Inference for Non-Homogeneous Poisson Processes: Models for Repairable Systems

Madhuchhanda Bhattacharjee (nee Mukhopadhyay)
c/o Prof. J.V. Deshpande,
Department of Statistics,
University of Pune,
Pune 411 007,
India.

1 Introduction:

Application of Non-homogeneous Poisson processes to model successive failures of a repairable system undergoing minimal repair is well known. If a system, upon failure, is repaired in such a way that it’s age remains the same as it was just prior to the failure, then we say that the system is minimally repaired. This occurs frequently in a very natural way, particularly when a system is large. In such a case, upon failure of the system, repair/replacement of only a small part may change the state of the system from failed to working without affecting the system materially. In this paper, we assume minimal repair, and that such repairs are instantaneous.

Let \( N(s) \) be the number of failures undergone by a repairable system under consideration, up to time \( s \). We assume \( N(s) \) to be a Non-homogeneous Poisson process (NHPP) with cumulative intensity function \( \Lambda(s) \) (also known as the mean value function), giving the expected number of failures up to time \( s \). It is well known that an NHPP is characterized by \( \Lambda(s) \) or the corresponding intensity function \( \lambda(s) \) (the time derivative of \( \Lambda(s) \)). The intensity function of a process gives the instantaneous rate of change of the expected number of failures with respect to time, and hence may be interpreted as a measure of the wear-out (or improvement) of the system. The shape of \( \lambda(s) \); increasing, decreasing or otherwise; provides information about the changes in the reliability of the system over time. Throughout the paper, we shall assume that \( \Lambda(s) \to \infty \) as \( s \to \infty \), and that it is differentiable, i.e., \( \lambda(s) \) exists. Such NHPPs are frequently used for modeling repairable systems (see e.g., Ascher and Feingold, 1984).

A survey of the literature relating to inference for repairable systems reveals that, the existing work in this area, in general, either presupposes a parametrically specified model, or, in case such a model is not specified, availability of infinite/large number of copies of the system is assumed. Further, in the latter case, the inferential procedures
derived for the system are applicable only over a finite time interval. This, in fact, is a limitation since, a repairable system unfolds itself as time passes, and hence it is important to study (and be able to infer about) its behavior over time. Moreover, in practice, a large number of identical copies of a repairable system may not often be available for analysis, due either to the nature of the system or the repair, or due to size and cost constraints.

A new approach to modeling repairable systems was recently proposed by Deshpande et. al. (1999), which takes care of some of these problems in an elegant and meaningful way. They considered, specifically, a testing problem involving two NHPPs wherein the solutions suggested by them have desirable properties valid for a large time \( t \), and do not require large number of copies, i.e., they are based on either a single realisation or only finitely many copies of the system.

In this work, we present an unified and comprehensive strategy for inference relating to repairable systems, broadly based on the approach given by Deshpande et. al. (1999). In particular, we consider some very general and frequently encountered testing problems relating to repairable systems, and propose non-parametric inferential procedures for the same. Such inference will obviously depend on the method of sampling, and we consider both the sampling scheme under which data for such a system are collected over a fixed time interval, giving rise to ”time truncated” data, and the scheme where a fixed number of failures are observed, which leads to ”failure truncated” data.

Further, the existing test procedures are typically derived after conditioning on the number of failures observed in \((0, t]\). Thus, under such conditioning, what is meant by asymptotics is not quite comprehensible, since for such regular processes, the number of failures in a finite interval cannot be increased at will by the experimenter. In our setup, we do not impose any such conditioning, and since we study the properties of the tests developed by us for large \( t \), we have a setup where asymptotics are more meaningful. We are not aware of any other work where unconditional tests (like the ones attempted here) have been developed for repairable systems. Though we derive our asymptotic inference as \( t \to \infty \), we also note that the above asymptotics actually depend on \( E(N(t)) = \Lambda(t) \) being large, and therefore our asymptotics can also be considered as \( \Lambda(t) \to \infty \) asymptotics.

The paper is divided into six sections. In section two we consider the typical testing problem for a single system, which is essentially a goodness-of-fit test. In section three, we present mathematical preliminaries which would be useful for the development in the rest of the paper. In the fourth and fifth sections, testing problems for a pair of systems are considered with the help of results in the third section. In the final section, we illustrate the inferential procedures developed by us, by applying them on a real life data set.
2 Unconditional tests of goodness-of-fit for a non-homogeneous Poisson process

Given a realization of an NHPP of unknown intensity $\lambda(s)$ observed over the interval $(0, t]$ and for a given/known (baseline) intensity function $\lambda_0(s)$, we first desire to test the hypothesis

$$H_0 : \lambda(s) \propto \lambda_0(s), 0 \leq u \leq t \text{ versus }$$

$$H_1 : \lambda(s)/\lambda_0(s) \text{ increases on } (0, t].$$

Though we would like to test this hypothesis for all $t$, observe that in the non-parametric setup it is actually not possible to extrapolate the inference beyond the interval of observation. Further, note that a suitable time transformation transforms the problem stated above, to the problem of testing that the transformed process is a Homogeneous Poisson process (HPP), with unknown constant intensity versus the intensity being a non-constant increasing function of time (i.e. it is an NHPP), which one realises, is a frequently encountered problem in many practical situations.

