajah (1996) and Berman (1981) for models in the modulated renewal process (Section 2.2.5) and trend renewal process (Section 2.2.3) classes, respectively.

Our approach to this problem is similar to the approach of Lewis & Robinson (1974). The Lewis-Robinson test is a modified Laplace test constructed such that the significance level under the RP null hypothesis remains asymptotically correct. By using resampling techniques the HPP null hypothesis of this and some of the other most common trend tests can be extended such that the significance level is correct under an RP null hypothesis, without developing new tests. This method has the advantage of being of general nature and work well particularly for small samples.

The resampling is in each case done under the chosen null hypothesis, using the original data as a basis for a resampling distribution. This gives approximately correct significance level under the extended null hypothesis. In the case of an RP null hypothesis it is possible to obtain correct significance levels for the case of failure truncated data by permutation methods. Other null hypotheses in the GSS class are also possible, restricted by the possibility of resampling the original data under the chosen null hypothesis.

This method of test construction is studied for the tests mentioned above, also in the case of several systems. A comparison of some standard tests with HPP and RP as null hypotheses to the resampled versions is performed in a simulation study. The simulation models include HPP, RP and processes with trend and serial correlation.

The tests are easily performed and the resampling algorithm for each case is outlined.

2 MODELS AND TESTS

2.1 Notation

Consider a single system which is observed from time $t = a$ ($\equiv T_0$). The successive failure times are denoted $T_1, T_2, \ldots$, where T_1 is the time of the first recorded failure.

The interfailure times, or times between failures, are denoted $X_1, X_2, \ldots$, thus

$$X_j = T_j - T_{j-1}, \quad j = 1, 2 \ldots$$

Another way of representing the same information is by the number of failures at any given time, thus:

$$N(t) = \text{number of failures in } (a, t]$$

We define a system observation as failure truncated if the system is observed until a given number of failures, n, has occurred. Likewise it is time truncated if observed until a given time, b. We will focus mostly on failure truncation which allows more straightforward resampling schemes.

We say that a system has no trend if the sequence of interarrival times $X_1, X_2, \ldots$ is a stationary sequence. Thus the interarrival times need not be independent. Ascher & Feingold (1984) discussed the difficulty of a general definition of a trend, but in our simulations we define it to be a non-constant trend function (Sections (2.2.3) and (2.2.4)).

For multiple systems datasets, the systems are numbered $i = 1, \ldots, m$ and so the notation becomes $T_{ij}, X_{ij}, a_i, b_i, N_i(t)$.

2.2 Models

2.2.1 Nonhomogeneous Poisson Process

Let $N(t)$ be the number of failures up to and including time t (as before). $N(t)$ is then a Poisson process with intensity function $\lambda(t)$ if 1) $N(0) = 0$, 2) the number of failures in disjoint time intervals are independent, 3) $P(N(t + \Delta t) - N(t) \geq 2) = o(\Delta t)$ and 4) $P(N(t + \Delta t) - N(t) = 1) = \lambda(t)\Delta t + o(\Delta t)$

If $\lambda(t) = \lambda$ (constant) we get the special case of the homogeneous Poisson process (HPP), otherwise we have a nonhomogeneous Poisson process (NHPP) and by definition a trend.

Two classes of intensity functions have been of special interest in the mathematical modelling of repairable systems. It is the so-called *power law*:

$$\lambda(t) = \alpha\beta t^{\beta-1}, \quad \alpha, \beta > 0, \quad t \geq 0 \tag{1}$$

and *log linear* intensity functions:

$$\lambda(t) = \delta e^{\gamma t}, \quad \delta > 0, \ -\infty < \gamma < \infty, \quad t \geq 0 \tag{2}$$

It is often useful to look at an NHPP in terms of its cumulative intensity function, defined as

$$\Lambda(t) = \int_0^t \lambda(u)du \tag{3}$$

If a stochastic process $T_1, T_2, \ldots$ is an NHPP, the time transformed process $\Lambda(T_1), \Lambda(T_2), \ldots$ is an HPP(1), a homogeneous Poisson process where $\lambda = 1$. This fact is helpful in simulation of NHPPs: if $S_1, S_2, \ldots$ are arrival times from an HPP(1), $\Lambda^{-1}(S_1), \Lambda^{-1}(S_2), \ldots$ are arrival times from an NHPP with intensity $\lambda(t)$. It will also help us to define other processes involving trend.

2.2.2 Renewal Process

$T_1, T_2, \ldots$ is said to constitute a renewal process if the interarrival times $X_1, X_2, \ldots$ are independent and identically distributed according to a distribution function, F. As negative interarrival times are not allowed F must be such that $F(0) = 0$. If F is the exponential distribution we get the special case of the homogeneous Poisson process (HPP).

2.2.3 Trend-Renewal Process

The trend-renewal process (TRP) is defined and studied in Lindqvist (1993) and Elvebakk, Lindqvist & Heggland (1999). It can be seen as an extension of the renewal process and the nonhomogeneous Poisson process, having both as special cases.

$T_1, T_2, \ldots$ is said to constitute a trend-renewal process if $\Lambda(T_1), \Lambda(T_2), \ldots$ is a renewal process. This property is an extension of the NHPP property where the transformed process is no longer restricted to a HPP(1). The process is now defined by the trend function $\lambda(t)$ (NHPP intensity function) and the distribution, F, of the interarrival times of the renewal process. To avoid ambiguities F may be required to have unit expectation.

The conditional intensity function of a TRP given the history $\mathcal{F}_{t-}$ up to, but not including time t may be written as (Lindqvist, 1993)

$$\lambda(t|\mathcal{F}_{t-}) = z(\Lambda(t) - \Lambda(T_{N(t-)}))\lambda(t) \tag{4}$$

where $z(t)$ is the hazard rate of the distribution of interarrival times of the underlying RP, F. In other words the intensity depends on both the age of the system and the transformed time since last failure.

Simulation from the trend-renewal process is similar to NHPP simulation. If $S_1, S_2, \ldots$ are arrival times from a RP with distribution F, then $\Lambda^{-1}(S_1), \Lambda^{-1}(S_2), \ldots$ are arrival times from a TRP with intensity given by the equation above.

The modulated gamma process of Berman (1981) is a TRP model with gamma distribution for F.

2.2.4 Serially Dependent Trend-Renewal Processes

$T_1, T_2, \ldots$ is said to constitute a serially dependent trend-renewal (STRP) process if $\Lambda(T_1), \Lambda(T_2), \ldots$ is a stationary sequence with serially dependent and identically marginally distributed interarrival times. In the special case where the marginal distribution of $\Lambda(T_i) - \Lambda(T_{i-1}),\ i = 1, \ldots$ is unit exponentially distributed we get a serially dependent NHPP.

Simulation from these processes is done in the same manner as TRP simulations. If $S_1, S_2, \ldots$ are arrival times from a serially dependent sequence with marginal distribution F for the interarrival times, $\Lambda^{-1}(S_1), \Lambda^{-1}(S_2), \ldots$ are arrival times from an STRP.

2.2.5 Modulated Renewal Processes

This process is a generalization by Cox (1972) of the renewal process and the nonhomogeneous Poisson process, having both as special cases. The conditional intensity function of a MRP given the history $\mathcal{F}_{t-}$ up to, but not including time t may be written as

$$\lambda(t|\mathcal{F}_{t-}) = z(t - T_{N(t-)})\lambda(t) \tag{5}$$

where $z(t)$ is the hazard rate of the renewal process. Also for this process the intensity depends on the age of the system and the time since last failure. Note that in the TRP process the hazard rate was a function of a transformed time, meaning that the underlying trend function $\lambda(\cdot)$ will affect the intensity also via $z(\cdot)$.

The model of Lawless & Thiagarajah (1996) is a generalization of modulated renewal process.

2.3 Trend Tests

The special tests for trend mentioned in the introduction are defined in this section. These tests will be used later in the simulation section. The tests are defined for one system observed either until n failures, or until time b is reached. To simplify notation we define $\hat{n}$ as

$$\hat{n} = \begin{cases} n & \text{if the process is time truncated} \\ n-1 & \text{if the process is failure truncated} \end{cases}$$

where n in case of time truncation is the observed number of failures, $N(b)$.

2.3.1 The Laplace Test

The null hypothesis of this test is HPP and the alternative is an NHPP with monotonic intensity function. It is well known (Bain, Engelhardt & Wright, 1985) (Cox & Lewis, 1966) that this test is an optimal test if the alternative hypothesis is an NHPP with log linear intensity function (2). The test statistic is

$$L = \frac{\sum_{j=1}^{\hat{n}} T_j - \frac{1}{2}\hat{n}(b+a)}{\sqrt{\frac{1}{12}\hat{n}(b-a)^2}}. \tag{6}$$

The test statistic is asymptotically standard normally distributed under the null hypothesis, and this approximation is said to be very good (Ascher & Feingold, 1984) if $\hat{n} \geq 3$. The null hypothesis is rejected for small or large values of L. If $L > 0$ ($L < 0$) it is an indication of an increasing (decreasing) trend.

2.3.2 The Military Handbook Test

The null hypothesis of this test is HPP and the alternative is an NHPP with monotonic intensity function. This test is an optimal test if the alternative hypothesis is an NHPP with increasing power law intensity (1) (Bain *et al.*, 1985). The test statistic is

$$MH = 2\sum_{j=1}^{\hat{n}} \ln\left(\frac{b-a}{T_j - a}\right) \tag{7}$$

The test statistic is $\chi^2_{2(n-1)}$-distributed under the null hypothesis (*MIL-HDBK-189, Reliability Growth Managment*, 1981). The null hypothesis is rejected for small or large values of Z. Large (small) MH is an indication of increasing (decreasing) trend.

2.3.3 The Anderson-Darling Test

The null hypothesis of this test is HPP and the alternative is a trend, possibly non-monotonic. The test is constructed to test for deviation from an HPP (Lewis, 1963), but may also be used as a trend test in Poisson processes (Kvaløy & Lindqvist, 1998). When used as a trend test it is therefore obviously not robust against renewal processes or serially dependent processes. The test statistic is

$$AD = -\frac{1}{\hat{n}}\left[\sum_{j=1}^{\hat{n}}(2j-1)(\ln\left(\frac{T_j - a}{b-a}\right) + \ln(1 - \frac{T_{\hat{n}+1-j} - a}{b-a}))\right] - \hat{n}. \tag{8}$$

If AD is large the null hypothesis is rejected. An asymptotic distribution was derived by Anderson & Darling (1952). This approximation to the exact distribution is said to be very good even for small samples (Kvaløy & Lindqvist, 1998). A visual inspection of the TTT transformation (Section 2.3.7) may give an indication of the type of trend.

2.3.4 The Lewis-Robinson Test (I)

The null hypothesis of this test is RP and the alternative is a monotonic trend. This test, derived by Lewis & Robinson (1974) is a modification of the Laplace test where L is divided by the estimated coefficient of variation of the interarrival times, $\widehat{CV(X)}$. The test statistic is

$$LR_1 = \frac{L}{\widehat{CV(X)}}. \tag{9}$$

Under the RP assumption LR_1 will be asymptotically standard normally distributed. Lewis & Robinson (1974) performed an extensive simulation study to investigate its properties for modulated renewal processes (Section 2.2.5). Based on a simulation study involving some examples of modulated renewal processes Lawless & Thiagarajah (1996) recommended this test based on its simplisity, robustness and power properties.

2.3.5 The Mann Test

This is a rank test developed by Mann (1945). The null hypothesis of this test is RP and the alternative is a monotonic trend, more precisely a sequence of independent random variables with monotonically increasing or decreasing expectation. The test statistic is computed by counting the number of reverse arrangements, M, among the interarrival times, $X_1, X_2, \ldots, X_n$. A reverse arrangement exists if $X_j < X_k$ for $j < k$:

$$M = \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} I(X_j < X_k) \tag{10}$$

M is approximately normally distributed for $n \geq 10$, there exist tables for $n < 10$.

2.3.6 Another Lewis-Robinson Test (II)

The null hypothesis of this test is GSS and the alternative is a monotonic trend. It is again a modification of the Laplace test. The test statistic is

$$LR_2 = \frac{L}{\sqrt{E(X)\widehat{V'}(\infty)}}. \tag{11}$$

$E(X)$ is the expectation of the interarrival times and $V'(\infty)$ is the asymptotic slope of the variance of $N(t)$. Lewis & Robinson (1974) argued that under the null hypothesis LR_2 is asymptotically standard normally distributed. According to Ascher & Feingold (1984) the number of failures needed to get acceptable power properties would probably exceed the output of most repairable systems.

2.3.7 Multiple Systems

So far all tests have been tests for trend in one system. When testing for trend simultaneously in several systems the null hypothesis must be further specified. Two natural extensions of the HPP null hypothesis:

I) H_0 : All systems follow a common HPP.

II) H_0 : All systems follow possibly different HPPs.

The null hypothesis I) leads to the development of the so-called TTT-based tests for trend (Kvaløy & Lindqvist, 1998). Since all systems follow a common HPP, the failure times of all individual systems may be replaced by the *total time on test* transformation for NHPP data (Barlow & Davis, 1977) for all systems. The resulting (single) process will follow the same HPP as the original systems. We will denote the corresponding tests by *TTT tests* for trend. For example the TTT version of the Laplace test is

$$L_T = \frac{\sum_{j=1}^{\hat{N}} \mathcal{T}_j - \frac{1}{2}\hat{N}(B+A)}{\sqrt{\frac{1}{12}\hat{N}(B-A)^2}} \tag{12}$$

where $\mathcal{T}_j$ is the TTT transformation corresponding to the j'th failure time of the superimposed system, $j = 1, \ldots, N$, and $\hat{N}$ corresponds to $\hat{n}$ for the superimposed system. A and B are the times of observation start and stop for the TTT transformed system. The TTT transformation for NHPP data is defined by:

$$\mathcal{T}(t) = \int_0^t p(u)du \tag{13}$$

where $p(u)$ is the number of systems alive at time u.

The null hypothesis II) suggests that the results of several single system tests could be combined to construct what could be called a *combined test* (Kvaløy & Lindqvist, 1998). There are several ways to achieve this, most commonly by combining test statistics or p-values from single systems. For the Laplace and military handbook tests there are natural generalizations to several systems by combining test statistics:

$$MH_c = \sum_{i=1}^{m} \sum_{j=1}^{\hat{n}_i} \ln \left(\frac{b_i - a_i}{T_{ij} - a_i} \right) \tag{14}$$

$$L_c = \frac{\sum_{i=1}^{m} \sum_{j=1}^{\hat{n}_i} T_{ij} - \sum_{i=1}^{m} \frac{1}{2} \hat{n}_i (b_i + a_i)}{\sqrt{\frac{1}{12} \sum_{i=1}^{m} \hat{n}_i (b_i - a_i)^2}} \tag{15}$$

Because of the less restrictive null hypothesis, tests for the hypothesis II) will generally be weaker than corresponding tests for hypothesis I), when used as tests for I).

3 TREND TESTS BY RESAMPLING

In this section resampling methods will be used to compute the distribution of the test statistic under the extension of the null hypothesis of a HPP to other non-trend situations. Focus will mostly be on the RP null hypothesis. For a general reference to resampling tests see Ripley (1987) or Davison & Hinkley (1997).

3.1 Extension of H_0 to RP

3.1.1 Failure Truncation

We will concentrate on the simplest case first: a single failure truncated system with a fixed number n of failure times, $T_1, T_2, \ldots, T_n$ or equivalently n interfailure times $X_1, X_2, \ldots, X_n$. Our null hypothesis is a renewal process (RP).

Suppose a test of HPP versus monotonic trend is given by a test statistic $W = W(X_1, X_2, \ldots, X_n)$. For simplicity we assume that H_0 is rejected for large values of W, for example the Anderson-Darling test statistic. From its distribution under the HPP assumption we know whether to reject it or not, in other words we know the critical value, k, for a given level of significance, α, when the HPP null hypothesis is true, i.e.

$$P(W \geq k) = \alpha. \tag{16}$$

Or equivalently we may compute a p-value and reject H_0 whenever it falls below the level of significance,

$$P(W \geq w_{obs}) \leq \alpha. \tag{17}$$

Here w_{obs} is the observed test statistic computed from the observations. Our goal is to use the well-known tests for the HPP null hypothesis as tests for the RP null hypothesis, i.e. we want to extend the null hypothesis to include the RP case.

We could have done this easily if we knew the distribution F of the RP under the extended null hypothesis. Monte Carlo simulations could then be applied to estimate the distribution of W under the extended H_0 and estimated p-values computed by:

$$p_{MC} = \frac{\#\{w^{*k} \geq w_{obs}\}}{K} \tag{18}$$

where w^{*k} is the test statistic for the kth Monte Carlo sample simulated under the null hypothesis, $k = 1, \ldots, K$. As K increases the p-value converges toward the correct value.

In our resampling tests we use the same approach, but since our RP null hypothesis does not completely define the distribution, F, of the interarrival times, this distribution must be estimated based on the observations. There are several ways to do this, we will describe two slightly different approaches.

Permutation Test

Under the RP null hypothesis the situation is ideally suited for permutation testing. The null hypothesis implies that the interarrival times $X_1, X_2, \ldots, X_n$ are i.i.d data, distributed according to an unknown distribution, F. The order statistic $S = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ will then be a statistically sufficient statistic, and by conditioning on S we eliminate the unknown F. For a permutation test with significance level α we reject H_0 if

$$P(W \geq w_{obs} | S = s) \leq \alpha \tag{19}$$

As all permutations of S are equally likely this probability can in principle be found by computing W for all permutations of S. In practice, however, since the number of permutations grows very fast as n increases, we will only compute a fraction and thus get approximate significance levels. A resampling scheme for the permutation test will be as follows, to simplify the algorithm we assume that the time of observation start is $a = 0$:

1. Record the interfailure times $\underline{x} = x_1, x_2, \ldots, x_n$.

2. Compute $w_{obs} = W(\underline{x})$.

3. Draw K randomly chosen permutations of $\underline{x} = x_1, x_2, \ldots, x_n$ and compute W for each of them. This gives $\underline{x}^{*k} = x_1^{*k}, x_2^{*k}, \ldots, x_n^{*k}$ and $w^{*k} = W(\underline{x}^{*k})$, $k = 1, \ldots, K$.

4. Calculate the p-value estimate by $p = \frac{\#\{w^{*k} \geq w_{obs}\}}{K}$.

Typical values of K in resampling tests are 1000 - 2000, in the simulation section (Section 4) we always use $K = 1000$.

Bootstrap Test

Let F be the distribution of the interarrival times under H_0. If we knew F we would reject H_0 if $P_F(W \geq w_{obs}) \leq \alpha$ for a given level of significance, α. Since F is not known it is reasonable to replace it by an estimate. The empirical distribution, $\hat{F}$, which will converge to F as $n \to \infty$ is a good candidate. This is equivalent to giving mass $\frac{1}{n}$ to each of the observed interarrival times $x_1, x_2, \ldots, x_n$. H_0 will then be rejected if

$$P_{\hat{F}}(W \geq w_{obs}) \leq \alpha. \tag{20}$$

which can be computed by simulation. The resampling scheme for this test will be a permutation-like bootstrap, the only difference being that the bootstrap test draws with replacements from $x_1, x_2, \ldots, x_n$, the permutation test without. Also the distribution function estimate must satisfy the null hypothesis, which it clearly does. Thus a resampling scheme for the bootstrap test will be as follows, to simplify the algorithm we assume that the time of observation start is $a = 0$:

1. Record the interfailure times $\underline{x} = x_1, x_2, \ldots, x_n$.
2. Compute $w_{obs} = W(\underline{x})$.
3. Draw with replacement K randomly chosen samples of n from $\underline{x} = x_1, x_2, \ldots, x_n$ and compute W for each of them. This gives $\underline{x}^{*k} = x_1^{*k}, x_2^{*k}, \ldots, x_n^{*k}$ and $w^{*k} = W(\underline{x}^{*k})$, $k = 1, \ldots, K$.
4. Calculate the p-value estimate by $p = \frac{\#\{w^{*k} \geq w_{obs}\}}{K}$

The bootstrap and permutation tests in this situation are very similar, in fact the permutation test is just the bootstrap test conditioned on the empirical distribution function $\hat{F}$. They will also give very similar results as will be shown in the simulations of Section 4. It should also be noted that the bootstrap test will have approximately correct level, while the level of the permutation test in principle is exact. Of course in practical situations the limited number of permutations performed will also make the permutation test approximate.

Note that the bootstrap method of this section does not condition on the (stochastic) total observation period. Doing this for the bootstrap test would mean resampling such that the sum of the interfailure times stays the same. A permutation test will have this property but it is difficult to do such a conditioning for bootstrap replications based on the empirical distribution, $\hat{F}$, other than just choosing the permutations as our conditional bootstrap samples. If $\hat{F}$ was chosen continuous (parametric, kernel estimate) it could be done by Markov chain Monte Carlo methods. However as we use the HPP versions of the tests, which are conditioned on the time of the nth failure to eliminate the nuisance scale parameter, it turns out to be of little importance as they are functions of $\frac{T_j}{T_n}$, $j = 1, \ldots, n$. If we instead of the Laplace and military handbook tests used the "basic" test statistics $W = \sum T_j$ and $W = \sum \log(T_j)$ it would be essential to condition on the total time.

To sum up, in the situation with a fixed number of failures, and where the observation period ends with a failure, it is of little importance whether we use bootstrap or permutation tests (see the simulation study of Section 4.1). In other situations it might be impossible to use a permutation test while bootstrap tests still apply. The next section is an example of this.

3.1.2 Time Truncation

In this setting the system is monitored from time a to time b. During that period n failures are observed. Note that n will be a realization of a random variable. Under H_0 we assume that the set of interfailure times is a realization of a renewal process. Since there is a censored interfailure time at the end of the observation period, a permutation test will not directly apply. A bootstrap approach, however, will still apply, and in this time truncation setting it will be slightly different from the method of the previous section.

The first problem is to estimate the distribution of the interarrival times, which now includes a censored observation. A natural choice is to use the Kaplan-Meier estimator of the cumulative distribution (Kaplan & Meier, 1958):

$$\hat{F}(x) = 1 - \prod_{j \in J_x} \frac{(n_j - 1)}{n_j} \tag{21}$$

where J_x is the set of all indices j where $x_{(j)} \leq x$ and $x_{(j)}$ represents an uncensored interarrival time. n_j is the number of units under observation immediately before $x_{(j)}$. This estimator reduces to the empirical distribution function for failure truncated data. If the censored interarrival time is the largest (21) will not be a proper distribution, i.e. $\hat{F}(\infty) < 1$. In this case we have therefore chosen to extrapolate the Kaplan-Meier estimator with a Weibull distribution passing through the point corresponding to the largest uncensored observation in the Kaplan-Meier estimator. The concept is explained in figure 1. Another possible method is to always regard the largest observation as uncensored (Efron, 1967). This problem diminishes as the number of failures in the system grows larger.

A naive resampling scheme would be to try to recreate the process that produced the original dataset without conditioning on the number of failures observed. We would then keep the start and endpoint fixed and draw new observations from the estimated distribution of the interfailure times until the given endpoint (b) is reached. Resampling in this manner means the number of observations is random, as in the original sample. However, the usual Poisson trend tests (e.g. Laplace, military handbook, Anderson-Darling) are conditioned on the observed number of failures, $N = n$, in the observation period. Also the distribution of the military handbook test statistic depends on the number of failures in the observed sample. (The distributions of the other two test statistics considered, the Laplace and Anderson-Darling tests, converge quickly towards their asymptotic distributions as n increases.) Thus when extending the use of such tests to the RP case it might be advantageous to construct the new tests as conditional tests given $N = n$.

In other words we would like to compute a p-value such that

$$p = P_F(W \geq w_{obs} | N = n) \tag{22}$$

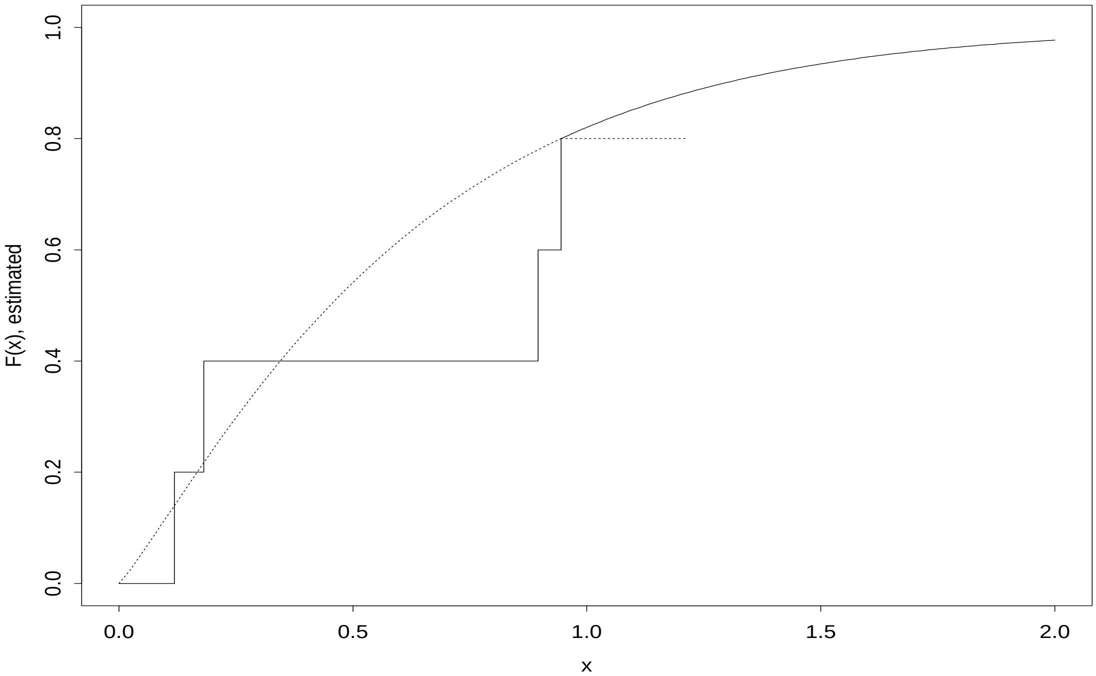

Figure 1: An estimate of the cumulative distribution function (solid line) of the dataset $0.118, 0.181, 0.896, 0.945, 1.212$ where the last observation is censored. Until the last uncensored observation the Kaplan-Meier estimator is used. From then on the estimate is based on a 2-parameter Weibull distribution, fitted by maximum likelihood under the restriction that the distribution function goes through the point of the Kaplan-Meier curve corresponding to the largest uncensored observation.

and reject the null hypothesis of a renewal process whenever it is smaller than our level of significance, α. Again since we do not know F we replace it by $\hat{F}$ and reject H_0 if

$$P_{\hat{F}}(W \geq w_{obs}|N = n) \leq \alpha. \tag{23}$$

The resampling scheme is conditioned on the number of failures observed, so each resampled process should have the same number of failures as the original sample. Thus for a time truncated system the resampling scheme for the bootstrap test may be as follows, to simplify the algorithm we assume that the time of observation start is $a = 0$:

1. Record the interfailure times and the censoring time $\underline{x} = x_1, x_2, \ldots, x_n, b$.

2. Compute $w_{obs} = W(\underline{x})$.

3. Draw randomly K new observation sequences from $\hat{F}$ (the (extrapolated) Kaplan-Meier estimator), each sequence observed until b, conditional on having n failures and

compute W for each of them. This gives $\underline{x}^{*k} = x_1^{*k}, x_2^{*k}, \ldots, x_n^{*k}, b$ and $w^{*k} = W(\underline{x}^{*k})$, $k = 1, \ldots, K$.

