Tests for Monotonic and Nonmonotonic Trend in Time Censored Recurrent Event Data

Jan Terje Kvaløy*
Department of Mathematics and Natural Sciences, University of Stavanger, Stavanger, Norway - jan.t.kvaloy@uis.no

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

Abstract

A class of statistical tests for trend in the event times in time censored recurrent event data based on the general null hypothesis of a renewal process is proposed. This class does in particular include a test which is attractive for general use by having good power properties against both monotonic and nonmonotonic trends.

Keywords: Trend testing; Renewal process; Trend renewal process; Robust tests.

1. Introduction

In recurrent event data it is often of interest to detect possible systematic changes in the pattern of events. An example is a repairable system for which it is important to detect changes in the pattern of failures. Such changes can for instance be caused by various aging effects or reliability growth. We say that there is a trend in the pattern of failures if the inter-failure times tend to alter in some systematic way, which means that the inter-failure times are not identically distributed. By using statistical trend tests it is possible to decide whether such an alteration is statistically significant or not.

In this paper we focus on the repairable system example, but the methods presented are relevant for any kind of recurrent event data which can be modeled by the same statistical models as used for repairable systems in this paper. A trend in the pattern of failures of a repairable system can be either monotonic, corresponding to an improving or deteriorating system, or non-monotonic, corresponding to for instance cyclic variations or a so called bathtub trend characterizing a system going through the three phases "burn-in frailty", "useful life" and "wear-out".

A common statistical model for systems without trend is the renewal process (RP), which assumes independent identically distributed interarrival times. There exist several tests for trend based on the null hypothesis of an RP, see for instance Lawless, Çiğşar & Cook (2012), but most tests are constructed for failure censored data and there are few tests with power against nonmonotonic trends. In Kvaløy and Lindqvist (2003) we presented a class of tests for trend for failure censored data based on the null hypothesis of an RP. In the present paper we extend that work to time censored data. This leads to a generalization of certain types of trend tests commented by Lawless, Çiğşar & Cook (2012) to be missing, and to be of greater practical usefulness than the corresponding tests for failure censored data.

2. Notation and Models

We assume that the failure time process of a repairable system is observed from time $t = 0$. The successive failure times are denoted $T_1, T_2, \ldots$, and the corresponding interfailure or interarrival times are denoted $X_1, X_2, \ldots$ where $X_i = T_i - T_{i-1},\ i = 1, 2, \ldots$. Further we let $N(t)$ be the number of failures in $(0, t]$ for all $t > 0$. Repair times are assumed to be negligible, or if not, we use operational time as time scale.

We say that a system inhibits no trend if the marginal distributions of all interarrival times are identical, otherwise there is a trend. If the expected length of the interarrival times is monotonically increasing or decreasing with time, corresponding to an improving or a deteriorating system, there is a monotone trend (a decreasing or an increasing trend), otherwise the trend is nonmonotone.

2.1 Renewal Process

The stochastic process $T_1, T_2, \ldots$ is an RP if the interarrival times $X_1, X_2, \ldots$ are independent and identically distributed. The conditional intensity of an RP given the history $\mathcal{F}_{t-}$ up to, but not including time t, can be written

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

where $z(t)$ is the hazard rate of the distribution of the interarrival times, and $T_{N(t-)}$ is the last failure time strictly before time t.

2.2 Trend-Renewal Process

The trend-renewal process (TRP) presented by Lindqvist, Elvebakk and Heggland (2003) is a generalization of the RP to a process with trend. The RP, the homogeneous Poisson process (HPP) and the nonhomogeneous Poisson process (NHPP) are all special cases of this process.

Let $\lambda(t)$ be a non-negative function, called the trend function, defined for $t \geq 0$ and let $\Lambda(t) = \int_0^t \lambda(u)du$. Then the process $T_1, T_2, \ldots$ is a TRP if $\Lambda(T_1), \Lambda(T_2), \ldots$ is an RP.

The conditional intensity of a TRP may be written (Lindqvist, Elvebakk & Heggland, 2003)

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

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

3. The Class of Tests for Trend

We shall now consider testing for trend in time censored processes. The derivation of the class of tests considered is based on an extension of Donsker's theorem to time censored processes, given e.g. in theorem 14.6 in Billingsley (1999).