Consider time truncated data, truncated at a fixed time point $t$. Let there be $N(t) = n$ failures by time $t$ starting from time $0$ and $T_1, T_2, \ldots, T_n$ be the failure times of the process. Then $0 < T_1 < T_2 < \ldots < T_n < t$.

Earlier Work: The problem of testing constant versus monotonic trend for NHPP’s has been addressed by several authors till now, (see. Ascher and Feingold, 1984; Bain et. al., 1985, Cohen and Sacrowitz, 1993) but in the conditional framework described earlier. The two most prominent such tests are described in the following.

(i) Laplace test statistic : $L = \sum_{i=1}^{n} T_i/t$. The test based on this statistic was studied by Cox (1955) and many others. Under $H_0$ of constant intensity and conditional on $N(t) = n$, $L$ is distributed as the sum of $n$ independent Uniform $(0,1)$ random variables. The test based on $L$ is known to be the uniformly most powerful unbiased (UMPU) test for the Log-Linear alternative with cumulative intensity $\Lambda(s) = e^{\alpha+\beta s}, s > 0, \beta > 0$.

(ii) Power Law test statistic : $Z = 2 \sum_{i=1}^{n} \ln(t/T_i)$. The test based on this statistic was studied by Crow (1974) and others. Under $H_0$ of constant intensity and conditional on $N(t) = n$, $Z$ has a Chi-squared distribution with $2n$ degrees of freedom. The test based on $Z$ is UMPU for the alternative of Power Law process with cumulative intensity function $\Lambda(s) = \lambda s^\beta, s > 0, \beta > 1, \lambda > 0$.

Proposed Statistics: We propose unconditional tests based on the original time truncated data which enables us to have meaningful asymptotics. In analogy with
the conditional problem we propose the use of the following test statistics for this problem.

\[ L^*(t) = \frac{\sqrt{12}}{\sqrt{N(t)\Lambda_0(t)}} \left[ \sum_{i=1}^{N(t)} \Lambda_0(t_i) - \frac{\Lambda_0(t)N(t)}{2} \right] \]

\[ = \frac{\sqrt{12}}{\sqrt{N(t)\Lambda_0(t)}} \left[ \int_0^t \left( \Lambda_0(s) - \frac{\Lambda_0(t)}{2} \right) dN(s) \right], \text{ and} \]

\[ Z^*(t) = \frac{-1}{\sqrt{N(t)}} \left[ \sum_{i=1}^{N(t)} \ln \left( \frac{\Lambda_0(t_i)}{\Lambda_0(t)} \right) + N(t) \right] \]

\[ = \frac{-1}{\sqrt{N(t)}} \left[ \int_0^t \left( \ln \left( \frac{\Lambda_0(s)}{\Lambda_0(t)} \right) + 1 \right) dN(s) \right], \]

where \( t_1, t_2, \ldots, t_{N(t)} \) are the observed failure times of the process.

Asymptotic Distribution under Null: In the following theorem we describe the asymptotic (as \( t \to \infty \)) distributions of the proposed test statistics.

Theorem 2.1 :
(i) Under \( H_0 \), \( L^*(t) \) has asymptotically (as \( t \to \infty \)), standard normal distribution.

(ii) Under \( H_0 \), \( Z^*(t) \) has asymptotically (as \( t \to \infty \)), standard normal distribution.

Proofs of (i) and (ii) above can be obtained by applying a slightly modified version of Theorem B.21 from Karr’s book (1991).

Finite Time Distributions under the Null Hypothesis: It was observed that \( L^*(t) \) approaches normality considerably faster than \( Z^*(t) \) and even for small \( t \) the percentile points of \( L^*(t) \) are quite comparable with those of a standard Normal distribution. Hence the asymptotic distribution may provide a good enough approximation for the finite truncation time distribution for most practical purposes. Therefore we consider only \( Z^*(t) \) here.

Using the fact that for an NHPP with cumulative intensity function \( \Lambda_0(t) \), the conditional distribution of \( 2 \sum_{i=1}^{N(t)} \ln(\Lambda_0(t_i)/\Lambda_0(t_i)) \) given \( N(t) = n \) is Chi-square with \( 2n \) degrees of freedom, the expression for the exact (finite time) distribution of \( Z^*(t) \), is for some constant \( c_0 \),

\[ P \left[ Z^*(t) \leq c_0 \right] = \sum_{n=1}^{\infty} P \left[ \chi^2_{2n} \leq 2(\sqrt{n}c_0 + n) \right] \frac{(\lambda\Lambda_0(t))^n e^{-\lambda\Lambda_0(t)}}{n!(1 - e^{-\lambda\Lambda_0(t)})}, \text{ under } H_0, \]