4. Calculate the p-value estimate by $p = \frac{\#\{w^{*k} \geq w_{obs}\}}{K}$

The easiest way to execute the third point of the above scheme is to sample with rejection. If a resampled sequence does not contain the correct number, n, of failures it is simply rejected, we continue until there are K correct sequences. It is of course possible to fit a continuous distribution (parametric, kernel estimate) to the interfailure times. Sequences could then be simulated from this distribution under the same conditions specified above.

We call this resampling scheme *conditional* bootstrap. If the third point of the above algorithm is executed without conditioning on the number of failures we call it *naive* or *unconditional* bootstrap.

3.1.3 Multiple Systems

We have so far only considered the situation of a single system. Simultaneous trend testing for a dataset consisting of several systems will increase our options with regard to the resampling schemes. By analogy with the Poisson process case we define two natural null hypotheses:

$$\begin{array}{rl} \text{I)} \;\; H_0: & \text{All } m \text{ systems follow a common RP.} \\ \text{II)} \;\; H_0: & \text{All } m \text{ systems follow possibly different RPs.} \end{array}$$

A consequence of the first hypothesis is that assuming H_0 is true we may use interfailure times from all systems to estimate a common $\hat{F}$. Assuming the second hypothesis is true the estimate of $\hat{F}_i$ for system i should be based on interfailure data from this system only.

Of course there are other possible generalizations of the renewal process hypothesis for a single system. One possibility is to let the distribution of the interfailure times for different systems differ only by a scale parameter (simple heterogeneity). We will however concentrate on the hypotheses I) and II) above.

I. All m systems follow a common RP

Assuming this hypothesis it is tempting to try to extend the TTT tests for trend to the RP case. The TTT transformation of realizations of m systems following a common HPP can be seen as a realization of a single system following the same HPP but observed over the sum of the observation periods of the individual systems. The TTT tests for trend under the common HPP assumption are therefore conditioned on the total number of failures in the sum of observation periods of all systems. So to construct a resampling algorithm for the common RP case it would seem important to try to preserve this property.

However, since the distributions of the TTT test statistics, in the case of a common renewal process, will not only depend on the distribution of the interarrival times, but also on the number of systems, $p(u)$, observed at each point in time, it might be advantageous to preserve the original observation periods of the individual systems in the resampled datasets. This is most easily achieved in the time truncated case, so we start by describing a possible resampling scheme for this situation.

For the time truncation case recall that the interarrival times of the resampled datasets are now drawn from $\hat{F}$, the (extrapolated) Kaplan-Meier estimate of the distribution of interfailure times from the whole dataset (all m systems). The observation period for each resampled system is kept the same as in the corresponding systems of the original dataset and the total number of failures are kept constant, to simplify the algorithm we assume that $a_i = 0,\ \forall i$:

1. Record the interfailure times and observation period of all m systems of the dataset. $\underline{x} = x_{11}, \ldots, x_{1n_1}, b_1, \ldots, x_{m1}, \ldots, x_{mn_m}, b_m$
2. Compute $w_{T,obs} = W_T(\underline{x})$
3. Simulate K sets of m sequences of interarrival times under H_0. The interarrival times are drawn from the (extrapolated) Kaplan-Meier estimate of the distribution of interfailure times of all systems. The observation periods are the same as in the original dataset. The simulation is conditional on observing totally $\sum_{i=1}^{m} n_i$ events and W_T is computed for each of them. This gives $\underline{x}^{*k} = x_{11}^{*k}, \ldots, x_{1n_{1*}}^{*k}, b_1, \ldots, x_{m1}^{*k}, \ldots, x_{mn_{m*}}^{*k}, b_m$ (where $\sum_{i=1}^{m} n_{i*} = \sum_{i=1}^{m} n_i$) and $w_T^{*k} = W_T(\underline{x}^{*k})$, $k = 1, \ldots, K$.
4. Calculate the p-value estimate by $p = \frac{\#\{w_T^{*k} \geq w_{T,obs}\}}{K}$

For failure truncated systems it is not possible to resample new failure truncated systems with the original observation periods for each individual system and same total number of failures. To accomplish this we would at least have to make the additional restriction that the number of failures in the individual systems of the resampled datasets should be the same as in the individual systems of the original dataset. Another possibility would be to resample the systems as if they were time truncated.

However, we have found that a variation of the permutation scheme seems to be a reasonable choice. So we propose a permutation test where the interfailure times of all systems are included in the permutation sample. The permutation is performed such that the individual systems of the resampled datasets will have the same number of failures as in the corresponding systems of the the original dataset. This means that observation period will be stochastic for each system, although the total observation period and total number of failures will be kept constant. Even if the number of systems under observation at each point in time is not exactly the same as in the original dataset this method seems to give quite good results (see Section 4.3.2). To simplify the algorithm we assume that $a_i = 0,\ \forall i$:

1. Record the interfailure times and observation period of all m systems. $\underline{x} = x_{11}, \ldots, x_{1n_1}, \ldots, x_{m1}, \ldots, x_{mn_m}$
2. Compute $w_{T,obs} = W_T(\underline{x})$

3. From the dataset $x_{11}\ldots, x_{1n_1}, \ldots, x_{m1}, \ldots, x_{mn_m}$ draw K randomly chosen permutation sets and compute W_T for each of them.
   This gives $\underline{x}^{*k} = x^*_{11}, \ldots, x^*_{1n_1}, \ldots, x^*_{m1}, \ldots, x^*_{mn_m}$ and $w_T^{*k} = W_T(\underline{x}^{*k})$, $k = 1, \ldots, K$.

4. Calculate the p-value estimate by $p = \frac{\#\{w_T^{*k} \geq w_{T,obs}\}}{K}$

II. All m systems follow possibly different RPs
Assuming this hypothesis a possible way to proceed would be to use the methods of Section 3.1.1 or 3.1.2 and compute a p-value for each systems, $P_1, P_2, \ldots, P_m$. Under H_0 these will be m realizations of a uniform (0,1) variable. We then have that

$$-2\ln P_i \sim \chi^2_2$$

so a possible test statistic would be

$$Z = -2\sum_{i=1}^{m} \ln P_i \sim \chi^2_{2m} \;\text{ under } H_0 \tag{24}$$

Thus H_0 would be rejected when $Z \geq \chi^2_{2m}(1-\alpha)$, the $1-\alpha$ quantile of the chi-square distribution with $2m$ degrees of freedom.

An alternative to this approach is to compute a combined test statistic, for example the combined military handbook test for possibly different homogeneous Poisson processes (14), and find the distribution of this under H_0 by individually resampling all systems with the methods of Section 3.1.1 or 3.1.2. This will probably give tests that have correct level under H_0 but will most likely have poor power properties when the distribution of the interarrival times vary considerably between systems, it is particularly sensible to variation in the coefficient of variation (CV). The reason for this is that the distribution of the standard test statistics under the renewal process null hypothesis depends on F, the distribution of the interarrival times. So an indication of trend in a low CV process could be obscured by the variation in a large CV process.

3.2 Serial Dependence

To extend the null hypothesis to include a serially dependent stationary sequence a resampling scheme designed to preserve this property must be employed. In the time domain there are two approaches to resampling of time series data, model based and model free. Two of the most well known examples of these approaches will be briefly described here. A detailed description can be found in Efron & Tibshirani (1993) or Davison & Hinkley (1997).

In a model based method a model is fitted to the data from which residuals can be computed. Resampled series are then generated by using the fitted model and random samples of the estimated residuals. A simple example of this approach is to assume a 1-step autoregressive model (AR1), where:

$$Z_j = X_j - E(X_j) \qquad \text{and} \qquad Z_j = \beta Z_{j-1} + \epsilon_j$$

where $\epsilon \sim F$ and $E(\epsilon_j) = 0$, i.e. white noise.

To bootstrap from this model, β and $E(X_j)$ must be estimated from the data. It is then possible to compute estimates for $\epsilon_j, \;\; j = 2, \ldots, n$, these will define an empirical distribution for F. Assuming the estimated model is correct new series can then be constructed by drawing a random sample of the residuals and using the computed estimates for β and $E(X_j)$.

The most common model free method for bootstrapping in stationary sequences is the so-called block bootstrap (Hall, 1985) (Carlstein, 1986). The interarrival times are split into as many overlapping or non-overlapping blocks of a certain size, d, as possible. A block then consists of d interarrival times in succession. Resampling is performed by sampling the number of blocks necessary to get n interarrival times in succession.

The idea of this scheme is that long enough blocks will preserve enough of the original dependence in the data. However, due to the joins between the blocks, the resampled series will be closer to white noise than the original series. Several methods have been devised to lessen the "whitening" effect of the joins between blocks. In one of these, the so-called "post-blackening" method (Davison & Hinkley, 1997), the block bootstrap is performed on the residuals of a fitted model.

Choice of block size is a complicated question as the optimal block size will depend on the dependence structure of the series as well as the statistic of interest. In addition, in a trend testing situation, long blocks will preserve more of the trend, and thus most likely be a disadvantage to the trend-detecting ability (power) of the test.

3.3 P-value Adjustments

In contrast to permutation tests, bootstrap tests only have asymptotically correct significance levels (when $n \to \infty$). For situations where there are no permutation tests available a resampling p-value adjustment technique (Beran, 1988) (Davison & Hinkley, 1997) may be applied. This is a very computer intensive method as it involves nested (double) bootstrap.

From the bootstrap replications of Section 3.1.1 a p-value is computed:

$$p = P_{\hat{F}}(W \geq w_{obs}) = \lim_{k \to \infty} \frac{\#\{w^{*k} > w_{obs}\}}{K} \tag{25}$$

The p-value may be considered a realization of a test statistic, P, which ideally should be uniformly distributed on [0,1] under H_0, i.e. thus:

$$P(P \leq p) = p \tag{26}$$

For approximate tests, however, this is not the case. In short, p-value adjustment involves estimation of the distribution of P under H_0 and assessment of the real significance of w_{obs} by comparing the computed p to this distribution. This is computer intensive because a resampling distribution for P is obtained by resampling the already resampled datasets. For each re-resampling a p-value of w^{*k} is computed by comparing w^{*k} to its resampling distribution under H_0. This can be thought of as a simulation study under H_0, where the result is an estimate of the distribution of the p-values of the statistic of interest. The adjusted algorithm for the bootstrap scheme of Section 3.1.1:

1. Record the interfailure times $\underline{x} = x_1, x_2, \ldots, x_n$.
2. Compute $w_{obs} = W(\underline{x})$.
3. Draw with replacement K randomly chosen samples of size n from $\underline{x} = x_1, x_2, \ldots, x_n$ and compute W for each of them. This gives $\underline{x}^{*k} = x_1^{*k}, x_2^{*k}, \ldots, x_n^{*k}$ and $w^{*k} = W(\underline{x}^{*k})$, $k = 1, \ldots, K$.
   Calculate the p-value estimate by $p = \frac{\#\{w^{*k} \geq w_{obs}\}}{K}$
   For each k:
   (a) Record the interfailure times $\underline{x}^{*k} = x_1^{*k}, x_2^{*k}, \ldots, x_n^{*k}$ and w^{*k}.
   (b) Draw with replacement L randomly chosen samples of size n from $\underline{x}^{*k} = x_1^{*k}, x_2^{*k}, \ldots, x_n^{*k}$ and compute W for each of them. This gives $\underline{x}^{**kl} = x_1^{**kl}, x_2^{**kl}, \ldots, x_n^{**kl}$ and $w^{**kl} = W(\underline{x}^{**kl})$, $l = 1, \ldots, L$.
   (c) Calculate the p-value estimate by $p_k^* = \frac{\#\{w^{**kl} \geq w^{*k}\}}{L}$
4. Calculate the adjusted p-value estimate by $p_{adj} = \frac{\#\{p_k^* \leq p\}}{K}$

Generally we do not need as many samples for the second loop as for the first, according to Davison & Hinkley (1997) 100 would generally be satisfactory and 250 would be safe.

In our experience this technique has been most useful for small samples where $\hat{F}$ generally is a poor estimate of F. Section 4.1.3 contains some results of small sample simulations with p-value adjustment.

4 SIMULATION STUDY

A simulation study was performed to investigate the properties of the resampling tests described in the previous section. The first part of the study deals with a single failure truncated system governed by a TRP process. We also consider time truncated systems, serially dependent processes and the several systems case. Very small samples are also discussed.

4.1 TRP with Failure Truncation

We have simulated trend renewal processes (TRPs) with trend function given by a power law or log linear law intensity, while the interarrival times of the corresponding RP follow a Weibull distribution. Certain choices of parameters will therefore produce the special cases of NHPP, a Weibull-distributed RP, or an HPP.

The one-parameter Weibull distribution with density

$$f(x) = wx^{w-1}e^{-(x)^w}, \qquad t > 0, \quad w > 0 \tag{27}$$

may be DFR (decreasing failure rate, $w < 1$) IFR (increasing failure rate, $w > 1$) or may degenerate to an exponential distribution ($w = 1$). This will lead to an RP that is an HPP ($w = 1$), overdispersed relative to an HPP ($w < 1$) or underdispersed relative to an HPP ($w > 1$). In the simulation study we have used three values of w: 0.75, 1.00 and 1.5.

For power law intensity (1) the parameter β controls the trend. If $\beta = 1$ there is no trend; if $\beta < 1$ ($\beta > 1$) there is a decreasing (increasing) trend. For the log linear intensity (2) $\gamma = 0$ corresponds to no trend, while for $\gamma > 0$ there is an increasing trend.

Simulations were performed for four different numbers of failures, 5, 15, 35 and 100. For each combination of parameters, 10,000 TRP realizations were simulated. For each simulated process five tests were used: Laplace, military handbook, Lewis-Robinson(I), Anderson-Darling and the Mann test. In addition we also used the resampled versions (bootstrap and permutation) of those tests. The number of resampling replications (K) was always 1000. The results of the study are the relative frequencies of rejection of H_0, interpreted as rejection probabilities, the nominal level is always 5%.

Some of the more important results are presented graphically in this section. In addition the results of all simulations are listed in Tables 2-25 in the appendix.

4.1.1 Standard versus Resampled Tests

In Figure 2 we compare the rejection probabilities of the standard tests to the resampled versions of the same tests. We present here the results of simulations with 15 failures, which we consider a small to moderate sample size for repairable systems data. The results for the Laplace test are plotted in the upper row, the military handbook test in the middle row, and the Anderson-Darling test in the lower row. Results for the Lewis-Robinson test are added to the plots for the Laplace test.

We will first concentrate on the results for the Laplace and Lewis-Robinson tests (upper row). For the NHPP simulation $\beta = 1$ means an HPP. In this case the rejection probabilities for all tests, standard as well as resampled, are close to the nominal level of 0.05. This is to be expected since all these tests have null hypotheses covering the HPP case. For increasing trend ($\beta > 1$) as well as decreasing trend ($\beta < 1$) we observe that the resampled tests lose some power compared to the standard Laplace test. This is the price we have to pay for the extension of H_0 to RP. The loss of power diminishes, however, as the number of failures increases. This can be seen from simulations with large number of failures (Figure 7). We also make the interesting observation that the resampled Laplace tests seem to be somewhat more powerful than the Lewis-Robinson test for moderately sized samples.

We will refer to the TRP simulations with Weibull parameters $w = 0.75$ or $w = 1.5$ as overdispersed and underdispersed TRPs, respectively. For overdispersed TRP the (standard) Laplace test has too high rejection probability for $\beta = 1$ (RP). The resampled tests and the Lewis-Robinson test are again close to 0.05. As a consequence the Laplace test will have higher rejection probability also in the case of increasing or decreasing trend. Again we observe that the resampled tests seem to be more powerful than the Lewis-Robinson test. For underdispersed TRP there is a similar pattern. But here the Laplace test has too low actual level, making it less powerful also for moderate trend alternatives.

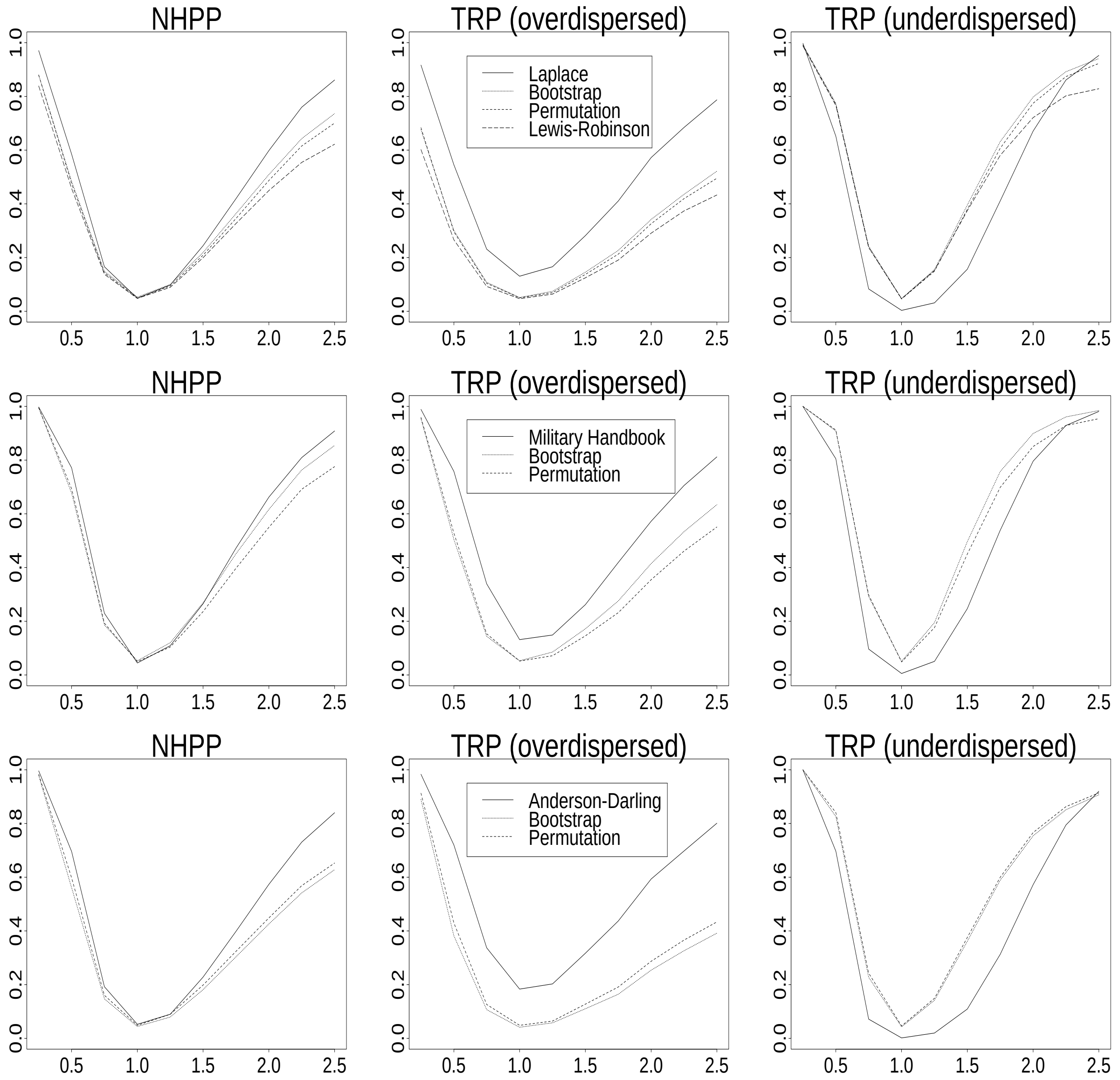

Figure 2: Upper row: The Laplace, bootstrapped Laplace, permuted Laplace and Lewis-Robinson tests for trend. Middle row: The military handbook, bootstrapped military handbook and permuted military handbook tests for trend. Lower row: The Anderson-Darling, bootstrapped Anderson-Darling and permuted Anderson-Darling tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (27) using three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 failure truncated processes with 15 failures are simulated, the resulting rejection probabilities are plotted.

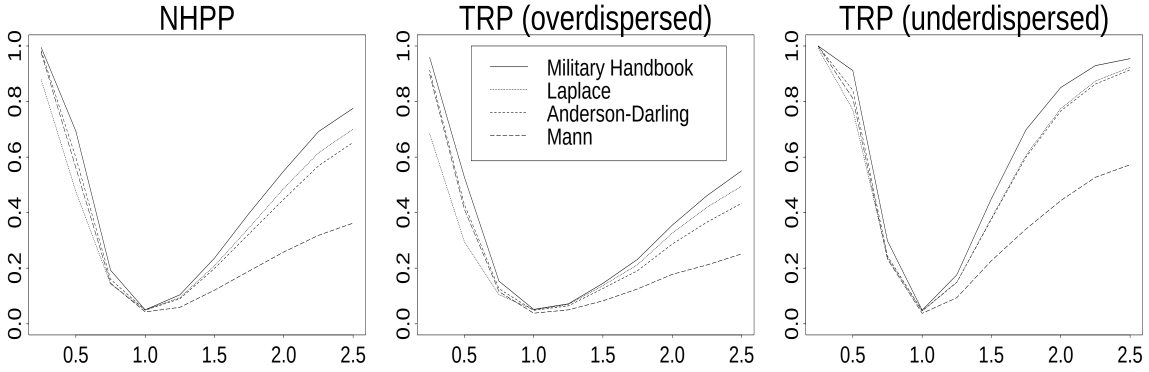

Figure 3: The permuted military handbook, permuted Laplace, permuted Anderson-Darling and Mann tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (27) using three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 failure truncated processes with 15 failures are simulated, the resulting rejection probabilities are plotted.

From the the middle and lower row of Figure 2 we observe the same pattern for the military handbook and Anderson-Darling tests as for the Laplace test. In the NHPP simulations all tests have correct level, but the resampled tests lose some power against trend alternatives. For overdispersed TRP the standard tests have too high actual level, while the resampled tests are close to the correct 5%. For underdispersed TRP the standard tests have too low actual level, while the resampled tests again seem to be correct.

4.1.2 Resampled Versions of Different Tests

As noted in Section 2.3.1, in a NHPP setting the standard Laplace and military handbook tests are optimal tests against log linear and power law trend respectively. In Figure 3 and 4 we compare the power of the permuted versions of these tests in a TRP setting. Thus we have plotted the estimated power functions of the permuted versions of those tests, and we have also added the permuted Anderson-Darling test and the non-parametric Mann (rank) test.

In Figure 3 we consider a power law TRP processes. For all three choices, NHPP, overdispersed TRP and underdispersed TRP the permuted military handbook test is the most powerful. This is expected since it is known to be the optimal test for HPP versus power law NHPP. We observe a similar pattern for log linear TRP processes in Figure 4 . The permuted Laplace is the most powerful test, just like the standard test in the NHPP situation.

The non-parametric Mann test is quite good for detecting decreasing trend, while for increasing trend it loses considerable power compared to the resampled tests (Figures 3 and 4).

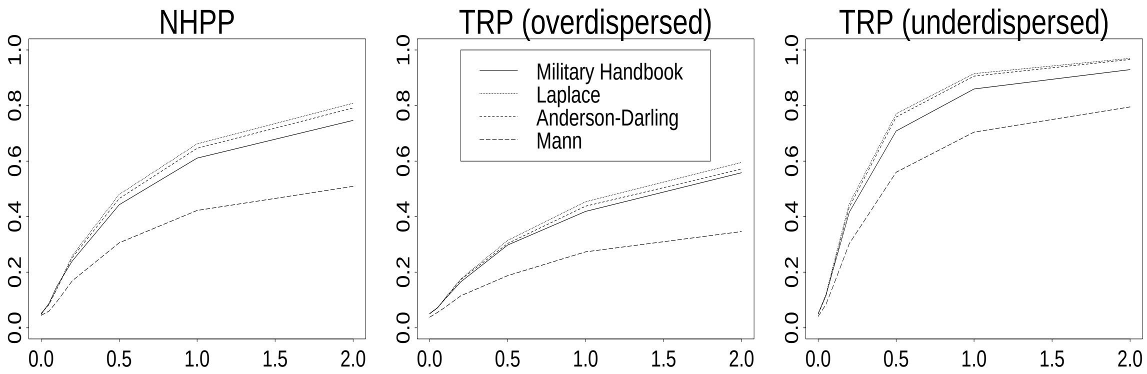

Figure 4: The permuted military handbook, permuted Laplace, permuted Anderson-Darling and Mann tests for trend. Simulations from a TRP model with log linear trend function $\lambda(t) = e^{\gamma t}$ using several values for γ (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (27) using three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 failure truncated processes with 15 failures are simulated, the resulting rejection probabilities are plotted.

4.1.3 Very Small Samples

As an example of test behaviour for very small samples we have simulated processes with five failures. From the tables in the appendix we can observe that the rejection probabilities of the bootstrapped tests under H_0 are not 5% anymore. They are in our simulations typically a bit low (2%-4%). In Figure 5 we have therefore compared the Laplace test and the Lewis-Robinson test with the permuted Laplace, which will always maintain correct level.

The plots in Figure 5 show the same pattern as before. The permuted test has correct rejection probability under the RP null hypothesis, while the standard Laplace test is only correct under the HPP null hypothesis. However, when using resampled tests with very few failures considerable power is lost for trend alternatives. Also in this situation the standard Lewis-Robinson test is practically useless, obviously the distribution of the test statistic has not converged to the asymptotic distribution. The Lewis-Robinson test could of course also have been permuted, but the permuted Lewis-Robinson test is exactly the same as the permuted Laplace or indeed a permuted test based on the test statistic $W = \sum T_j$.

For situations where there are no permutation tests available the resampling p-value adjustment technique of Section 3.3 may be applied. This method seems to work well in order to correct rejection probabilities under H_0 for small sample situations.

From Tables 2-7 in the appendix the Anderson-Darling test seems to be the most biased, the estimated rejection probability under H_0 is about 0.02. Figure 6 displays histograms of p-values under H_0 from bootstrap tests of 10,000 simulations of underdispersed power law TRPs with and without p-value adjustment.

From the left plot of Figure 6 we see that while the bootstrap test has too low rejection probability under H_0, the p-value adjusted bootstrap tests are close to the nominal level of 5%. This also propagates to the trend cases: The adjusted bootstrap tests are more

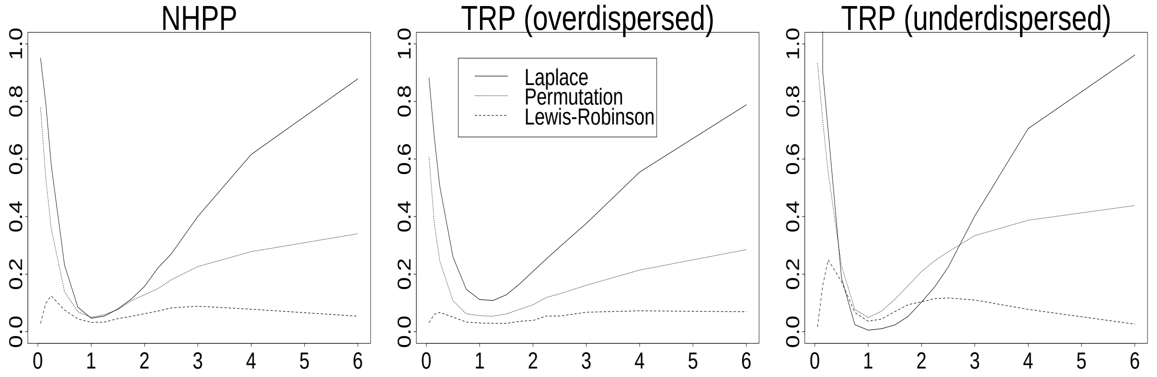

Figure 5: The Laplace, permuted Laplace and Lewis-Robinson (I) tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (27) using three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 failure truncated processes with five failures are simulated, the resulting rejection probabilities are plotted.

