NTNU
Innovation and Creativity

Exact Statistical Inference in Nonhomogeneous Poisson Processes, based on Simulation

Bjarte Rannestad

Master of Science in Physics and Mathematics
Submission date: July 2007
Supervisor: Bo Henry Lindqvist, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

-Give an introduction to the term of sufficiency
-Demonstrate how sufficiency is applied in construction of exact goodness-of-fit tests, and in particular by implementation for parametric nonhomogeneous Poisson processes
-Study the two parametrizations the nonhomogeneous Poisson process, with power law and log linear intensity functions, both by simulation and analysis of specific datasets

Assignment given: 01. February 2007
Supervisor: Bo Henry Lindqvist, MATH

Abstract

We present a general approach for Monte Carlo computation of conditional expectations of the form $E[\phi(T)|S = s]$ given a sufficient statistic S.

The idea of the method was first introduced by Lillegård and Engen [4], and has been further developed by Lindqvist and Taraldsen [7, 8, 9].

If a certain pivotal structure is satisfied in our model, the simulation could be done by direct sampling from the conditional distribution, by a simple parameter adjustment of the original statistical model. In general it is shown by Lindqvist and Taraldsen [7, 8] that a weighted sampling scheme needs to be used.

The method is in particular applied to the non-homogeneous Poisson process, in order to develop exact goodness-of-fit tests for the null hypothesis that a set of observed failure times follow the NHPP of a specific parametric form. In addition exact confidence intervals for unknown parameters in the NHPP model are considered [6].

Different test statistics $W \equiv W(T)$ designed in order to reveal departure from the null model are presented [1, 10, 11]. By the method given in the following, the conditional expectation of these test statistics could be simulated in the absence of the pivotal structure mentioned above. This extends results given in [10, 11], and answers a question stated in [1].

We present a power comparison of 5 of the test statistics considered under the nullhypothesis that a set of observed failure times are from a NHPP with log linear intensity, under the alternative hypothesis of power law intensity.

Finally a convergence comparison of the method presented here and an alternative approach of Gibbs sampling is given.

Preface

The Master Thesis presented is performed during the spring 2007, and completes the 5 year programme "Sivilingeniør, Fysikk og Matematikk", at the Norwegian University of Science and Technology (NTNU). The thesis was written at the Department of Mathematical Sciences, with Professor Bo Henry Lindqvist as the professional supervisor.

I would especially like to thank Professor Bo Lindqvist for all the help and guidance throughout the semester, which has been highly appreciated.

I would also like to thank my fellow students for 5 good years together in Trondheim.

Contents

Chapter 1

Introduction

Assume a repairable system is observed in a time interval [0,t]. The system may fail many times during this time interval resulting in the failure times $T = (T_1, ..., T_n)$.

Understanding and ability to model the systems failure behaviour might be of great interest from an economical, manufacturing, planning and/or other viewpoints.

Assume that the observed failure times T comes from a specific model with unknown parameters θ. A test statistic $W \equiv W(T)$ is designed in order to reveal departure from the model under the assumption. Now $S \equiv S(T)$ is the sufficient statistics for the unknown parameters of our model. Then if w_{obs} is the value of W based on the observed failure times T, the aim is to determine the conditional probability:

$$P_{H_0}(W(T) \geq w_{obs}|S = s) \qquad \text{(for all s)} \tag{1.1}$$

in order to make statistical inference concerning our assumed model. In the following we will see that equation (1.1), by the proper definition of a function $\phi(T)$, could be expressed as:

$$P_{H_0}(W(T) \geq w_{obs}|S = s) = E[\phi(T)|S = s]$$

This conditional extectation is by sufficiency independent of the unknown parameters of the assumed model, and in principle it could be found. However, in practical cases this might be very difficult or even impossible, and hence one has to rely on simulations.

We present a method of how to simulate such conditional expectations conditioned on the value of the sufficient statistic S, and in particular we apply it to the nonhomogeneous Poisson process.

The method could be applied to make exact statistical inference, including goodness-of-fit testing and parameter estimation, concerning the model under the assumption. This could provide better results than existing asymptotic methods, especially when the observed number of failures is small.

Chapter 2 : In this chapter we give a short introduction to the term sufficient statistics. In particular the Factorization Theorem is stated, followed by an example of how to apply it.

Chapter 3 : The definition of the non-homogeneous Poisson process (NHPP) is stated together with the two popular parametrizations power law and log linear law, of this model. The Factorization Theorem is then applied to identify the sufficient statistics for these two parametrizations.

Chapter 4 : In chapter 4 we outline a framework for an α-level conditional test of the nullhypothesis that a set of observed failure times $T = (T_1, ..., T_n)$ comes from a NHPP model of a specific parametric type.

Chapter 5 : A general method of Monte Carlo simulation of conditional expectations based on sufficient statistics is presented. The method was first introduced by Lillegård and Engen [4], and was further developed by Lindqvist and Taraldsen [7, 8, 9]. In the following this method is referred to as **LT (2006)**.

Chapter 6 : We adjust the LT (2006) method to the two parametrizations of the NHPP model considered in chapter 3. An alternative procedure for simulating conditional expectations based on sufficient statistics is considered for the log linear parametrization. This is done by Gibbs sampling, and 3 alternative Gibbs algorithms are presented.

Chapter 7 : In this chapter different test statistics $W \equiv W(T)$ are considered. The statistics are from [1, 10, 11]. We extend the test statistics to a more general situation in the absence of a certain pivotal structure in our assumed model, which answers a question given in [1]. The test statistics given in [10, 11] are also extended to the new situation of time censoring.

Chapter 8 : The LT (2006) method and Gibbs sampling is applied in practice, to make exact statistical inference in NHPP models, with power law and log linear intensity function. In [10, 11] a power comparison of the test statistics are given under the nullhypothesis of power law NHPP with alternative hypothesis that data comes from a NHPP with log linear intensity. In this chapter we present a power comparison under the nullhypothesis of log linear intensity NHPP with alternative hypothesis that failure data comes from power law NHPP, which is a new result. Finally a convergence comparison of the Gibbs sampling and the LT (2006) method is presented.

Chapter 2

Sufficient Statistics

In this chapter we introduce the term sufficient statistic. The material is from [3]

Suppose a random sample $X = (X_1, X_2, ..., X_n)$ is drawn from a population. The population contains one or more unknown parameters θ=$(\theta_1, \theta_2, ..., \theta_k)$ which could be of interest to estimate for various reasons. A statistic is a function $T = r(X_1, X_2, ..., X_n)$ of the random sample drawn from the population. Examples of such functions are:

$$T_1 = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{(sample mean)}$$
$$T_2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \qquad \text{(sample variance)}$$
$$T_3 = \max\{X_1, X_2, ..., X_n\}$$

A statistic T is said to be an estimator of the population parameter θ if T is usually close to θ. If we look at the statistics given above the sample mean and the sample variance are estimators of the population mean and the population variance, respectively.

Obviously there are lots of functions and so lots of statistics of the random sample that could be computed, in order to estimate unknown parameters of a population. Now the challenge is to find a small set of statistics, if they exist, which by themselves exctracts all the information the sample contains about the population. Suppose one calculates the mean and variance of a sample. Does the sample contain any more information about the population than this?

It is important to remember that the population is always assumed to be described by a certain distribution. This could be the normal distribution, the exponential distribution, the gamma distribution or one of the other known distributions with one or more unknown parameters. Hence, the answer to the above question depends on what family of distributions we assume describes the population.