\( \lambda \) being the constant of proportionality.
From this relation we observe that the finite time distribution depends on the unknown $\lambda$ and time $t$, but only through the product $\lambda\Lambda_0(t)$ (which is also the mean function under $H_0$). The practitioner may well have some idea about the mean number of failures to be observed up to time $t$ even if $\lambda$, the constant of proportionality is unknown. In which case, exact critical points for that value of $\lambda\Lambda_0(t)$ can then be obtained from Table 2.1 of Mukhopadhyay (1999). Otherwise the asymptotic critical point may be used. Actually the critical points monotonically decrease (a fact which is suggested by Theorem 2.2 stated below) to the corresponding normal critical point. First we state two interesting lemmas, which are needed to prove theorem 2.2.

**Lemma 2.1** Truncated Poisson random variables are stochastically ordered.

**Lemma 2.2** Standardised $\chi^2_{2n}$ are lower tail ordered, i.e. $P[\chi^2_{2n} \leq 2(\sqrt{n}c_0 + n)]$ increases with $n$, $\forall c_0 < -1.5$.

**Theorem 2.2** $P[Z^*(t) \leq c_0]$ is increasing in $\Lambda(t)$ for given constant $c_0$, $\exists c_0 < -1.5$.

Proof of the Lemma 2.1 is straightforward. But proof of Lemma 2.2 is quite involved and for that see Mukhopadhyay (1999). The result stated in Theorem 2.2 is a simple application of Lemma 2.1 and 2.2. With the help of these results we conclude the following. If the critical points of the limiting distribution is used then such an approximation will result in an actual level less than the corresponding nominal level $\alpha$, i.e. we will always have a conservative test. Also this actual size increases monotonically to the nominal size as $\Lambda(t)$ becomes large.

Asymptotic behaviour under alternative: Note that the asymptotic variances of both the statistics are finite. Then a sufficient condition for consistency of the proposed tests may be obtained by saying that the asymptotic means diverge to infinity. Therefore, if $\sqrt{\Lambda(t)} \to \infty$ as $t \to \infty$, a sufficient condition for the divergence of $L^*(t)$ or $Z^*(t)$ to $\infty$ is

$$\lim_{t \to \infty} \frac{\sqrt{12}}{\Lambda(t)\Lambda_0(t)} \left[ \int_0^t (\Lambda_0(s) - \frac{\Lambda_0(t)}{2})d\Lambda(s) \right] > 0.$$

Using this condition it can be shown that the tests based on the proposed statistics $Z^*(t)$ and $L^*(t)$ are consistent against several large classes of alternatives.

A well-established method of comparing two tests is to rank them according to their relative abilities to detect an alternative hypothesis. Thus if the power function of one size-\(\alpha\) test lies always above the power function of another size-\(\alpha\) test, then the first one is definitely preferred over the second one. However, it is also known that in the non-parametric setup, the underlying hypothesis not being so well structured, it is seldom the case that one distribution free test procedure is uniformly more powerful than another competitor.
As an alternative measure, the comparison is made through the properties of their asymptotic distribution. The concept of Pitman Asymptotic Relative Efficiency (A.R.E.) is well defined for the i.i.d. sample case. In our situation we have non-i.i.d. observations and also have time truncated data yielding a random number of observations. We extend the concept of A.R.E. to our setup by defining it as the “limiting ratio of the truncation times” required to achieve the same limiting power against the same sequence of alternatives (converging to the null hypothesis) when the limiting significance levels of the two tests are equal.

Theorem 2.3 Let \( \{S_i\} \) and \( \{T_i\} \) be two sequences of tests with associated sequences of numbers \( \{\mu_{S_i}(\theta)\}, \{\mu_{T_i}(\theta)\}, \{\sigma_{S_i}^2(\theta)\} \) and \( \{\sigma_{T_i}^2(\theta)\} \) satisfying following assumptions (1) to (6).