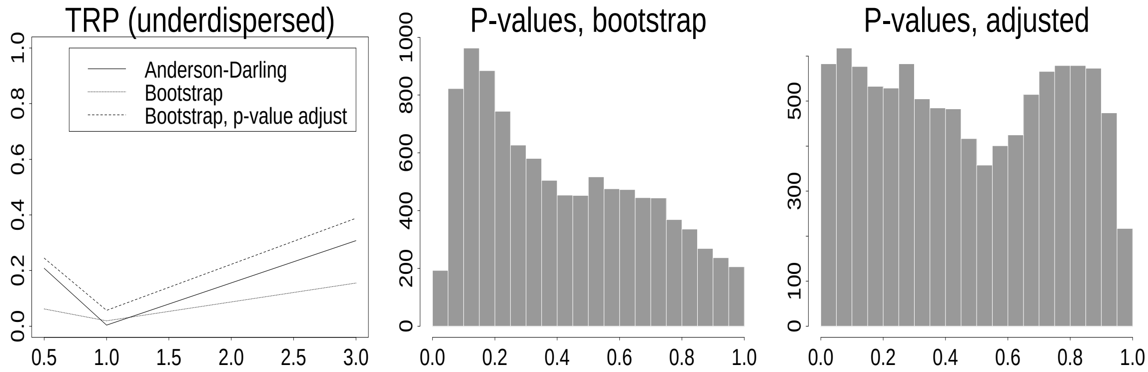

Figure 6: The left plot shows Anderson-Darling, bootstrapped Anderson-Darling and bootstrapped p-value adjusted Anderson-Darling test for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using three different values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (27) with with $w = 1.50$ (underdispersed). For each parameter combination 10,000 failure truncated processes with five failures are simulated, the resulting rejection probabilities are plotted. The plot in the middle shows the estimated p-values of the bootstrap test for the 10,000 simulations in the H_0 case ($\beta = 1$). The right plot shows the estimated p-values of the p-value adjusted bootstrap. For the second loop of the p-value adjustment 100 replications are used.

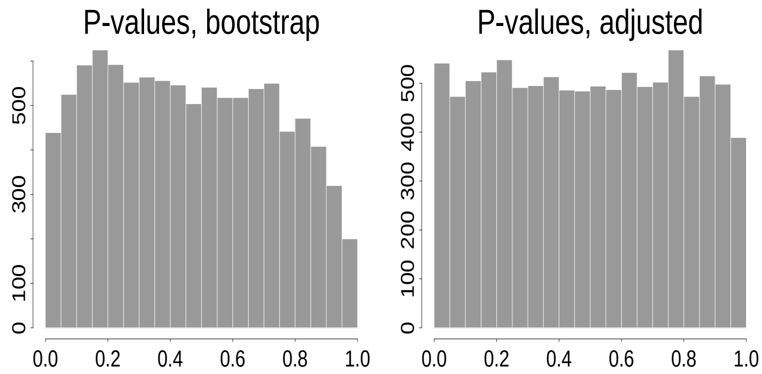

Figure 7: Simulations from an RP model, the interarrival times are Weibull distributed (27) with $w = 1.50$ (underdispersed). 10,000 failure truncated processes with 15 failures are simulated. The left plot shows the estimated p-values of the Anderson-Darling bootstrap test for the 10,000 simulations. The right plot shows the estimated p-values of the p-value adjusted bootstrap. For the second loop of the p-value adjustment 100 replications are used.

powerful than the standard bootstrap. In fact if we compare them to the permutation tests (Table 6 in the appendix) we see that the rejection probabilities of the p-value adjusted bootstrap in this case is quite close to the permutation rejection probabilities.

Histograms of p-values under H_0 are shown in the middle and right plot in Figure 6. The distribution of p-values in the bootstrap case is far from uniform, which explains the low rejection probability. The adjustment makes a large difference, but it is still not perfectly uniform. We have used 100 replications in the second loop of the p-value adjusted bootstrap, the results may have been even better for a larger number. The procedure may also be iterated.

P-values for bootstrap tests are more uniformly distributed for larger number of failures. For a comparison, we have repeated the simulations for p-values and adjusted p-values for the bootstrapped Anderson-Darling test when the number of failures is increased to 15 (Figure 7). It is clear from the plots that the distribution of p-values for the unadjusted bootstrap test is more uniform for this sample size. Still, even here the p-value adjustment seems to make it more uniform.

4.1.4 Large Samples

As an example of large sample behaviour of the tests we have simulated processes with 100 failures. In this situation bootstrapped and permuted tests give approximately the same result, so in Figure 8 we again compare the standard Laplace test and the Lewis-Robinson test with the permuted Laplace test.

From the plots we see that the resampled (permuted) Laplace test is now almost indistinguishable from the Lewis-Robinson test for all three cases. Also the power of the standard Laplace test is very close to the power of the two resampling tests in the case of NHPP. For TRP (over or underdispersed) we observe the same pattern as before, the

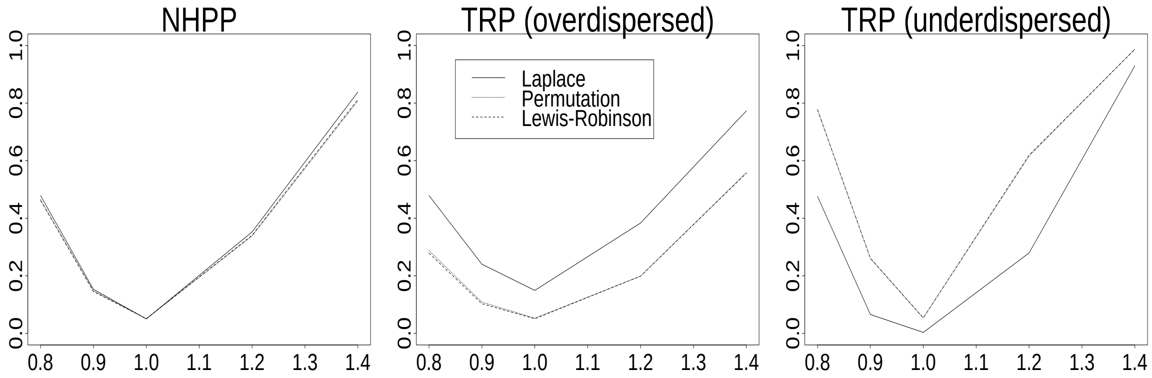

Figure 8: The Laplace, permuted Laplace and Lewis-Robinson tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (27) using three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 failure truncated processes with 100 failures are simulated, the resulting rejection probabilities are plotted.

standard Laplace test has incorrect level under the RP null hypothesis

4.2 TRP with Time Truncation

We have simulated trend renewal processes (TRPs) with trend function given by a power law function, while the interarrival times of the corresponding RP follow a Weibull distribution. Certain choices of parameters will therefore produce the special cases of NHPP, a Weibull distributed RP, or a HPP.

The Weibull distribution is slightly modified from (27). To simplify the computation of expected number of failures in the observation period we choose to let $E_F(X) = 1$ for all possible parameter values. The density function becomes:

$$f(x|w) = w(\Gamma(\frac{1}{w}+1))^w x^{w-1} e^{-(\Gamma(\frac{1}{w}+1)x)^w}, \qquad x > 0, \quad w > 0 \tag{28}$$

In the simulation study we have used three values of w: 0.75, 1.00 and 1.5, resulting in overdispersed RP, HPP and underdispersed RP, respectively.

We performed simulations for expected number of failures equalling 10, 15 and 35. For each simulated process we tested for trend by three tests: Laplace, military handbook and Anderson-Darling. In addition we also used the resampled versions (unconditional and conditional bootstrap) of those tests. Some of the results are presented graphically, all are listed in Tables 26-34 in the appendix.

4.2.1 Standard versus Resampled Tests

The results mostly follow the same pattern as for failure truncated data. For NHPPs the resampled tests are weaker than the standard tests for trend alternatives, the loss of

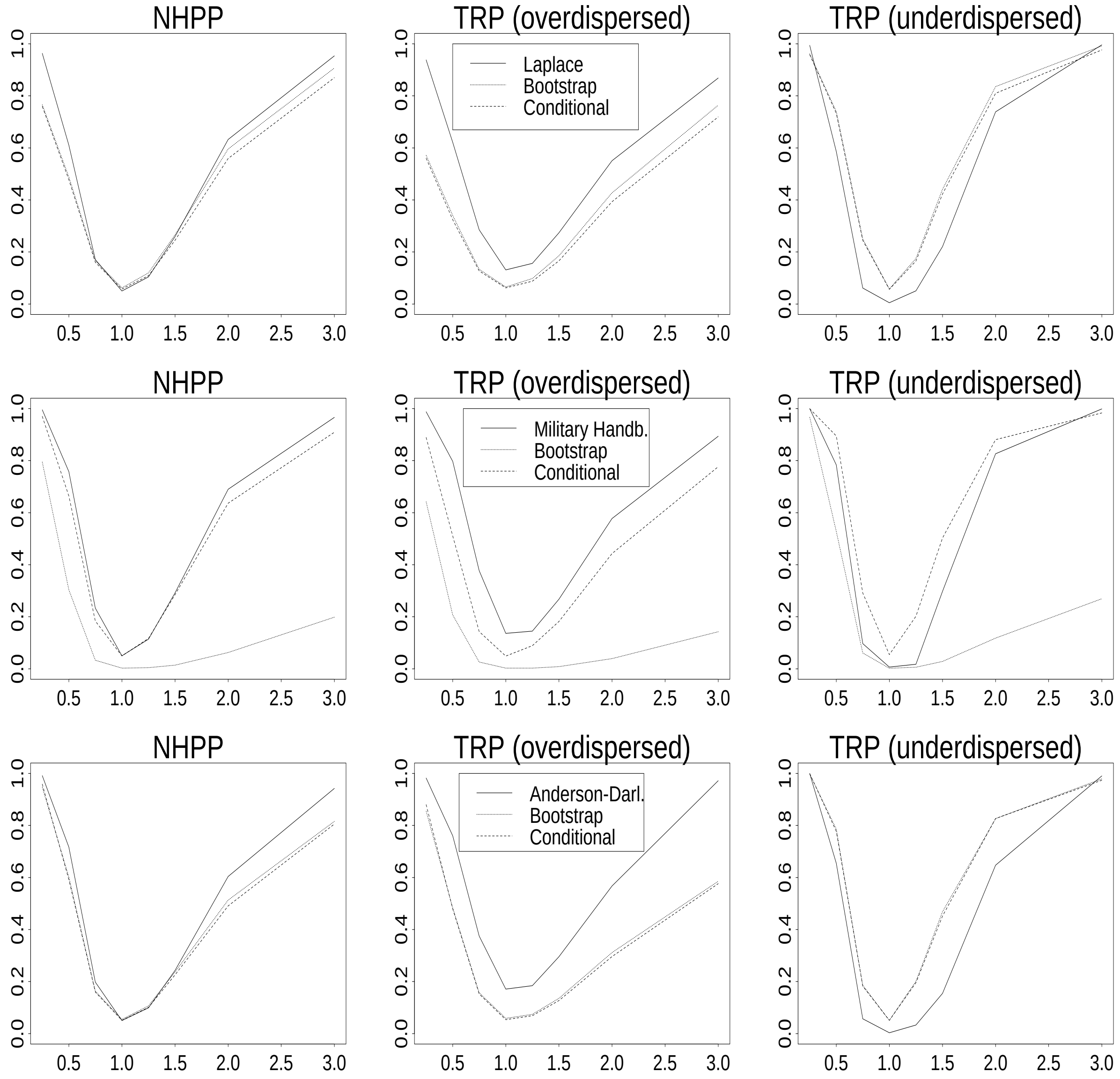

Figure 9: Upper row: The Laplace, unconditional bootstrapped Laplace and conditional bootstrapped Laplace tests for trend. Middle row: The military handbook, unconditional bootstrapped military handbook and conditional bootstrapped military handbook tests for trend. Lower row: The Anderson-Darling, unconditional bootstrapped Anderson-Darling and conditional bootstrapped Anderson-Darling tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ and several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (28) with three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 time truncated processes are simulated until expected number of failures is 15 (fewer than five are not allowed). The resulting rejection probabilities are plotted.

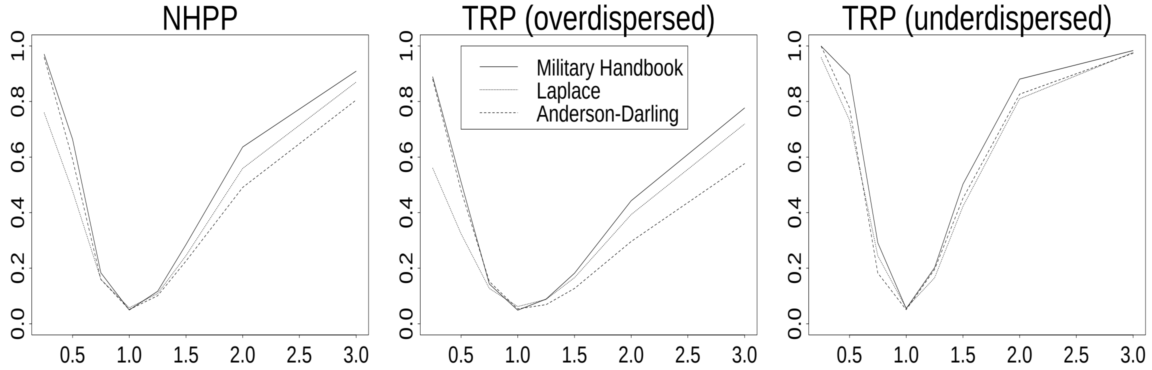

Figure 10: The conditional bootstrapped military handbook, Laplace and Anderson-Darling tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ and several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (28) with three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 time truncated processes are simulated until expected number of failures is 15 (fewer than five are not allowed). The resulting rejection probabilities are plotted.

power diminishes as the number of failures increases. For under- and overdispersed TRPs, the resampled tests still maintain close to correct rejection probabilities for the no trend situation ($\beta = 1$). The results for the Laplace, military handbook and Anderson-Darling tests for 15 failures are displayed in the upper, middle and lower row of Figure 9.

For the military handbook test we note a main difference from failure truncated data. As expected the unconditional or "naive" bootstrap have very poor properties for the military handbook since the test statistic does not converge in distribution to a fixed distribution independent of the number of failures.

For the other two tests we observe that the actual level of the unconditional bootstrap test generally is a bit higher than for the conditional test, typically about 0.005 higher for the smallest samples in our simulations. Probably as a result of this, the unconditional bootstrap test is generally also slightly stronger against trend alternatives, for large samples the tests become indistinguishable. This convergence is expected for the Laplace test as we know from results for the Lewis-Robinson test (Lewis & Robinson, 1974) that the Laplace test statistic for a renewal process converges towards an asymptotic (normal) distribution as the number of failures increases. From these simulations the same seems to be true for the Anderson-Darling test.

4.2.2 Resampled Versions of Different Tests

In Figure 10 we compare the results of the conditional bootstrap test for the Laplace, military handbook and Anderson-Darling tests. Again the results of the time truncated data follow the results from the failure truncated situation: For TRPs with power law trend function the military handbook test is the most powerful, which is expected as this is known to be the optimal test for HPP versus power law NHPP.

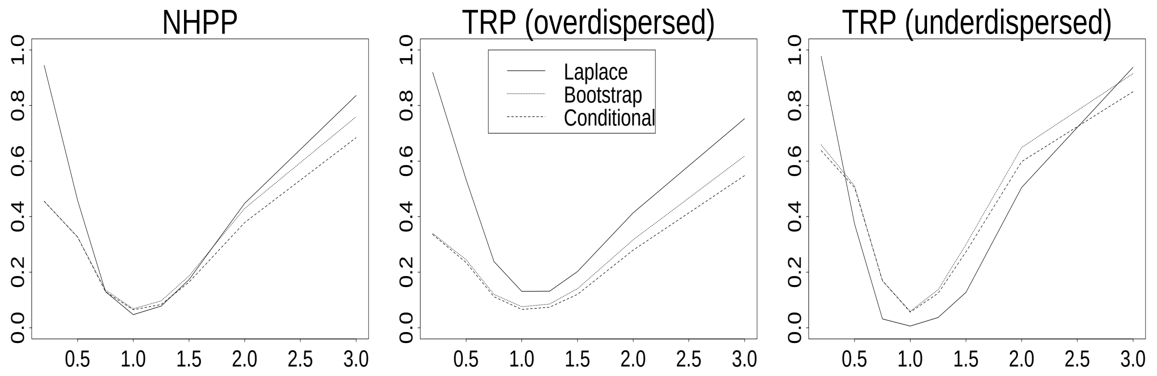

Figure 11: The Laplace, unconditional bootstrapped Laplace and conditional bootstrapped Laplace tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ and several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (28) with three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 time truncated processes are simulated until expected number of failures is 10 (fewer than five are not allowed). The resulting rejection probabilities are plotted.

4.2.3 Very Small Samples

For the shortest observation periods used in the simulations the expected number of failures is about 10. Figure 11 shows the results for the Laplace tests. From the plots in the figure and the Tables 26-28 in the appendix we again observe that the rejection probabilities for the resampled tests under H_0 differ a bit from the nominal level of 0.05. They are typically too high, 5%-8%, higher for the unconditional than the conditional bootstrap test. P-value adjustment (Section 3.3) may be applied to try to correct this.

4.3 Multiple Systems, TTT Tests

In the multiple systems case we have only considered the situation where all systems follow the same renewal process under the null hypothesis I in Section 3.1.3. The intent of the simulations is to explore the properties of the resampling techniques of that section. Hopefully it is possible to extend the null hypothesis to a common RP also for TTT tests.

There are of course a number of possibilities regarding "configurations", i.e. number of systems, observations periods and number of failures for each system. We have restricted the simulations to only two cases. For the time truncated situation five time truncated systems with expected number of observation equalling five in each system and where the observation periods start at different times. The configuration for the time truncated simulations is explained in Figure 12.

For the failure truncated situation there are five failure truncated systems observed until the number of failures is 3, 4, 5, 6 and 7 respectively. The configuration for the failure truncated simulations is explained in Figure 13.

In the simulations all systems follow a (common) TRP model with trend function given

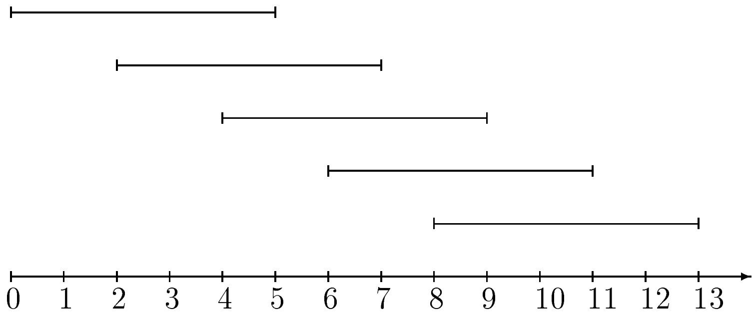

Figure 12: Expected number of failures for the five systems of the simulation of time truncated multiple systems. A renewal will always take place at the time of observation start.

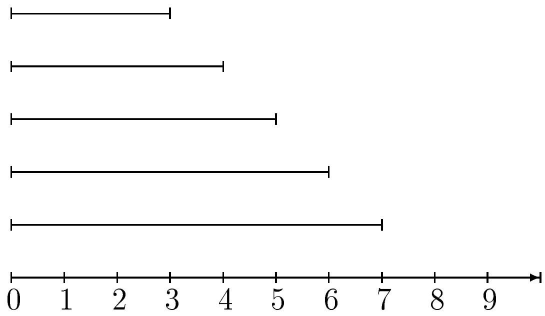

Figure 13: The number of failures for the five systems of the simulation of failure truncated multiple systems

by a power law function, while the interarrival times of the corresponding RP follow a Weibull distribution (28). Certain choices of parameters will therefore produce the special cases of NHPP, a Weibull-distributed RP, or an HPP.

For each simulated dataset we test for trend by three TTT tests: Laplace, military handbook and Anderson-Darling. In addition we also use the resampling tests described in Section 3.1.3 (conditional bootstrap or permutation) of those tests, and the standard combined Laplace and military handbook tests. Some of the results are presented graphically in this section, all are listed in Tables 35-40 in the appendix.

4.3.1 Time Truncation

For the time truncation situation we have used the configuration of Figure 12. In this situation where the systems only partly overlap we would expect the TTT tests to excel compared to the combined tests. The results for the Laplace, military handbook and

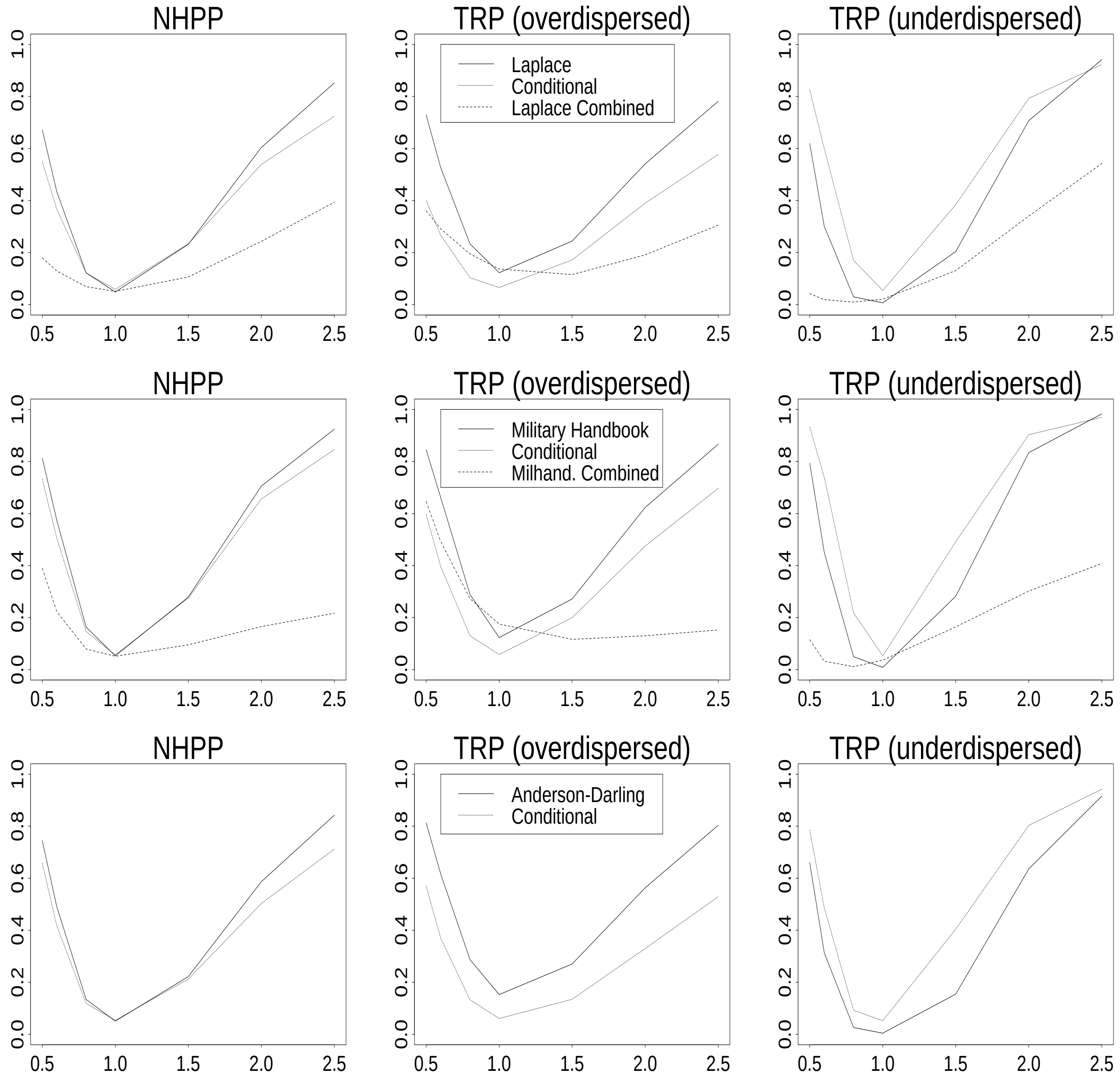

Figure 14: Upper row: The Laplace TTT test, conditional bootstrapped Laplace TTT and combine Laplace tests for trend. Middle row: The military handbook TTT test, conditional bootstrapped military handbook TTT and combined military handbook tests for trend. Lower row: The Anderson-Darling TTT test and conditional bootstrapped Anderson-Darling TTT test for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ and several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (28) with three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 simulations, each consisting of five systems, expected number of failures is five in each system, the observation period is governed by the configuration of Figure 12. The resulting rejection probabilities are plotted.

Anderson-Darling tests for the configuration of Figure 12 are shown in Figure 14.

In the NHPP case the actual level of all tests seems to be approximately correct, and as expected the TTT Laplace and military handbook tests are considerably stronger than the corresponding combined tests. The TTT versions of all three tests are also slightly stronger than the conditional bootstrap versions of the same tests.

For the overdispersed case we have the familiar pattern of the standard (TTT) tests rejecting too many under H_0, while for the underdispersed case they reject too few. The conditional bootstrap tests have reasonably correct actual level in all cases. Like the TTT tests the combined test have too high (low) actual level for overdispersed (underdispersed) situation. In addition the power functions seem to be biased.

4.3.2 Failure Truncation

For the failure truncation situation we have used the configuration of Figure 13. In this situation there is much overlap in the observation periods and we would thus expect the TTT tests to be only slightly better than the combined tests. The results for the Laplace, military handbook and Anderson-Darling tests are displayed in Figure 15.

In the NHPP case the actual level of all tests seems to be approximately correct. The difference in power between the TTT Laplace and military handbook tests and the corresponding combined tests are, as expected, reduced compared to the time truncated case. The TTT versions of all three tests are also in this situation slightly stronger than the permuted versions of the same tests.

For the overdispersed case the TTT tests reject too often under H_0. In addition the power functions of the TTT tests are clearly biased too the right. For the underdispersed case the situation is the opposite: the TTT tests reject too seldom and are biased too the left.

For all three tests the permuted versions maintain correct actual level under the null hypothesis. But also for these tests the power functions are slightly biased. This is clearly seen in the case of permuted Anderson-Darling test, hardly noticeable for the Laplace and military handbook tests. This effect, due to the “superposition” of several renewal processes all started at 0, will probably increase for increasing number of systems and decreasing number of failures per system.

4.4 Serial Dependence

4.4.1 Simulation of an STRP Model

To simulate serially dependent data with trend we simulate an STRP model, i.e. the serial dependency is in the time transformed process. Since this is a minor generalization of the TRP and NHPP models it seems appropriate to allow the marginal distribution, F, of the time transformed process to attain the same distributions. In particular we want F to be exponentially and Weibull distributed (with expectation 1), the simulation results may then be directly compared with the results of the TRP/NHPP simulations.

