Exact Statistical Inference for Nonhomogeneous Poisson Processes

Bo Henry Lindqvist
bo@math.ntnu.no

Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7491 Trondheim

1. Introduction and Summary

Nonhomogenous Poisson processes (NHPP) are widely used as models for events occuring in time, for example failures of a repairable system. In practice it may be of interest to check the NHPP property by statistical tests given failure data. The main purpose of this paper is to develop exact tests for null hypotheses that failure times follow NHPPs of particular parametric forms.

One way of testing NHPP models is to embed the models in more general parametric failure models and use likelihood ratio tests based on asymptotic chi-square distributions. A problem is here that the observed number of failures is often too small to justify the use of asymptotic distributions. This motivates the use of exact statistical inference whenever possible (Baker 1996, Gaudoin 1999).

In the present paper we consider exact (Monte Carlo) tests obtained by conditioning on a sufficient statistic under the NHPP model of the null hypothesis. A new method for simulating conditionally on sufficient statistics (Engen and Lillegård 1997, Lindqvist and Taraldsen 2000) will be used. A closely related way of obtaining exact confidence intervals in parametric models (Lillegård and Engen 1999) is also briefly considered.

2. Observations, Likelihood Functions and Sufficient Statistics

We assume that a repairable system is observed from time $t = 0$ and until n failures have occurred. This is known as failure truncation. Most of our results may be modified to the case of time truncation, where the process is observed until a given time t_0.

Suppose we model the failure process by an NHPP with intensity function $\lambda(t)$ and failure times $T = (T_1, \ldots, T_n)$. The log likelihood function resulting from observed failure times $\{t_j\}$ is then given as $\sum_{j=1}^{n} \log \lambda(t_j) - \int_0^{t_n} \lambda(u)du$ (Crowder et al. 1991).

Two popular parametrizations of NHPPs are the power law, with intensity function given by $\lambda_{\text{pow}}(t) = abt^{b-1}$ and the log linear law, with intensity function $\lambda_{\text{log}}(t) = \exp(a + bt)$. Substituting each of these functions into the above log likelihood and using the factorization criterion for sufficiency, gives that the statistic $S = \left(\log T_n, \sum_{j=1}^{n-1} \log T_j\right)$ is sufficient in the power law case, while $S = \left(T_n, \sum_{j=1}^{n-1} T_j\right)$ is sufficient in the log linear case.

3. Conditional Testing Given a Sufficient Statistic

Let T be the vector of failure times as described in the previous section, and consider testing of the null hypothesis H_0 that these data come from an NHPP of a particular parametric type. Let $W \equiv W(T)$ be any test statistic for revealing departure from the null model. Suppose that $S \equiv S(T)$ is a sufficient statistic for the unknown parameters under the null hypothesis. Then if w_{obs} is the observed value of W we can obtain a conditional p-value by computing the conditional probability $P_{H_0}(W \leq w_{\text{obs}} | S = s)$. For this we need to know the conditional distribution of W given $S = s$ under H_0, which by sufficiency is independent of the unknown parameters and thus in principle can be found. This may, however, be difficult in many practical cases and we shall thus rely on simulations. Our tests will thus be Monte Carlo tests.

4. Monte Carlo Conditioning on a Sufficient Statistic

In general, let the model for the observation T (vector) under the null hypothesis be specified in terms of a parameter (vector) θ. Suppose that for a given value of θ we can simulate realizations of T by $T = \chi(U, \theta)$ for some function χ and a random vector U with known distribution. Further, suppose $S(T) \equiv S$ is sufficient for θ. Then S can be simulated by the function $\tau(U, \theta) \equiv S(\chi(U, \theta))$.

Consider now computation of conditional expectations of the form $\mathrm{E}(\phi(T)|S = s)$ for given functions ϕ, where s is the observed value of S. By sufficiency, this conditional expectation is independent of θ

and the clue of the approach is that it can be expressed in terms of ordinary expectations of functions of U (Theorems 1 and 2 below). The basic idea comes from Engen and Lillegård (1997) who, however, erroneously claimed that the conclusion of Theorem 2 will also hold under the more general assumptions of Theorem 1.