We give two definitions of sufficiency [3]:

Definition 1. (Heuristic definition) *"We say $T_1, T_2, ..., T_k$ are jointly sufficient statistics if the statistician who knows the values of $T_1, T_2, ..., T_k$ can do just as good a job of estimating the unknown parameters θ as a statistician who knows the entire random sample. In this setting θ typically represents several parameters and the number of statistics, k, is equal to the number of unknown parameters".*

Definition 2. (Mathematical definition) *"The statistics $T_1, T_2, ..., T_k$ are jointly sufficient if for each $t_1, t_2, ..., t_k$, the conditional distribution of $X_1, X_2, ..., X_n$ given $T_i = t_i$ for $i = 1, ..., k$ and θ does not depend on θ".*

To motivate the mathematical definition we consider an "experiment". Suppose there are two statisticians, we call them A and B. A random sample $X = (X_1, X_2, ..., X_n)$ is drawn from a population. Statistician A knows this entire random sample, while statistician B only knows the values $t_1, t_2, ..., t_k$ of the sufficient statistics $T_i = r_i(X_1, X_2, ..., X_n)$ for $i = 1, ..., k$. Now the conditional distribution of $X_1, X_2, ..., X_n$ given $(T_1, T_2, ..., T_k) = (t_1, t_2, ..., t_k)$ and θ does not depend on θ. This implies that statistician B knows this distribution. Hence by the use of a computer statistician B is able to generate a random sample $X_t = (X_1', X_2', ..., X_n')$ which has this conditional distribution. Then X$_t$ has the same distribution as a random sample drawn from the population, and statistician B can use this random sample to compute whatever statistician A computes by the random sample X drawn from the population. On average statistician B will do as well as A on estimating unknown parameters of the population. Thus the mathematical definition of sufficient statistics implies the heuristic definition [3].

We illustrate the fact that the conditional distribution of X given $T = t$, is independent of θ by an example:

Example 2.1 Suppose one has conducted n Bernoulli trials with outcome $X = (X_1, X_2, ..., X_n)\sim$Bernoulli($\theta$). Statistician A knows the outcome of each of these trials, while statistician B only knows the value of the sufficient statistic $T = \sum_{i=1}^{n} X_i$. Now the conditional distribution of $X = (X_1, X_2, ..., X_n)$ given $T = \sum_{i=1}^{n} X_i = t$ and θ is seen to be independent of θ by the following:

$$
\begin{aligned}
P(X = x|T = t) &= P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n| \sum_{i=1}^{n} X_i = t) \\
&= \frac{P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n, \sum_{i=1}^{n} X_i = t)}{P(\sum_{i=1}^{n} X_i = t)} \\
&= \frac{P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n)}{\binom{n}{t}\theta^t (1-\theta)^{n-t}} \\
&= \frac{\theta^{x_1}(1-\theta)^{x_1} ... \theta^{x_n}(1-\theta)^{x_n}}{\binom{n}{t}\theta^t (1-\theta)^{n-t}}
\end{aligned}
$$

$$= \frac{\theta^t(1-\theta)^{n-t}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}}$$
$$= \frac{1}{\binom{n}{t}}$$

It is difficult to check if a set of statistics are sufficient or to find such a set in terms of the mathematical definition. But there is a theorem that could be applied [3]:

Theorem 1. (Factorization Theorem) *Let $X_1, X_2, ..., X_n$ be a random sample with joint density $f(x_1, x_2, ..., x_n|\theta)$. The statistics*

$$T_i = r_i(X_1, X_2, ..., X_n) \qquad (for \quad i = 1, ..., k)$$

are jointly sufficient if and only if the joint density can be factored as follows:

$$f(x_1, x_2, ..., x_n|\theta) = u(x_1, x_2, ..., x_n)v(r_1(x_1, x_2, ..., x_n), r_2(x_1, x_2, ..., x_n), ..., r_k(x_1, x_2, ..., x_n), \theta)$$

where u and v are non-negative functions. The function u can depend on the full random sample $x_1, x_2, ..., x_n$ but not the unknown parameter θ. The function v can depend on θ, but can depend on the random sample only through the values of $r_i(x_1, x_2, ..., x_n)$, for $i = 1, ..., k$.

"Let $g(t_1, t_2, ..., t_k)$ be a function whose values are in R^k and which is one to one. Also let $g_i(t_1, t_2, ..., t_k), i = 1, ..., k$ be the component functions of g. Then if $T_1, T_2, ..., T_k$ are jointly sufficient, then $g_i(T_1, T_2, ..., T_k)$ are jointly sufficient" [3]. This result will be used in the next chapter when we identify sufficient statistics of the non-homogeneous Poisson process.

We now apply Theorem 1 to an example concerning a normal distributed population, with a single unknown population parameter [3]:

Example 2.2 We consider a normal population for which the mean μ is unknown, but the variance σ^2 is known. The joint density is

$$\begin{aligned} f(x_1, x_2, ..., x_n|\mu) =& (2\pi)^{-n/2}\sigma^{-n}exp(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(t_i-\mu)^2) \\ =& (2\pi)^{-n/2}\sigma^{-n}exp(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{n\mu^2}{2\sigma^2}) \end{aligned} \tag{2.1}$$

Since σ^2 is known, we can let

$$u(x_1, x_2, ..., x_n) = (2\pi)^{-n/2}\sigma^{-n}exp(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}x_i^2)$$

and

$$v(r(x_1, x_2, ..., x_n), \mu) = exp(\frac{-n\mu^2}{2\sigma^2} + \frac{\mu}{\sigma^2}r(x_1, x_2, ..., x_n))$$

where

$$r(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i$$

By the Factorization Theorem this shows that $T = \sum_{i=1}^{n} X_i$ is a sufficient statistic. It follows that the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is also a sufficient statistic in this case.

Now if both the mean and the variance where unknown parameters of the population, the sample mean is not a sufficient statistic. In this case we need to use more than one statistic to get sufficiency. It is seen from equation(2.1) that:

$$T_1 = \sum_{i=1}^{n} X_i \qquad T_2 = \sum_{i=1}^{n} X_i^2$$

are jointly sufficient statistics in this case.

Chapter 3

Definition & Sufficient Statistics of NHPP Models

In this chapter we identify the sufficient statistic in the NHPP model. This is done in accordance with the definitions of the previous chapter and in particular by applying the Factorization Theorem.

3.1 Definition NHPP

[13] A counting process $\{N(t), t \geq 0\}$ is said to be a non-homogeneous Poisson process with intensity function $\lambda(t), t \geq 0$, if:

(i) $N(0) = 0$
(ii) $\{N(t), t \geq 0\}$ has independent increments.
(iii) $P\{N(t + \Delta t) - N(t) = 1\} = \lambda(t)\Delta t + o(\Delta t)$
(iv) $P\{N(t + \Delta t) - N(t) \geq 2\} = o(\Delta t)$

The CROCOF ("cumulative rate of occurence of failures") for a NHPP is usually denoted $\Lambda(t)$ and defined as [12]:

$$\Lambda(t) = \int_0^t \lambda(u)du$$

We are concerned with two different parametrizations of the non-homogeneous Poisson process, which is the power law process and the log linear process with intensity functions and belonging CROCOF functions given below [2]:

Power Law Intensity

$$\lambda(t) = abt^{b-1}, \qquad (a, b > 0, t \geq 0) \tag{3.1}$$

$$\Lambda(t) = at^b \tag{3.2}$$

Log Linear Intensity

$$\lambda(t) = e^{a+bt} \qquad (-\infty < a, b < \infty, t \geq 0) \tag{3.3}$$

$$\Lambda(t) = \begin{cases} \frac{e^a(e^{bt}-1)}{b} & b \neq 0 \\ e^a t & b=0 \end{cases} \tag{3.4}$$

3.2 Sufficient Statistics of NHPP Models

Suppose a failure process $T_1, T_2, ..., T_n$ is modelled by a NHPP with intensity function $\lambda(t)$. In accordance with the Factorization Theorem (given in the previous chapter) we will identify the joint denstity function resulting from the observed failure times $\{T_j\}$ in order to find the sufficient statistics for both power law intensity and log linear intensity. We will take into account that the failure times could be failure censored or time censored:

1. **Failure censoring** : A system is observed from time $t = 0$ until the n'th failure at time T_n, resulting in failure times $T_1, T_2, ..., T_n$. In this case the length of the observation interval is stochastic.

2. **Time censoring** : A system is observed from time $t = 0$ until a predetermined time τ, with failures occuring at $T_1, T_2, ..., T_n$. In this case the number of observed failures, $N(\tau) = n$, in the time interval [0,τ] is stochastic.