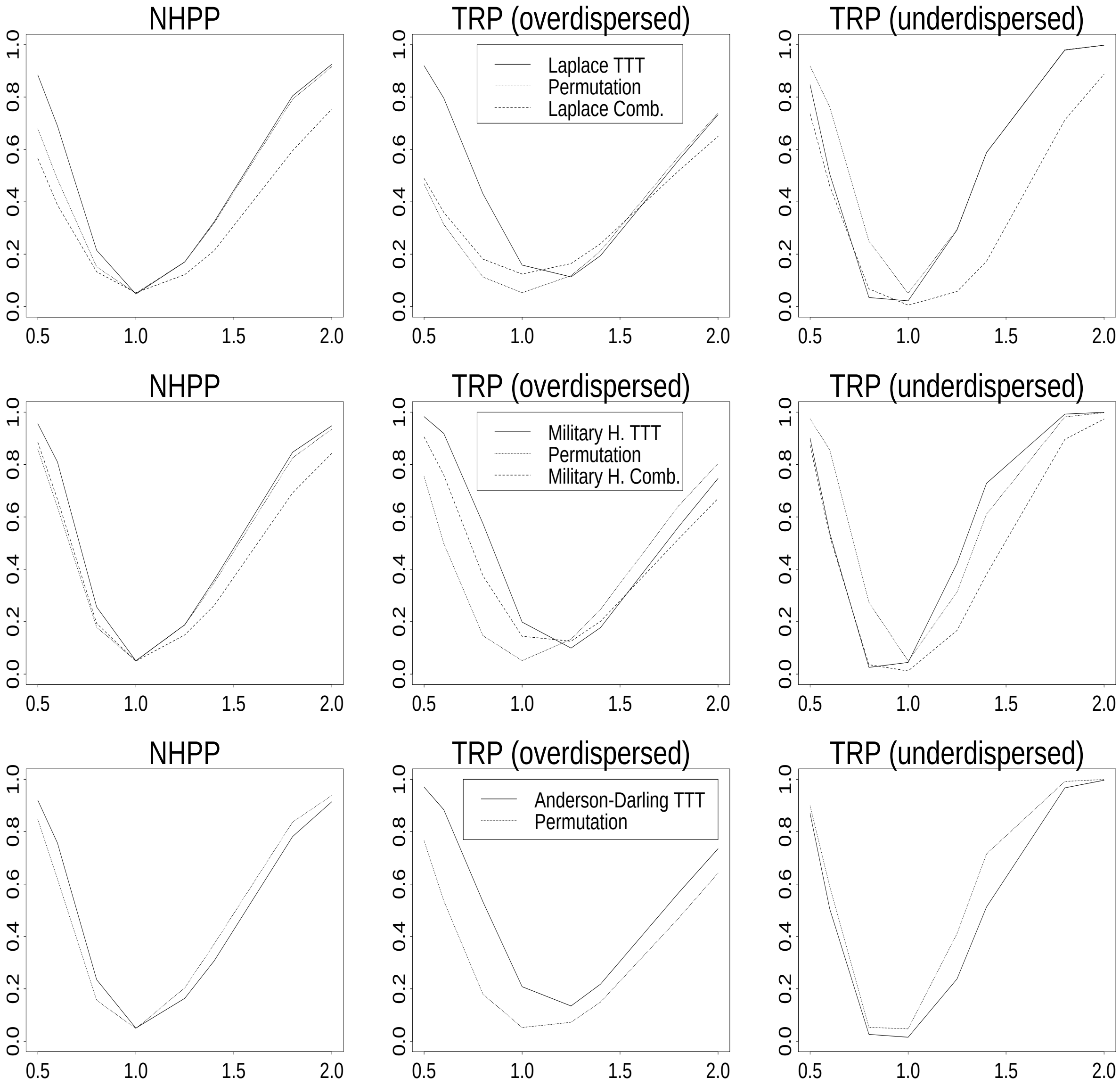

Figure 15: Upper row: The Laplace TTT test, permuted Laplace TTT and combined Laplace tests for trend. Middle row: The military handbook TTT test, permuted military handbook TTT and combined military handbook tests for trend. Lower row: The Anderson-Darling TTT test, permuted Anderson-Darling TTT test for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ and several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (28) with three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 simulations, each consisting of five failure truncated systems with number of failures 3, 4, 5, 6 and 7 respectively, all observation periods start at zero. The resulting rejection probabilities are plotted.

There exist several proposals for bi-exponential distributions, but a simpler and general method is to simulate from a suitable bivariate distribution with uniform marginals (copula). These may then be transformed to obtain the correct marginal distribution. An algorithm that will produce $U_1, U_2, \cdots$ such that U_i is $uniform[0,1]$ distributed and U_{i+1} depends on U_i is the following:

1. Draw U_1 from the $uniform[0,1]$ distribution.
2. Let $V = aU_1 + (1-a)W$ where W is drawn from the $uniform[0,1]$ distribution.
3. Let $U_2 = F_V(V)$, where F_V is the distribution of V. U_2 will then be uniformly distributed on [0,1].

The sequence can then be repeated with U_2 taking the role of U_1 etc. The parameter a controls the amount of dependence between subsequent observations, in particular the correlation between U_i and U_{i+1} is a strictly increasing function of a:

$$Corr(U_i, U_{i+1}) = \begin{cases} sgn(a)\frac{17|a|^2-50|a|+30}{10(|a|-1)^2} - 3 & 0 \le |a| \le \frac{1}{2} \\ sgn(a)\frac{33|a|^3+16|a|^2-11|a|+2}{10|a|^3} - 3 & \frac{1}{2} \le |a| \le 1 \end{cases} \tag{29}$$

where $sgn(a)$ equals 1 for $a \ge 0$, -1 for $a < 0$. When transforming the sequence $U_1, U_2, \cdots$ to, for instance, marginally exponentially distributed variables the correlation will of course be affected. Trend is as usual introduced by time transforming with the inverse cumulative trend function.

4.4.2 Bootstrapping

Three different bootstrap methods were employed in the simulations. The first two, model based assuming an AR1 model and block bootstrap were described in Section 3.2. The AR1 model is chosen because of the simulation routine, and the fact that more complicated dependence models seem less likely for repairable systems data.

The third is model based bootstrap assuming the model simulated from in Section 4.4.1 without trend, where the parameter a and the marginal distribution of the interfailure times are unknown. The parameter a is estimated by transforming the interfailure times with the empirical distribution of the interfailure times, and then computing the correlation coefficient of the transformed series. Then new series are generated using the estimate for a and the inverse of the empirical distribution of the interfailure times. We will call this method SRP model bootstrap.

4.4.3 Simulation Results

Data were simulated from an STRP model with power law trend function. The underlying serially dependent sequence is marginally Weibull distributed (28) with parameter values $w = 0.75, 1.00, 1.50$. Simulation is performed using the algorithm of Section 4.4.1 with three different values of the parameter a: 0 (renewal process), 0.35 (positive correlation)

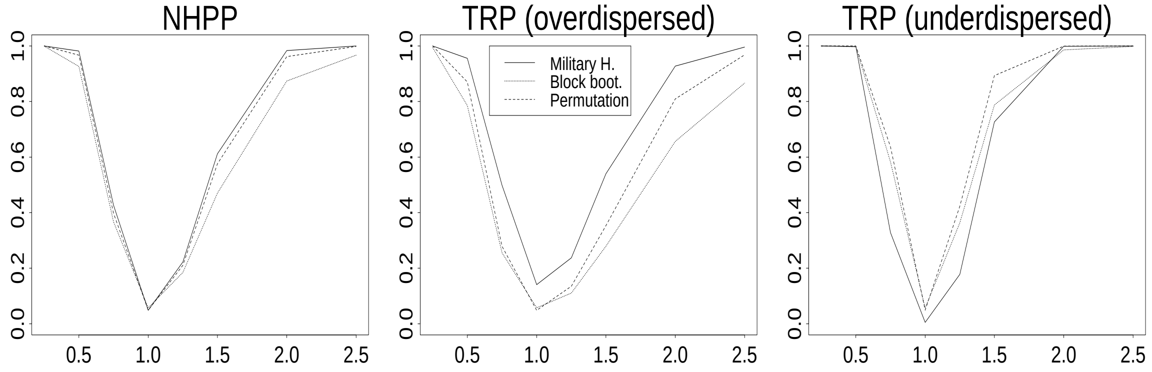

Figure 16: The military handbook, block bootstrapped military handbook (block size 7) and permuted military handbook tests for trend. Simulations from a TRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The interarrival times of the underlying RP are Weibull distributed (28) using three different values for w: 1.00 (NHPP), 0.75 (overdispersed TRP) and 1.50 (underdispersed TRP). For each parameter combination 10,000 failure truncated processes with 15 failures are simulated, the resulting rejection probabilities are plotted.

and -0.35 (negative correlation). This will result in a correlation coefficient for marginally exponentially distributed interarrival times of about 0.4 ($a = 0.35$) and -0.3 ($a = -0.35$). For Weibull distribution with parameter $w = 0.75$ and $w = 1.50$ the results are similar. The reason for using these relatively moderate values is that interfailure times from repairable systems seldom exhibit strong correlation. But we consider these values to be within what is normal.

For serially dependent processes more failures are needed to reliably detect a trend than for the TRP case. Because of this, and because we wished to compare the result with the results of the TRP simulations, simulations were performed for failure truncated processes with $n = 35$ failures. For each simulated process three tests were used to test for trend: Laplace, military handbook and Anderson-Darling. The resampled versions (block bootstrap and the two model based techniques) of these tests were also employed. For the block bootstrap we use overlapping blocks with a block size of seven, chosen on the basis of some preliminary simulations.

Based on the simulation results of Tables 41-49 in the appendix, we choose to present graphically in Figures (16-18) the results of the military handbook test with block bootstrap. This test seems to be better suited for robustification against serial dependence than the other two tests of interest. In the simulations the resampled military handbook test is generally superior to the resampled Laplace test in terms of power. It also maintains correct level better than the block bootstrapped Anderson-Darling test. Of the two model based bootstrap methods the results reveal that the AR1-method is inferior to the other two in this situation, that may be due to bad model fit or more generally the trend testing situation. The results for the SRP method is close to the results of the block method, but focus will be on the latter since it is is a more general method.

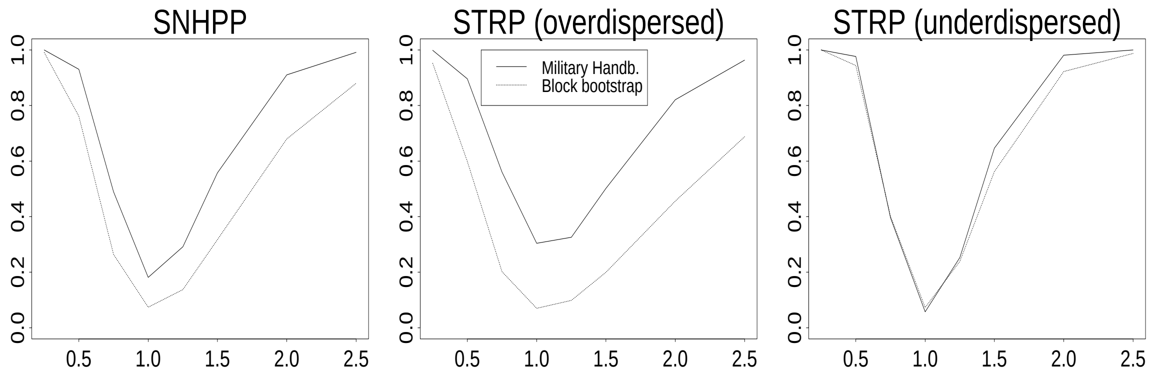

Figure 17: The military handbook and block bootstrapped military handbook (block size 7) tests for trend. Simulations from an STRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The underlying dependent sequence is generated by the algorithm of Section 4.4.1 with a parameter value of $a = 0.35$ (positive lag 1 autocorrelation). The marginal distribution of the interarrival times are Weibull (28) distributed using three different values for w: 1.00 (SNHPP), 0.75 (overdispersed) and 1.50 (underdispersed). For each parameter combination 10,000 failure truncated processes with 35 failures are simulated, the resulting rejection probabilities are plotted.

Figure 16 presents the results for the TRP case (independent transformed process) including the results of the permuted military handbook test (Tables 14, 16 and 18 in the appendix). The actual level of the block bootstrap military handbook test is close to the nominal level of 5% for both the HPP and RP simulations, it is typically around 6%. However, as expected, the block bootstrap test loses some power compared to the permutation test. For this particular choice of block size the loss is larger than the typical loss of power for permutation tests compared to standard tests in the NHPP case.

Figure 17 presents the results for STRP with positively autocorrelated transformed process. This leads to too high actual level of the standard power law test. It is especially noticeable in the plot in the middle where we get the added effect of the "overdispersed" Weibull distribution. In the right plot the two effects work in opposite direction, and thus nearly cancel each other out. The block bootstrap test is quite close to the nominal level for all cases, typically around 7%.

Figure 18 presents the results for STRP with negatively autocorrelated transformed process. This leads to too low actual level of the standard power law test, especially noticeable in the "underdispersed" case. The block bootstrap tests are again close to the nominal level for all cases.

As seen from the results under the null hypothesis, the p-values of the block bootstrap method was slightly above the nominal level of 5%. P-value adjustment via double bootstrap methods may remedy this, but has not been tried in these simulations. Longer blocks might also be of help, but would have other disadvantages.

An important observation from the simulation results is that a very strong trend may be harder to detect than a slightly weaker one using resampling tests. This is particularly

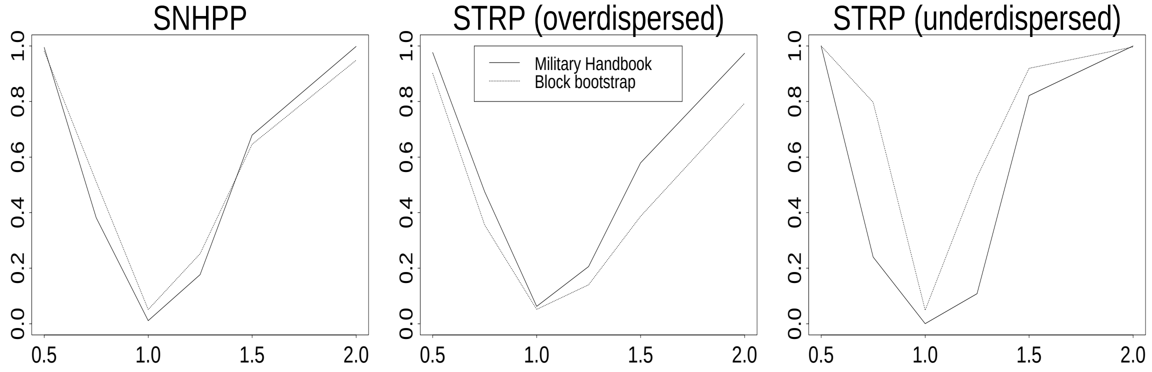

Figure 18: The military handbook and block bootstrapped military handbook (block size t) tests for trend. Simulations from an STRP model with power law trend function $\lambda(t) = \beta t^{\beta-1}$ using several values for β (horizontal axis). The underlying dependent sequence is generated by the algorithm of Section 4.4.1 with a parameter value of $a = -0.35$ (positive lag 1 autocorrelation). The marginal distribution of the interarrival times are Weibull (28) distributed using three different values for w: 1.00 (SNHPP), 0.75 (overdispersed) and 1.50 (underdispersed). For each parameter combination 10,000 failure truncated processes with 35 failures are simulated, the resulting rejection probabilities are plotted.

the case for the Laplace test, and to a lesser degree for the Anderson-Darling test. The explanation is probably that a strong trend is interpreted as evidence of a strong positive correlation under the H_0 assumption of a stationary sequence. In, for instance, the AR1 bootstrap, the parameter of the fitted model could then be so large that the variation of the resampled series would make it less likely to reject the null hypothesis. For block bootstrap a remedy for this is using shorter blocks, longer sequences will allow longer blocks before this problem occurs.

5 EXAMPLE

The compressor data set, consisting of 49 critical failures for a single gas compressor, was analysed in Elvebakk *et al.* (1999). They concluded that a reasonable model was a renewal process overdispersed compared to an HPP. In particular they argued that the bimodal exponential distribution gave a very good fit, and that a possible interpretation was two types of repair actions: With probability p a successful repair is carried out and with probability $1 - p$ an unsuccessful repair is carried out. The time to next failure is exponentially distributed for both cases, but the expectations differ. Elvebakk *et al.* (1999) concluded that there were no indications of monotone (power law) trend in the process.

When standard HPP tests are used for overdispersed data they typically tend to reject too often, so a resampling test may be useful. Tests for trend were performed for this time truncated dataset using the Laplace, military handbook and Anderson-Darling tests for trend and the conditional bootstrap (Section 3.1.2) versions of the same tests with 1500 bootstrap replications. The results are given as p-values against trend in Table 1. Using

Test	Laplace	Military Handbook	Anderson-Darling
Standard	0.039	0.018	0.032
Bootstrap	0.200	0.198	0.210

Table 1: P-values against trend for the compressor dataset. Standard tests are tests for trend in Poisson processes, bootstrap tests are conditional bootstrapped tests using the resampling method of Section 3.1.2 and 1500 replications.

a 5% significance level the standard HPP tests all conclude that the HPP null hypothesis should be rejected. The resampled tests, however, do not reject the RP null hypothesis, so the conclusion is in agreement with the choice of model in Elvebakk *et al.* (1999).

6 CONCLUSIONS

The simulation study shows that the three tests for trend in Poisson processes, military handbook, Laplace and Anderson-Darling are all good at detecting trend departures from the HPP null hypothesis. However, they are also sensitive to non-trend departures from Poisson processes, like renewal processes and serially dependent processes, and may thus give misleading results when used as test for trend in these situations. The Anderson-Darling test is particularly affected by these departures since this is a test against a non-monotonic trend. Simulation studies performed in Lindqvist *et al.* (1994), Kvaløy (1994) and Lawless & Thiagarajah (1996) also point to the lack of robustness against renewal processes for several tests for trend in Poisson processes.

The other main point of the simulation study is that resampling methods may be employed to estimate the distribution of the test statistics under a more general null hypothesis, like general renewal process or processes with simple dependence structures. The resulting resampling tests are robust against these deviations from the null hypothesis of homogeneous Poisson processes. The simulations also indicate that in the case of renewal processes the resulting resampling tests seem to have good power properties, the loss of power compared to the non resampled tests is small for nonhomogeneous Poisson process alternatives for moderate sample sizes, negligible for larger samples.

6.1 H_0: General Renewal Process

To extend the null hypothesis to "renewal process" for a failure truncated system the two resampling schemes "bootstrap testing" and "permutation testing" can be applied. Both achieved approximately correct (nominal) level for sample sizes of at least fifteen in our simulations. The permutation test is an exact test and will thus always achieve correct level. For time truncated systems a conditional bootstrap testing scheme can be applied, the actual level using this method were also very close to the nominal for similar sample sizes. The standard Poisson tests could not maintain correct level when simulating from renewal processes different from the homogeneous Poisson process. In the Poisson case there is a loss of power for the resampling tests compared to the Poisson test, quite substantial for very few (five) failures. This loss diminishes rapidly in our simulations as the number of failures increases, in our opinion they are reasonably good from as few as 10-15 failures.

For very small samples the permutation methods will still maintain correct rejection probability while the simulations showed that the bootstrap tests may give somewhat misleading results. Another advantage of resampling tests is that resampling p-value correction methods, like double bootstrap, may be applied to (partially) correct this. This method gave very good results in some limited simulations.

Another approach to extending the null hypothesis to include a general renewal process is to modify the test statistic of the standard tests. In case of the Laplace test there already exists a modified test for the RP null hypothesis, the Lewis-Robinson test which is asymptotically standard normally distributed. The simulation study showed that the resampled versions of the Laplace test are more powerful than this test for moderate sample sizes, the difference may be quite large for smaller sample sizes. This is probably because the distribution of the Lewis-Robinson test statistic has not yet converged to the asymptotic distribution, the results from the resampled Laplace tests and the Lewis-Robinson test are very similar for a sample size of at least one hundred.

To choose among the different resampling tests the simulation results offer some general advice: Tests which are optimal for HPP versus a NHPP of a special functional form seem to have good properties as tests for RP versus a TRP with trend function of the same functional form. Whether this holds for more extreme choices for renewal processes (less like an HPP), or other models including trend and renewal properties, like the modulated renewal process, we do not know. But generally, unlike the corresponding Laplace and military handbook tests in the Poisson situation, it is very unlikely that the resampling tests are optimal for any specified alternative hypotheses. Alternative tests like rank tests and likelihood based tests might be stronger in many situations.

Limited simulations in the several systems case showed that tests for a common HPP may give very misleading results when used as test for a common RP. For the several systems case two resampling schemes were proposed for the failure and time truncated cases. Both gave reasonably good results in testing the null hypothesis of identical renewal processes, although especially the Anderson-Darling test seemed to be somewhat biased in situations with many systems and much overlap in the observation periods.

The simulation study was limited to simulations from a TRP model with monotonic trend function of the power law (1) or log linear law (2) type and a corresponding renewal process with Weibull distributed interarrivals. For the parameter of the Weibull distribution three values were used. The effects of using resampling tests to detect non-monotonic trends are thus not explored, neither are the effects of other choices for interarrival distribution.

6.2 H_0: Serial Dependence

From a very limited simulation study it was shown that it is sometimes possible to employ resampling to extend the null hypothesis of HPP to a more general null hypothesis of dependent interarrival times. To get powerful tests for this hypothesis we would ideally like to have a resampling method that would preserve the dependence structure of a sequence (under the null hypothesis of identically marginally distributed interfailure times) without being influenced by any trends in the sequence. A compromiseq is the block bootstrap method, where we assume that the dependence in the series acts mostly on short lags. Thus much of the dependence structure will be preserved, but hopefully most of the trend will be lost. For data from repairable systems this assumption may be reasonable.

The simulation study was performed with a sample size of 35 failures and a block size of seven. The block bootstrapped tests maintained close to correct (nominal) level also for the simulated serially dependent sequences, in contrast to the standard tests and permutation tests. In the NHPP and TRP cases there was a certain loss of power compared to the permutation test. The loss was greater than the loss of power for permutation tests compared to standard tests in the NHPP situation. Of the block bootstrapped tests the military handbook test seemed to be superior to the other. This effect may at least partly be due to simulations exclusively with power law trend functions.

Longer blocks will preserve more of the dependence structure, while shorter blocks will give more powerful tests. Too long blocks may lead too paradoxical behaviour of the tests, very strong trends may be more difficult to detect than slightly weaker ones. The critical block size is probably dependent on dependence structure as well as test statistic and the sample size. For larger samples the block size may be increased.

The block bootstrap is a standard method for bootstrapping stationary time series. It may be possible to suggest variants more suited to the testing situation. Block permutation of non-overlapping blocks is a possibility. Small scale simulations indicate that there is less problem for very strong trends, but a loss in power. By analogy with permutation and conditional bootstrap tests for the renewal process null hypothesis, it may be possible to construct conditional bootstrap methods, for example conditioned on the total time (with a necessary flexibility) of the original sample.

Specifically for the simulations of STRP models only two values of the dependence parameter were employed, both resulting in rather weak correlation between subsequent interfailure times. The results of the block bootstrap may not be similar for other choices of parameters or other dependence structures in the interfailure times. In addition only one choice of failure number (35) and block size (seven) was used.

6.3 Alternative Methods

Obvious alternatives to the resampling tests described in this paper are modifications to the Poisson tests statistics like the Lewis-Robinson test, rank tests and likelihood based tests in model families including trend and renewal processes (and serial dependence). The latter are especially interesting as they, like the Lewis-Robinson and resampling tests, can make use of the numerical values of the interfailure times, not only the ranks.

Such likelihood based tests have been constructed by Berman (1981) in a TRP model with gamma distribution of the interarrival times of the renewal process and by Lawless & Thiagarajah (1996) based on a modulated renewal process. Elvebakk *et al.* (1999) used likelihood ratio tests in a modelling procedure for several different models in the TRP class. These tests are powerful when the specified trend and renewal process are appropriate, they may not be effective outside this family. In a simulation study by Kvaløy (1994) the actual level of a test of Berman (1981) were shown to be affected by using a renewal process with Weibull distributed interarrival times. On these grounds Lawless & Thiagarajah (1996) recommend that these tests are best used in the modelling of the failure process. Assuming there are good models, likelihood based tests may be especially attractive for data from STRP processes where there is a greater loss of power for resampling tests compared to the standard tests.

A more ad hoc alternative may be to transform the interfailure times to exponential distribution, before employing the standard tests for trend in Poisson processes. To do this the distribution of the interfailure times must be known. It may be estimated by for example parametric or kernel estimation, or most easily by the empirical distribution. In the latter case we would then really be testing on the transformed ranks as the original values of the interfailure times are lost, only the ranking determines the transformed values. No simulation have been performed to assess the properties of this method.

6.4 Concluding Remarks

The resampling tests described in this paper were in a simulation study shown to have close to correct actual level under a renewal process null hypothesis. Power properties were acceptable from about 10-15 failures. These tests are due to their simplicity and generality good alternatives to other trend tests for the renewal process null hypothesis, especially for smaller samples. Similar tests can also be constructed for a null hypothesis of serial dependence, for which simulations from a particular model gave reasonable results.

References

AIYAR, R. J., GUILLIER, C. L. & ALBERS, W. (1979). Asymptotic relative efficiencies of rank tests for trend alternatives, *Journal of The American Statistical Association* **74**(365): 226–231.

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

ASCHER, H. & FEINGOLD, H. (1984). *Repairable Systems Reliability. Modeling, Inference, Misconceptions and their Causes*, Marcel Dekker, New York.

BAIN, M. J., ENGELHARDT, F. & WRIGHT, F. T. (1985). Tests for increasing trend in the intensity of a poisson process: A power study, *Journal of The American Statistical Association* **80**: 419–422.

BARLOW, R. E. & DAVIS, B. (1977). Analysis of time between failures of repairable components, *in* J. B. Fussel & G. R. Burdick (eds), *Nuclear Systems Reliability Engineering and Risk Asessment*, SIAM, Philadelphia, pp. 543–561.

BERAN, R. (1988). Prepivoting test statistics: A bootstrap view of asymptotic refinements, *Journal of the American Statistical Association* **83**(403): 687–697.

BERMAN, M. (1981). Inhomogeneous and modulated gamma processes, *Biometrika* **68**: 143–152.

CARLSTEIN, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence, *Annals of Statistics* pp. 1171–1179.

COX, D. R. (1972). The statistical analysis of point processes, *in* P. A. Lewis (ed.), *Stochastic Point Processes*, Wiley, New York, pp. 55–66.

COX, D. R. & LEWIS, P. A. (1966). *The Statistical Analysis of Series of Events*, Methuen, London.

DAVISON, A. C. & HINKLEY, D. V. (1997). *Bootstrap Methods and their Application*, Cambridge University Press, Cambridge, UK.

EFRON, B. (1967). The two sample problem with cencored data, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4, Prentice-Hall, New York, pp. 831–853.

EFRON, B. & TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*, Chapman & Hall, New York.

ELVEBAKK, G., LINDQVIST, B. H. & HEGGLAND, K. (1999). The trend renewal process for statistical analysis of repairable systems, Paper I of this thesis.

HALL, P. (1985). Resampling a coverage pattern, *Stochastic Processes and their Applications* **20**: 231–246.

KAPLAN, E. L. & MEIER, P. (1958). Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association* pp. 457–481.

KVALØY, J. T. (1994). *Analysis of repairable systems: Trend testing*, Master's thesis, Norwegian Institute of Technology. In Norwegian.

KVALØY, J. T. & LINDQVIST, B. H. (1998). TTT-based tests for trend in repairable systems data, *Reliability Engeneering and System Safety* **60**: 13–28.

LAWLESS, J. F. & THIAGARAJAH, K. (1996). A point-process model incorporating renewals and time trends, with applications to repairable systems, *Technometrics* **38**: 131–138.