Let $X_1, X_2, \ldots X_{N(\tau)}$ be the interarrival times of an RP observed from time $t = 0$ and censored at time τ. I.e. the number of events is random, and the time $\tau - T_{N(\tau)}$ is a censored interarrival time. Let μ and σ be the mean and standard deviation of the distribution generating the interarrival times. Then theorem 14.6 in Billingsley (1999) gives that as $\tau \to \infty$

$$V_{\tau,\mu,\sigma}(s) = \mu^{3/2} \frac{N(s\tau) - s\tau/\mu}{\sigma\sqrt{\tau}} \xrightarrow{d} W$$

for $0 \leq s \leq 1$, where W is a Brownian motion and $\xrightarrow{d}$ denotes convergence in distribution. Further defining $W^0(t) = W(t) - sW(1)$, such that W^0 is a Brownian bridge, we will have that as $\tau \to \infty$

$$V^0_{\tau,\mu,\sigma}(s) = V_{\tau,\mu,\sigma}(s) - sV_{\tau,\mu,\sigma}(1) = \mu^{3/2} \frac{N(s\tau) - sN(\tau)}{\sigma\sqrt{\tau}} \xrightarrow{d} W^0 \tag{1}$$

Replacing μ and σ by consistent estimators $\hat{\mu}$ and $\hat{\sigma}$, we conclude that $V^0_{\tau,\hat{\mu},\hat{\sigma}} \xrightarrow{d} W^0$.

Thus, under the null hypothesis of a renewal process, $V^0_{\tau,\hat{\mu},\hat{\sigma}}$ will be approximately a Brownian bridge. On the other hand, if there is a trend in the data, $V^0_{\tau,\hat{\mu},\hat{\sigma}}$ is likely to deviate from the Brownian bridge. Thus tests for trend can be based on measures of deviation from a Brownian bridge of $V^0_{\tau,\hat{\mu},\hat{\sigma}}$. Some suggestions for such measures are the Kolmogorov-Smirnov type statistic $\sup_s V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)$, the Anderson-Darling type statistic $\int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)^2/(s(1-s))ds$, the Cramér-von Mises type statistic $\int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)^2 ds$, or simply $\int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)ds$. The Kolmogorov-Smirnov type statistic is not using all available information and with the Anderson-Darling statistic there are some practical challenges with division by zero at the endpoints so we will here focus on the two last statistics.

3.1 A Cramér-von Mises Type Test for Trend

With the Cramér-von Mises type measure we have that

$$CV = \int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)^2 ds \xrightarrow{d} \int_0^1 W^0(s)^2 ds$$

which has the usual limit distribution of the Cramér-von Mises statistic (Anderson & Darling, 1952). Due to the squaring, a test which rejects the null hypothesis of renewal process for large values of CV should have sensitivity against both monotonic and nonmonotonic trends. Straightforward but tedious calculations give that

$$CV = \frac{\hat{\mu}^3}{\hat{\sigma}^2\tau}\left\{\sum_{i=0}^{N(\tau)-1}\left[i^2\frac{X_{i+1}}{\tau} - iN(\tau)\frac{T_{i+1}^2 - T_i^2}{\tau}\right] + N(\tau)^2\left[\frac{T_{N(\tau)}^2}{\tau^2} - \frac{T_{N(\tau)}}{\tau} + \frac{1}{3}\right]\right\}$$

As estimators of μ and σ we have used the sample mean and sample standard deviation of the completely observed interarrival times (ignoring the last censored interarrival time). These estimators are consistent also in the current setting with a random number of observations, see e.g. Kremers and Robson (1987).

3.2 A Lewis-Robinson Type Test

For the statistic $\int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)ds$ we have that $\int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(s)ds \xrightarrow{d} \int_0^1 W^0(t)dt$ which is normally distributed with expectation 0 and variance 1/12. This statistic will primarily have power against deviations from a Brownian bridge caused by monotonic trends. For non-monotonic trends the positive and negative deviations will tend to cancel. By scaling the test statistic to be asymptotically standard normally distributed, straightforward calculations yield the test statistic

$$LR = \sqrt{12}\int_0^1 V^0_{\tau,\hat{\mu},\hat{\sigma}}(t)dt = \sqrt{\frac{12\hat{\mu}^3}{\hat{\sigma}^2\tau}}\left[\sum_{i=0}^{N(\tau)-1} i\frac{X_{i+1}}{\tau} + N(\tau)\frac{\tau - T_{N(\tau)}}{\tau} - \frac{N(\tau)}{2}\right]$$

This can be considered a Lewis-Robinson type test (Lewis & Robinson, 1974) for time censored data.

4. Simulation Study

We have done various simulations to study and compare the properties of CV and LR for time censored data. In these simulations we estimate rejection probabilities by simulating 100 000 data sets for each choice of model and parameter values, and recording the relative number of rejections of each test. The simulation error is then ≤ 0.0016. All simulations are done in R. The nominal significance level was set to 5%.

4.1 Level Properties

First the level properties of the tests are studied by generating data sets from Weibull RPs with shape parameters respectively 0.75 and 1.5, corresponding respectively to a process which is overdispersed and a process which is underdispersed relative to an HPP. In Figure 1 the simulated level of the tests for systems with the expected number of failures ranging from 10 to 80 are reported.