Theorem 1 (Lindqvist and Taraldsen, 2000) *Suppose θ and $S(T)$ take values in R^k and suppose the equation $\tau(u,\theta)=s$ has the unique solution $\theta=\hat{\theta}(u,s)$ for each fixed u,s. Let $f(\theta)$ be a nonnegative function defined on the parameter space, and let $\det \partial_\theta \tau(u,\theta)$ be the determinant of the matrix of partial derivatives of $\tau(u,\theta)$ for fixed u. Then*

$$E(\phi(T)|S=s) = \frac{E_U\left[\phi(\chi(U,\hat{\theta}(U,s)))|\frac{f(\theta)}{\det \partial_\theta \tau(U,\theta)}|_{\theta=\hat{\theta}(U,s)}\right]}{E_U\left[|\frac{f(\theta)}{\det \partial_\theta \tau(U,\theta)}|_{\theta=\hat{\theta}(U,s)}\right]}$$

It is tacitly assumed above that f is such that the expectations exist and the denominator is positive, but f may otherwise be arbitrarily chosen. The idea is that the expectations can be computed by simulation by drawing an i.i.d. sample of U and averaging the expressions inside the expectations. It follows that if a function $f(\cdot)$ can be chosen so that $|f(\theta)/\det \partial_\theta \tau(U,\theta)|_{\theta=\hat{\theta}(U,s)}$ does not depend on U, then we will have $\mathrm{E}(\phi(T)|S=s) = \mathrm{E}_U\left[\phi(\chi(U,\hat{\theta}(U,s)))\right]$ which means that the function $\chi(U,\hat{\theta}(U,s))$ can be used to sample directly from the conditional distribution of T given $S=s$. Unfortunately, it is not always possible to find such an f, but the following sufficient condition can be given:

Theorem 2 (Lindqvist and Taraldsen, 2000) *Let the situation be as in the previous theorem. Assume that there exist functions r and $\tilde{\tau}$ with $\tau(u,\theta)=\tilde{\tau}(r(u),\theta)$, such that the equation $\tilde{\tau}(v,\theta)=s$ has a unique solution $v=\tilde{v}(\theta,s)$ for all (θ,s). Then $\chi(U,\hat{\theta}(U,s))$ is distributed as the conditional distribution of T given $S=s$.*

The new assumption of Theorem 2 means that $\tau(u,\theta)$ depends on u only through $r(u)$, which usually has a much lower dimension than u, and has the property that for given θ, $r(u)$ is uniquely determined by s. Note that $\tilde{v}(\theta,S)$ is a pivotal quantity in the classical meaning.

5. Conditional Simulation for Parametric NHPP Models

The first n events of an NHPP with intensity function $\lambda(\cdot)$ can be simulated by letting $U=(U_1,U_2,\ldots,U_n)$ be the first n events of a homogeneous Poisson process with unit intensity, and then letting

$$T_j = \Lambda^{-1}(U_j)\ ;\ j=1,\ldots,n \tag{1}$$

where Λ^{-1} is the inverse function of the cumulative intensity function $\Lambda(t)\equiv\int_0^t \lambda(u)du$.

5.1 Power Law Intensity

From Section 2 follows that $T_j=(U_j/a)^{1/b}$. With notation from the previous section we simulate T by

$$\chi(u;a,b) = ((u_1/a)^{1/b},\ldots,(u_n/a)^{1/b})$$

while the sufficient statistic $S=(\log T_n, \sum_{j=1}^{n-1}\log T_j)$ is simulated by

$$\tau(u;a,b) = ((\log u_n - \log a)/b,\ (\sum_{j=1}^{n-1}\log u_j - (n-1)\log a)/b)$$

Thus the pivotal condition in Theorem 2 holds with $r(u)=(\log u_n, \sum_{j=1}^{n-1}\log u_j)$. Letting the observed times be $t=(t_1,\ldots,t_n)$ and letting $s=(t_n,\sum_{j=1}^{n-1}t_j)$, it is straightforward to obtain $\hat{\theta}(u,s)\equiv(\hat{a}(u,s),\hat{b}(u,s))$ by solving $\tau(u;a,b)=s$ for a and b.

Samples $\tilde{t}=(\tilde{t}_1,\ldots,\tilde{t}_n)$ from the conditional distribution of T given $S=s$ can now be obtained by first sampling $u=(u_1,\ldots,u_n)$ and then computing $\tilde{t}=\chi(u;\hat{\theta}(u,s))$. We get

$$\tilde{t}_j = (u_j/u_n)^{1/\hat{b}(u,s)}t_n\ ;\ j=1,\ldots,n$$

Note that $\tilde{t}$ can be easily simulated by noting that the u_j/u_n are distributed as the order statistic of $n-1$ independent uniforms on $[0,1]$.