In case of a failure censored dataset the joint density function is [2]:

$$f(t_1, t_2, ..., t_n|\theta) = \{\prod_{j=1}^{n} \lambda(t_j)\} e^{-\Lambda(t_n)} \tag{3.5}$$

And in the case of a time censored dataset the joint density of $T_1, T_2, ..., T_n$ is [2]:

$$f(t_1, t_2, ..., t_n, n|\theta) = \{\prod_{j=1}^{n} \lambda(t_j)\} e^{-\Lambda(\tau)} \tag{3.6}$$

3.2.1 Power Law Intensity

Failure censoring

We start with the case of failure censored data, and put equations (3.1) and (3.2) into equation (3.5) to obtain the joint density function:

$$\begin{aligned} f(t_1, t_2, ..., t_n|\theta) =& \{\prod_{j=1}^{n} \lambda(t_j)\} e^{-\Lambda(t_n)} \\ =& \prod_{j=1}^{n} abt_j^{b-1} e^{-at_n^b} \end{aligned}$$

Then by applying the Factorization Theorem and let $u(t_1, t_2, ..., t_n)$=1, we see that the statistics $(T_n, \prod_{j=1}^{n-1} T_j)$ are jointly sufficient. This implies (since $log(T)$ is a 1-1 function) that we can take

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

as the sufficient statistics.

Time Censoring

If we turn to the case of time censored data and put equations (3.1) and (3.2) into equation(3.6) to obtain the the joint density function which is now:

$$\begin{aligned} f(t_1, t_2, ..., t_n, n|\theta) =& \{\prod_{j=1}^{n} \lambda(t_j)\} e^{-\Lambda(\tau)} \\ =& \prod_{j=1}^{n} abt_j^{b-1} e^{-a\tau^b} \end{aligned}$$

Again we apply the Factorization Theorem and let $u(t_1, t_2, ..., t_n)$=1, to obtain the jointly sufficient statistics $(N(\tau) = n, \prod_{j=1}^{n} T_j)$. This implies by the same argument as above that

$$S = (N(\tau) = n, \sum_{j=1}^{n} \log T_j)$$

are jointly sufficient in this case.

3.2.2 Log Linear Intensity

Failure Censoring

We start with the failure censored data and put equation (3.3) and (3.4) into equation(3.5), assuming $b \neq 0$, to obtain the joint density function:

$$\begin{aligned} f(t_1, t_2, ..., t_n|\theta) =& \{\prod_{j=1}^{n} \lambda(t_j)\} e^{-\Lambda(t_n)} \\ =& \prod_{j=1}^{n} (e^{a+bt_j}) e^{-(\frac{e^a(e^{bt_n}-1)}{b})} \end{aligned}$$

By applying the Factorization Theorem we obtain the jointly sufficient statistics

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

Time Censoring

In order to obtain the sufficient statistics for the time censored data we put equations (3.3) and (3.4) into equation(3.6), assuming $b \neq 0$, and obtain the joint density function:

$$\begin{aligned} f(t_1, t_2, ..., t_n, n|\theta) &= \{\prod_{j=1}^{n} \lambda(t_j)\} e^{-\Lambda(\tau)} \\ &= \prod_{j=1}^{n} (e^{a+bt_j}) e^{-(\frac{e^a(e^{b\tau}-1)}{b})} \end{aligned}$$

Hence by the Factorization Theorem the jointly sufficient statistics in this case is:

$$S = (N(\tau) = n, \sum_{j=1}^{n} T_j)$$

Table(3.1) displays the sufficient statistic for these four cases.

Table 3.1: Sufficient Statistics for power law and log linear NHPP

Intensity	Power-law	Log linear
Failure Censoring	$(\log T_n, \sum_{j=1}^{n-1} \log T_j)$	$(T_n, \sum_{j=1}^{n-1} T_j)$
Time Censoring	$(N(\tau) = n, \sum_{j=1}^{n} \log T_j)$	$(N(\tau) = n, \sum_{j=1}^{n} T_j)$

Chapter 4

Conditional Testing given a Sufficient Statistic

This chapter is based on [6]. Suppose $T = (T_1, T_2, ..., T_n)$ is a vector of observed failure times of a system. One is interested in testing the nullhypothesis H_0 that these data come from a non-homogeneous Poisson process of a specific parametric type:

H_0 :Observed failure times $T_1, T_2, ..., T_n$ comes from a non-homogeneous Poisson process of a specific parametric type.

We design a test statistic $W \equiv W(T)$ which is a function of the failures times $T_1, T_2, ..., T_n$ and aims to reveal departure from the nullhypothesis. Suppose that $S \equiv S(T)$ are the sufficient statistics for the unknown parameters in the assumed model of H_0.

An α-level conditional test of the nullhypothesis is obtained if H_0 is rejected for given $S = s$, when $W \geq k(s)$,where $k(s)$ is a critical value chosen such that:

$$P_{H_0}(W \geq k(s)|S = s) = \alpha \qquad \text{(for all } s)$$

Thus, in order to find the critical value $k(s)$ and being able to make statisitcal inference concerning the nullhypothesis one needs to know the conditional distribution of W given $S = s$. As we have demonstrated in chapter 2 this distribution is by sufficiency independent of the unknown parameters of the model, and in principle it can be found. However, this might be very difficult or even impossible in many practical cases and we have to rely on simulations.

Suppose we are able to simulate the conditional distribution of W given $S = s$. We are then able to calculate the conditional p-value:

$$p_{obs} = P_{H_0}(W \geq w_{obs}|S = s) \tag{4.1}$$

where w_{obs} is the value of the test stasticic $W(T)$ based on the observed failure times $T_1, T_2, ..., T_n$. We reject the nullhypothesis if $p_{obs} \leq \alpha$, since if

$$p_{obs} = P_{H_0}(W \geq w_{obs}|S = s) \leq \alpha$$

and

$$P_{H_0}(W \geq k(s)|S = s) = \alpha$$

then

$$w_{obs} \geq k(s)$$

which would lead to rejection of H_0.

We give an algorithm for this procedure:

Algorithm 1 : Conditional Testing of T given S = s

(1) Start with observed failure times $T = (T_1, T_2, ..., T_n)$ of a system.
(2) Choose a suitable test statistic $W \equiv W(T)$, and calculate w_{obs}.
(3) Calculate the sufficient statistics $S \equiv S(T) = s$
(4) Simulate the conditional distribution of W given $S = s$.
(5) Calculate p_{obs} as in equation(4.1) .
(6) Reject H_0 if $p_{obs} \leq \alpha$
where α is a predetermined level of significance.

In order to complete algorithm 1, and make statistical inference concerning our nullhypothesis, we need a method for simulating the conditional distribution of W given $S = s$. In the next chapter we present a general method for how one can simulate this conditional distribution, and then in chapter 6 we apply this method to our specific problem concerning the non-homogeneous Poisson processes.

Notice:

The conditional test described above is also an unconditional α-level test, since we must have

$$\begin{aligned} P_{H_0}(\text{reject H}_0) &= \int P_{H_0}(\text{reject H}_0|S = s) f_s(s) ds \\ &= \int P_{H_0}(W \geq k(s)|S = s) f_s(s) ds \\ &= \int \alpha f_s(s) ds = \alpha \end{aligned}$$

where the last equality follows from $\int f_s(s)ds{=}1$ (integral of density). Thus if H_0 holds, the probability of rejection is α.

Chapter 5

Conditional Monte Carlo Based on Sufficient Statistics

In this chapter we present a general approach for Monte Carlo computations of conditional expectations given a sufficient statistic.

The test statistic $W = W(X)$ introduced in the previous chapter, is a function of the observations $X = (X_1, X_2, ..., X_n)$. We define a function $\phi(X)$ such that

$$\phi(X) = \begin{cases} 1 & \text{if} \quad W(X) \geq w_{obs} \\ 0 & \text{if} \quad W(X) < w_{obs} \end{cases}$$

where w_{obs} is the value of W based on the observed values of X. This implies that the conditional p-value in algorithm 1 (previous chapter) could be expressed as:

$$p_{obs} = P_{H_0}(W(X) \geq w_{obs}|T = t) = E[\phi(X)|T = t]$$

In this chapter we present a general method for how one can simulate such conditional expectations given the sufficient statistic T, in order to complete algorithm 1.

The method presented in this chapter was first introduced by Lillegård and Engen (1997) [4], and was further developed by Lindqvist and Taraldsen [7, 8, 9].