Both tests have generally good level properties. The CV test is slightly conservative in the overdispersed case, while both tests are slightly non-conservative for small samples in the underdispersed case.

4.2 Power Properties - Monotonic trend

Data sets with a monotonic trend are generated by simulating data from TRPs with the underlying RP distribution being Weibull and the trend function $\lambda(t)$, governing the trend of the process, being on the power law form $\lambda(t) = bt^{b-1}$. The conditional intensity of the process can then be written

$$\gamma(t|\mathcal{F}_{t-}) = \alpha\beta(t^b - T^b_{N(t-)})^{\beta-1}bt^{b-1}. \qquad (2)$$

Rejection probabilities are computed by simulation systems with the censoring time adjusted such that the expected number of failures is 30. Here $b < 1$ corresponds to a decreasing trend, $b = 1$ corresponds to no trend and $b > 1$ corresponds to an increasing trend. This is repeated for two different values of the shape parameter β of the underlying Weibull distribution, $\beta = 0.75$ and $\beta = 1.5$. The results are displayed in Figure 2. We see in this figure that the two tests have fairly similar power properties, with the LR test having slightly higher power.

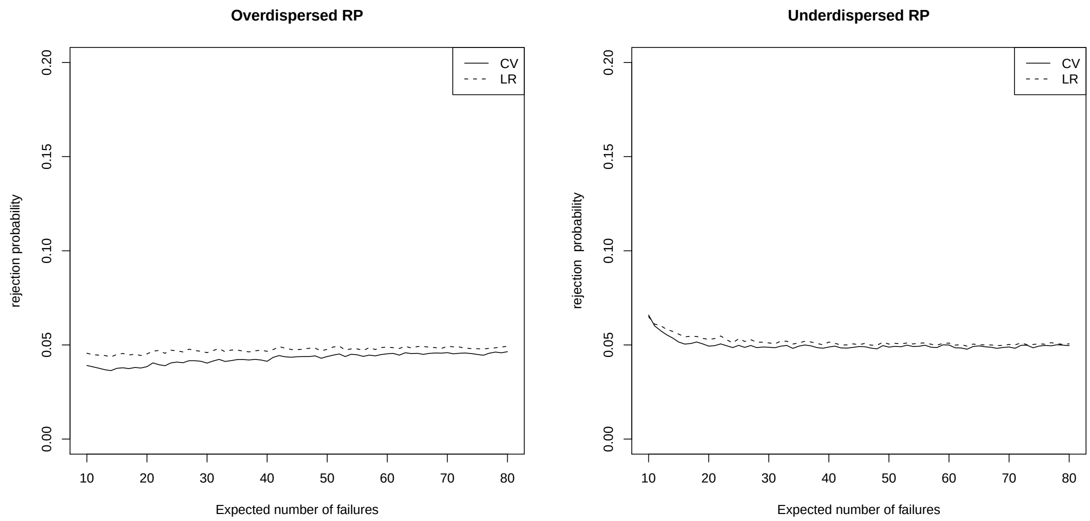

Figure 1: Simulations of Weibull RPs with shape parameters respectively 0.75 (overdispersed RP) and 1.5 (underdispersed RP).

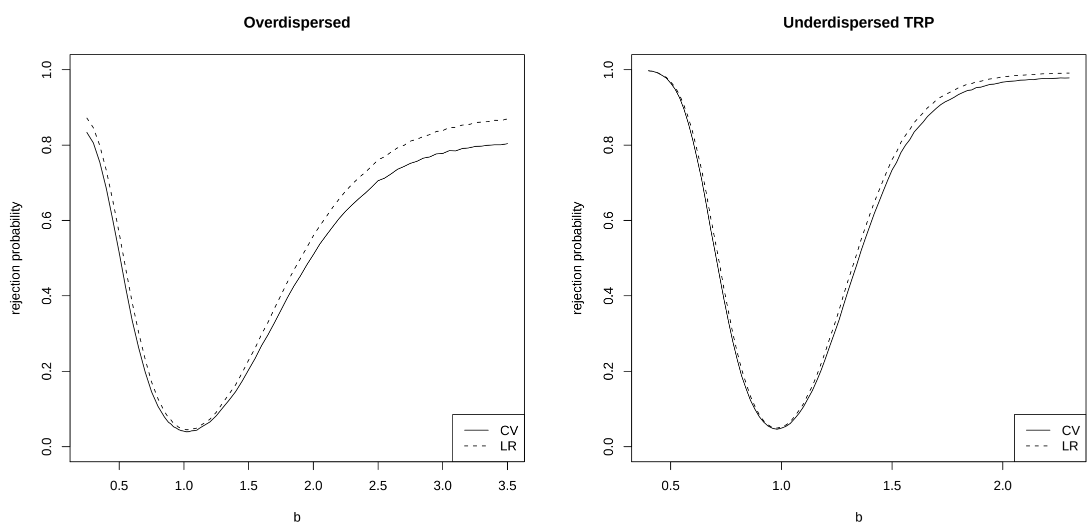

Figure 2: Simulations of TRP with underlying Weibull RPs with shape parameters respectively $\beta = 0.75$ (overdispersed TRP) and $\beta = 1.5$ (underdispersed TRP). The censoring time is adjusted such that the expected number of failures is 30