5.2 Log Linear Intensity

Now (1) becomes $T_j = \log(1+be^{-a}U_j)/b$, so T is simulated by

$$\chi(u;a,b) = (\log(1+be^{-a}u_1)/b, \ldots, \log(1+be^{-a}u_n)/b)$$

and $S = (T_n, \sum_{j=1}^{n-1} T_j)$ by

$$\tau(u;a,b) = (\log(1+be^{-a}u_n)/b, \sum_{j=1}^{n-1}\log(1+be^{-a}u_j)/b)$$

It is clear that the pivotal condition of Theorem 2 is not satisfied here. This should be no surprise since it is well known that the log linear NHPP model has no interesting pivotal statistic. In order to compute p-values of conditional tests, we thus have to use Theorem 1 with some arbitrarily chosen function $f(a,b)$.

Now we can first find $\hat{\theta}(u,s) \equiv (\hat{a}(u,s), \hat{b}(u,s))$ by solving for a, b the equations

$$t_n = \frac{1}{b}\log(1+be^{-a}u_n) \tag{2}$$

$$\sum_{j=1}^{n-1} t_j = \frac{1}{b}\sum_{j=1}^{n-1}\log(1+be^{-a}u_j) \tag{3}$$

Solving (2) for a and substituting into (3) gives

$$b\sum_{j=1}^{n-1} t_j = \sum_{j=1}^{n-1}\log\left[1+\frac{u_j}{u_n}(e^{t_n b}-1)\right] \tag{4}$$

which is an equation in b only. By differentiating twice it is seen that the right hand side of (4) is convex in b. Further consideration leads to the conclusion that the equation (4) in addition to the trivial solution $b=0$ has a unique additional solution, which is the one that solves our problem, and which is easily found by numerical methods. The solution for a is then finally found from (2).

Now $\tilde{t} = \chi(u;\hat{a}(u,s),\hat{b}(u,s))$ is given by $\tilde{t}_j = \log\left(1+(e^{t_n\hat{b}(u,s)}-1)(u_j/u_n)\right)/\hat{b}(u,s)$ and a straightforward computation shows that the determinant $\det\partial_{a,b}\tau(u;a,b)$ with $(\hat{a}(u,s),\hat{b}(u,s))$ substituted for (a,b) is given by

$$\left(t_n\sum_{j=1}^{n-1} h_j - h_n\sum_{j=1}^{n-1} t_j\right)/\hat{b}(u,s)^2$$

where $h_j = ((e^{t_n\hat{b}}-1)(u_j/u_n))/(1+(e^{t_n\hat{b}}-1)(u_j/u_n))$ and $\hat{b} \equiv \hat{b}(u,s)$. As for the power law case we simulate the u_j/u_n for $j=1,\ldots,n-1$ as the order statistic from a set of $n-1$ i.i.d. uniforms on $[0,1]$.

6. Statistical Inference in NHPP Models

6.1 Goodness-of-Fit Testing

The identity (1) is equivalent to $U_j = \Lambda(T_j)$. It follows that if $\Lambda(\cdot)$ is the cumulative intensity function of the NHPP $T_1, T_2, \ldots$, then $\Lambda(T_1), \Lambda(T_2), \ldots$ is a homogeneous Poisson process with unit intensity. Thus the transformed times $V_j = \Lambda(T_j)/\Lambda(T_n)$ for $j=1,\ldots,n-1$ are distributed as the order statistic of $n-1$ i.i.d. uniform variables on $[0,1]$. If $\Lambda^*(\cdot)$ is an estimate of $\Lambda(\cdot)$ based on data $t=(t_1,\ldots,t_n)$, then we shall define estimated transformed times $v_1^*,\ldots,v_{n-1}^*$ by $v_j^* = \Lambda^*(t_j)/\Lambda^*(t_n)$. One then anticipates these to behave much similar to uniform variables, and goodness-of-fit testing may thus be based on comparing the behaviour of the estimated transformed times to that of uniform variates.

Baker (1996) shows for the power law process that when Λ^* is based on the maximum likelihood estimates for the parameters, then the estimated transformed times are pivots, i.e. have distributions which do not depend on the unknown parameters. This follows in fact from the representation $t_j =$

$(u_j/a)^{1/b}$ in Section 5.1, noting that $b^* = -n/\sum_{j=1}^{n-1} \log(t_j/t_n)$ is the maximum likelihood estimate of b based on the data (Crowder et al. 1991). Then we have $v_j^* = (t_j/t_n)^{b^*} = (u_j/u_n)^{-n/\sum_{j=1}^{n-1}\log(u_j/u_n)}$ which is independent of the parameters.