1. \( \frac{S_i - \mu_{S_i}(\theta_i)}{\sigma_{S_i}(\theta_i)} \) and \( \frac{T_i - \mu_{T_i}(\theta_i)}{\sigma_{T_i}(\theta_i)} \) have the same interval of support when \( \theta_i \) is the true value of \( \theta \) and they have the same continuous limiting distribution with c.d.f. \( H(.) \).
2. same as (1) but \( \theta_i \) replaced by \( \theta_0 \) (the null hypothesis value of \( \theta \)) throughout.
3. \( \lim_{i \to \infty} \frac{\sigma_{S_i}(\theta_i)}{\sigma_{S_i}(\theta_0)} = \lim_{i \to \infty} \frac{\sigma_{T_i}(\theta_i)}{\sigma_{T_i}(\theta_0)} = 1, \)
4. \( \frac{d}{d\theta} \left[ \mu_{S_i}(\theta) \right] = \mu_{S_i}'(\theta) \) and \( \frac{d}{d\theta} \left[ \mu_{T_i}(\theta) \right] = \mu_{T_i}'(\theta) \) are assumed to exist and be continuous in some closed interval about \( \theta = \theta_0 \) with \( \mu_{S_i}'(\theta_0) \) and \( \mu_{T_i}'(\theta_0) \) both being nonzero.
5. \( \lim_{i \to \infty} \frac{\mu_{S_i}'(\theta_i)}{\mu_{S_i}(\theta_0)} = \lim_{i \to \infty} \frac{\mu_{T_i}'(\theta_i)}{\mu_{T_i}(\theta_0)} = 1, \)
6. \( \lim_{i \to \infty} \frac{\mu_{S_i}'(\theta_0)}{\sqrt{\sigma_{S_i}^2(\theta_0)}} = K(s) \) and \( \lim_{i \to \infty} \frac{\mu_{T_i}'(\theta_0)}{\sqrt{\sigma_{T_i}^2(\theta_0)}} = K(t) \), where \( K(s) \) and \( K(t) \) are positive constants. Then ARE(S, T) = \( K(s)^2 / K(t)^2 \).

Proof and motivation of this theorem is almost same as that of Noether’s theorem (Randles and Wolfe, 1979, Theorem 5.2.7), hence is omitted.

We illustrate this concept, through computing the A.R.E. of \( L^*(t) \) and \( Z^*(t) \) against the Power-Law alternative. Let \( \Lambda_0(s) = u \) (i.e. \( \lambda_0(s) = 1 \)) and \( \lambda(s) = \beta u^{\beta-1} \). Based on the discussion on consistancy, it can be shown that \( L^*(t) \) and \( Z^*(t) \) are both consistent against the Power-Law alternative. Then we see that conditions (1)-(6) of the above Theorem can easily be verified and the A.R.E. is given by

\[
\frac{K_L^2}{K_Z^2} = 3/4,
\]
which shows that asymptotically $Z^*(t)$ is more efficient in this sense than $L^*(t)$.

Power of the tests: we know that the tests are consistent against a very large set of alternatives implying that the power will be approximately 1 for large number of failures, even though the power actually depends on $\Lambda(t)$ for finite $t$. The unconditional powers of the existing tests were computed and compared with the corresponding powers of the proposed tests. It was found that the unconditional tests (with asymptotics as $t \to \infty$) are quite comparable to the conditional tests and may even be preferred in certain cases.

Analysis of Failure Truncated Data: It is easily understandable now, that similar asymptotic inference procedure will hold, for failure truncated data from a single system also and hence detailed discussion for such data sets is omitted here.

3 Mathematical preliminaries for comparison of two or more processes

From the statistics constructed for testing the hypothesis of interest in section two, it was observed that the test statistics can be expressed in the form of stochastic integrals. Therefore in the third section we will derive certain results regarding some such stochastic integrals, which will help us in obtaining the asymptotic behaviour of the statistics for the tests to be considered in the fourth and fifth sections. Before that some basic results related to NHPPs will be stated.

As mentioned in the first section, data from a repairable system are obtained usually in two forms, viz., time truncated and failure truncated. When two or more processes are to be compared non-parametrically, based on data collected from these processes, then it is evident that the time interval common to all the processes only can be used for the purpose of such a comparison. In that case, if the processes are observed from time “0” till a fixed time point, say, $t$, then data from all the concerned processes become time truncated at $t$ and the overall data may be called “time truncated data”.

But if we try to use the existing concept of failure truncation, i.e. the systems are observed till their respective n-th failure, then we see that to consider the common observation interval, a considerable part of the data might have to be wasted. Since with high probability the processes will stop at different time points. In this paper we have defined the concept of “failure truncated data” for more than one system in the following manner. Suppose that the total number of failures arising out of the processes is fixed before hand, (starting from time “0” say), then the data collection stops at the “n-th” failure of the superimposed process. Then this will provide a common interval of observation for all the processes without any wastage of data.
Next we state two results from Thompson (1981), in the notation used in section two. These are

\[
\frac{N(s) - \Lambda(s)}{\Lambda(s)} \overset{p}{\to} 0, \text{ as } s \to \infty, \text{ and }
\]

\[
\frac{N(s) - \Lambda(s)}{\sqrt{\Lambda(s)}} \overset{d}{\to} N(0,1) \text{ as } s \to \infty.
\]

Regarding the first of the two above results, it can be shown that we actually have a stronger mode of convergence for both type of sampling, viz., time and failure truncation. The results are respectively,

Lemma 3.1 Let \( N(t) \) be a non-homogeneous Poisson process with mean function (i.e. cumulative intensity function) as \( \Lambda(t) \).

\[
\frac{N(s)}{\Lambda(s)} \overset{a.s.}{\to} 1, \text{ as } s \to \infty,
\]

and if \( t_n \) denotes the time till the \( n \)-th failure from the process, then

\[
\frac{N(t_n^-)}{\Lambda(t_n^-)} \overset{a.s.}{\to} 1, \text{ as } n \to \infty.
\]

These results can be obtained using existing similar results for the Homogeneous Poisson processes and the properties of \( \Lambda(s) \), for a detailed proof see Mukhopadhyay (1999). These results help us in avoiding technical complexities in proving the main theorems involving the necessary stochastic integrals.