4.3 Power Properties - Nonmonotonic trend

Data sets with a bathtub trend are generated by simulating data from TRPs with trend function $\lambda(t)$ on the form displayed in Figure 3. Here d is the average of $\lambda(t)$ over $[0, \tau]$. The degree of bathtub shape can be expressed by the parameter c, with $c = 0$ corresponding to a horizontal line (no trend).

The rejection probability as a function of c is simulated with e and τ in each case set to values such that the expected number of failures in each phase (decreasing, no, increasing trend) are equal to 20. The shape

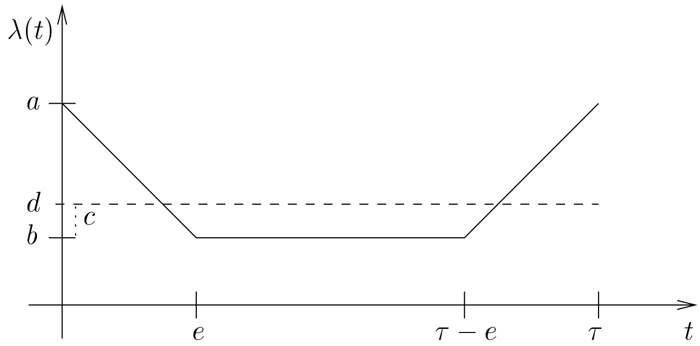

Figure 3: Bathtub-shaped trend function.

parameter of the underlying Weibull distribution is set to respectively $\beta = 0.75$ and $\beta = 1.5$. The results are displayed in Figure 4.

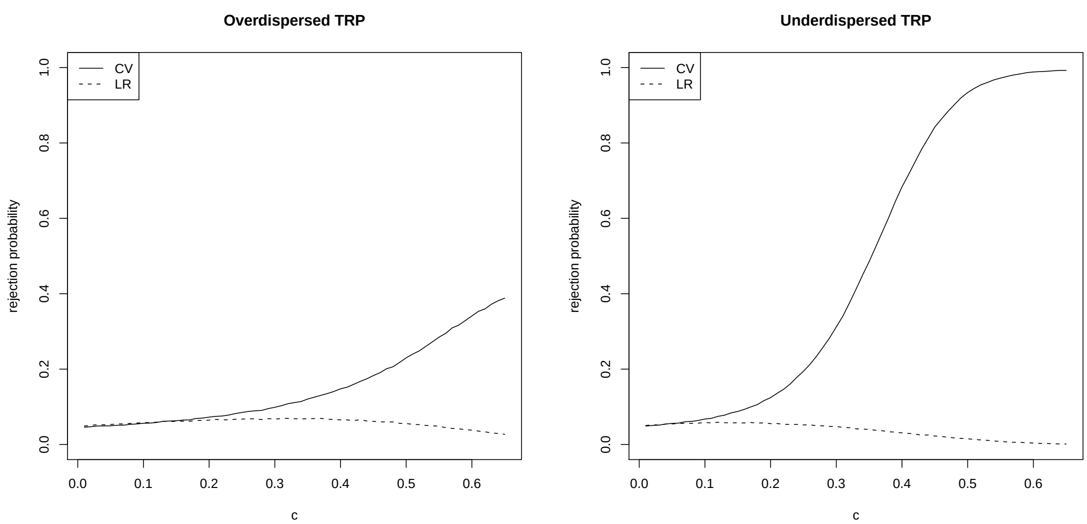

Figure 4: Simulations of TRPs with underlying Weibull RPs with shape parameter $\beta = 0.75$ (overdispersed) and $\beta = 1.5$ (underdispersed TRP), trend function displayed in Figure 3 and expected number of failures in each phase equal to 20.

We see in Figure 4 that the CV test has clearly better power properties than the LR test which is unable to detect this type of trend. Not surprisingly, the trend is easier to detect in the underdispersed case.

5. Conclusion

A class of tests for trend in repairable systems, based on the general null hypothesis of an RP, has been presented. In particular a Cramér-von Mises type test in this class turns out to be an attractive test for general use by having good power properties against both monotonic and nonmonotonic trends.

References

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.

Billingsley, P. (1999). *Convergence of Probability Measures.* Second edition. John Wiley, New York.

Kremers, W.K. & Robson, D.S. (1987). Unbiased estimation when sampling from renewal processes: The single sample and k-sample random means cases, *Biometrika*, 74: 329–336.

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

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

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

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