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

LEWIS, P. A. W. (1963). A branching poisson process model for the analysis of computer failure patterns, *Journal of the Royal Statistical Society, Series B* **26**: 398–456.

LINDQVIST, B. H. (1993). The trend renewal process, a useful model for repairable systems, *SRE-93, Tillförlitlighet i reparerbara system*, Society of Reliability Engineers, Scandinavian Chapter, Annual Conference, Malmö, Sweden.

LINDQVIST, B. H., KJØNSTAD, G. A. & MELAND, N. (1994). Testing for trend in repairable systems data, *Proceedings of ESREL'94*, The European Safety and Reliability Association, La Baule, France.

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

RIPLEY, B. D. (1987). *Stochastic Simulation*, Wiley, New York.

MIL-HDBK-189, Reliability Growth Managment (1981). Headquaters. U.S. Army Communications Research and Development Command, ATTN: DRDCO-PT, Fort Monmouth, NJ 07702.

APPENDIX: SIMULATION RESULTS

w = 1.00, n = 5	resampling	milhand	laplace	lew-rob	and-dar	mann
β = 0.05	standard	0.9982	0.9507	0.0292	0.9985	0.7167
	bootstrap	0.1862	0.0786	0.0957	0.0519	0.9557
	permutation	0.9575	0.7790	0.7788	0.4646	0.9560
β = 0.15	standard	0.9556	0.7933	0.0993	0.9449	0.3872
	bootstrap	0.1694	0.1196	0.1863	0.0626	0.7579
	permutation	0.7423	0.5328	0.5327	0.6812	0.7593
β = 0.25	standard	0.8222	0.5796	0.1246	0.7859	0.2310
	bootstrap	0.1419	0.1223	0.2033	0.0556	0.5548
	permutation	0.5202	0.3590	0.3589	0.4740	0.5558
β = 0.50	standard	0.3607	0.2317	0.0758	0.3237	0.0641
	bootstrap	0.0835	0.0815	0.1138	0.0314	0.2367
	permutation	0.1947	0.1392	0.1395	0.1759	0.2330
β = 0.75	standard	0.1174	0.0855	0.0448	0.1107	0.0272
	bootstrap	0.0481	0.0483	0.0648	0.0223	0.1209
	permutation	0.0851	0.0698	0.0699	0.0829	0.1065
β = 1.00	standard	0.0481	0.0470	0.0328	0.0523	0.0179
	bootstrap	0.0375	0.0389	0.0495	0.0185	0.0782
	permutation	0.0507	0.0502	0.0502	0.0543	0.0528
β = 1.25	standard	0.0525	0.0547	0.0336	0.0564	0.0168
	bootstrap	0.0410	0.0414	0.0485	0.0183	0.0771
	permutation	0.0494	0.0596	0.0596	0.0616	0.0317
β = 1.50	standard	0.0840	0.0796	0.0457	0.0813	0.0243
	bootstrap	0.0520	0.0498	0.0614	0.0250	0.0951
	permutation	0.0605	0.0774	0.0774	0.0765	0.0303
β = 1.75	standard	0.1280	0.1135	0.0534	0.1114	0.0296
	bootstrap	0.0693	0.0687	0.0824	0.0352	0.1240
	permutation	0.0804	0.1075	0.1075	0.1017	0.0323
β = 2.00	standard	0.1752	0.1579	0.0627	0.1514	0.0380
	bootstrap	0.0839	0.0814	0.0974	0.0405	0.1400
	permutation	0.0982	0.1283	0.1283	0.1228	0.0400
β = 2.25	standard	0.2297	0.2216	0.0719	0.2109	0.0466
	bootstrap	0.0993	0.0981	0.1019	0.0509	0.1637
	permutation	0.1162	0.1499	0.1499	0.1439	0.0472
β = 2.50	standard	0.2865	0.2700	0.0825	0.2554	0.0522
	bootstrap	0.1213	0.1176	0.1206	0.0571	0.1942
	permutation	0.1379	0.1802	0.1802	0.1712	0.0527
β = 3.00	standard	0.4038	0.4006	0.0891	0.3729	0.0720
	bootstrap	0.1565	0.1553	0.1400	0.0771	0.2369
	permutation	0.1748	0.2264	0.2264	0.2117	0.0724
β = 4.00	standard	0.6403	0.6157	0.0785	0.5743	0.0850
	bootstrap	0.2006	0.1951	0.1331	0.1052	0.2836
	permutation	0.2158	0.2783	0.2783	0.2620	0.0858
β = 6.00	standard	0.8905	0.8788	0.0539	0.8484	0.1082
	bootstrap	0.2633	0.2592	0.1027	0.1439	0.3357
	permutation	0.2623	0.3408	0.3408	0.3203	0.1087
β = 10.0	standard	0.9952	0.9933	0.0138	0.9877	0.1202
	bootstrap	0.2866	0.2807	0.0379	0.1797	0.3465
	permutation	0.2855	0.3649	0.3649	0.3477	0.1215
β = 20.0	standard	1.0000	1.0000	0.0004	1.0000	0.1249
	bootstrap	0.2972	0.2822	0.0016	0.2220	0.3614
	permutation	0.2962	0.3753	0.3753	0.3579	0.1261

Table 2: Results from simulations of failure truncated TRP processes with $n=5$ failures. The trend function is given by a power law function $\lambda(t)=\beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w=1.0$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00 n = 5	resampling	milhand	laplace	lew-rob	and-dar	mann
γ = -0.0300	standard	0.0490	0.0463	0.0339	0.0517	0.0144
	bootstrap	0.0354	0.0372	0.0491	0.0169	0.0731
	permutation	0.0509	0.0507	0.0507	0.0549	0.0518
γ = -0.0175	standard	0.0495	0.0511	0.0305	0.0551	0.0167
	bootstrap	0.0384	0.0389	0.0460	0.0182	0.0784
	permutation	0.0520	0.0493	0.0493	0.0556	0.0540
γ = -0.0100	standard	0.0473	0.0484	0.0323	0.0557	0.0150
	bootstrap	0.0349	0.0347	0.0483	0.0148	0.0762
	permutation	0.0518	0.0479	0.0481	0.0542	0.0514
γ = -0.0050	standard	0.0486	0.0479	0.0361	0.0510	0.0171
	bootstrap	0.0387	0.0383	0.0491	0.0174	0.0792
	permutation	0.0549	0.0543	0.0542	0.0603	0.0524
γ = 0.0000	standard	0.0499	0.0448	0.0312	0.0502	0.0156
	bootstrap	0.0371	0.0373	0.0461	0.0177	0.0784
	permutation	0.0506	0.0492	0.0492	0.0560	0.0511
γ = 0.0500	standard	0.0561	0.0443	0.0335	0.0525	0.0174
	bootstrap	0.0380	0.0374	0.0480	0.0181	0.0790
	permutation	0.0518	0.0545	0.0546	0.0583	0.0450
γ = 0.1000	standard	0.0564	0.0542	0.0318	0.0595	0.0165
	bootstrap	0.0363	0.0360	0.0461	0.0170	0.0745
	permutation	0.0474	0.0525	0.0524	0.0555	0.0381
γ = 0.2000	standard	0.0706	0.0650	0.0393	0.0706	0.0200
	bootstrap	0.0445	0.0437	0.0571	0.0181	0.0864
	permutation	0.0573	0.0669	0.0669	0.0692	0.0374
γ = 0.5000	standard	0.0999	0.0985	0.0503	0.1027	0.0295
	bootstrap	0.0597	0.0566	0.0727	0.0257	0.1084
	permutation	0.0719	0.0899	0.0899	0.0876	0.0400
γ = 1.0000	standard	0.1518	0.1481	0.0532	0.1515	0.0330
	bootstrap	0.0740	0.0718	0.0787	0.0348	0.1300
	permutation	0.0834	0.1108	0.1108	0.1074	0.0394
γ = 2.0000	standard	0.2283	0.2195	0.0623	0.2215	0.0428
	bootstrap	0.0959	0.0946	0.0955	0.0473	0.1551
	permutation	0.1091	0.1432	0.1432	0.1399	0.0456
γ = 5.0000	standard	0.3597	0.3562	0.0818	0.3491	0.0652
	bootstrap	0.1380	0.1367	0.1241	0.0676	0.2134
	permutation	0.1569	0.2042	0.2042	0.1920	0.0665
γ = 10.000	standard	0.4746	0.4703	0.0912	0.4617	0.0804
	bootstrap	0.1658	0.1630	0.1400	0.0835	0.2529
	permutation	0.1875	0.2444	0.2444	0.2311	0.0812
γ = 20.000	standard	0.5761	0.5799	0.0849	0.5675	0.0891
	bootstrap	0.1912	0.1906	0.1397	0.1028	0.2910
	permutation	0.2156	0.2775	0.2775	0.2647	0.0897
γ = 40.000	standard	0.6702	0.6779	0.0890	0.6614	0.1049
	bootstrap	0.2188	0.2153	0.1457	0.1172	0.3091
	permutation	0.2440	0.3139	0.3139	0.3026	0.1056
γ = 100.00	standard	0.7881	0.7775	0.0747	0.7632	0.1093
	bootstrap	0.2402	0.2407	0.1328	0.1384	0.3376
	permutation	0.2684	0.3430	0.3430	0.3229	0.1103

Table 3: Results from simulations of failure truncated TRP processes with $n=5$ failures. The trend function is given by a log-linear law function $\lambda(t)=e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w=1.0$ (NHPP). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 5	resampling	milhand	laplace	lew-rob	and-dar	mann
β = 0.05	standard	0.9935	0.8819	0.0316	0.9922	0.5413
	bootstrap	0.1661	0.0807	0.0975	0.0454	0.8700
	permutation	0.8683	0.6066	0.6060	0.3443	0.8706
β = 0.15	standard	0.9215	0.6754	0.0609	0.9057	0.2571
	bootstrap	0.1244	0.0927	0.1260	0.0559	0.5984
	permutation	0.5774	0.3786	0.3787	0.4975	0.5982
β = 0.25	standard	0.7915	0.5069	0.0671	0.7792	0.1475
	bootstrap	0.0974	0.0826	0.1207	0.0404	0.4197
	permutation	0.3818	0.2462	0.2461	0.3465	0.4179
β = 0.50	standard	0.4511	0.2594	0.0511	0.4276	0.0515
	bootstrap	0.0567	0.0590	0.0827	0.0275	0.1886
	permutation	0.1516	0.1078	0.1077	0.1382	0.1836
β = 0.75	standard	0.2301	0.1473	0.0339	0.2323	0.0233
	bootstrap	0.0399	0.0377	0.0520	0.0158	0.1013
	permutation	0.0763	0.0623	0.0624	0.0758	0.0867
β = 1.00	standard	0.1347	0.1120	0.0304	0.1557	0.0186
	bootstrap	0.0325	0.0333	0.0431	0.0149	0.0808
	permutation	0.0544	0.0560	0.0560	0.0585	0.0526
β = 1.25	standard	0.1138	0.1085	0.0295	0.1406	0.0159
	bootstrap	0.0352	0.0340	0.0444	0.0160	0.0718
	permutation	0.0459	0.0544	0.0544	0.0555	0.0327
β = 1.50	standard	0.1333	0.1285	0.0294	0.1591	0.0195
	bootstrap	0.0358	0.0359	0.0464	0.0160	0.0806
	permutation	0.0490	0.0621	0.0620	0.0627	0.0286
β = 1.75	standard	0.1643	0.1664	0.0366	0.1992	0.0239
	bootstrap	0.0462	0.0455	0.0580	0.0222	0.0953
	permutation	0.0597	0.0782	0.0782	0.0758	0.0305
β = 2.00	standard	0.2088	0.2102	0.0390	0.2388	0.0284
	bootstrap	0.0545	0.0530	0.0597	0.0234	0.1105
	permutation	0.0708	0.0936	0.0936	0.0895	0.0306
β = 2.25	standard	0.2503	0.2532	0.0545	0.2791	0.0338
	bootstrap	0.0669	0.0651	0.0840	0.0314	0.1363
	permutation	0.0911	0.1189	0.1189	0.1146	0.0353
β = 2.50	standard	0.2860	0.2947	0.0548	0.3194	0.0417
	bootstrap	0.0786	0.0748	0.0868	0.0356	0.1477
	permutation	0.1007	0.1316	0.1316	0.1283	0.0425
β = 3.00	standard	0.3794	0.3764	0.0678	0.3976	0.0460
	bootstrap	0.0914	0.0879	0.1065	0.0408	0.1759
	permutation	0.1203	0.1614	0.1614	0.1519	0.0466
β = 4.00	standard	0.5541	0.5549	0.0726	0.5660	0.0645
	bootstrap	0.1214	0.1175	0.1201	0.0559	0.2290
	permutation	0.1643	0.2143	0.2143	0.1998	0.0648
β = 6.00	standard	0.8080	0.7885	0.0698	0.7833	0.0885
	bootstrap	0.1725	0.1700	0.1296	0.0812	0.2865
	permutation	0.2282	0.2855	0.2855	0.2649	0.0895
β = 10.0	standard	0.9764	0.9683	0.0251	0.9626	0.0990
	bootstrap	0.2036	0.1969	0.0665	0.1116	0.3078
	permutation	0.2518	0.3198	0.3198	0.2979	0.1000
β = 20.0	standard	0.9980	0.9999	0.0012	0.9997	0.1048
	bootstrap	0.2336	0.2229	0.0061	0.1557	0.3270
	permutation	0.2637	0.3406	0.3406	0.3209	0.1057

Table 4: Results from simulations of failure truncated TRP processes with $n = 5$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 5	resampling	milhand	laplace	lew-rob	and-dar	mann
γ = -0.0300	standard	0.1336	0.1163	0.0291	0.1632	0.0162
	bootstrap	0.0323	0.0345	0.0451	0.0132	0.0832
	permutation	0.0549	0.0539	0.0539	0.0575	0.0568
γ = -0.0175	standard	0.1316	0.1090	0.0271	0.1549	0.0153
	bootstrap	0.0301	0.0297	0.0408	0.0125	0.0757
	permutation	0.0486	0.0463	0.0463	0.0518	0.0502
γ = -0.0100	standard	0.1290	0.1053	0.0287	0.1513	0.0156
	bootstrap	0.0322	0.0330	0.0441	0.0140	0.0777
	permutation	0.0527	0.0535	0.0535	0.0557	0.0507
γ = -0.0050	standard	0.1242	0.1087	0.0273	0.1574	0.0156
	bootstrap	0.0309	0.0338	0.0419	0.0155	0.0785
	permutation	0.0512	0.0500	0.0500	0.0537	0.0522
γ = 0.0000	standard	0.1301	0.1101	0.0292	0.1537	0.0162
	bootstrap	0.0330	0.0340	0.0457	0.0145	0.0841
	permutation	0.0521	0.0532	0.0532	0.0571	0.0521
γ = 0.0500	standard	0.1321	0.1093	0.0323	0.1584	0.0178
	bootstrap	0.0346	0.0339	0.0467	0.0169	0.0804
	permutation	0.0527	0.0564	0.0564	0.0593	0.0448
γ = 0.1000	standard	0.1331	0.1189	0.0298	0.1672	0.0172
	bootstrap	0.0331	0.0328	0.0429	0.0147	0.0782
	permutation	0.0525	0.0543	0.0544	0.0594	0.0454
γ = 0.2000	standard	0.1372	0.1281	0.0294	0.1762	0.0180
	bootstrap	0.0352	0.0346	0.0454	0.0168	0.0782
	permutation	0.0522	0.0595	0.0596	0.0602	0.0375
γ = 0.5000	standard	0.1688	0.1539	0.0349	0.2009	0.0250
	bootstrap	0.0436	0.0420	0.0525	0.0203	0.0920
	permutation	0.0600	0.0722	0.0722	0.0707	0.0382
γ = 1.0000	standard	0.2148	0.1922	0.0390	0.2445	0.0270
	bootstrap	0.0501	0.0478	0.0612	0.0230	0.1097
	permutation	0.0713	0.0884	0.0884	0.0856	0.0373
γ = 2.0000	standard	0.2578	0.2567	0.0445	0.3042	0.0317
	bootstrap	0.0620	0.0583	0.0695	0.0281	0.1269
	permutation	0.0864	0.1085	0.1085	0.1070	0.0369
γ = 5.0000	standard	0.3611	0.3573	0.0490	0.4025	0.0414
	bootstrap	0.0766	0.0744	0.0811	0.0370	0.1581
	permutation	0.1062	0.1368	0.1368	0.1303	0.0450
γ = 10.000	standard	0.4354	0.4328	0.0581	0.4766	0.0490
	bootstrap	0.0946	0.0931	0.0944	0.0453	0.1870
	permutation	0.1276	0.1691	0.1691	0.1609	0.0505
γ = 20.000	standard	0.5196	0.5158	0.0595	0.5554	0.0580
	bootstrap	0.1099	0.1051	0.1043	0.0514	0.2103
	permutation	0.1484	0.1973	0.1973	0.1879	0.0594
γ = 40.000	standard	0.5920	0.6051	0.0653	0.6373	0.0707
	bootstrap	0.1272	0.1246	0.1133	0.0674	0.2390
	permutation	0.1754	0.2274	0.2274	0.2148	0.0717
γ = 100.00	standard	0.6933	0.6837	0.0719	0.7084	0.0793
	bootstrap	0.1480	0.1468	0.1233	0.0750	0.2612
	permutation	0.1956	0.2570	0.2569	0.2423	0.0798

Table 5: Results from simulations of failure truncated TRP processes with $n = 5$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 5	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.9999	0.9917	0.0174	0.9998	0.9057
β = 0.05	bootstrap	0.2030	0.0555	0.0665	0.0433	0.9946
	permutation	0.9951	0.9331	0.9330	0.5799	0.9947
	standard	0.9910	0.9011	0.1617	0.9801	0.6240
β = 0.15	bootstrap	0.2244	0.1524	0.2767	0.0652	0.9164
	permutation	0.9152	0.7316	0.7318	0.8786	0.9164
	standard	0.8929	0.6978	0.2500	0.8381	0.3945
β = 0.25	bootstrap	0.2448	0.2041	0.3622	0.0831	0.7609
	permutation	0.7434	0.5523	0.5528	0.6874	0.7621
	standard	0.2515	0.1867	0.1743	0.2127	0.1165
β = 0.50	bootstrap	0.1589	0.1452	0.2363	0.0638	0.3588
	permutation	0.3052	0.2294	0.2297	0.2780	0.3577
	standard	0.0290	0.0250	0.0675	0.0250	0.0305
β = 0.75	bootstrap	0.0713	0.0652	0.0905	0.0292	0.1372
	permutation	0.1022	0.0779	0.0780	0.0948	0.1291
	standard	0.0056	0.0056	0.0364	0.0049	0.0161
β = 1.00	bootstrap	0.0421	0.0430	0.0498	0.0186	0.0721
	permutation	0.0497	0.0500	0.0500	0.0532	0.0473
	standard	0.0119	0.0109	0.0440	0.0065	0.0219
β = 1.25	bootstrap	0.0547	0.0526	0.0618	0.0268	0.0851
	permutation	0.0563	0.0713	0.0712	0.0690	0.0346
	standard	0.0293	0.0236	0.0701	0.0165	0.0336
β = 1.50	bootstrap	0.0863	0.0856	0.0950	0.0435	0.1254
	permutation	0.0858	0.1133	0.1133	0.1103	0.0376
	standard	0.0615	0.0534	0.0942	0.0371	0.0490
β = 1.75	bootstrap	0.1277	0.1252	0.1250	0.0662	0.1687
	permutation	0.1248	0.1606	0.1606	0.1576	0.0498
	standard	0.1123	0.1032	0.1040	0.0705	0.0620
β = 2.00	bootstrap	0.1774	0.1668	0.1459	0.0872	0.2048
	permutation	0.1583	0.2096	0.2096	0.2002	0.0634
	standard	0.1844	0.1563	0.1152	0.1062	0.0778
β = 2.25	bootstrap	0.2125	0.2041	0.1627	0.1128	0.2497
	permutation	0.1950	0.2471	0.2471	0.2391	0.0784
	standard	0.2730	0.2245	0.1176	0.1615	0.0872
β = 2.50	bootstrap	0.2514	0.2357	0.1696	0.1298	0.2769
	permutation	0.2160	0.2776	0.2776	0.2667	0.0879
	standard	0.4484	0.4022	0.1104	0.3075	0.1074
β = 3.00	bootstrap	0.3113	0.3032	0.1686	0.1630	0.3263
	permutation	0.2611	0.3339	0.3339	0.3223	0.1082
	standard	0.7493	0.7054	0.0777	0.6050	0.1313
β = 4.00	bootstrap	0.3697	0.3623	0.1331	0.2123	0.3779
	permutation	0.3080	0.3874	0.3874	0.3730	0.1321
	standard	0.9720	0.9610	0.0264	0.9282	0.1472
β = 6.00	bootstrap	0.4160	0.4170	0.0590	0.2701	0.4216
	permutation	0.3475	0.4383	0.4383	0.4260	0.1478
	standard	0.9996	0.9997	0.0023	0.9988	0.1619
β = 10.0	bootstrap	0.4031	0.3983	0.0091	0.3081	0.4363
	permutation	0.3737	0.4595	0.4595	0.4440	0.1631
	standard	1.0000	1.0000	0.0000	1.0000	0.1689
β = 20.0	bootstrap	0.3915	0.3742	0.0001	0.3403	0.4487
	permutation	0.3857	0.4748	0.4748	0.4576	0.1700

Table 6: Results from simulations of failure truncated TRP processes with $n = 5$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50 n = 5	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.0050	0.0066	0.0397	0.0048	0.0171
γ = -0.0300	bootstrap	0.0464	0.0431	0.0531	0.0205	0.0805
	permutation	0.0558	0.0495	0.0496	0.0587	0.0603
	standard	0.0054	0.0055	0.0393	0.0045	0.0163
γ = -0.0175	bootstrap	0.0421	0.0418	0.0515	0.0172	0.0765
	permutation	0.0521	0.0503	0.0503	0.0558	0.0522
	standard	0.0050	0.0057	0.0363	0.0050	0.0162
γ = -0.0100	bootstrap	0.0427	0.0417	0.0508	0.0186	0.0760
	permutation	0.0524	0.0516	0.0518	0.0585	0.0548
	standard	0.0047	0.0070	0.0410	0.0051	0.0176
γ = -0.0050	bootstrap	0.0449	0.0431	0.0558	0.0197	0.0783
	permutation	0.0515	0.0540	0.0541	0.0566	0.0513
	standard	0.0056	0.0059	0.0403	0.0043	0.0167
γ = 0.0000	bootstrap	0.0446	0.0434	0.0547	0.0211	0.0786
	permutation	0.0559	0.0539	0.0539	0.0600	0.0516
	standard	0.0074	0.0068	0.0389	0.0059	0.0163
γ = 0.0500	bootstrap	0.0416	0.0402	0.0540	0.0177	0.0748
	permutation	0.0495	0.0540	0.0541	0.0575	0.0407
	standard	0.0086	0.0085	0.0418	0.0069	0.0179
γ = 0.1000	bootstrap	0.0493	0.0479	0.0590	0.0225	0.0794
	permutation	0.0528	0.0617	0.0619	0.0641	0.0389
	standard	0.0139	0.0129	0.0504	0.0087	0.0225
γ = 0.2000	bootstrap	0.0581	0.0570	0.0676	0.0265	0.0903
	permutation	0.0633	0.0774	0.0775	0.0770	0.0345
	standard	0.0374	0.0324	0.0758	0.0241	0.0376
γ = 0.5000	bootstrap	0.0892	0.0887	0.1027	0.0459	0.1389
	permutation	0.0946	0.1252	0.1252	0.1203	0.0423
	standard	0.0815	0.0742	0.1004	0.0544	0.0548
γ = 1.0000	bootstrap	0.1419	0.1357	0.1408	0.0712	0.1946
	permutation	0.1378	0.1819	0.1819	0.1756	0.0580
	standard	0.1798	0.1627	0.1278	0.1217	0.0800
γ = 2.0000	bootstrap	0.1907	0.1892	0.1777	0.1034	0.2511
	permutation	0.1967	0.2469	0.2469	0.2349	0.0810
	standard	0.3735	0.3441	0.1344	0.2771	0.1113
γ = 5.0000	bootstrap	0.2835	0.2813	0.2035	0.1643	0.3317
	permutation	0.2707	0.3377	0.3377	0.3252	0.1123
	standard	0.5477	0.5097	0.1265	0.4284	0.1344
γ = 10.000	bootstrap	0.3213	0.3248	0.1959	0.1901	0.3877
	permutation	0.3114	0.3900	0.3900	0.3729	0.1354
	standard	0.6839	0.6587	0.1152	0.5854	0.1540
γ = 20.000	bootstrap	0.3661	0.3680	0.1879	0.2218	0.4244
	permutation	0.3497	0.4320	0.4320	0.4176	0.1549
	standard	0.7958	0.7763	0.0894	0.7100	0.1623
γ = 40.000	bootstrap	0.3960	0.4007	0.1538	0.2495	0.4260
	permutation	0.3672	0.4524	0.4524	0.4390	0.1631
	standard	0.8947	0.8837	0.0728	0.8342	0.1781
γ = 100.00	bootstrap	0.4278	0.4324	0.1285	0.2827	0.4594
	permutation	0.4000	0.4901	0.4901	0.4749	0.1793

Table 7: Results from simulations of failure truncated TRP processes with $n = 5$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, n = 15	resampling	milhand	laplace	lew-rob	and-dar	mann
β = 0.25	standard	0.9976	0.9705	0.8392	0.9957	0.9758
	bootstrap	0.9935	0.8796	0.8664	0.9784	0.9776
	permutation	0.9947	0.8791	0.8791	0.9830	0.9782
β = 0.50	standard	0.7711	0.5853	0.4572	0.6955	0.5613
	bootstrap	0.6771	0.4836	0.4684	0.5610	0.5858
	permutation	0.6932	0.4756	0.4756	0.5978	0.5912
β = 0.75	standard	0.2297	0.1660	0.1377	0.1923	0.1465
	bootstrap	0.1862	0.1487	0.1411	0.1470	0.1614
	permutation	0.1929	0.1431	0.1431	0.1603	0.1658
β = 1.00	standard	0.0450	0.0487	0.0480	0.0525	0.0427
	bootstrap	0.0530	0.0532	0.0472	0.0444	0.0542
	permutation	0.0500	0.0500	0.0500	0.0489	0.0516
β = 1.25	standard	0.1092	0.0983	0.0896	0.0902	0.0591
	bootstrap	0.1202	0.0996	0.0914	0.0794	0.0819
	permutation	0.1043	0.0941	0.0941	0.0898	0.0747
β = 1.50	standard	0.2649	0.2442	0.1996	0.2280	0.1210
	bootstrap	0.2682	0.2189	0.2020	0.1811	0.1601
	permutation	0.2359	0.2087	0.2087	0.1987	0.1481
β = 1.75	standard	0.4722	0.4191	0.3268	0.3986	0.1899
	bootstrap	0.4513	0.3638	0.3301	0.3035	0.2420
	permutation	0.3975	0.3454	0.3454	0.3242	0.2267
β = 2.00	standard	0.6624	0.5981	0.4497	0.5736	0.2586
	bootstrap	0.6164	0.5093	0.4651	0.4261	0.3120
	permutation	0.5499	0.4872	0.4872	0.4486	0.2988
β = 2.25	standard	0.8093	0.7599	0.5542	0.7307	0.3185
	bootstrap	0.7629	0.6427	0.5745	0.5414	0.3805
	permutation	0.6915	0.6153	0.6153	0.5683	0.3637
β = 2.50	standard	0.9084	0.8609	0.6219	0.8403	0.3622
	bootstrap	0.8543	0.7361	0.6540	0.6279	0.4247
	permutation	0.7756	0.7005	0.7005	0.6529	0.4048