In the following a theorem regarding the asymptotic behaviour (as \( t \to \infty \)), of some stochastic integrals based on time truncated data from two independent NHPPs is stated. This result will help us in obtaining the asymptotic distributions of the tests statistics proposed by us for time truncated data, In the fourth section.

Theorem 3.1 Let \( N_1 \) and \( N_2 \) be two independent counting processes with continuous compensators \( \Lambda_1, \Lambda_2 \) and innovation martingales \( M_1 \) and \( M_2 \) respectively.

Suppose \( Y_1 \) and \( Y_2 \) are two \( \{\mathcal{F}_s\} \) predictable processes such that

\[
\frac{1}{\Lambda_1(t)} \int_0^t Y_1^2(s) d\Lambda_1(s) \overset{a.s.}{\to} c_1 \text{ and }
\]

\[
\frac{1}{\Lambda_1(t)} \int_0^t Y_2^2(s) d\Lambda_2(s) \overset{a.s.}{\to} c_2, \text{ as } t \to \infty.
\]

Then

\[
\frac{1}{\sqrt{\Lambda_1(t)}} \left[ \int_0^t \alpha_1 Y_1(s) dM_1(s) + \int_0^t \alpha_2 Y_2(s) dM_2(s) \right]
\]
\[ \xrightarrow{D} N(0, \sigma^2), \text{ as } t \to \infty, \]
where \( \sigma^2 = \alpha_1^2 c_1 + \alpha_2^2 c_2, \forall \alpha_1, \alpha_2. \)

Details of the proof may be found in Deshpande et. al. (1999). The corresponding result for failure truncated data is as following

Theorem 3.2 Let \( Y_1(s) \) and \( Y_2(s) \) be two predictable processes. Assume that as \( n \to \infty, \)
\[ Z_1 = \frac{1}{n} \int_0^{t_n} Y_1^2(s) d\Lambda_1(s) \xrightarrow{a.s.} c_1, \]
and
\[ Z_2 = \frac{1}{n} \int_0^{t_n} Y_2^2(s) d\Lambda_2(s) \xrightarrow{a.s.} c_2. \]
Then as \( n \to \infty, \)
\[ \frac{1}{\sqrt{n}} \left[ \int_0^{t_n} \alpha_1 Y_1(s) dM_1(s) - \int_0^{t_n} \alpha_2 Y_2(s) dM_2(s) \right] \xrightarrow{D} N(0, \sigma^2), \]
where \( \sigma^2 = \alpha_1^2 c_1 + \alpha_2^2 c_2, \forall \text{ finite } \alpha_1, \alpha_2. \)

Proof of this theorem can be obtained similarly as the proof of Theorem 3.1 as given in Deshpande et. al. (1999). Proofs are based on the use of stopping times, exponential of semi-martingales and characteristic functions. Proofs can also be obtained from a similar (and more general) theorem given in Hutton and Nelson (1994).

4 Applications to time truncated data

For two independent NHPPs with unknown intensities it is of importance to test for the proportionality of intensities against the alternative of monotonic ratio of intensities, particularly in the situation of choosing intensities. Here we describe the test procedure following Deshpande et. al. (1999), for testing this hypothesis, i.e., on \((0, t] \).

\[ H_0 : \frac{\lambda_2(s)}{\lambda_1(s)} = c, \text{ a constant} \]

against \( H_1 : \frac{\lambda_2(s)}{\lambda_1(s)} \text{ increasing in } s \text{ (but not constant for all } s). \)

Earlier work: Some tests for this problem had been developed by Boyett and Saw (1980) and Lee and Pirie (1981). But their approach was to consider the behavior of the two processes in a conditional setup. Therefore it was not very clear as what is meant by “asymptotic” in such a situation, since they first assume that within the
“fixed” interval of observation, the number of failures is “fixed”, say \( n \), and then let \( n \to \infty \). The problem arises because regular Poisson processes have the property that in a finite interval one observes with probability one only finitely many events. But to apply any kind of asymptotic result one must have large number of observations and such an event will have negligible probability, in a finite time interval.