The ability of direct sampling from the conditional distiribution would be particularly useful. It is demonstrated by Lindqvist and Taraldsen that if a certain condition, the pivotal condition, is satisfied this could be done by a simple parameter adjustment of the original statistical model. But in general one needs to apply a weighting scheme in order to obtain the correct conditional expectation. The coming sections are based on work by Lindqvist and Taraldsen [7, 8, 9].

5.1 Setup and basic algorithm

We consider a general pair (X, T) of a random vector consisting of the observations $X = (X_1, X_2, ..., X_n)$ and a sufficient statistic T, with joint distribution indexed by a parameter θ. Suppose there is a given random vector U, with known distribution function $f(u)$ such that (X, T) could be simulated by means of U. "More precisely one assumes the existence of functions χ and τ, such that, for each θ, the joint distribution of $(\chi(U, \theta), \tau(U, \theta))$ equals the joint distribution of (X, T)" [8]. We consider an example [8]:

Example 5.1:

Exponential distribution. Suppose $X = (X_1, X_2, ..., X_n)$ are i.i.d. from the exponential distribution with hazard rate θ, denoted Exp(θ). Then $T = \sum_{i=1}^{n} X_i$ is sufficient for θ. Letting $U = (U_1, U_2, ..., U_n)$ be i.i.d. Exp(1) variables we can put

$$\chi(U, \theta) = (U_1/\theta, ..., U_n/\theta)$$
$$\tau(U, \theta) = \sum_{j=1}^{n} U_j/\theta$$

The aim is to obtain a sample X_t of the conditional distribution of X given $T = t$. Since this distribution is independent of θ, it is reasonable to assume that it can be described in some simple way by means of U. According to [8] a suggestive method for this would be to first draw U from its known distribution, then to determine a parameter value $\hat{\theta}$ such that $\tau(U, \hat{\theta}) = t$. Then finally $X_t(U) = \chi(U, \hat{\theta})$ could be used as the desired sample. By applying this procedure we obtain a sample of X with the corresponding T having the correct value t. The question whether or not $X_t(U)$ really is a sample from the conditional distribution of X given $T = t$ then remains.

Example 5.1 (continued):

For given t and U there is a unique $\hat{\theta} \equiv \hat{\theta}(U, t)$ with $\tau(U, \hat{\theta}) = t$, namely

$$\hat{\theta}(U, t) = \frac{\sum_{i=1}^{n} U_i}{t}$$

This leads to the sample

$$X_t(U) = \chi\{U, \hat{\theta}(U, t)\} = (\frac{tU_1}{\sum_{i=1}^{n} U_i}, ..., \frac{tU_n}{\sum_{i=1}^{n} U_i})$$

and it is well known [8] that the distribution of $X_t(U)$ indeed coincides with the conditional distribution of X given $T = t$.

The algorithm used in the example above could more generally be described as:

Algorithm 1 : Conditional sampling of X given T = t.

(1) Generate U from its density $f(u)$
(2) Solve $\tau(U,\theta) = t$ for θ. The (unique) solution is $\hat{\theta}(U,t)$.
(3) Return $X_t(U) = \chi\{U, \hat{\theta}(U,t)\}$

The Pivotal Condition

In addtion to uniqueness of $\hat{\theta}(U,t)$ in step 2, a pivotal condition needs to be satisfied to ensure that algorithm 1 produce a sample $X_t(U)$ from the conditional distribution of X given $T = t$. "Assume that $\tau(U,\theta)$ depends on u only through a function $r(u)$, where the value of $r(u)$ can be uniquely recovered from the equation $\tau(U,\theta) = t$ for given θ and t. This means that there is a function $\tilde{\tau}$ such that $\tau(U,\theta) = \tilde{\tau}\{r(u),\theta\}$ for all (u,θ) and a function $\tilde{v}$ such that $\tilde{\tau}\{r(u),\theta\} = t$ implies $r(u) = \tilde{v}(\theta,t)$. Note that in this case $\tilde{v}(\theta,T)$ is a pivotal quantity in the classical meaning that its distribution does not depend on θ" [8].

Example 5.1 (**Continued**) :

The pivotal condition is satisfied here with $r(U) = \sum_{i=1}^{n} U_i$. Thus Algorithm 1 is valid, as verified earlier by a direct method [8].

Hence, when the pivotal condition is satisfied the conditional expectation of $\phi(X)$ given the sufficient statistic $T = t$ could be simulated by means of U using:

$$E[\phi(X)|T = t] = E[\phi(X_t)]$$

since the samples X_t are known functions of U.

5.2 General algorithm for unique $\hat{\theta}(u,t)$, Eucledian Case

In general algorithm 1 will not produce samples from the correct conditional distribution, even if the solution $\hat{\theta}$ of $\tau(U,\theta) = t$ is unique. This fact is proven by a counterexample in [9]

Hence a modified algorithm has to be constructed when the pivotal condition is not satisfied in the model we consider. The main idea is to consider the parameter θ as a random variable Θ, independent of U, and with some conveniently chosen distribution π [7].

This idea leads to the result that the conditional distribution of X given $T = t$ could be simulated by:

$$E\{\phi(X)|T = t\} = \frac{E[\phi(X_t(U))W_t(U)]}{E[W_t(U)]}$$

where $W_t(U)$ acts as a weight function for a sample U from $f(u)$.

In the Eucledian case the weight function $W_t(U)$ is given by [8]:

$$W_t(U) = |\frac{\pi(\theta)}{det\partial_\theta\tau(U,\theta)}|_{\theta=\hat{\theta}(U,t)}$$

Chapter 6

Conditional Simulation for Parametric NHPP Models

An argument (somewhat heuristic) [13] shows that given n events of a non-homogeneous Poisson process with intensity function $\lambda(t)$ by time τ, the n events are distributed as the ordering of n independent observations with a common density function:

$$f(t) = \frac{\lambda(t)}{\Lambda(\tau)} \qquad (\text{for} \quad 0 \leq t \leq \tau) \tag{6.1}$$

Similarly in the case of failure censoring, given n events of a NHPP with intensity function $\lambda(t)$, with its last failure occuring at time T_n, the $(n-1)$ earlier events will be distributed as the ordering of $(n-1)$ independent observations with a common density function:

$$f(t) = \frac{\lambda(t)}{\Lambda(T_n)} \qquad (\text{for} \qquad 0 \leq t \leq T_n) \tag{6.2}$$

Hence by applying these results together with the method presented in the previous chapter (algorithm 1, chapter 5) we are able to simulate samples of a non-homogeneous Poisson process with intensity function $\lambda(t)$, given the sufficient statistic S. In accordance with the previous chapter these samples are then to be used to determine the conditional expectation of $\phi(T)$ given the sufficient statistic $S = s$.

Both parametrizations of the NHPP, considered in chapter 3, have a parameter vector of two unknowns $\theta = (a, b)$, with jointly sufficient statistics $S = (s_1, s_2)$. The sufficient statistics are given in table(3.1). The method we present in this chapter is divided in two steps concerning $S = (s_1, s_2)$.

Failure Censoring

1. We condition on the first element of the sufficient statistic, s_1. From table(3.1) it is seen that $s_1 = \log T_n$ for the power law parametrization and $s_1 = T_n$ for the log linear parametrization.

2. Then by the argument above it is known that the $(n-1)$ first failures are distributed as the ordering of $(n-1)$ independent observations with common density function given by equation (6.2). Hence one is able to simulate these $(n-1)$ observations by the inversion method, conditioned on the second element of S, s_2, which is now seen to be sufficient for this model. From table (3.1) we see that $s_2 = \sum_{j=1}^{n-1} \log T_j$ for the power law parametrization, and $s_2 = \sum_{j=1}^{n-1} T_j$ for the log linear parametrization.

Time Censoring

1. Again we condition on the first element of the sufficient statistic, s_1. From table(3.1) we see that $s_1 = N(\tau) = n$, for both parametrizations.

2. Then by the argument above it is known that these n failures are distributed as the ordering of n independent observations with a common density function given by equation (6.1). Hence one is able to simulate these n observations by the inversion method, conditioned on the second element of S, s_2, which is now seen to be sufficient for this model. From table(3.1) we see that $s_2 = \sum_{j=1}^{n} \log T_j$ for the power law parametrization, and $s_2 = \sum_{j=1}^{n} T_j$ for the log linear parametrization.