Table 8: Results from simulations of failure truncated TRP processes with $n = 15$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.0$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, n = 15	resampling	milhand	laplace	lew-rob	and-dar	mann
γ = -0.0175	standard	0.0568	0.0594	0.0572	0.0622	0.0492
	bootstrap	0.0532	0.0646	0.0597	0.0511	0.0591
	permutation	0.0542	0.0597	0.0597	0.0569	0.0607
γ = -0.0100	standard	0.0535	0.0558	0.0553	0.0563	0.0467
	bootstrap	0.0540	0.0588	0.0548	0.0475	0.0568
	permutation	0.0546	0.0555	0.0555	0.0545	0.0577
γ = -0.0050	standard	0.0511	0.0504	0.0484	0.0511	0.0418
	bootstrap	0.0500	0.0513	0.0488	0.0434	0.0533
	permutation	0.0469	0.0492	0.0492	0.0484	0.0528
γ = -0.0020	standard	0.0481	0.0494	0.0515	0.0513	0.0431
	bootstrap	0.0530	0.0553	0.0512	0.0453	0.0543
	permutation	0.0493	0.0524	0.0524	0.0517	0.0530
γ = 0.0000	standard	0.0507	0.0477	0.0501	0.0482	0.0448
	bootstrap	0.0501	0.0532	0.0488	0.0436	0.0559
	permutation	0.0482	0.0519	0.0519	0.0515	0.0559
γ = 0.0500	standard	0.0874	0.0891	0.0856	0.0907	0.0596
	bootstrap	0.0927	0.0892	0.0826	0.0750	0.0804
	permutation	0.0826	0.0850	0.0850	0.0829	0.0754
γ = 0.1000	standard	0.1485	0.1622	0.1371	0.1608	0.0936
	bootstrap	0.1553	0.1472	0.1361	0.1268	0.1215
	permutation	0.1347	0.1402	0.1402	0.1389	0.1133
γ = 0.2000	standard	0.2821	0.3073	0.2487	0.3018	0.1693
	bootstrap	0.2660	0.2698	0.2503	0.2317	0.2094
	permutation	0.2409	0.2602	0.2602	0.2528	0.1982
γ = 0.5000	standard	0.5242	0.5829	0.4480	0.5674	0.3056
	bootstrap	0.4913	0.4995	0.4598	0.4427	0.3628
	permutation	0.4429	0.4800	0.4800	0.4651	0.3427
γ = 1.0000	standard	0.7300	0.7876	0.6018	0.7720	0.4223
	bootstrap	0.6576	0.6869	0.6286	0.6225	0.4857
	permutation	0.6103	0.6620	0.6620	0.6457	0.4646
γ = 2.0000	standard	0.8727	0.9214	0.7219	0.9138	0.5089
	bootstrap	0.7942	0.8326	0.7613	0.7723	0.5733
	permutation	0.7465	0.8079	0.8079	0.7910	0.5544

Table 9: Results from simulations of failure truncated TRP processes with $n = 15$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.0$ (NHPP). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 15	resampling	milhand	laplace	lew-rob	and-dar	mann
β = 0.25	standard	0.9888	0.9168	0.6024	0.9831	0.8982
	bootstrap	0.9540	0.6766	0.6524	0.8898	0.9074
	permutation	0.9590	0.6837	0.6837	0.9126	0.9089
β = 0.50	standard	0.7580	0.5452	0.2661	0.7198	0.4123
	bootstrap	0.5046	0.3009	0.2847	0.3816	0.4370
	permutation	0.5281	0.2960	0.2960	0.4296	0.4404
β = 0.75	standard	0.3392	0.2316	0.0928	0.3373	0.1157
	bootstrap	0.1437	0.1082	0.1004	0.1069	0.1311
	permutation	0.1533	0.1045	0.1045	0.1261	0.1322
β = 1.00	standard	0.1316	0.1308	0.0468	0.1837	0.0378
	bootstrap	0.0531	0.0516	0.0470	0.0412	0.0512
	permutation	0.0517	0.0504	0.0504	0.0484	0.0482
β = 1.25	standard	0.1491	0.1662	0.0643	0.2030	0.0501
	bootstrap	0.0854	0.0740	0.0676	0.0578	0.0654
	permutation	0.0715	0.0694	0.0694	0.0645	0.0604
β = 1.50	standard	0.2619	0.2816	0.1246	0.3171	0.0827
	bootstrap	0.1720	0.1458	0.1302	0.1105	0.1105
	permutation	0.1453	0.1374	0.1374	0.1272	0.1018
β = 1.75	standard	0.4189	0.4102	0.1907	0.4368	0.1251
	bootstrap	0.2757	0.2263	0.1994	0.1645	0.1582
	permutation	0.2323	0.2128	0.2128	0.1911	0.1501
β = 2.00	standard	0.5725	0.5725	0.2907	0.5935	0.1779
	bootstrap	0.4149	0.3420	0.3067	0.2542	0.2228
	permutation	0.3550	0.3264	0.3264	0.2871	0.2087
β = 2.25	standard	0.7052	0.6843	0.3735	0.6976	0.2117
	bootstrap	0.5333	0.4354	0.3887	0.3261	0.2615
	permutation	0.4612	0.4194	0.4194	0.3659	0.2465
β = 2.50	standard	0.8117	0.7874	0.4333	0.8007	0.2514
	bootstrap	0.6342	0.5219	0.4613	0.3916	0.3028
	permutation	0.5509	0.4951	0.4951	0.4337	0.2874

Table 10: Results from simulations of failure truncated TRP processes with $n = 15$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 15	resampling	milhand	laplace	lew-rob	and-dar	mann
γ = -0.0175	standard	0.1568	0.1514	0.0546	0.2010	0.0477
	bootstrap	0.0526	0.0656	0.0578	0.0502	0.0587
	permutation	0.0533	0.0619	0.0619	0.0593	0.0601
γ = -0.0100	standard	0.1483	0.1320	0.0483	0.1870	0.0440
	bootstrap	0.0521	0.0560	0.0511	0.0433	0.0537
	permutation	0.0513	0.0536	0.0536	0.0529	0.0548
γ = -0.0050	standard	0.1419	0.1313	0.0467	0.1844	0.0412
	bootstrap	0.0546	0.0536	0.0489	0.0427	0.0543
	permutation	0.0537	0.0512	0.0512	0.0510	0.0508
γ = -0.0020	standard	0.1366	0.1237	0.0441	0.1774	0.0418
	bootstrap	0.0532	0.0529	0.0455	0.0404	0.0516
	permutation	0.0505	0.0490	0.0490	0.0489	0.0506
γ = 0.0000	standard	0.1267	0.1269	0.0446	0.1772	0.0382
	bootstrap	0.0526	0.0530	0.0473	0.0423	0.0486
	permutation	0.0512	0.0505	0.0505	0.0501	0.0476
γ = 0.0500	standard	0.1542	0.1703	0.0665	0.2242	0.0551
	bootstrap	0.0815	0.0764	0.0688	0.0601	0.0731
	permutation	0.0718	0.0710	0.0710	0.0718	0.0699
γ = 0.1000	standard	0.2037	0.2420	0.0985	0.2961	0.0737
	bootstrap	0.1234	0.1138	0.1057	0.0916	0.0959
	permutation	0.1040	0.1079	0.1079	0.1054	0.0896
γ = 0.2000	standard	0.2980	0.3640	0.1572	0.4182	0.1153
	bootstrap	0.1944	0.1836	0.1659	0.1514	0.1476
	permutation	0.1650	0.1758	0.1758	0.1730	0.1362
γ = 0.5000	standard	0.4899	0.5731	0.2762	0.6105	0.1877
	bootstrap	0.3442	0.3340	0.2928	0.2731	0.2351
	permutation	0.2977	0.3140	0.3140	0.3037	0.2216
γ = 1.0000	standard	0.6438	0.7355	0.3942	0.7643	0.2731
	bootstrap	0.4763	0.4799	0.4257	0.3983	0.3274
	permutation	0.4189	0.4536	0.4536	0.4375	0.3118
γ = 2.0000	standard	0.7866	0.8693	0.5008	0.8862	0.3464
	bootstrap	0.6199	0.6304	0.5511	0.5281	0.4107
	permutation	0.5578	0.5945	0.5945	0.5714	0.3937

Table 11: Results from simulations of failure truncated TRP processes with $n = 15$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 15	resampling	milhand	laplace	lew-rob	and-dar	mann
β = 0.25	standard	0.9999	0.9977	0.9871	0.9998	0.9993
	bootstrap	1.0000	0.9911	0.9891	0.9996	0.9994
	permutation	1.0000	0.9916	0.9916	0.9995	0.9994
β = 0.50	standard	0.8046	0.6517	0.7663	0.6969	0.8115
	bootstrap	0.9082	0.7739	0.7643	0.8255	0.8274
	permutation	0.9117	0.7699	0.7699	0.8406	0.8314
β = 0.75	standard	0.0966	0.0832	0.2378	0.0718	0.2323
	bootstrap	0.2919	0.2430	0.2353	0.2277	0.2521
	permutation	0.2981	0.2381	0.2381	0.2433	0.2587
β = 1.00	standard	0.0056	0.0035	0.0472	0.0019	0.0374
	bootstrap	0.0506	0.0476	0.0465	0.0432	0.0490
	permutation	0.0487	0.0456	0.0456	0.0463	0.0467
β = 1.25	standard	0.0509	0.0314	0.1524	0.0199	0.0939
	bootstrap	0.1954	0.1554	0.1441	0.1420	0.1223
	permutation	0.1762	0.1495	0.1495	0.1493	0.1144
β = 1.50	standard	0.2470	0.1567	0.3746	0.1094	0.2268
	bootstrap	0.4964	0.3956	0.3708	0.3612	0.2755
	permutation	0.4503	0.3797	0.3797	0.3758	0.2568
β = 1.75	standard	0.5388	0.4111	0.5806	0.3128	0.3412
	bootstrap	0.7573	0.6316	0.5865	0.5888	0.4061
	permutation	0.6987	0.6071	0.6071	0.6001	0.3879
β = 2.00	standard	0.7960	0.6709	0.7212	0.5724	0.4431
	bootstrap	0.8990	0.7977	0.7369	0.7539	0.5081
	permutation	0.8506	0.7746	0.7746	0.7660	0.4878
β = 2.25	standard	0.9285	0.8612	0.8022	0.7945	0.5273
	bootstrap	0.9607	0.8919	0.8290	0.8512	0.5880
	permutation	0.9290	0.8739	0.8739	0.8627	0.5671
β = 2.50	standard	0.9809	0.9526	0.8283	0.9198	0.5723
	bootstrap	0.9841	0.9409	0.8696	0.9066	0.6351
	permutation	0.9540	0.9226	0.9226	0.9138	0.6142

Table 12: Results from simulations of failure truncated TRP processes with $n = 15$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 15	resampling	milhand	laplace	lew-rob	and-dar	mann
γ = -0.0175	standard	0.0047	0.0092	0.0730	0.0055	0.0584
	bootstrap	0.0625	0.0778	0.0737	0.0647	0.0698
	permutation	0.0626	0.0711	0.0711	0.0685	0.0715
γ = -0.0100	standard	0.0051	0.0057	0.0587	0.0032	0.0478
	bootstrap	0.0562	0.0612	0.0579	0.0515	0.0589
	permutation	0.0538	0.0593	0.0593	0.0576	0.0588
γ = -0.0050	standard	0.0059	0.0051	0.0568	0.0037	0.0458
	bootstrap	0.0547	0.0562	0.0550	0.0484	0.0563
	permutation	0.0533	0.0539	0.0539	0.0539	0.0576
γ = -0.0020	standard	0.0047	0.0048	0.0518	0.0030	0.0418
	bootstrap	0.0487	0.0533	0.0496	0.0436	0.0529
	permutation	0.0485	0.0501	0.0501	0.0490	0.0516
γ = 0.0000	standard	0.0054	0.0036	0.0509	0.0017	0.0409
	bootstrap	0.0507	0.0497	0.0481	0.0437	0.0524
	permutation	0.0493	0.0504	0.0504	0.0493	0.0517
γ = 0.0500	standard	0.0247	0.0207	0.1237	0.0120	0.0839
	bootstrap	0.1298	0.1234	0.1195	0.1079	0.1075
	permutation	0.1163	0.1196	0.1196	0.1148	0.1013
γ = 0.1000	standard	0.0798	0.0623	0.2342	0.0418	0.1549
	bootstrap	0.2365	0.2375	0.2288	0.2152	0.1957
	permutation	0.2168	0.2323	0.2323	0.2234	0.1825
γ = 0.2000	standard	0.2294	0.2213	0.4475	0.1617	0.3030
	bootstrap	0.4420	0.4608	0.4405	0.4247	0.3555
	permutation	0.4160	0.4467	0.4467	0.4348	0.3427
γ = 0.5000	standard	0.5973	0.6176	0.7533	0.5298	0.5597
	bootstrap	0.7429	0.7908	0.7583	0.7545	0.6212
	permutation	0.7081	0.7702	0.7702	0.7583	0.6035
γ = 1.0000	standard	0.8458	0.8663	0.8859	0.8169	0.7041
	bootstrap	0.8834	0.9247	0.8972	0.8996	0.7543
	permutation	0.8592	0.9151	0.9151	0.9057	0.7383
γ = 2.0000	standard	0.9573	0.9764	0.9316	0.9609	0.7952
	bootstrap	0.9513	0.9775	0.9505	0.9658	0.8357
	permutation	0.9291	0.9697	0.9697	0.9659	0.8250

Table 13: Results from simulations of failure truncated TRP processes with $n = 15$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, n = 35	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	1.0000	0.9999	0.9991	1.0000	0.9999
β = 0.25	bootstrap	1.0000	0.9998	0.9995	1.0000	0.9999
	permutation	1.0000	0.9996	0.9996	1.0000	0.9999
	standard	0.9799	0.9109	0.8619	0.9549	0.9144
β = 0.50	bootstrap	0.9666	0.8705	0.8640	0.9308	0.9150
	permutation	0.9668	0.8673	0.8673	0.9322	0.9153
	standard	0.4302	0.3152	0.2808	0.3577	0.3017
β = 0.75	bootstrap	0.3926	0.2960	0.2862	0.3177	0.3042
	permutation	0.3977	0.2872	0.2872	0.3236	0.3072
	standard	0.1753	0.1365	0.1286	0.1491	0.1332
β = 0.85	bootstrap	0.1649	0.1347	0.1299	0.1355	0.1354
	permutation	0.1665	0.1329	0.1329	0.1390	0.1342
	standard	0.0488	0.0499	0.0491	0.0495	0.0478
β = 1.00	bootstrap	0.0495	0.0526	0.0495	0.0478	0.0515
	permutation	0.0486	0.0503	0.0503	0.0479	0.0510
	standard	0.1157	0.1020	0.0962	0.0965	0.0751
β = 1.15	bootstrap	0.1199	0.1004	0.0938	0.0893	0.0813
	permutation	0.1133	0.0955	0.0955	0.0910	0.0802
	standard	0.2214	0.1881	0.1772	0.1783	0.1345
β = 1.25	bootstrap	0.2276	0.1818	0.1755	0.1643	0.1451
	permutation	0.2106	0.1788	0.1788	0.1702	0.1430
	standard	0.6145	0.5272	0.4679	0.5116	0.3087
β = 1.50	bootstrap	0.6019	0.4871	0.4701	0.4511	0.3231
	permutation	0.5759	0.4738	0.4738	0.4590	0.3181
	standard	0.8846	0.8323	0.7459	0.8201	0.4777
β = 1.75	bootstrap	0.8814	0.7700	0.7501	0.7350	0.4981
	permutation	0.8608	0.7591	0.7591	0.7432	0.4927
	standard	0.9838	0.9530	0.8801	0.9515	0.5942
β = 2.00	bootstrap	0.9725	0.9050	0.8863	0.8801	0.6120
	permutation	0.9619	0.8982	0.8982	0.8877	0.6065
	standard	0.9970	0.9931	0.9412	0.9929	0.6840
β = 2.25	bootstrap	0.9962	0.9684	0.9522	0.9538	0.6981
	permutation	0.9928	0.9644	0.9644	0.9598	0.6926
	standard	0.9998	0.9989	0.9695	0.9993	0.7397
β = 2.50	bootstrap	0.9995	0.9887	0.9787	0.9821	0.7532
	permutation	0.9983	0.9864	0.9864	0.9846	0.7487

Table 14: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.0$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00 n = 35	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.0968	0.1142	0.1106	0.1096	0.0920
γ = -0.010	bootstrap	0.0849	0.1169	0.1125	0.1018	0.0958
	permutation	0.0862	0.1133	0.1133	0.1050	0.0947
	standard	0.0581	0.0597	0.0589	0.0610	0.0575
γ = -0.005	bootstrap	0.0571	0.0632	0.0609	0.0570	0.0598
	permutation	0.0569	0.0615	0.0615	0.0587	0.0589
	standard	0.0528	0.0488	0.0505	0.0509	0.0464
γ = -0.002	bootstrap	0.0499	0.0549	0.0520	0.0488	0.0503
	permutation	0.0490	0.0503	0.0503	0.0496	0.0486
	standard	0.0521	0.0521	0.0543	0.0525	0.0498
γ = -0.001	bootstrap	0.0536	0.0593	0.0555	0.0511	0.0519
	permutation	0.0516	0.0561	0.0561	0.0525	0.0534
	standard	0.0503	0.0488	0.0494	0.0521	0.0478
γ = 0.000	bootstrap	0.0531	0.0519	0.0499	0.0492	0.0525
	permutation	0.0505	0.0512	0.0512	0.0494	0.0522
	standard	0.1259	0.1438	0.1355	0.1392	0.1079
γ = 0.020	bootstrap	0.1358	0.1410	0.1335	0.1281	0.1156
	permutation	0.1256	0.1353	0.1353	0.1289	0.1137
	standard	0.3320	0.3935	0.3563	0.3826	0.2693
γ = 0.050	bootstrap	0.3252	0.3677	0.3544	0.3445	0.2842
	permutation	0.3079	0.3589	0.3589	0.3512	0.2796
	standard	0.5964	0.6892	0.6390	0.6735	0.4849
γ = 0.100	bootstrap	0.5763	0.6538	0.6371	0.6256	0.5031
	permutation	0.5522	0.6434	0.6434	0.6307	0.4977
	standard	0.8439	0.9115	0.8682	0.9042	0.7060
γ = 0.200	bootstrap	0.8000	0.8772	0.8666	0.8618	0.7162
	permutation	0.7793	0.8754	0.8754	0.8665	0.7145
	standard	0.9810	0.9949	0.9795	0.9941	0.8812
γ = 0.500	bootstrap	0.9582	0.9885	0.9835	0.9828	0.8886
	permutation	0.9449	0.9859	0.9859	0.9838	0.8866
	standard	0.9974	0.9997	0.9933	0.9997	0.9319
γ = 1.000	bootstrap	0.9882	0.9991	0.9965	0.9972	0.9378
	permutation	0.9826	0.9977	0.9977	0.9976	0.9358
	standard	0.9998	1.0000	0.9942	1.0000	0.9511
γ = 2.000	bootstrap	0.9970	0.9998	0.9981	0.9994	0.9547
	permutation	0.9938	0.9998	0.9998	0.9996	0.9516

Table 15: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.0$ (NHPP). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 35	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.9999	0.9986	0.9665	0.9999	0.9989
β = 0.25	bootstrap	0.9997	0.9811	0.9744	0.9991	0.9987
	permutation	0.9999	0.9782	0.9782	0.9992	0.9985
	standard	0.9539	0.8372	0.6274	0.9291	0.7686
β = 0.50	bootstrap	0.8666	0.6567	0.6393	0.7704	0.7682
	permutation	0.8706	0.6469	0.6469	0.7798	0.7703
	standard	0.5084	0.3504	0.1703	0.4690	0.2158
β = 0.75	bootstrap	0.2759	0.1880	0.1759	0.2122	0.2173
	permutation	0.2788	0.1823	0.1823	0.2181	0.2164
	standard	0.2969	0.2102	0.0875	0.2973	0.1002
β = 0.85	bootstrap	0.1242	0.0989	0.0922	0.0982	0.1042
	permutation	0.1278	0.0914	0.0914	0.1011	0.1037
	standard	0.1372	0.1399	0.0463	0.1988	0.0493
β = 1.00	bootstrap	0.0497	0.0538	0.0481	0.0451	0.0543
	permutation	0.0480	0.0490	0.0490	0.0502	0.0541
	standard	0.1584	0.1828	0.0716	0.2285	0.0653
β = 1.15	bootstrap	0.0924	0.0794	0.0749	0.0677	0.0712
	permutation	0.0847	0.0762	0.0762	0.0716	0.0692
	standard	0.2427	0.2626	0.1166	0.3054	0.0971
β = 1.25	bootstrap	0.1518	0.1259	0.1182	0.1075	0.1016
	permutation	0.1351	0.1208	0.1208	0.1110	0.1006
	standard	0.5296	0.5146	0.2865	0.5531	0.2027
β = 1.50	bootstrap	0.3832	0.3113	0.2943	0.2650	0.2171
	permutation	0.3531	0.2970	0.2970	0.2707	0.2100
	standard	0.7931	0.7533	0.4972	0.7807	0.3251
β = 1.75	bootstrap	0.6494	0.5335	0.5055	0.4697	0.3434
	permutation	0.6138	0.5166	0.5166	0.4808	0.3364
	standard	0.9271	0.8899	0.6777	0.9075	0.4289
β = 2.00	bootstrap	0.8394	0.7232	0.6906	0.6527	0.4445
	permutation	0.8090	0.7054	0.7054	0.6644	0.4425
	standard	0.9799	0.9658	0.7902	0.9751	0.5080
β = 2.25	bootstrap	0.9408	0.8385	0.8085	0.7834	0.5254
	permutation	0.9190	0.8243	0.8243	0.7902	0.5191
	standard	0.9971	0.9897	0.8570	0.9928	0.5547
β = 2.50	bootstrap	0.9796	0.9100	0.8786	0.8589	0.5695
	permutation	0.9678	0.8979	0.8979	0.8696	0.5666

Table 16: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 35	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.2341	0.2340	0.0990	0.2918	0.0923
γ = -0.010	bootstrap	0.0870	0.1117	0.1044	0.0979	0.0948
	permutation	0.0890	0.1067	0.1067	0.1002	0.0938
	standard	0.1624	0.1478	0.0527	0.2039	0.0543
γ = -0.005	bootstrap	0.0554	0.0606	0.0565	0.0541	0.0565
	permutation	0.0543	0.0565	0.0565	0.0564	0.0575
	standard	0.1442	0.1432	0.0531	0.2030	0.0503
γ = -0.002	bootstrap	0.0534	0.0570	0.0540	0.0523	0.0518
	permutation	0.0515	0.0545	0.0545	0.0552	0.0520
	standard	0.1404	0.1441	0.0541	0.1990	0.0533
γ = -0.001	bootstrap	0.0574	0.0604	0.0554	0.0518	0.0551
	permutation	0.0547	0.0550	0.0550	0.0544	0.0551
	standard	0.1405	0.1384	0.0479	0.1958	0.0469
γ = 0.000	bootstrap	0.0564	0.0538	0.0497	0.0473	0.0514
	permutation	0.0527	0.0512	0.0512	0.0500	0.0515
	standard	0.2000	0.2472	0.1011	0.3056	0.0941
γ = 0.020	bootstrap	0.1102	0.1098	0.1019	0.0996	0.1021
	permutation	0.1014	0.1046	0.1046	0.1046	0.0999
	standard	0.3618	0.4533	0.2429	0.5074	0.1957
γ = 0.050	bootstrap	0.2331	0.2624	0.2497	0.2410	0.2065
	permutation	0.2172	0.2543	0.2543	0.2476	0.2036
	standard	0.5740	0.6775	0.4354	0.7155	0.3378
γ = 0.100	bootstrap	0.4094	0.4669	0.4420	0.4345	0.3529
	permutation	0.3791	0.4551	0.4551	0.4426	0.3462
	standard	0.7778	0.8824	0.6754	0.8983	0.5110
γ = 0.200	bootstrap	0.6328	0.7133	0.6867	0.6753	0.5263
	permutation	0.5987	0.6989	0.6989	0.6849	0.5251
	standard	0.9454	0.9829	0.8765	0.9869	0.6979
γ = 0.500	bootstrap	0.8592	0.9202	0.8949	0.8937	0.7114
	permutation	0.8285	0.9081	0.9081	0.9010	0.7096
	standard	0.9866	0.9973	0.9395	0.9976	0.7819
γ = 1.000	bootstrap	0.9386	0.9759	0.9595	0.9610	0.7913
	permutation	0.9147	0.9703	0.9703	0.9649	0.7888
	standard	0.9968	0.9998	0.9609	0.9998	0.8265
γ = 2.000	bootstrap	0.9777	0.9951	0.9800	0.9870	0.8351
	permutation	0.9591	0.9924	0.9924	0.9903	0.8333

Table 17: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 35	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	1.0000	1.0000	1.0000	1.0000	1.0000
β = 0.25	bootstrap	1.0000	1.0000	1.0000	1.0000	1.0000
	permutation	1.0000	1.0000	1.0000	1.0000	1.0000
	standard	0.9975	0.9807	0.9932	0.9901	0.9953
β = 0.50	bootstrap	0.9997	0.9932	0.9931	0.9979	0.9948
	permutation	0.9997	0.9929	0.9929	0.9978	0.9953
	standard	0.3222	0.2430	0.5069	0.2165	0.5037
β = 0.75	bootstrap	0.6329	0.5120	0.5051	0.5313	0.5056
	permutation	0.6371	0.5070	0.5070	0.5335	0.5094
	standard	0.0604	0.0558	0.2239	0.0389	0.2123
β = 0.85	bootstrap	0.2677	0.2297	0.2248	0.2169	0.2148
	permutation	0.2691	0.2259	0.2259	0.2220	0.2153
	standard	0.0044	0.0026	0.0495	0.0017	0.0475
β = 1.00	bootstrap	0.0517	0.0500	0.0487	0.0477	0.0523
	permutation	0.0488	0.0499	0.0499	0.0497	0.0520
	standard	0.0513	0.0296	0.1675	0.0168	0.1262
β = 1.15	bootstrap	0.2118	0.1684	0.1639	0.1632	0.1358
	permutation	0.2032	0.1644	0.1644	0.1661	0.1320
	standard	0.1827	0.1091	0.3477	0.0721	0.2369
β = 1.25	bootstrap	0.4379	0.3514	0.3439	0.3417	0.2519
	permutation	0.4234	0.3436	0.3436	0.3468	0.2467
	standard	0.7196	0.5612	0.7914	0.4741	0.5583
β = 1.50	bootstrap	0.9038	0.8027	0.7907	0.8036	0.5759
	permutation	0.8932	0.7931	0.7931	0.8079	0.5722
	standard	0.9734	0.9201	0.9575	0.8852	0.7565
β = 1.75	bootstrap	0.9936	0.9660	0.9595	0.9690	0.7715
	permutation	0.9912	0.9637	0.9637	0.9708	0.7674
	standard	0.9993	0.9937	0.9916	0.9898	0.8643
β = 2.00	bootstrap	0.9996	0.9948	0.9932	0.9950	0.8707
	permutation	0.9992	0.9946	0.9946	0.9956	0.8677
	standard	1.0000	0.9998	0.9968	0.9994	0.9090
β = 2.25	bootstrap	1.0000	0.9993	0.9982	0.9987	0.9160
	permutation	1.0000	0.9993	0.9993	0.9997	0.9116
	standard	1.0000	1.0000	0.9980	1.0000	0.9373
β = 2.50	bootstrap	1.0000	1.0000	0.9994	0.9998	0.9429
	permutation	1.0000	1.0000	1.0000	1.0000	0.9419