Proposed Statistics: The tests proposed by Deshpande et. al. (1999) are unconditional and are based on the original time truncated data which enables to have meaningful asymptotics. Suppose that the two processes are observed from time 0 to time \( t \). Let \( X_1 < X_2 < \ldots < X_{N_1(t)} \) and \( Y_1 < Y_2 < \ldots < Y_{N_2(t)} \) be the event times of the two processes. Let \( H(u) \) be defined by

\[
H(u) = \Lambda_2(\Lambda_1^{-1}(u)).
\]

Then the stated testing problem is equivalent to testing \( H(u) = cu \), for some unknown \( c \), i.e. a straight line through the origin with slope \( c \), versus that \( H(u) \) is convex. Consider an estimate of \( H(u) \) as

\[
\hat{H}(u) = N_2(N_1^{-1}(u)),
\]

where \( N_1^{-1}(u) = \inf \{ s : N_1(s) = [u] \} \) and \( [u] \) = largest integer less than or equal to \( u \). Then \( \hat{H}(u) = N_2(X_{[u]}) \), i.e., the number of events observed from the second process upto the time to the \( [u] \)th event from the first process.

The observed deviation of \( \hat{H}(u) \) from the line \( cu \) can give us some idea about the reasonability of the null hypothesis. Consider the estimated area below the curve \( H(u) \) and the estimated area below the line \( cu \) bounded between \( u = 0 \) and \( u = t \). Different statistics can be formed using these and similar quantities, e.g. consider the following,

1. \( S_1(t) = \frac{1}{(N_1(t)N_2(t))^{3/4}} \left[ \frac{N_1(t)N_2(t)}{2} - \int_0^t N_2(s-)dN_1(s) \right] \),

2. \( S_2(t) = \frac{1}{(N_1(t)N_2(t))^{3/4}} \left[ \int_0^t N_1(s-)dN_2(s) - \int_0^t N_2(s-)dN_1(s) \right] \), and

3. \( S_3(t) = \frac{1}{(N_1(t)N_2(t))^{1/4}} \left[ \frac{1}{N_2t} \int_0^t (N_1(s-) + N_2(s-))dN_2(s) - \frac{1}{N_1(t)} \int_0^t (N_1(s-) + N_2(s-))dN_1(s) \right] \).

Using integration by parts and assuming that the two processes do not jump simultaneously, it can be shown that the above three statistics are related.

Asymptotic distribution under Null: Using the results from the third section, the asymptotic null distributions of the proposed tests can be derived as
Theorem 4.1 Under $H_0$ and as $t \to \infty$,

$$S_1(t) = \frac{1}{(N_1(t)N_2(t))^{3/4}} \left[ \frac{N_1(t)N_2(t)}{2} - \int_0^t N_2(s-)dN_1(s) \right] \xrightarrow{D} N(0, \sigma_1^2),$$

where $\sigma_1^2 = \frac{c^{1/2} + c^{-1/2}}{12} = \frac{c + 1}{12\sqrt{c}}$.

In the above Theorem, $\sigma_1^2$, the variance of the limiting distribution is a function of the unknown ‘c’. However it can be consistently estimated as follows. Let

$$w(t) = \left\{ \frac{N_2(t)}{N_1(t)} + 1 \right\} \frac{1}{12 \left( \frac{N_2(t)}{N_1(t)} \right)^{1/2}}.$$

Then using Lemma 3.1, we get that under $H_0$

$$w(t) \xrightarrow{a.s.} \sigma_1^2 \text{ as } t \to \infty.$$

This together with Theorem 3.1 leads to the following Studentization result, and it is suggested that the Studentized statistic $\left( S_1(t)/\sqrt{w(t)} \right)$ be used for carrying out the test.

Theorem 4.2 Under $H_0$ and as $t \to \infty$,

$$\frac{S_1(t)}{\sqrt{w(t)}} \xrightarrow{D} N(0, 1).$$

Similarly the asymptotic null distributions of $S_2(t)$ and $S_3(t)$ can be shown to be

$$S_2(t) \xrightarrow{D} N(0, \sigma_2^2), \text{ where } \sigma_2^2 = \frac{c^{1/2} + c^{-1/2}}{3} = \frac{c + 1}{3\sqrt{c}},$$

$$S_3(t) \xrightarrow{D} N(0, \sigma_3^2), \text{ where } \sigma_3^2 = \frac{(c + 1)^3}{12c^{3/2}}.$$

The corresponding Studentized versions after estimating $c$ as suggested above be used as test statistics.

Asymptotic behaviour under alternative: Similar to the statistics proposed in section two, here also it can be shown that the proposed tests are consistent against several large classes of alternatives. Apart from the known situations where such tests are required, an interesting application of these tests where these tests may prove to be very useful, is described in the following. It was observed by Park and Kim (1992) that though Power-Law process (PLP) and Log-Linear process (LLP) are
two different classes of NHPPs, for many choices of the parameters and up to small values of \( t \), it is difficult to distinguish between the two models. For such values of the parameters even if the observations come from two different processes, initially they appear to satisfy the null hypothesis of proportionality or even equal intensity and hence might lead to a wrong conclusion of accepting the null hypothesis. It may be shown that the proposed tests are consistent for such LLP/PLP alternatives also.