Pivotal Condition

In accordance with chapter 5, we then have to check if the pivotal condition is satisfied, for both parametrizations, in order to simulate the conditional expectation

$$E[\phi(T)|S = s]$$

6.1 Power Law Intensity

Failure Censoring

Suppose a system is observed from time $t = 0$, until the n'th failure occurs at time T_n. Assuming that the observed failure times $T = (T_1, T_2, ..., T_n)$ are NHPP with power law intensity, we present a method for simulating a sample $T_s = (T_1', T_2', ..., T_n')$ of T given the sufficient statistic $S = (\log T_n, \sum_{j=1}^{n-1} T_j)$. The method is divided in two steps:

1. We condition on the first element of S, $s_1 = \log T_n = \log T_n'$.

2. From the argument above we know that the $(n-1)$ first failures are distributed as the ordering of $(n-1)$ independent observations with common density function given by equation (6.2):

$$f(t) = \frac{\lambda(t)}{\Lambda(T_n)} = \frac{bt^{b-1}}{T_n^b} \qquad (\text{for} \quad 0 \leq t \leq T_n)$$

with one unknown parameter b. This implies that we have to do $(n-1)$ simulations from the distribution $f(t)$ conditioned on $s_2 = \sum_{j=1}^{n-1} \log T_j$, which is seen to be sufficient for b in this model. Following the method presented in the previous chapter (algorithm 1, chapter 5) we search for functions $\chi(U, b)$ and $\tau(U, b)$ such that $((T_1, ..., T_{n-1}), s_2)$ could be simulated by means of a random vector U with known distribution, for given values of b. The cumulative distribution function of $f(t)$ is given by:

$$F(t) = \int_0^t f(y)dy = (\frac{t}{T_n})^b = U$$

with inverse function

$$F^{-1}(U) = T_n U^{\frac{1}{b}} = t$$

We let $U = (U_1, U_2, ..., U_{n-1})$ ~uniform[0,1], and

$$\chi(U,b) = ((T_n U_1^{\frac{1}{b}}), (T_n U_2^{\frac{1}{b}}), ..., (T_n U_{n-1}^{\frac{1}{b}}))$$

$$\begin{aligned} \tau(U,b) =& \sum_{j=1}^{n-1} \log(\chi(U_j,b)) = (\sum_{j=1}^{n-1} \log T_n + \frac{1}{b} \log U_j) \\ =& (n-1)\log T_n + \frac{1}{b}\sum_{j=1}^{n-1} \log U_j \end{aligned} \tag{6.3}$$

Now by solving the equation

$$(n-1)\log T_n + \frac{1}{b}\sum_{j=1}^{n-1} \log U_j = \sum_{j=1}^{n-1} \log T_j = s_2$$

with respect to b yields:

$$\hat{b} = \frac{\sum_{j=1}^{n-1} \log U_j}{s_2 - (n-1)\log T_n}$$

and we are able to simulate the $(n-1)$ first events of T, which by sufficiency is independent of the value of b, by means of U:

$$T'_j = \chi(U_j, \hat{b}) = T_n U_j^{\frac{1}{\hat{b}}} \quad \text{(for j=1,..,n-1)}$$

which then satisfies $\sum_{j=1}^{n-1} \log T'_j = s_2$.

Pivotal Condition :

From equation (6.3) we see that the pivotal condition holds with $r(u) = \sum_{j=1}^{n-1} \log U_j$ and the method does indeed produce a sample T_s from the conditional distribution of T given $S = (s_1, s_2)$.

Hence the conditional expectation of $\phi(T)$ given the sufficient statistic $S = (s_1, s_2)$ could now be simulated by means of U, using:

$$E[\phi(T)|S = (s_1, s_2)] = E[\phi(T_s)]$$

Time Censoring:

Suppose a system is observed with $N(\tau) = n$ failures in the time interval [0,τ]. Assuming the observed failure times $T = (T_1, T_2, ..., T_n)$ are NHPP with power law intensity, we present a method for simulating a sample $T_s = (T_1', T_2', ..., T_n')$ of T given the sufficient statistic $S = (N(\tau) = n, \sum_{j=1}^{n} \log T_j)$. The method is divided in two steps:

1. We condition on the first element of S, $s_1 = N(\tau) = n$, which means that our sample should contain n failures in the time interval [0,τ].

2. By the argument given above we know that the n events should be the ordering of n independent observations with a common density function given by equation (6.1):

$$f(t) = \frac{\lambda(t)}{\Lambda(\tau)} = \frac{bt^{b-1}}{\tau^b} \quad (\text{for } 0 \leq \text{t} \leq \tau)$$

with one unknown parameter b. This implies that we have to do n simulations from the distribution $f(t)$ conditioned on $s_2 = \sum_{j=1}^{n} \log T_j$, which is seen to be sufficient for b in this model. The cumulative distribution function is now:

$$F(t) = \int_0^t f(y)dy = (\frac{t}{\tau})^b = U$$

with inverse function

$$F^{-1}(U) = \tau U^{\frac{1}{b}} = t$$

We let $U = (U_1, U_2, ..., U_n)$ ~uniform[0,1], and

$$\chi(U, b) = ((\tau U_1^{\frac{1}{b}}), (\tau U_2^{\frac{1}{b}}), ..., (\tau U_n^{\frac{1}{b}}))$$

$$\begin{aligned} \tau(U, b) =& \sum_{j=1}^{n} \log(\chi(U_j, b)) = (\sum_{j=1}^{n} \log \tau + \frac{1}{b} \log U_j) \\ =& n \log \tau + \frac{1}{b} \sum_{j=1}^{n} \log U_j \end{aligned} \tag{6.4}$$

By solving the equation:

$$n \log \tau + \frac{1}{b} \sum_{j=1}^{n} \log U_j = \sum_{j=1}^{n} \log T_j = s_2$$

with respect to b yields:

$$\hat{b} = \frac{\sum_{j=1}^{n} \log U_j}{s_2 - n \log \tau}$$

and we are able to simulate the n events of T, which by sufficiency is independent of the value of b, by means of U:

$$T_j' = \chi(U_j, \hat{b}) = \tau U_j^{\frac{1}{\hat{b}}} \quad (\text{for j=1,..,n})$$

which then satisfies $\sum_{j=1}^{n} \log T_j = s_2$.

Pivotal Condition :

From equation (6.4) we see that the pivotal condition holds with $r(u) = \sum_{j=1}^{n} \log U_j$ and the method produces a sample T_s from the conditional distribution of T given $S = (s_1, s_2)$.

Hence the conditional expectation of $\phi(T)$ given the sufficient statistic $S = (s_1, s_2)$ could now be simulated by means of U, using:

$$E[\phi(T)|S = (s_1, s_2)] = E[\phi(T_s)]$$

6.2 Log Linear Intensity

Failure Censoring

Suppose a system is observed from time $t = 0$, until the n'th failure occurs at time T_n. Assuming that the observed failure times $T = (T_1, T_2, ..., T_n)$ are NHPP with log linear intensity we present a method for simulating a sample T_s of T given the sufficient statistic $S = (T_n, \sum_{j=1}^{n-1} T_j)$. Again the method is divided in two steps:

1. We condition on the first element of S, $s_1 = T_n = T_n'$.

2. By the same argument as used previously we know that the $(n-1)$ first failure times are distributed as the ordering of $(n-1)$ independent observations with common density function:

$$f(t) = \frac{\lambda(t)}{\Lambda(T_n)} = \frac{be^{bt}}{e^{bT_n} - 1} \qquad (\text{for} \quad 0 \leq t \leq T_n)$$

with one unknown parameter b. This implies that we have to do $(n-1)$ simulations from the distribution $f(t)$ conditioned on $s_2 = \sum_{j=1}^{n-1} T_j$, which is seen to be sufficient for b in this model. The cumulative distribution function $F(t)$ is:

$$F(t) = \int_0^t f(y)dy = \frac{e^{bt} - 1}{e^{bT_n} - 1} = U$$

with inverse function

$$F^{-1}(U) = \frac{\log(1 + U(e^{bT_n} - 1))}{b} = t$$

We let $U = (U_1, U_2, ..., U_{n-1})$ ~uniform[0,1], and

$$\chi(U, b) = [(\frac{\log(1 + U_1(e^{bT_n} - 1))}{b}), ..., (\frac{\log(1 + U_{n-1}(e^{bT_n} - 1))}{b})]$$

$$\tau(U, b) = \sum_{j=1}^{n-1} \chi(U_j, b) = \sum_{j=1}^{n-1} \frac{\log(1 + U_j(e^{bT_n} - 1))}{b} \tag{6.5}$$

To obtain an estimate for b, we have to solve the equation:

$$\sum_{j=1}^{n-1} \frac{\log\left(1 + U_j(e^{bT_n} - 1)\right)}{b} = \sum_{j=1}^{n-1} T_j = s_2$$

which means that we have to solve the equation:

$$g(b) = \sum_{j=1}^{n-1} \log\left(1 + U_j(e^{bT_n} - 1)\right) - bs_2 = 0$$

for b. The equation is seen to convex by differentiating it twice with respect to b, and in addition to the trivial solution $b = 0$, it has a unique additional solution $b = \hat{b}$, which is the one we look for. This solution could be found by different numerical techniques, such as the bisection method for instance.