Table 18: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 35	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.0139	0.0272	0.1565	0.0162	0.1309
γ = -0.010	bootstrap	0.1122	0.1601	0.1578	0.1477	0.1328
	permutation	0.1128	0.1577	0.1577	0.1490	0.1362
	standard	0.0046	0.0078	0.0709	0.0049	0.0648
γ = -0.005	bootstrap	0.0617	0.0723	0.0712	0.0674	0.0666
	permutation	0.0593	0.0703	0.0703	0.0694	0.0667
	standard	0.0054	0.0042	0.0524	0.0017	0.0487
γ = -0.002	bootstrap	0.0519	0.0550	0.0528	0.0484	0.0518
	permutation	0.0538	0.0535	0.0535	0.0501	0.0518
	standard	0.0051	0.0046	0.0497	0.0031	0.0456
γ = -0.001	bootstrap	0.0490	0.0519	0.0502	0.0457	0.0500
	permutation	0.0474	0.0489	0.0489	0.0468	0.0485
	standard	0.0047	0.0041	0.0513	0.0023	0.0475
γ = 0.000	bootstrap	0.0523	0.0537	0.0519	0.0491	0.0508
	permutation	0.0502	0.0519	0.0519	0.0514	0.0488
	standard	0.0549	0.0492	0.2123	0.0311	0.1662
γ = 0.020	bootstrap	0.2022	0.2133	0.2088	0.2034	0.1804
	permutation	0.1941	0.2105	0.2105	0.2027	0.1763
	standard	0.2876	0.2939	0.5958	0.2219	0.4741
γ = 0.050	bootstrap	0.5321	0.5979	0.5907	0.5776	0.4917
	permutation	0.5184	0.5908	0.5908	0.5788	0.4889
	standard	0.6700	0.7233	0.9010	0.6435	0.7863
γ = 0.100	bootstrap	0.8320	0.9036	0.8980	0.8889	0.7995
	permutation	0.8189	0.8985	0.8985	0.8913	0.7951
	standard	0.9330	0.9681	0.9907	0.9487	0.9439
γ = 0.200	bootstrap	0.9644	0.9910	0.9903	0.9891	0.9484
	permutation	0.9595	0.9906	0.9906	0.9893	0.9473
	standard	0.9983	1.0000	0.9999	0.9995	0.9924
γ = 0.500	bootstrap	0.9986	0.9999	0.9999	0.9999	0.9933
	permutation	0.9976	1.0000	1.0000	0.9999	0.9932
	standard	1.0000	1.0000	1.0000	1.0000	0.9974
γ = 1.000	bootstrap	0.9998	1.0000	1.0000	1.0000	0.9977
	permutation	0.9996	1.0000	1.0000	1.0000	0.9978
	standard	1.0000	1.0000	1.0000	1.0000	0.9982
γ = 2.000	bootstrap	1.0000	1.0000	1.0000	1.0000	0.9981
	permutation	1.0000	1.0000	1.0000	1.0000	0.9979

Table 19: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, n = 100	resampling	milhand	laplace	lew-rob	and-dar	mann
$\beta = 0.8$	standard	0.6213	0.4790	0.4615	0.5282	0.4471
	bootstrap	0.6060	0.4688	0.4660	0.5078	0.4458
	permutation	0.6063	0.4650	0.4650	0.5089	0.4489
$\beta = 0.9$	standard	0.1909	0.1533	0.1448	0.1601	0.1379
	bootstrap	0.1933	0.1523	0.1496	0.1545	0.1400
	permutation	0.1915	0.1479	0.1479	0.1548	0.1394
$\beta = 1.0$	standard	0.0499	0.0505	0.0512	0.0502	0.0481
	bootstrap	0.0506	0.0533	0.0513	0.0502	0.0486
	permutation	0.0482	0.0509	0.0509	0.0495	0.0500
$\beta = 1.2$	standard	0.4237	0.3528	0.3410	0.3481	0.2384
	bootstrap	0.4207	0.3434	0.3366	0.3319	0.2405
	permutation	0.4127	0.3384	0.3384	0.3303	0.2425
$\beta = 1.4$	standard	0.9128	0.8384	0.8127	0.8450	0.5830
	bootstrap	0.9079	0.8156	0.8096	0.8150	0.5889
	permutation	0.9039	0.8082	0.8082	0.8183	0.5893

Table 20: Results from simulations of failure truncated TRP processes with $n = 100$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.0$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, n = 100	resampling	milhand	laplace	lew-rob	and-dar	mann
γ = -0.005	standard	0.3963	0.5131	0.4990	0.5030	0.3845
	bootstrap	0.3737	0.5087	0.5054	0.4845	0.3841
	permutation	0.3719	0.5038	0.5038	0.4860	0.3855
γ = -0.002	standard	0.0892	0.0953	0.0931	0.0947	0.0840
	bootstrap	0.0815	0.0970	0.0954	0.0921	0.0850
	permutation	0.0817	0.0966	0.0966	0.0927	0.0848
$\gamma = 0.000$	standard	0.0495	0.0510	0.0512	0.0522	0.0479
	bootstrap	0.0529	0.0539	0.0525	0.0494	0.0505
	permutation	0.0506	0.0510	0.0510	0.0503	0.0501
$\gamma = 0.003$	standard	0.1014	0.1118	0.1093	0.1123	0.0991
	bootstrap	0.1014	0.1128	0.1116	0.1074	0.1007
	permutation	0.0983	0.1091	0.1091	0.1083	0.1010
$\gamma = 0.020$	standard	0.7843	0.8747	0.8607	0.8662	0.7416
	bootstrap	0.7655	0.8617	0.8588	0.8498	0.7445
	permutation	0.7608	0.8578	0.8578	0.8536	0.7458
$\gamma = 0.200$	standard	1.0000	1.0000	1.0000	1.0000	0.9998
	bootstrap	1.0000	1.0000	1.0000	1.0000	0.9998
	permutation	1.0000	1.0000	1.0000	1.0000	0.9998

Table 21: Results from simulations of failure truncated TRP processes with $n = 100$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.0$ (NHPP). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 100	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.6307	0.4786	0.2786	0.5952	0.3050
$\beta = 0.8$	bootstrap	0.4118	0.2892	0.2843	0.3200	0.3054
	permutation	0.4137	0.2876	0.2876	0.3223	0.3060
	standard	0.3096	0.2403	0.1032	0.3235	0.1075
$\beta = 0.9$	bootstrap	0.1364	0.1100	0.1087	0.1122	0.1080
	permutation	0.1362	0.1084	0.1084	0.1136	0.1091
	standard	0.1494	0.1496	0.0510	0.2115	0.0512
$\beta = 1.0$	bootstrap	0.0520	0.0536	0.0511	0.0500	0.0526
	permutation	0.0518	0.0531	0.0531	0.0509	0.0526
	standard	0.3904	0.3836	0.1992	0.4400	0.1526
$\beta = 1.2$	bootstrap	0.2516	0.2060	0.1997	0.1919	0.1596
	permutation	0.2407	0.1990	0.1990	0.1905	0.1575
	standard	0.8253	0.7729	0.5562	0.8156	0.3868
$\beta = 1.4$	bootstrap	0.6786	0.5671	0.5583	0.5385	0.3935
	permutation	0.6667	0.5587	0.5587	0.5403	0.3891

Table 22: Results from simulations of failure truncated TRP processes with $n = 100$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, n = 100	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.6116	0.6758	0.4710	0.7174	0.3799
γ = -0.005	bootstrap	0.3674	0.4854	0.4761	0.4685	0.3813
	permutation	0.3701	0.4798	0.4798	0.4668	0.3810
	standard	0.2039	0.2040	0.0833	0.2667	0.0713
γ = -0.002	bootstrap	0.0751	0.0913	0.0886	0.0834	0.0719
	permutation	0.0765	0.0887	0.0887	0.0837	0.0731
	standard	0.1408	0.1411	0.0482	0.2036	0.0479
$\gamma = 0.000$	bootstrap	0.0513	0.0523	0.0510	0.0505	0.0483
	permutation	0.0501	0.0502	0.0502	0.0500	0.0500
	standard	0.1901	0.2227	0.0965	0.2907	0.0872
$\gamma = 0.003$	bootstrap	0.0901	0.1008	0.0966	0.0934	0.0907
	permutation	0.0872	0.0968	0.0968	0.0937	0.0904
	standard	0.7622	0.8615	0.7000	0.8834	0.5683
$\gamma = 0.020$	bootstrap	0.6106	0.7074	0.7000	0.6884	0.5725
	permutation	0.5973	0.7004	0.7004	0.6926	0.5706
	standard	0.9998	1.0000	1.0000	1.0000	0.9929
$\gamma = 0.200$	bootstrap	0.9995	1.0000	1.0000	1.0000	0.9930
	permutation	0.9992	1.0000	1.0000	1.0000	0.9934

Table 23: Results from simulations of failure truncated TRP processes with $n = 100$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 0.75$ (overdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 100	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.6131	0.4759	0.7766	0.4484	0.7434
$\beta = 0.8$	bootstrap	0.8865	0.7826	0.7810	0.8070	0.7427
	permutation	0.8841	0.7769	0.7769	0.8075	0.7438
	standard	0.0797	0.0657	0.2611	0.0478	0.2308
$\beta = 0.9$	bootstrap	0.3213	0.2658	0.2640	0.2619	0.2312
	permutation	0.3234	0.2600	0.2600	0.2625	0.2317
	standard	0.0047	0.0034	0.0538	0.0019	0.0508
$\beta = 1.0$	bootstrap	0.0546	0.0534	0.0521	0.0509	0.0538
	permutation	0.0523	0.0540	0.0540	0.0512	0.0540
	standard	0.4475	0.2794	0.6191	0.2160	0.4598
$\beta = 1.2$	bootstrap	0.7367	0.6161	0.6149	0.6358	0.4647
	permutation	0.7339	0.6160	0.6160	0.6396	0.4647
	standard	0.9824	0.9295	0.9868	0.9124	0.8888
$\beta = 1.4$	bootstrap	0.9967	0.9877	0.9873	0.9913	0.8873
	permutation	0.9969	0.9876	0.9876	0.9910	0.8906

Table 24: Results from simulations of failure truncated TRP processes with $n = 100$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, n = 100	resampling	milhand	laplace	lew-rob	and-dar	mann
	standard	0.2007	0.3688	0.6974	0.2859	0.5842
γ = -0.005	bootstrap	0.5415	0.7024	0.6976	0.6822	0.5845
	permutation	0.5443	0.6998	0.6998	0.6841	0.5851
	standard	0.0124	0.0210	0.1316	0.0117	0.1093
γ = -0.002	bootstrap	0.1019	0.1355	0.1333	0.1256	0.1097
	permutation	0.1009	0.1348	0.1348	0.1274	0.1109
	standard	0.0040	0.0033	0.0477	0.0015	0.0489
$\gamma = 0.000$	bootstrap	0.0492	0.0469	0.0466	0.0445	0.0483
	permutation	0.0482	0.0470	0.0470	0.0461	0.0494
	standard	0.0351	0.0340	0.1710	0.0197	0.1430
$\gamma = 0.003$	bootstrap	0.1498	0.1699	0.1683	0.1665	0.1482
	permutation	0.1474	0.1689	0.1689	0.1682	0.1473
	standard	0.8657	0.9235	0.9860	0.8861	0.9483
$\gamma = 0.020$	bootstrap	0.9538	0.9863	0.9858	0.9839	0.9500
	permutation	0.9513	0.9858	0.9858	0.9838	0.9485
	standard	1.0000	1.0000	1.0000	1.0000	1.0000
$\gamma = 0.200$	bootstrap	1.0000	1.0000	1.0000	1.0000	1.0000
	permutation	1.0000	1.0000	1.0000	1.0000	1.0000

Table 25: Results from simulations of failure truncated TRP processes with $n = 100$ failures. The trend function is given by a log-linear law function $\lambda(t) = e^{\gamma t}$ and the interarrival times of the corresponding renewal process are Weibull distributed (27) with $w = 1.50$ (underdispersed). For each value of γ 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6), Lewis-Robinson (9), Anderson-Darling (8) and Mann (10). In addition the bootstrapped and permuted variants of those tests (Section 3.1.1) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00	resampling	milhand	laplace	and-dar
	standard	0.9916	0.9448	0.9839
$\beta = 0.20$	naive	0.5058	0.4553	0.7773
	conditional	0.8536	0.4555	0.8674
	standard	0.6338	0.4589	0.5637
$\beta = 0.50$	naive	0.1093	0.3269	0.3974
	conditional	0.4443	0.3260	0.4007
	standard	0.1822	0.1301	0.1584
$\beta = 0.75$	naive	0.0119	0.1357	0.1175
	conditional	0.1298	0.1301	0.1159
	standard	0.0473	0.0469	0.0514
$\beta = 1.00$	naive	0.0006	0.0684	0.0553
	conditional	0.0515	0.0647	0.0533
	standard	0.0841	0.0779	0.0753
$\beta = 1.25$	naive	0.0010	0.0966	0.0879
	conditional	0.0864	0.0839	0.0808
	standard	0.1910	0.1730	0.1642
$\beta = 1.50$	naive	0.0033	0.1862	0.1679
	conditional	0.1806	0.1640	0.1560
	standard	0.4872	0.4479	0.4194
$\beta = 2.00$	naive	0.0107	0.4295	0.3666
	conditional	0.4177	0.3789	0.3437
	standard	0.8675	0.8361	0.8138
$\beta = 3.00$	naive	0.0442	0.7597	0.6373
	conditional	0.7263	0.6843	0.6207
	standard	0.9915	0.9882	0.9830
$\beta = 5.00$	naive	0.0939	0.9020	0.7711
	conditional	0.8481	0.8359	0.8004

Table 26: Results from simulations of time truncated TRP processes with expected number of failures equalling 10 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.00$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75	resampling	milhand	laplace	and-dar
	standard	0.9822	0.9193	0.9781
$\beta = 0.20$	naive	0.4351	0.3395	0.6754
	conditional	0.7579	0.3360	0.7698
	standard	0.7079	0.5334	0.6079
$\beta = 0.50$	naive	0.0890	0.2454	0.3469
	conditional	0.3465	0.2347	0.3492
	standard	0.3273	0.2382	0.3204
$\beta = 0.75$	naive	0.0108	0.1200	0.1246
	conditional	0.1130	0.1117	0.1205
	standard	0.1372	0.1312	0.1675
$\beta = 1.00$	naive	0.0009	0.0759	0.0590
	conditional	0.0526	0.0659	0.0546
	standard	0.1275	0.1318	0.1549
$\beta = 1.25$	naive	0.0013	0.0851	0.0624
	conditional	0.0724	0.0738	0.0585
	standard	0.1966	0.2024	0.2176
$\beta = 1.50$	naive	0.0031	0.1409	0.1031
	conditional	0.1284	0.1204	0.0971
	standard	0.4293	0.4136	0.4265
$\beta = 2.00$	naive	0.0105	0.3168	0.2206
	conditional	0.3018	0.2793	0.2109
	standard	0.7839	0.7528	0.7573
$\beta = 3.00$	naive	0.0425	0.6178	0.4404
	conditional	0.5936	0.5476	0.4309
	standard	0.9662	0.9558	0.9549
$\beta = 5.00$	naive	0.1162	0.8360	0.6131
	conditional	0.7869	0.7663	0.6455

Table 27: Results from simulations of time truncated TRP processes with expected number of failures equalling 10 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50	resampling	milhand	laplace	and-dar
$\beta = 0.20$	standard	0.9997	0.9768	0.9969
	naive	0.6664	0.6587	0.8830
	conditional	0.9583	0.6382	0.9648
$\beta = 0.50$	standard	0.5597	0.3738	0.4442
	naive	0.1531	0.5108	0.5122
	conditional	0.6659	0.5042	0.5102
$\beta = 0.75$	standard	0.0505	0.0317	0.0305
	naive	0.0115	0.1688	0.0964
	conditional	0.1899	0.1698	0.0981
$\beta = 1.00$	standard	0.0072	0.0061	0.0039
	naive	0.0006	0.0590	0.0461
	conditional	0.0483	0.0564	0.0437
$\beta = 1.25$	standard	0.0475	0.0370	0.0256
	naive	0.0019	0.1370	0.1574
	conditional	0.1355	0.1240	0.1433
$\beta = 1.50$	standard	0.1676	0.1275	0.0877
	naive	0.0037	0.3001	0.3267
	conditional	0.3170	0.2729	0.3093
$\beta = 2.00$	standard	0.5866	0.5047	0.4115
	naive	0.0154	0.6488	0.6474
	conditional	0.6610	0.5981	0.6182
$\beta = 3.00$	standard	0.9680	0.9377	0.8971
	naive	0.0370	0.9147	0.8857
	conditional	0.8700	0.8495	0.8612
$\beta = 5.00$	standard	0.9997	0.9997	0.9988
	naive	0.0679	0.9721	0.9340
	conditional	0.9314	0.9265	0.9449

Table 28: Results from simulations of time truncated TRP processes with expected number of failures equalling 10 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	0.9953	0.9634	0.9915
	naive	0.7952	0.7657	0.9473
	conditional	0.9700	0.7594	0.9590
$\beta = 0.50$	standard	0.7581	0.6103	0.7165
	naive	0.3036	0.4882	0.6014
	conditional	0.6640	0.4777	0.5920
$\beta = 0.75$	standard	0.2330	0.1712	0.1980
	naive	0.0331	0.1665	0.1632
	conditional	0.1837	0.1600	0.1595
$\beta = 1.00$	standard	0.0501	0.0498	0.0499
	naive	0.0023	0.0620	0.0545
	conditional	0.0496	0.0576	0.0513
$\beta = 1.25$	standard	0.1141	0.1044	0.0992
	naive	0.0043	0.1197	0.1067
	conditional	0.1167	0.1085	0.1007
$\beta = 1.50$	standard	0.2920	0.2594	0.2429
	naive	0.0139	0.2665	0.2366
	conditional	0.2849	0.2453	0.2257
$\beta = 2.00$	standard	0.6902	0.6321	0.6035
	naive	0.0624	0.5955	0.5127
	conditional	0.6365	0.5591	0.4906
$\beta = 3.00$	standard	0.9664	0.9538	0.9420
	naive	0.1986	0.9071	0.8156
	conditional	0.9094	0.8698	0.8052

Table 29: Results from simulations of time truncated TRP processes with expected number of failures equalling 15 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.00$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	0.9879	0.9384	0.9823
	naive	0.6419	0.5716	0.8547
	conditional	0.8889	0.5602	0.8799
$\beta = 0.50$	standard	0.7971	0.6212	0.7599
	naive	0.2072	0.3397	0.4830
	conditional	0.5109	0.3239	0.4793
$\beta = 0.75$	standard	0.3773	0.2849	0.3748
	naive	0.0261	0.1334	0.1565
	conditional	0.1432	0.1273	0.1523
$\beta = 1.00$	standard	0.1364	0.1311	0.1708
	naive	0.0025	0.0656	0.0583
	conditional	0.0491	0.0620	0.0531
$\beta = 1.25$	standard	0.1448	0.1562	0.1842
	naive	0.0028	0.0977	0.0740
	conditional	0.0887	0.0877	0.0690
$\beta = 1.50$	standard	0.2676	0.2737	0.2956
	naive	0.0083	0.1853	0.1355
	conditional	0.1819	0.1659	0.1269
$\beta = 2.00$	standard	0.5775	0.5506	0.5669
	naive	0.0389	0.4271	0.3117
	conditional	0.4427	0.3931	0.2961
$\beta = 3.00$	standard	0.8937	0.8686	0.9719
	naive	0.1423	0.7639	0.5857
	conditional	0.7765	0.7191	0.5762

Table 30: Results from simulations of time truncated TRP processes with expected number of failures equalling 15 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	0.9999	0.9940	0.9992
	naive	0.9670	0.9618	0.9984
	conditional	0.9996	0.9585	0.9992
$\beta = 0.50$	standard	0.7837	0.5858	0.6548
	naive	0.5279	0.7397	0.7856
	conditional	0.8948	0.7325	0.7770
$\beta = 0.75$	standard	0.0966	0.0613	0.0570
	naive	0.0606	0.2500	0.1851
	conditional	0.2917	0.2449	0.1822
$\beta = 1.00$	standard	0.0064	0.0047	0.0031
	naive	0.0020	0.0582	0.0511
	conditional	0.0545	0.0564	0.0504
$\beta = 1.25$	standard	0.0176	0.0505	0.0326
	naive	0.0059	0.1751	0.2009
	conditional	0.2020	0.1654	0.1956
$\beta = 1.50$	standard	0.2980	0.2206	0.1545
	naive	0.0284	0.4409	0.4685
	conditional	0.5028	0.4225	0.4518
$\beta = 2.00$	standard	0.8262	0.7384	0.6469
	naive	0.1177	0.8357	0.8262
	conditional	0.8804	0.8100	0.8262
$\beta = 3.00$	standard	0.9989	0.9957	0.9905
	naive	0.2692	0.9919	0.9787
	conditional	0.9835	0.9764	0.9736

Table 31: Results from simulations of time truncated TRP processes with expected number of failures equalling 15 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00	resampling	milhand	laplace	and-dar
$\beta = 0.50$	standard	0.9779	0.9098	0.9545
	naive	0.8497	0.8571	0.9316
	conditional	0.9651	0.8506	0.9290
$\beta = 0.65$	standard	0.7396	0.5782	0.6523
	naive	0.3885	0.5262	0.5999
	conditional	0.6908	0.5192	0.5932
$\beta = 0.90$	standard	0.1107	0.0919	0.0962
	naive	0.0196	0.0912	0.0897
	conditional	0.1024	0.0889	0.0873
$\beta = 1.00$	standard	0.0499	0.0511	0.0530
	naive	0.0039	0.0584	0.0544
	conditional	0.0547	0.0571	0.0523
$\beta = 1.15$	standard	0.1196	0.1059	0.1031
	naive	0.0103	0.1082	0.1005
	conditional	0.1257	0.1062	0.0981
$\beta = 1.25$	standard	0.2316	0.1994	0.1887
	naive	0.0248	0.2002	0.1850
	conditional	0.2364	0.1937	0.1811
$\beta = 1.50$	standard	0.6166	0.5338	0.5201
	naive	0.1412	0.5109	0.4743
	conditional	0.6096	0.5032	0.4671
$\beta = 2.00$	standard	0.9741	0.9493	0.9474
	naive	0.5525	0.9184	0.8873
	conditional	0.9654	0.9116	0.8822

Table 32: Results from simulations of time truncated TRP processes with expected number of failures equalling 35 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.00$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75	resampling	milhand	laplace	and-dar
$\beta = 0.50$	standard	0.9545	0.8568	0.9358
	naive	0.6385	0.6540	0.8030
	conditional	0.8560	0.6390	0.7931
$\beta = 0.65$	standard	0.7555	0.5911	0.7150
	naive	0.2291	0.3501	0.4439
	conditional	0.5037	0.3440	0.4357
$\beta = 0.90$	standard	0.0124	0.1936	0.2624
	naive	0.0124	0.0803	0.0872
	conditional	0.0825	0.0794	0.0841
$\beta = 1.00$	standard	0.1370	0.1408	0.1909
	naive	0.0033	0.0608	0.0549
	conditional	0.0495	0.0579	0.0530
$\beta = 1.15$	standard	0.1480	0.1611	0.2064
	naive	0.0062	0.0791	0.0622
	conditional	0.0825	0.0762	0.0593
$\beta = 1.25$	standard	0.2365	0.2416	0.2820
	naive	0.0147	0.1379	0.1051
	conditional	0.1499	0.1294	0.1025
$\beta = 1.50$	standard	0.5180	0.4880	0.5297
	naive	0.0655	0.3343	0.2661
	conditional	0.3912	0.3229	0.2562
$\beta = 2.00$	standard	0.9081	0.8704	0.8903
	naive	0.2965	0.7436	0.6429
	conditional	0.8248	0.7241	0.6289

Table 33: Results from simulations of time truncated TRP processes with expected number of failures equalling 35 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50	resampling	milhand	laplace	and-dar
$\beta = 0.50$	standard	0.9975	0.9669	0.9845
	naive	0.9937	0.9916	0.9980
	conditional	0.9999	0.9910	0.9977
$\beta = 0.65$	standard	0.7506	0.5770	0.5951
	naive	0.7214	0.8211	0.8459
	conditional	0.9254	0.8173	0.8423
$\beta = 0.90$	standard	0.0180	0.0154	0.0102
	naive	0.0248	0.1272	0.0978
	conditional	0.1349	0.1261	0.0956
$\beta = 1.00$	standard	0.0042	0.0035	0.0018
	naive	0.0037	0.0539	0.0498
	conditional	0.0497	0.0519	0.0499
$\beta = 1.15$	standard	0.0644	0.0425	0.0261
	naive	0.0277	0.1679	0.1879
	conditional	0.2046	0.1643	0.1867
$\beta = 1.25$	standard	0.1983	0.1312	0.0886
	naive	0.0850	0.3510	0.3781
	conditional	0.4377	0.3426	0.3750
$\beta = 1.50$	standard	0.7459	0.5976	0.5116
	naive	0.4203	0.8077	0.8290
	conditional	0.8975	0.8036	0.8273
$\beta = 2.00$	standard	0.9994	0.9953	0.9918
	naive	0.8762	0.9963	0.9952
	conditional	0.9996	0.9969	0.9959

Table 34: Results from simulations of time truncated TRP processes with expected number of failures equalling 35 (fewer than 5 is not allowed). The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the unconditional (naive) and conditional bootstrap variants of those tests (Section 3.1.2) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, m = 5	resampling	mil-ttt	lap-ttt	and-dar	mil-com	lap-com
	standard	0.9562	0.8848	0.9199	0.8850	0.5663
$\beta = 0.50$	bootstrap	0.7466	0.5882	0.7707	0.5502	0.3411
	permutation	0.8585	0.6793	0.8472	0.6101	0.3285
	standard	0.8099	0.6897	0.7565	0.6673	0.3882
$\beta = 0.60$	bootstrap	0.4976	0.3965	0.5281	0.3809	0.2593
	permutation	0.6368	0.4859	0.6191	0.4296	0.2439
	standard	0.2558	0.2151	0.2348	0.1934	0.1325
$\beta = 0.80$	bootstrap	0.1134	0.1145	0.1136	0.1148	0.1075
	permutation	0.1787	0.1515	0.1559	0.1343	0.0988
	standard	0.0508	0.0478	0.0498	0.0496	0.0527
$\beta = 1.00$	bootstrap	0.0343	0.0399	0.0361	0.0453	0.0554
	permutation	0.0505	0.0511	0.0482	0.0505	0.0514
	standard	0.1883	0.1702	0.1634	0.1495	0.1219
$\beta = 1.25$	bootstrap	0.1679	0.1670	0.1823	0.1594	0.1539
	permutation	0.1873	0.1707	0.2037	0.1517	0.1516
	standard	0.3591	0.3240	0.3071	0.2617	0.2140
$\beta = 1.40$	bootstrap	0.3267	0.3153	0.3383	0.2798	0.2667
	permutation	0.3484	0.3192	0.3713	0.2697	0.2597
	standard	0.8467	0.8044	0.7814	0.6915	0.5952
$\beta = 1.80$	bootstrap	0.8119	0.7914	0.8132	0.7110	0.6834
	permutation	0.8255	0.7916	0.8366	0.7002	0.6773
	standard	0.9480	0.9254	0.9144	0.8431	0.7541
$\beta = 2.00$	bootstrap	0.9294	0.9176	0.9252	0.8571	0.8241
	permutation	0.9355	0.9159	0.9384	0.8469	0.8218