Baker (1996) derives a class of score tests based on the estimated transformed times. A special case, which we shall use for illustration, is $W = \sum_{j=1}^{n}(v_j^* - v_{j-1}^*)^2$ where $v_0^* = 0, v_n^* = 1$. The null hypothesis of NHPP is rejected for either too small or too large values of this statistic. Note that since the v_j^* are pivots, we can in the power law case compute (by simulation) the unconditional probabilities $P_{H_0}(W \leq w_{\text{obs}})$.

In the case of log linear intensity (Section 5.2) we get $v_j^* = (e^{b^* t_j} - 1)/(e^{b^* t_n} - 1)$ where the maximum likelihood estimate b^* is given as the solution to an equation given in Crowder et al. (1991). Conditional p-values for the test based on W can then be computed by using Theorem 1 with $\phi(T) = I(W \leq w_{\text{obs}})$.

6.2 Exact Confidence Intervals

Lillegård and Engen (1999) show how the $\hat{b}(u, s)$ of Section 5 can be used to obtain exact confidence intervals for the parameter b, both for the power law and the log linear law cases: Draw a (large) number of independent realizations $u_1, \ldots, u_m$ of U. Let s be the observed value of the sufficient statistic and let $\tilde{b}_{(1)} < \cdots < \tilde{b}_{(m)}$ be the ordered values of the $\hat{b}(u_j, s)$. Then $(\tilde{b}_{(k)}, \tilde{b}_{(m-k+1)})$ is an exact $1 - 2k/(m+1)$ confidence interval for b.

For the power law case it can be seen that the above interval (for $m \to \infty$) is the same as the classical one based on the pivotal statistic $2nb/b^*$, which is known to be chi-square distributed with $2(n-1)$ degrees of freedom. The interval obtained for the log linear case, however, appears to be new.

6.3 Application to Data Set

We apply the above results to data from a reliability growth program, taken from Leitch (1995, p. 98). There are $n = 10$ failures, at times 103, 315, 801, 1183, 1345, 2957, 3909, 5702, 7261, 8245. Suppose first that it is of interest to know if the data are consistent with a power law or a log linear NHPP.

In the power law case we get $w_{\text{obs}} = 0.1263$. Simulating the distribution of W we find $P_{H_0}(W \leq 0.1263) \approx 0.024$, which implies some evidence against the power law NHPP. In the log linear case we get $w_{\text{obs}} = 0.1466$, $s = (8245, 23576)$ and $P_{H_0}(W \leq w_{\text{obs}}|S = s) \approx 0.217$, which would not lead to rejection of the log linear NHPP assumption.

An exact 90% confidence interval for b in the power law model, using the chi-square distribtuion as explained above, follows by first computing the maximum likelihood estimate $b^* = 0.6249$. The resulting interval is $(0.2933, 0.9020)$. Since the interval does not contain 1, we can conclude that we have reliability growth. For the log linear case the maximum likelihood estimate is $b^* = -1.75 \cdot 10^{-4}$ and an exact 90% confidence interval is $(-5.60 \cdot 10^{-4}, -3.83 \cdot 10^{-5})$, again indicating reliability growth.

References

1. Baker, R. D. (1996). Some new tests of the power law process. *Technometrics*, **38**, 256-265.

2. Crowder, M. J., Kimber, A. C., Smith, R. L. and Sweeting, T. J. (1991). *Statistical Analysis of Reliability Data.* Chapman & Hall, Great Britain.

3. Engen, S. and Lillegård, M. (1997). Stochastic simulations conditioned on sufficient statistics. *Biometrika*, **84**, 235-240.

4. Gaudoin, O. (1999). CPIT goodness-of-fit tests for reliability growth models. In: Ionescu, V. and Limnios, N. (editors): *Statistical and Probabilistic Methods in Reliability*, Birkhauser Boston, pp. 27-37.

5. Leitch, R. D. (1995). *Reliability Analysis for Engineers.* Oxford University Press, New York.

6. Lillegård, M. and Engen, S. (1999). Exact confidence intervals generated by conditional parametric bootstrapping. *Journal of Applied Statistics*, **26**, 447-459.

7. Lindqvist, B. H. and Taraldsen, G. (2000). Monte Carlo conditioning on a sufficient statistic. (In preparation.)