When an estimate $\hat{b}$ is obatined we are able to simulate the $(n-1)$ first events of T, which by sufficiency is independent of the value of b, by means of U:

$$T_j' = \chi(U_j, \hat{b}) = \frac{\log\left(1 + U_j(e^{\hat{b}T_n} - 1)\right)}{\hat{b}} \quad \text{(for j=1,..,n-1)}$$

which then satisfies $\sum_{j=1}^{n-1} T_j' = s_2$.

Pivotal Condition and Weights

It is clear from the equation (6.5) that the pivotal condition is not satisfied here. Hence, in order to simulate the conditional expectation of $\phi(T)$ given $S = (s_1, s_2)$, we need to evaluate the weights given by:

$$\begin{aligned} W_s(U) &= \left| \frac{\pi(\theta)}{det \partial_\theta \tau(U, \theta)} \right|_{\theta = \hat{\theta}(U,s)} \\ &= \left| \frac{\pi(b)}{\partial_b \tau(U, b)} \right|_{b = \hat{b}(U,s_2)} \end{aligned}$$

where $\pi(b)$ is some arbitrarily chosen function of b, and

$$\partial_b \tau(U, b) = \sum_{j=1}^{n-1} \frac{\frac{bU_jT_ne^{bT_n}}{(1+U_j(e^{bT_n}-1))} - \log(1 + U_j(e^{bT_n} - 1))}{b^2}$$

Hence the conditional expectation of $\phi(T)$ given the sufficient statistic $S = (s_1, s_2)$ could now be simulated by means of U, using:

$$E[\phi(T)|S = (s_1, s_2)] = \frac{E[\phi(T_s)|\frac{\pi(b)}{\partial_b \tau(U,b)}|_{b=\hat{b}(U,s_2)}]}{E[|\frac{\pi(b)}{\partial_b \tau(U,b)}|_{b=\hat{b}(U,s_2)}]}$$

Time Censoring

Suppose a system is observed with $N(\tau) = n$ failures in the time interval [0,τ]. Assuming the failure times $T = (T_1, T_2, ..., T_n)$ are NHPP with log linear intensity we present a method for simulating a sample T_s of T given the sufficient statistic $S = (N(\tau) = n, \sum_{j=1}^{n} T_j)$. Again the method is divided in two steps:

1. We condition on the first element of S, $s_1 = N(\tau) = n$, which means that our sample should contain n failures in the time interval [0,τ]

2. By the same argument as used previously we know that the n failure times are distributed as the ordering of n independent observations with common density function:

$$f(t) = \frac{\lambda(t)}{\Lambda(\tau)} = \frac{be^{bt}}{e^{b\tau} - 1} \quad (\text{for } 0 \leq \text{t} \leq \tau)$$

with one unknown parameter b. This implies that we have to do n simulations from the distribution of $f(t)$ conditioned on $s_2 = \sum_{j=1}^{n} T_j$, which is seen to be sufficient for b in this model. The cumulative distribution function $F(t)$ is in this case:

$$F(t) = \int_0^t f(y)dy = \frac{e^{bt} - 1}{e^{b\tau} - 1} = U$$

with inverse function

$$F^{-1}(U) = \frac{\log(1 + U(e^{b\tau} - 1))}{b} = t$$

We let $U = (U_1, U_2, ..., U_n)$ ~uniform[0,1],and

$$\chi(U, b) = [(\frac{\log(1 + U_1(e^{b\tau} - 1))}{b}), ..., (\frac{\log(1 + U_n(e^{b\tau} - 1))}{b})]$$

$$\tau(U, b) = \sum_{j=1}^{n} \chi(U_j, b) = \sum_{j=1}^{n} \frac{\log(1 + U_j(e^{b\tau} - 1))}{b} \tag{6.6}$$

To obtain an estimate for b, we have to solve the equation:

$$\sum_{j=1}^{n} \frac{\log(1 + U_j(e^{b\tau} - 1))}{b} = \sum_{j=1}^{n} T_j = s_2$$

which means that we have to solve the equation:

$$g(b) = \sum_{j=1}^{n} \log(1 + U_j(e^{b\tau} - 1)) - bs_2 = 0$$

for b. The equation is seen to convex by differentiating it twice with respect to b, and in addition to the trivial solution $b = 0$, it has a unique additional solution $b = \hat{b}$, which is the one we look for. This solution could be found by different numerical techniques, as the bisection method for instance.

When an estimate $\hat{b}$ is obatined we are able to simulate the n events of T, which by sufficiency is independent of the value of b, by means of U:

$$T'_j = \chi(U_j, \hat{b}) = \frac{\log(1 + U_j(e^{\hat{b}\tau} - 1))}{\hat{b}} \quad \text{(for j=1,..,n)}$$

which then satisfies $\sum_{j=1}^{n} T'_j = s_2$.

Pivotal Condition and Weights

It is clear from the equation (6.6) that the pivotal condition is not satisfied here. Hence, in order to simulate the conditional expectation of $\phi(T)$ given $S = (s_1, s_2)$, we need to evaluate the weights given by:

$$\begin{aligned} W_s(U) &= \left|\frac{\pi(\theta)}{det\partial_\theta\tau(U,\theta)}\right|_{\theta=\hat{\theta}(U,s)} \\ &= \left|\frac{\pi(b)}{\partial_b\tau(U,b)}\right|_{b=\hat{b}(U,s_2)} \end{aligned}$$

where $\pi(b)$ is some arbitrarily chosen function of b, and

$$\partial_b\tau(U,b) = \sum_{j=1}^{n} \frac{\frac{bU_j\tau e^{b\tau}}{(1+U_j(e^{b\tau}-1))} - \log(1 + U_j(e^{b\tau} - 1))}{b^2}$$

Hence the conditional expectation of $\phi(T)$ given the sufficient statistic $S = (s_1, s_2)$ could now be simulated by means of U, using:

$$E[\phi(T)|S = (s_1, s_2)] = \frac{E[\phi(T_s)|\frac{\pi(b)}{\partial_b\tau(U,b)}|_{b=\hat{b}(U,s_2)}]}{E[|\frac{\pi(b)}{\partial_b\tau(U,b)}|_{b=\hat{b}(U,s_2)}]}$$

6.3 Gibbs Sampling, Log Linear Intensity

We consider an alternative method for simulating samples T_s of a NHPP with log-linear intensity function given by $\lambda(t) = e^{a+bt}$. These samples could then be applied to determine the conditional expectation:

$$E[\phi(T)|S = s]$$

Failure Censoring

Suppose a system is observed from time $t = 0$, until the n'th failure occurs at time T_n. Assuming that the observed failure times $T = (T_1, T_2, ..., T_n)$ are NHPP with log linear intensity we present a method for simulating a sample T_s of T given the sufficient statistic $S = (T_n, \sum_{j=1}^{n-1} T_j)$. The key to the new approach is that the conditional distribution of

T given $S = s$ does not depend on the unknown parameter $\theta = (a, b)$ of the model, which implies that we are able to set the parameters a=b=0. This gives intensity function:

$$\lambda(t) = 1$$

and cumulative intensity function:

$$\Lambda(t) = t$$

We follow the same procedure as before by dividing our simulation in two steps concerning the two elements of the sufficient statistic $S = (s_1, s_2)$.