Table 35: Results from simulations of failure truncated TRP processes simulated according to the configuration of Figure 13. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.00$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: TTT military handbook TTT Laplace, TTT Anderson-Darling, combined military handbook and combined Laplace. In addition bootstrap and permutation variants of those tests (Section 3.1.3) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, m = 5	resampling	mil-ttt	lap-ttt	and-dar	mil-com	lap-com
	standard	0.9827	0.9194	0.9701	0.9051	0.4887
$\beta = 0.50$	bootstrap	0.5784	0.3508	0.6388	0.3770	0.1964
	permutation	0.7539	0.4686	0.7650	0.4428	0.1837
	standard	0.9187	0.7963	0.8838	0.7602	0.3601
$\beta = 0.60$	bootstrap	0.3278	0.2306	0.4027	0.2339	0.1508
	permutation	0.4984	0.3141	0.5363	0.2846	0.1415
	standard	0.5745	0.4308	0.5326	0.3756	0.1810
$\beta = 0.80$	bootstrap	0.0796	0.0755	0.1179	0.0874	0.0763
	permutation	0.1463	0.1126	0.1796	0.1099	0.0715
	standard	0.1983	0.1587	0.2080	0.1439	0.1240
$\beta = 1.00$	bootstrap	0.0312	0.0350	0.0325	0.0465	0.0530
	permutation	0.0512	0.0529	0.0521	0.0497	0.0508
	standard	0.0988	0.1136	0.1346	0.1253	0.1640
$\beta = 1.25$	bootstrap	0.1046	0.0983	0.0558	0.1067	0.1064
	permutation	0.1330	0.1177	0.0723	0.1031	0.1023
	standard	0.1770	0.1943	0.2177	0.2013	0.2402
$\beta = 1.40$	bootstrap	0.2026	0.1947	0.1232	0.1870	0.1824
	permutation	0.2471	0.2120	0.1491	0.1815	0.1734
	standard	0.5639	0.5601	0.5675	0.5170	0.5193
$\beta = 1.80$	bootstrap	0.6017	0.5589	0.4064	0.4971	0.4709
	permutation	0.6430	0.5751	0.4707	0.4856	0.4622
	standard	0.7469	0.7327	0.7349	0.6703	0.6497
$\beta = 2.00$	bootstrap	0.7731	0.7254	0.5771	0.6531	0.6114
	permutation	0.8040	0.7382	0.6417	0.6437	0.6033

Table 36: Results from simulations of failure truncated TRP processes simulated according to the configuration of Figure 13. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: TTT military handbook TTT Laplace, TTT Anderson-Darling, combined military handbook and combined Laplace. In addition bootstrap and permutation variants of those tests (Section 3.1.3) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, m = 5	resampling	mil-ttt	lap-ttt	and-dar	mil-com	lap-com
	standard	0.9003	0.8469	0.8698	0.8729	0.7363
$\beta = 0.50$	bootstrap	0.9456	0.8857	0.8761	0.8542	0.6958
	permutation	0.9740	0.9180	0.8988	0.8822	0.6823
	standard	0.5389	0.5061	0.5050	0.5264	0.4603
$\beta = 0.60$	bootstrap	0.7838	0.7092	0.5488	0.6741	0.5491
	permutation	0.8559	0.7617	0.5904	0.7117	0.5371
	standard	0.0255	0.0345	0.0261	0.0356	0.0674
$\beta = 0.80$	bootstrap	0.2105	0.2186	0.0432	0.1931	0.1777
	permutation	0.2740	0.2498	0.0526	0.2148	0.1730
	standard	0.0126	0.0073	0.0055	0.0080	0.0177
$\beta = 0.90$	bootstrap	0.0726	0.0779	0.0146	0.0788	0.0820
	permutation	0.1011	0.0935	0.0172	0.0885	0.0778
	standard	0.0443	0.0218	0.0152	0.0119	0.0054
$\beta = 1.00$	bootstrap	0.0408	0.0453	0.0469	0.0478	0.0525
	permutation	0.0498	0.0504	0.0471	0.0510	0.0486
	standard	0.4220	0.2924	0.2379	0.1667	0.0572
$\beta = 1.25$	bootstrap	0.3127	0.2997	0.4118	0.2771	0.2500
	permutation	0.3119	0.2943	0.4109	0.2622	0.2460
	standard	0.7281	0.5879	0.5125	0.3804	0.1727
$\beta = 1.40$	bootstrap	0.6096	0.5972	0.7113	0.5330	0.4938
	permutation	0.6103	0.5884	0.7150	0.5161	0.4874
	standard	0.9927	0.9795	0.9675	0.8958	0.7128
$\beta = 1.80$	bootstrap	0.9820	0.9807	0.9923	0.9465	0.9304
	permutation	0.9818	0.9807	0.9920	0.9415	0.9287
	standard	0.9992	0.9980	0.9964	0.9736	0.8876
$\beta = 2.00$	bootstrap	0.9980	0.9981	0.9993	0.9883	0.8838
	permutation	0.9980	0.9978	0.9996	0.9874	0.9836

Table 37: Results from simulations of failure truncated TRP processes simulated according to the configuration of Figure 13. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: TTT military handbook TTT Laplace, TTT Anderson-Darling, combined military handbook and combined Laplace. In addition bootstrap and permutation variants of those tests (Section 3.1.3) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, m = 5	resampling	mil-ttt	lap-ttt	and-dar	mil-com	lap-com
$\beta = 0.50$	standard	0.8134	0.6717	0.7442	0.3894	0.1793
	conditional	0.7324	0.5494	0.6576	0.1133	0.0707
$\beta = 0.60$	standard	0.5727	0.4329	0.4896	0.2224	0.1286
	conditional	0.5044	0.3657	0.4163	0.0873	0.0717
$\beta = 0.80$	standard	0.1630	0.1218	0.1339	0.0789	0.0687
	conditional	0.1467	0.1214	0.1184	0.0517	0.0623
$\beta = 1.00$	standard	0.0529	0.0482	0.0507	0.0511	0.0506
	conditional	0.0563	0.0586	0.0541	0.0437	0.0492
$\beta = 1.50$	standard	0.2782	0.2313	0.2216	0.0952	0.1058
	conditional	0.2751	0.2330	0.2119	0.0921	0.1074
$\beta = 2.00$	standard	0.7053	0.6037	0.5867	0.1652	0.2427
	conditional	0.6555	0.5385	0.5025	0.1865	0.2343
$\beta = 2.50$	standard	0.9239	0.8520	0.8421	0.2165	0.3936
	conditional	0.8455	0.7239	0.7122	0.2628	0.3123

Table 38: Results from simulations of time truncated TRP processes simulated according to the configuration of Figure 12. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.00$ (NHPP). For each value of β 10,000 simulated processes are tested for trend by the following tests: TTT military handbook TTT Laplace, TTT Anderson-Darling, combined military handbook and combined Laplace. In addition the conditional bootstrap variant of those tests (Section 3.1.3) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, m = 5	resampling	mil-ttt	lap-ttt	and-dar	mil-com	lap-com
$\beta = 0.50$	standard	0.8455	0.7294	0.8134	0.6467	0.3606
	conditional	0.5971	0.4005	0.5700	0.1119	0.0748
$\beta = 0.60$	standard	0.6613	0.5258	0.6162	0.4922	0.2911
	conditional	0.3957	0.2664	0.3689	0.0851	0.0760
$\beta = 0.80$	standard	0.2885	0.2333	0.2868	0.2724	0.1946
	conditional	0.1298	0.1028	0.1330	0.0485	0.0601
$\beta = 1.00$	standard	0.1228	0.1224	0.1526	0.1751	0.1367
	conditional	0.0577	0.0654	0.0610	0.0372	0.0466
$\beta = 1.50$	standard	0.2717	0.2438	0.2703	0.1162	0.1148
	conditional	0.1996	0.1716	0.1344	0.0734	0.0815
$\beta = 2.00$	standard	0.6231	0.5409	0.5633	0.1297	0.1910
	conditional	0.4746	0.3895	0.3295	0.1277	0.1567
$\beta = 2.50$	standard	0.8665	0.7814	0.8034	0.1520	0.3050
	conditional	0.6970	0.5769	0.5291	0.1902	0.2197

Table 39: Results from simulations of time truncated TRP processes simulated according to the configuration of Figure 12. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: TTT military handbook TTT Laplace, TTT Anderson-Darling, combined military handbook and combined Laplace. In addition the conditional bootstrap variant of those tests (Section 3.1.3) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, m = 5	resampling	mil-ttt	lap-ttt	and-dar	mil-com	lap-com
$\beta = 0.50$	standard	0.7937	0.6189	0.6600	0.1153	0.0419
	conditional	0.9329	0.8256	0.7845	0.1357	0.0790
$\beta = 0.60$	standard	0.4503	0.3001	0.3133	0.0321	0.0190
	conditional	0.7389	0.5991	0.4856	0.1045	0.0820
$\beta = 0.80$	standard	0.0493	0.0296	0.0260	0.0112	0.0094
	conditional	0.2178	0.1701	0.0921	0.0588	0.0610
$\beta = 1.00$	standard	0.0083	0.0065	0.0041	0.0360	0.0209
	conditional	0.0529	0.0541	0.0522	0.0448	0.0506
$\beta = 1.50$	standard	0.2818	0.2035	0.1544	0.1639	0.1303
	conditional	0.4915	0.3839	0.4038	0.1413	0.1706
$\beta = 2.00$	standard	0.8336	0.7072	0.6355	0.3016	0.3401
	conditional	0.9026	0.7927	0.8030	0.3088	0.3804
$\beta = 2.50$	standard	0.9829	0.9411	0.9152	0.4080	0.5426
	conditional	0.9697	0.9226	0.9421	0.4256	0.4541

Table 40: Results from simulations of time truncated TRP processes simulated according to the configuration of Figure 12. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: TTT military handbook TTT Laplace, TTT Anderson-Darling, combined military handbook and combined Laplace. In addition the conditional bootstrap variant of those tests (Section 3.1.3) are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, a = 0.00	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	1.0000	0.9998	1.0000
	block	0.9993	0.5832	0.9838
	AR1	0.9998	0.8121	0.0558
	SRP	0.9998	0.6805	0.9487
$\beta = 0.50$	standard	0.9816	0.9128	0.9595
	block	0.9263	0.6053	0.8563
	AR1	0.9623	0.7592	0.2093
	SRP	0.0443	0.6869	0.8228
$\beta = 0.75$	standard	0.4265	0.3123	0.3548
	block	0.3666	0.2342	0.3079
	AR1	0.3868	0.2578	0.0960
	SRP	0.3495	0.2405	0.2589
$\beta = 1.00$	standard	0.0485	0.0498	0.0474
	block	0.0572	0.0514	0.0633
	AR1	0.0469	0.4802	0.0145
	SRP	0.0481	0.0508	0.0430
$\beta = 1.25$	standard	0.2222	0.1891	0.1800
	block	0.1849	0.1766	0.2065
	AR1	0.1473	0.1565	0.0557
	SRP	0.2166	0.1737	0.1501
$\beta = 1.50$	standard	0.6117	0.5292	0.5123
	block	0.4700	0.4445	0.5022
	AR1	0.3495	0.3968	0.1409
	SRP	0.5642	0.4493	0.3968
$\beta = 2.00$	standard	0.9831	0.9565	0.9544
	block	0.8739	0.8595	0.9069
	AR1	0.4593	0.7051	0.2593
	SRP	0.9559	0.8753	0.8200
$\beta = 2.50$	standard	0.9999	0.9986	0.9988
	block	0.9667	0.9620	0.9846
	AR1	0.4082	0.7339	0.2671
	SRP	0.9975	0.9737	0.9426

Table 41: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed (28) with $w = 1.00$ (exponential). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, a = 0.00	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	1.0000	0.9980	0.9998
	block	0.9942	0.4639	0.9241
	AR1	0.9974	0.7874	0.0642
	SRP	0.9980	0.6236	0.9405
$\beta = 0.50$	standard	0.9555	0.8353	0.9294
	block	0.7852	0.4045	0.6562
	AR1	0.8732	0.5465	0.0913
	SRP	0.8168	0.4871	0.6253
$\beta = 0.75$	standard	0.4989	0.3529	0.4634
	block	0.2556	0.1438	0.2104
	AR1	0.2883	0.1664	0.0255
	SRP	0.2354	0.1586	0.1645
$\beta = 1.00$	standard	0.1407	0.1452	0.1968
	block	0.0574	0.0535	0.0663
	AR1	0.0468	0.0520	0.0056
	SRP	0.0526	0.0542	0.0441
$\beta = 1.25$	standard	0.2373	0.2517	0.2928
	block	0.1107	0.1132	0.1342
	AR1	0.0883	0.1084	0.0150
	SRP	0.1346	0.1120	0.0910
$\beta = 1.50$	standard	0.5400	0.5221	0.5639
	block	0.2788	0.2717	0.3065
	AR1	0.2023	0.2477	0.0421
	SRP	0.3543	0.2838	0.2328
$\beta = 2.00$	standard	0.9276	0.8925	0.9104
	block	0.6562	0.6374	0.6827
	AR1	0.3634	0.5331	0.1013
	SRP	0.7969	0.6649	0.5637
$\beta = 2.50$	standard	0.9960	0.9888	0.9923
	block	0.8662	0.8553	0.8937
	AR1	0.3640	0.6536	0.1491
	SRP	0.9627	0.8849	0.7988

Table 42: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, a = 0.00	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	1.0000	1.0000	1.0000
	block	1.0000	0.7214	0.9994
	AR1	1.0000	0.7996	0.0270
	SRP	1.0000	0.5820	0.9012
$\beta = 0.50$	standard	0.9980	0.9790	0.9905
	block	0.9965	0.8572	0.9867
	AR1	0.9979	0.9396	0.3062
	SRP	0.9981	0.9005	0.9698
$\beta = 0.75$	standard	0.3279	0.2466	0.2232
	block	0.5828	0.4174	0.5103
	AR1	0.5902	0.4548	0.3332
	SRP	0.5827	0.4254	0.4255
$\beta = 1.00$	standard	0.0049	0.0032	0.0020
	block	0.0568	0.0528	0.0632
	AR1	0.0461	0.0444	0.0302
	SRP	0.0480	0.0484	0.0429
$\beta = 1.25$	standard	0.1778	0.1038	0.0691
	block	0.3634	0.3282	0.3973
	AR1	0.3540	0.2952	0.2149
	SRP	0.4127	0.3192	0.3033
$\beta = 1.50$	standard	0.7265	0.5558	0.4723
	block	0.7874	0.7500	0.8290
	AR1	0.6181	0.6698	0.4397
	SRP	0.8707	0.7441	0.7243
$\beta = 2.00$	standard	0.9986	0.9938	0.9907
	block	0.9859	0.9839	0.9961
	AR1	0.5946	0.8190	0.4074
	SRP	0.9984	0.9878	0.9774
$\beta = 2.50$	standard	1.0000	1.0000	1.0000
	block	0.9981	0.9981	0.9997
	AR1	0.5623	0.8166	0.4007
	SRP	1.0000	0.9985	0.9945

Table 43: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the interarrival times of the corresponding renewal process are Weibull distributed with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, a = 0.35	resampling	milhand	laplace	and-dar
$\beta = 0.25$	standard	0.9998	0.9930	0.9997
	block	0.9914	0.4545	0.9018
	AR1	0.9992	0.5225	0.0177
	SRP	0.9847	0.3359	0.6056
$\beta = 0.50$	standard	0.9302	0.8085	0.9081
	block	0.7612	0.3881	0.6361
	AR1	0.8470	0.3510	0.0307
	SRP	0.6795	0.3075	0.3937
$\beta = 0.75$	standard	0.4894	0.3803	0.5013
	block	0.2639	0.1529	0.2265
	AR1	0.3059	0.1176	0.0195
	SRP	0.1961	0.1236	0.1197
$\beta = 1.00$	standard	0.1815	0.1949	0.2628
	block	0.0737	0.0645	0.0939
	AR1	0.0504	0.0362	0.0089
	SRP	0.0601	0.0498	0.0405
$\beta = 1.25$	standard	0.2916	0.2929	0.3474
	block	0.1374	0.1318	0.1627
	AR1	0.0316	0.0608	0.0162
	SRP	0.1323	0.1035	0.0767
$\beta = 1.50$	standard	0.5573	0.5192	0.5620
	block	0.3174	0.2862	0.3286
	AR1	0.0845	0.1432	0.0517
	SRP	0.3172	0.2301	0.1742
$\beta = 2.00$	standard	0.9103	0.8579	0.8833
	block	0.6795	0.6178	0.6752
	AR1	0.1956	0.3727	0.1265
	SRP	0.6966	0.5210	0.4055
$\beta = 2.50$	standard	0.9915	0.9765	0.9831
	block	0.8803	0.8325	0.8737
	AR1	0.2907	0.5678	0.2132
	SRP	0.8998	0.7210	0.5660

Table 44: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the transformed interfailure times are simulated using the algorithm of Section 4.4.1 with $a = 0.35$ and a marginal distribution, F, which is Weibull (28) with $w = 1.00$ (exponential). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, a = 0.35	resampling	milhand	laplace	and-dar
	standard	0.9986	0.9748	0.9978
$\beta = 0.25$	block	0.9523	0.3436	0.7727
	AR1	0.9933	0.4548	0.0216
	SRP	0.9380	0.3003	0.5357
	standard	0.8961	0.7471	0.9008
$\beta = 0.50$	block	0.5986	0.2507	0.4613
	AR1	0.5986	0.2507	0.4613
	SRP	0.5204	0.2201	0.2753
	standard	0.5604	0.4323	0.6515
$\beta = 0.75$	block	0.2020	0.1108	0.1738
	AR1	0.3141	0.0945	0.0083
	SRP	0.1490	0.0971	0.0895
	standard	0.3038	0.3094	0.4835
$\beta = 1.00$	block	0.0698	0.0625	0.0878
	AR1	0.0742	0.0390	0.0061
	SRP	0.0606	0.0532	0.0397
	standard	0.3256	0.3657	0.5089
$\beta = 1.25$	block	0.0984	0.0972	0.1161
	AR1	0.0258	0.0587	0.0079
	SRP	0.0970	0.0811	0.0579
	standard	0.5017	0.5157	0.6276
$\beta = 1.50$	block	0.1995	0.1881	0.2222
	AR1	0.0389	0.0951	0.0247
	SRP	0.2125	0.1608	0.1133
	standard	0.8207	0.7926	0.8526
$\beta = 2.00$	block	0.4557	0.4197	0.4624
	AR1	0.0870	0.2291	0.0585
	SRP	0.4925	0.3723	0.2665
	standard	0.9633	0.9368	0.9610
$\beta = 2.50$	block	0.6879	0.6358	0.6703
	AR1	0.1552	0.3828	0.1133
	SRP	0.7301	0.5534	0.4028

Table 45: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the transformed interfailure times are simulated using the algorithm of Section 4.4.1 with $a = 0.35$ and a marginal distribution, F, which is Weibull (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, a = 0.35	resampling	milhand	laplace	and-dar
	standard	1.0000	0.9996	1.0000
$\beta = 0.25$	block	0.9997	0.5894	0.9838
	AR1	1.0000	0.5787	0.0194
	SRP	0.9981	0.3109	0.6159
	standard	0.9766	0.8985	0.9452
$\beta = 0.50$	block	0.9437	0.6340	0.8718
	AR1	0.9501	0.5260	0.0896
	SRP	0.8915	0.4633	0.5853
	standard	0.3959	0.3246	0.3532
$\beta = 0.75$	block	0.4015	0.2551	0.3440
	AR1	0.2252	0.1878	0.0744
	SRP	0.3076	0.1878	0.1844
	standard	0.0571	0.0670	0.0675
$\beta = 1.00$	block	0.0730	0.0680	0.0949
	AR1	0.0406	0.0359	0.0177
	SRP	0.0575	0.0492	0.0392
	standard	0.2528	0.2074	0.1938
$\beta = 1.25$	block	0.2391	0.2101	0.2615
	AR1	0.1043	0.1012	0.0545
	SRP	0.2252	0.1577	0.1286
	standard	0.6464	0.5386	0.5170
$\beta = 1.50$	block	0.5622	0.4905	0.5669
	AR1	0.2720	0.2846	0.1752
	SRP	0.5547	0.3779	0.3173
	standard	0.9809	0.9443	0.9422
$\beta = 2.00$	block	0.9222	0.8743	0.9247
	AR1	0.4807	0.6500	0.2477
	SRP	0.9279	0.7532	0.6370
	standard	0.9998	0.9982	0.9981
$\beta = 2.50$	block	0.9871	0.9730	0.9890
	AR1	0.5893	0.8193	0.4517
	SRP	0.9918	0.9030	0.7776

Table 46: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the transformed interfailure times are simulated using the algorithm of Section 4.4.1 with $a = 0.35$ and a marginal distribution, F, which is Weibull (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.00, a = -0.35	resampling	milhand	laplace	and-dar
	standard	1.0000	1.0000	1.0000
$\beta = 0.25$	block	0.9999	0.6734	0.9974
	AR1	0.9992	0.9761	0.0892
	SRP	1.0000	0.9400	0.9973
	standard	0.9948	0.9707	0.9856
$\beta = 0.50$	block	0.9814	0.7860	0.9692
	AR1	0.9964	0.9802	0.2665
	SRP	0.9961	0.9558	0.9906
	standard	0.3817	0.2598	0.2538
$\beta = 0.75$	block	0.5112	0.3502	0.4338
	AR1	0.6228	0.4866	0.0575
	SRP	0.5768	0.4316	0.4598
	standard	0.0114	0.0110	0.0076
$\beta = 1.00$	block	0.0510	0.0465	0.0571
	AR1	0.0579	0.0669	0.0034
	SRP	0.0477	0.0482	0.0426
	standard	0.1765	0.1239	0.0933
$\beta = 1.25$	block	0.2521	0.2603	0.3050
	AR1	0.3419	0.3416	0.0499
	SRP	0.3397	0.3015	0.2828
	standard	0.6789	0.5575	0.4906
$\beta = 1.50$	block	0.6457	0.6649	0.7325
	AR1	0.8012	0.7627	0.2166
	SRP	0.8243	0.7554	0.7516
	standard	0.9982	0.9894	0.9873
$\beta = 2.00$	block	0.9488	0.9597	0.9863
	AR1	0.7772	0.9393	0.4172
	SRP	0.9989	0.9924	0.9931
	standard	1.0000	1.0000	1.0000
$\beta = 2.50$	block	0.9892	0.9930	0.9993
	AR1	0.5435	0.8585	0.3511
	SRP	1.0000	0.9997	0.9996

Table 47: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the transformed interfailure times are simulated using the algorithm of Section 4.4.1 with $a = -0.35$ and a marginal distribution, F, which is Weibull (28) with $w = 1.00$ (exponential). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 0.75, a = -0.35	resampling	milhand	laplace	and-dar
	standard	0.9999	1.0000	1.0000
$\beta = 0.25$	block	0.9986	0.5476	0.9686
	AR1	0.9952	0.9833	0.9686
	SRP	0.9999	0.9010	0.9990
	standard	0.9763	0.9004	0.9532
$\beta = 0.50$	block	0.9017	0.5591	0.8190
	AR1	0.9633	0.8390	0.0746
	SRP	0.9577	0.7711	0.8988
	standard	0.4754	0.3181	0.3798
$\beta = 0.75$	block	0.3556	0.2045	0.2784
	AR1	0.4415	0.3055	0.0063
	SRP	0.3865	0.2583	0.2813
	standard	0.0623	0.0693	0.0793
$\beta = 1.00$	block	0.0514	0.0471	0.0610
	AR1	0.0594	0.0724	0.0012
	SRP	0.0504	0.0492	0.0429
	standard	0.2056	0.2062	0.2068
$\beta = 1.25$	block	0.1404	0.1532	0.1771
	AR1	0.1965	0.2155	0.0100
	SRP	0.1959	0.1711	0.1454
	standard	0.5789	0.5308	0.5264
$\beta = 1.50$	block	0.3867	0.4024	0.4525
	AR1	0.5233	0.5098	0.0376
	SRP	0.5452	0.4726	0.4309
	standard	0.9738	0.9430	0.9500
$\beta = 2.00$	block	0.7937	0.8008	0.8552
	AR1	0.7957	0.8459	0.1524
	SRP	0.9584	0.8966	0.8757
	standard	0.9999	0.9989	0.9992
$\beta = 2.50$	block	0.9341	0.9409	0.9753
	AR1	0.6162	0.8628	0.1970
	SRP	0.9996	0.9854	0.9782

Table 48: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the transformed interfailure times are simulated using the algorithm of Section 4.4.1 with $a = -0.35$ and a marginal distribution, F, which is Weibull (28) with $w = 0.75$ (overdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.

w = 1.50, a = -0.35	resampling	milhand	laplace	and-dar
	standard	1.0000	1.0000	1.0000
$\beta = 0.25$	block	1.0000	0.8203	1.0000
	AR1	1.0000	0.9484	0.0575
	SRP	1.0000	0.8422	0.9820
	standard	1.0000	0.9989	0.9991
$\beta = 0.50$	block	1.0000	0.9654	0.9998
	AR1	1.0000	0.9999	0.5993
	SRP	1.0000	0.9990	1.0000
	standard	0.2400	0.1567	0.1016
$\beta = 0.75$	block	0.7973	0.6581	0.7454
	AR1	0.8960	0.8291	0.4739
	SRP	0.8856	0.7981	0.8081
	standard	0.0001	0.0000	0.0000
$\beta = 1.00$	block	0.0490	0.0434	0.0490
	AR1	0.0516	0.0577	0.0175
	SRP	0.0468	0.0460	0.0420
	standard	0.1081	0.0326	0.0113
$\beta = 1.25$	block	0.5283	0.5373	0.6201
	AR1	0.6808	0.6380	0.3999
	SRP	0.6898	0.6229	0.6448
	standard	0.8210	0.6013	0.4427
$\beta = 1.50$	block	0.9193	0.9366	0.9776
	AR1	0.9857	0.9748	0.8268
	SRP	0.9927	0.9861	0.9909
	standard	1.0000	1.0000	1.0000
$\beta = 2.00$	block	0.9966	0.9988	1.0000
	AR1	0.7477	0.9511	0.5266
	SRP	1.0000	1.0000	1.0000
	standard	1.0000	1.0000	1.0000
$\beta = 2.50$	block	0.9994	1.0000	1.0000
	AR1	0.5745	0.8506	0.4036
	SRP	1.0000	1.0000	1.0000

Table 49: Results from simulations of failure truncated TRP processes with $n = 35$ failures. The trend function is given by a power law function $\lambda(t) = \beta t^{\beta-1}$ and the transformed interfailure times are simulated using the algorithm of Section 4.4.1 with $a = -0.35$ and a marginal distribution, F, which is Weibull (28) with $w = 1.50$ (underdispersed). For each value of β 10,000 simulated processes are tested for trend by the following tests: military handbook (7), Laplace (6) and Anderson-Darling (8). In addition the bootstrapped variants of those tests of Sections 3.2 and 4.4.2 are also employed. The results are the relative frequencies of rejection of H_0 for a nominal level of 5%.