5 Applications to failure truncated data

In continuation to our discussion in the fourth section, in the fifth section also, we will address certain testing problems for two NHPPs. We will develop testing procedure for testing equality of intensities of two NHPPs, i.e.

\[ H_0 : \Lambda_1(s) = \Lambda_2(s), \]

based on “failure truncated data” from single realisations of the two processes, with failure truncation as defined in the third section. We shall investigate the consistency and asymptotic relative efficiency, of the proposed tests, against the following classes of alternatives,

\[ H_{1a} : \frac{\Lambda_1(s)}{\Lambda_2(s)} = c, \text{ an unknown constant } \neq 1, \text{ and } \]

\[ H_{1b} : \frac{\log(\Lambda_1(s))}{\log(\Lambda_2(s))} = c, \text{ an unknown constant } \neq 1. \]

We propose the following two statistics for testing the above null hypothesis of equality. We are not aware of any earlier work for this setup and with this definition of failure truncation.

Proposed Statistics:

(1) Consider the sum of the ranks of the failures of the first kind in the superimposed process and similar sum of the ranks of the failures of the second kind. If the null hypothesis of equality is true then their difference can be expected to be small. Based on this difference we propose the first statistics, \( S_1(t_n) \), which is given below.

\[ S_1(t_n) = \frac{1}{n^{3/2}} \left[ \int_0^{t_n} N(s-)dN_1(s) - \int_0^{t_n} N(s-)dN_2(s) \right]. \]

(2) Consider the sum of number of failures observed from the first process before each failure from the second process and the similar sum of the number of failures observed from the second process, before each failure from the first process. The difference between these two also gives some idea about the reasonability of the assumption of the null hypothesis. This gives rise to the following statistics.

\[ S_2'(t_n) = \frac{1}{n^{3/2}} \left[ \int_0^{t_n} N_1(s-)dN_2(s) - \int_0^{t_n} N_2(s-)dN_1(s) \right]. \]
Though the above statistic seems easy to handle it is not consistent under the alternative \( H_1: \Lambda_2(s) = c\Lambda_1(s) \). Hence we consider the modified statistic

\[
S_2(t_n) = \frac{1}{[N_1(t_n-)N_2(t_n-)]^{1/4}} \left[ \frac{1}{N_2(t_n-)} \int_0^{t_n} N_1(s-)dN_2(s) - \frac{1}{N_1(t_n-)} \int_0^{t_n} N_2(s-)dN_1(s) \right].
\]

Asymptotic distribution under Null: In the following we describe the asymptotic null distributions of the proposed tests. The proof may be obtained using the results stated in the third section.

**Theorem 5.1** Let \( S_1(t_n) \) be as described above. Then under the null hypothesis of equality,

\[ S_1(t_n) \xrightarrow{D} N(0, \sigma_1^2), \quad \text{as } n \to \infty, \quad \text{where } \sigma_1^2 = 2/3. \]  

**Theorem 5.2** Let \( S_2(t_n) \) be as described above. Then under the null hypothesis of equality,

\[ S_2(t_n) \xrightarrow{D} N(0, \sigma_2^2), \quad \text{as } n \to \infty, \quad \text{where } \sigma_2^2 = 2/3. \]  

Asymptotic behaviour under alternative: As has been shown in section two and four, here also we will check for consistency of the proposed tests against standard classes of alternatives. It may be checked that the tests are consistent against the alternative of proportional intensities and also under that of proportionality of the log of the cumulative intensity functions (which is equivalently, \( \Lambda_2(t) = (\Lambda_1(t))^\beta \) for some \( \beta < 1 \)).

Apart from checking for consistency of the proposed tests, the proposed statistics were compared in terms of their “Asymptotic Relative Efficiency” (A.R.E.) for the classes of alternatives mentioned above. For this the proposed concept of A.R.E. was used. It was found that when the two tests are judged from the “A.R.E.”-point of view, they perform equally for both above the classes of alternatives.

6 Analysis Of Boring Machine Failure Data Using Competing Risk Theory

We will illustrate an application of the testing procedures developed in sections four and five to the problem of competing risks, through analysis of a real data set. A system is said to be subject to “competing risks” if it is exposed/vulnerable to more than one cause of failures. In the context of repairable systems we assume that each time the system fails, it fails due to exactly one of these causes or “risks” and then it is subsequently repaired and placed back in service.
The data considered here contains details of two hundred and sixty-two (262) successive failures of a vertical boring machine belonging to a large engineering industry, manufacturing hydraulic excavators.

The failures can be of various types. In fact there are approximately seventy (70) different modes of failures but following Majumder(1993), for the purpose of analysis these 262 failure epochs will be broadly catagorised in to six different categories numbered 1 to 6.

Apparently no information is available on the nature of repairs done to the system upon its failure. But earlier analyses by several people results in the following observations.

a) Simple birth process model or a Homogeneous Poisson process model or Exponential distribution for the inter-failure times does not fit well.

b) Renewal process with non-constant hazard rates also does not fit well. Distributions attempted - Uniform, Weibull, Log-Normal, etc.

c) Correlogram analysis shows systematic dependence of the inter failure times on earlier failures.

d) Some authors conclude that this (i.e. c above) may be due to a systematic change in the parameter value of the distribution of the inter-failure times.

e) Evidence of a positive time-stationary Markovian dependence of the sequence of failures were found.