1. We condition on the first element of S, $s_1 = T_n = T'_n$

2. Then the $(n-1)$ first failures are distributed as the ordering of $(n-1)$ independent observations with common density function:

$$f(t) = \frac{\lambda(t)}{\Lambda(T_n)} = \frac{1}{T_n}$$

which is seen to be uniform on [0,T_n]. Then one has to do $(n-1)$ simulations from the distribution $f(t)$ given the statistic $s_2 = \sum_{j=1}^{n-1} T_j$. The cumulative distribution function is now given by:

$$F(t) = \int_0^t f(t) = \frac{1}{T_n} t = U$$

with inverse function:

$$F^{-1}(U) = T_n U = t$$

Thus we let $U = (U_1, ..., U_{n-1})$ ∼uniform [0,1], and

$$\chi(U) = (T_n U_1, ..., T_n U_{n-1})$$
$$\tau(U) = T_n \sum_{j=1}^{n-1} U_j$$

Hence the $(n-1)$ first events of T given $\sum_{j=1}^{n-1} T_j = s_2$ may be simulated by drawing $T_s = (T'_1, ..., T'_{n-1})$ ∼uniform [0,T_n] conditoned that $\sum_{j=1}^{n-1} T'_j = s_2$.

Time Censoring

Suppose a system is observed with $N(\tau) = n$ failures in the time interval [0,τ]. Assuming the failure times $T = (T_1, T_2, ..., T_n)$ are NHPP with log linear intensity we present a method for simulating a sample T_s of T given the sufficient statistic $S = (N(\tau) = n, \sum_{j=1}^{n} T_j)$. The key to the new approach is that the conditional distribution of T given $S = s$ does not depend on the unknown parameter $\theta = (a, b)$ of the model, which implies that we are able to set the parameters $a = b = 0$. This gives intensity function:

$$\lambda(t) = 1$$

and cumulative intensity function:

$$\Lambda(t) = t$$

We follow the same procedure as before by dividing our simulation in two steps concerning the two elements of the sufficient statistic $S = (s_1, s_2)$.

1. We condition that $N(\tau) = n$ failures should occur in the time interval [0,τ].

2. Then the n failures are distributed as the ordering of n independent observations with a common density function:

$$f(t) = \frac{\lambda(t)}{\Lambda(\tau)} = \frac{1}{\tau}$$

which is seen to be uniform on [0,τ]. Then one has to do n simulations from the distribution $f(t)$ given the statistic $s_2 = \sum_{j=1}^{n} T_j$. The cumulative distribution function is now given by:

$$F(t) = \int_0^t f(t) = \frac{1}{\tau} t = U$$

with inverse function:

$$F^{-1}(U) = \tau U = t$$

Thus we let $U = (U_1, ..., U_n)$ ~uniform [0,1], and

$$\chi(U) = (\tau U_1, ..., \tau U_n)$$
$$\tau(U) = \tau \sum_{j=1}^{n} U_j$$

Hence the n events of T given $\sum_{j=1}^{n} T_j = s_2$ may be simulated by drawing $T_s = (T_1', ..., T_n')$ ~uniform [0,τ] conditoned that $\sum_{j=1}^{n} T_j' = s_2$.

We present 3 algorithms for how one can simulate $T = (T_1, ..., T_n)$ ~uniform [0,τ] given the statistic $S = \sum_{j=1}^{n} T_j = s$.

Algorithm 1 : Rue's Algorithm:

(1) Start with $T_i = \frac{s}{n}$ for $i = 1, ..., n$
(2) Draw integers j_1 and j_2 between 1 and n, $j_1 \neq j_2$.
(3) Draw d ~uniform [0,τ].
(4) Calculate

$$T_{j_1}' = T_{j_1} + d \qquad \text{and} \qquad T_{j_2}' = T_{j_2} - d$$

if

$$T_{j_1}' \leq \tau, \qquad \text{and} \qquad 0 \leq T_{j_2}'$$

then

$$T_{j_1}^{\text{new}} = T'_{j_1}, \qquad \text{and} \qquad T_{j_2}^{\text{new}} = T'_{j_2}$$

else

$$T_{j_1}^{\text{new}} = T_{j_1}, \qquad \text{and} \qquad T_{j_2}^{\text{new}} = T_{j_2}$$

endif
(5) Return to step 2.

Algorithm 2 : Gibbs Algorithm:

(1) Start with $T_i = \frac{s}{n}$ for $i = 2, ..., n$
(2) Given $(T_2^m, ..., T_n^m)$ with $s - \tau \leq \sum_{i=2}^n T_i^m \leq s$.
(3) Draw $j \in \{2, ..., n\}$ randomly.
(4) Calculate

$$z = \sum_{i=2, i \neq j}^{n} T_i^m$$

Replace T_j^m by

$$T_j^{m+1} = \begin{cases} \text{uniform}[0, s - \tau] & \text{if} \quad z \geq s - \tau \\ \text{uniform}[s - \tau - z, \tau] & \text{if} \quad z < s - \tau \end{cases}$$

(5) Return to step 2.

Algorithm 3 : Gibbs Block Algorithm:

(1) Start with $T_i = \frac{s}{n}$ for $i = 1, ..., n$
(2) Given $(T_1^m, ..., T_n^m)$ with $\sum_{i=1}^n T_i^m = s$.
(3) Draw integers $i < j$ randomly
(4) Draw

$$T_i^{m+1} = \begin{cases} \text{uniform}[0, T_i^m + T_j^m] & \text{if} \quad T_i^m + T_j^m \leq \tau \\ \text{uniform}[T_i^m + T_j^m - \tau, \tau] & \text{if} \quad T_i^m + T_j^m > \tau \end{cases}$$

Let

$$T_j^{m+1} = T_i^m + T_j^m - T_i^{m+1}$$

(5) Return to step 2.

Chapter 7

Statistical Inference in NHPP Models

In this chapter we present different test statistics developed in order to reveal departure from the nulllhypothesis that a given set of observed failure times $T = (T_1, ..., T_n)$ are from a NHPP with power law or log linear intensity function.

7.1 Test Statistics Failure Censoring

In the previous chapter it was seen that the $(n-1)$ first failures were distributed as the ordering of $(n-1)$ independent observations with a common density function, and cumulative density function given by:

$$f(t) = \frac{\lambda(t)}{\Lambda(T_n)} \qquad (\text{for} \quad 0 \leq t \leq T_n)$$

$$F(t) = \int_0^t f(t) = \frac{\Lambda(t)}{\Lambda(T_n)}$$

which has one unknown parameter b. Hence if we let

$$V_j = F(T_j) = \frac{\Lambda(T_j)}{\Lambda(T_n)} \qquad (\text{for} \quad j = 1, ..., n-1)$$

it is seen that the vector V is distributed as the order statistic of $(n-1)$ independent and identically distributed variables on [0,1].

Now the V_j are, however, nonobservable since Λ depends on unknown parameters. Thus, suppose Λ^* is an estimate of Λ. Then we define

$$V_j^* = \frac{\Lambda^*(T_j)}{\Lambda^*(T_n)} \qquad (\text{for} \quad j = 1, ..., n-1)$$

One then anticipates $V^* = (V_1^*, ..., V_{n-1}^*)$ to behave much similar to uniform variables on [0,1].

In the following V^* is based on the maximum likelihood estimate for the unkonwn parameter b. We now present 7 different test statistics, one of these is a two-sided test statistic from [1], while the 6 others are all one-sided and is found in [10, 11]. All these test statistics are based on the assumed uniform behaviour of V^*.

Greenwood Test Statistic [1]

$$G = \sum_{j=1}^{n} (V_j^* - V_{j-1}^*)^2$$

where V_n^*=1 and V_0^*=0. This is a two-sided test statistic and the null hypothesis of NHPP is rejected for either too small or too large values of this statistic.