We propose an NHPP model for the data based on the several observations made by us and some of the earlier analysis carried out by others. By comparing the graph of the observed number of failures and the expected number of failures by time t, it was observed that the Power-Law process model gave a reasonably good fit.

The failure epochs of the NHPP are marked as 1, 2, 3, 4, 5 & 6 according to the type of failure. Of the 262 failure epochs considered here, 45 were caused by risk 1. Of the remaining, the number of failures due to risks 2, 3, 4, 5 and 6 are respectively 47, 28, 32, 46 and 64. Clearly this is a situation where the machine/system is subject to six competing risks, occurrence of any of these causes a failure of the system, upon which the system is repaired (minimally, as per our assumption) and put back in the service. We use this data to carry out pairwise comparisons of pairs of competing failure types.

Note that even though the methods developed in section four are for time truncated data but it can be seen easily that the (asymptotic) distributional properties of the proposed test statistics remain unchanged under failure truncation and hence may be used for this data also. Therefore, the 6 different failure modes were compared one pair at a time for proportional intensity, using the tests developed in the fourth section. Of the 15 pairs 10 rejected the null hypothesis.
The remaining 5 pairs, as indicated in the first column of Table 6.1, accepting the proportionality hypothesis, were further tested for equality using the statistics proposed in section five. The values of the two test statistics and the decisions based on them are given in the subsequent columns.

Table 6.1: Results of pairwise testing for equality for pairs already found to be proportional

<table>
<thead>
<tr>
<th>Failure types</th>
<th>Value of $S_1(t_n)$</th>
<th>Decision</th>
<th>Value of $S_2(t_n)$</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td>0.251329</td>
<td>Accept $H_0$</td>
<td>0.789432</td>
<td>Accept $H_0$</td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>3.938008</td>
<td>Reject $H_0$</td>
<td>-1.342384</td>
<td>Accept $H_0$</td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>3.469719</td>
<td>Reject $H_0$</td>
<td>-2.418497</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>4 &amp; 5</td>
<td>-3.983770</td>
<td>Reject $H_0$</td>
<td>0.158236</td>
<td>Accept $H_0$</td>
</tr>
<tr>
<td>4 &amp; 6</td>
<td>-4.586200</td>
<td>Reject $H_0$</td>
<td>4.534417</td>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>

It is observed that test based on $S_1$ seems to reject the $H_0$ more often than that based on $S_2$ even though the A.R.E. is 1 for proportional alternative.

References


Summary:

Application of Non-homogeneous Poisson processes to model successive failures of a repairable system undergoing minimal repair is well known. Let \( N(s) \) be the number of failures undergone by a repairable system under consideration up to time \( s \). The repairs are either instantaneous or the repair times are ignored for the purpose of analysis in this paper. Then we assume \( N(s) \) to be a Non-homogeneous Poisson process (NHPP) with cumulative intensity function \( \Lambda(s) \) giving the expected number of failures up to time \( s \). It is well known that an NHPP is characterized by \( \Lambda(s) \) or its time derivative \( \lambda(s) \) which is known as the intensity function. The shape of \( \lambda(s) \); increasing, decreasing or otherwise; provides information about the changes in the reliability of the system over time. Throughout the paper, we shall assume that \( \Lambda(s) \to \infty \) as \( s \to \infty \), and that it is differentiable, i.e., \( \lambda(s) \) exists. Such NHPPs are frequently used for modeling repairable systems (see e.g., Ascher and Feingold, 1984).

A survey of the literature relating to inference for such repairable systems reveals that, the existing work in this area either presupposes a parametrically specified model, or, in case such a model is not specified, availability of infinite/large number of copies of the system is assumed. Further, the existing test procedures are typically derived after conditioning on the number of failures observed in \((0, t]\). Thus, under such conditioning, what is meant by asymptotics is not quite comprehensible; since for such regular processes, the number of failures in a finite interval cannot be increased at will by the experimenter.
For testing proportional intensity assumption of two non-homogeneous Poisson processes, a new approach was proposed by Deshpande et. al. (1999), which takes care of some of these problems in an elegant and meaningful way.

In this work, we consider some very general and frequently encountered testing problems relating to repairable systems and present a unified and comprehensive strategy broadly based on this new approach. Such inference will obviously depend on the method of sampling. Testing procedures with desirable properties, for different data types, and for one or more than one systems, have been described in this paper. In our setup, we do not impose any such conditioning, and since we study the properties of the tests developed by us for large $t$, we have a setup where asymptotics are more meaningful.

From the solutions derived herein it can be seen, that this approach is applicable not only to NHPPs, but also towards solving many other inferential problems regarding several other point processes, under different sampling schemes in univariate as well as in multivariate setup.