Modified Cramer − von Mises Statistic [10]

$$W^2 = \sum_{j=1}^{n-1} [V_j^* - \frac{(2j-1)}{2(n-1)}]^2 + \frac{1}{12(n-1)}$$

Modified Kolmogorov − Smirnov Statistic [10]

$$D = \max \quad \{D^+, D^-\}$$

where

$$D^+ \equiv \max_{1 \leq j \leq (n-1)} (\frac{j}{n-1} - V_j^*)$$

and

$$D^- \equiv \max_{1 \leq j \leq (n-1)} (V_j^* - \frac{(j-1)}{(n-1)})$$

Modified Anderson Darlington Statistic [10]

$$A^2 = \frac{-\{\sum_{j=1}^{n-1} (2j-1)[\log V_j^* + \log(1 - V_{n-j}^*)]\}}{(n-1)} - (n-1)$$

Modified Kuiper Statistic [11]

$$V = D^+ + D^-$$

where D^+ and D^- are defined as above.

Modified Watson Statistic [11]

$$U^2 = \frac{1}{12(n-1)} + \sum_{j=1}^{n-1} [\frac{(2j-1)}{2(n-1)} - V_j^*]^2 - (n-1)(\bar{V} - 0.5)^2$$

where

$$\bar{V} \equiv \frac{1}{(n-1)} \sum_{j=1}^{n-1} V_j^*$$

Modified Weighted Watson Statistic [11]

$$J = (n-1)^2 \sum_{j=1}^{n-1} d_j^2 - (n-1)(\sum_{j=1}^{n-1} d_j)^2$$

where

$$d_j \equiv \frac{[V_j^* - \frac{j}{n}]}{[j(n-j)]^{\frac{1}{2}}}$$

7.2 Test Statistics Time Censoring

In the previous section it was seen that the $N(\tau) = n$ failures in the time interval $[0,\tau]$ were distributed as the ordering of n independent observations with a common density function, and cumulative density function given by:

$$f(t) = \frac{\lambda(t)}{\Lambda(\tau)} \qquad (\text{for} \quad 0 \leq t \leq \tau)$$

$$F(t) = \int_0^t f(t) = \frac{\Lambda(t)}{\Lambda(\tau)}$$

which has one unknown parameter b. Hence if we let

$$V_j = F(T_j) = \frac{\Lambda(T_j)}{\Lambda(\tau)} \qquad (\text{for} \quad j = 1, ..., n)$$

it is seen that the vector V is distributed as the order statistic of n independent and identically distributed variables on [0,1].

Now the V_j are, however, nonobservable since Λ depends on unknown parameters. Thus, suppose Λ^* is an estimate of Λ. Then we define

$$V_j^* = \frac{\Lambda^*(T_j)}{\Lambda^*(\tau)} \qquad (\text{for} \quad j = 1, ..., n)$$

One then anticipates $V^* = (V_1^*, ..., V_n^*)$ to behave much similar to uniform variables on [0,1]. In the following V^* is based on the maximum likelihood estimate for the unkonwn parameter b, and we are able to adjust the test statistics given above for the new situation of time censoring:

Greenwood Test Statistic

$$G = \sum_{j=1}^{n+1} (V_j^* - V_{j-1}^*)^2$$

where V_{n+1}^{*}=1 and V_0^{*}=0. This is a two-sided test statistic and the null hypothesis of NHPP is rejected for either too small or too large values of this statistic.

Modified Cramer – von Mises Statistic

$$W^2 = \sum_{j=1}^{n}[V_j^{*} - \frac{(2j-1)}{2n}]^2 + \frac{1}{12n}$$

Modified Kolmogorov – Smirnov Statistic

$$D = \max\ [D^{+}, D^{-}]$$

where

$$D^{+} \equiv \max_{1\leq j\leq n}(\frac{j}{n} - V_j^{*})$$

and

$$D^{-} \equiv \max_{1\leq j\leq n}(V_j^{*} - \frac{(j-1)}{n})$$

Modified Anderson Darlington Statistic

$$A^2 = \frac{-\{\sum_{j=1}^{n}(2j-1)[\log V_j^{*} + \log(1 - V_{n+1-j}^{*})]\}}{n} - n$$

Modified Kuiper Statistic

$$V = D^{+} + D^{-}$$

where D^{+} and D^{-} are defined as above.

Modified Watson Statistic

$$U^2 = \frac{1}{12n} + \sum_{j=1}^{n}[\frac{(2j-1)}{2n} - V_j^{*}]^2 - n(\bar{V} - 0.5)^2$$

where

$$\bar{V} \equiv \frac{1}{n}\sum_{j=1}^{n} V_j^{*}$$

Modified Weighted Watson Statistic

$$J = n^2\sum_{j=1}^{n} d_j^2 - n(\sum_{j=1}^{n} d_j)^2$$

where

$$d_j \equiv \frac{[V_j^{*} - \frac{j}{n+1}]}{[j(n+1-j)]^{\frac{1}{2}}}$$

7.3 Power Law Intensity

Failure Censoring

In this situation the vector V is defined by:

$$V_j = F(T_j) = \frac{\Lambda(T_j)}{\Lambda(T_n)} = (\frac{T_j}{T_n})^b \qquad (\text{for} \quad j = 1, ..., n-1)$$

The maximum likelihood estimate b^* based on the observed failure times T is:

$$b^* = \frac{-n}{\sum_{j=1}^{n-1} \log \frac{T_j}{T_n}}$$

and our estimated transformed times V_j^* becomes:

$$V_j^* = (\frac{T_j}{T_n})^{b^*} \qquad (\text{for} \quad j = 1, ..., n-1)$$

Notice :

By the theoretical representation $T_j = T_n U_j^{\frac{1}{b}}$ it is seen that our estimated transformed times V_j^* in this particular situation becomes:

$$V_j^* = (\frac{T_j}{T_n})^{b^*} = U_j^{\frac{b^*}{b}} = U_j^{\frac{-n}{\sum_{j=1}^{n-1} \log U_j}} \qquad (\text{for} \quad j = 1, ..., n-1)$$

which is seen to independent of the unknown parameters of the model. This implies that for the power law NHPP, the V_j^* have a distribution which is independent of the unknown parameters (a, b), and could be simulated by means of a random vector $U \sim$uniform[0,1]. It turns out that this is not the case for the log linear NHPP.

Time Censoring

In this situation the vector V is defined by:

$$V_j = F(T_j) = \frac{\Lambda(T_j)}{\Lambda(\tau)} = (\frac{T_j}{\tau})^b \qquad (\text{for} \quad j = 1, ..., n)$$

The maximum likelihood estimate b^* based on the observed failure times T is:

$$b^* = \frac{-n}{\sum_{j=1}^{n} \log \frac{T_j}{\tau}}$$

and our estimated transformed times V_j^* becomes:

$$V_j^* = (\frac{T_j}{\tau})^{b^*} \qquad (\text{for} \quad j = 1, ..., n)$$

Notice :

As above it is seen by the theoretical representation $T_j = \tau U_j^{\frac{1}{b}}$ that our estimated transformed times V_j^* becomes:

$$V_j^* = (\frac{T_j}{\tau})^{b^*} = U_j^{\frac{b^*}{b}} = U_j^{\frac{-n}{\sum_{j=1}^{n} \log U_j}} \qquad (\text{for} \quad j = 1, ..., n)$$

which as above is seen to independent of the parameters a and b, and could be simulated by means of a random vector U ∼uniform[0,1].

7.4 Log Linear Intensity

Failure Censoring

In this situation the vector V is defined by:

$$V_j = F(T_j) = \frac{\Lambda(T_j)}{\Lambda(T_n)} = \frac{e^{bT_j} - 1}{e^{bT_n} - 1} \qquad (\text{for} \quad j = 1, ..., n-1)$$

The maximum likelihood estimate b^* is found by solving the equation [2]:

$$\sum_{j=1}^{n} T_j + \frac{n}{b} - nT_n(\frac{1}{1 - e^{-bT_n}}) = 0$$

with respect to b. This needs to be done by a numerical method such as the repeated bisection method. Then our estimated transformed times V_j^* becomes:

$$V_j^* = \frac{e^{b^*T_j} - 1}{e^{b^*T_n} - 1} \qquad (\text{for} \quad j = 1, ..., n-1)$$

Time Censoring

In this situation the vector V is defined by:

$$V_j = F(T_j) = \frac{\Lambda(T_j)}{\Lambda(\tau)} = \frac{e^{bT_j} - 1}{e^{b\tau} - 1} \qquad (\text{for} \quad j = 1, ..., n)$$

The maximum likelihood estimate b^* is found by solving the equation [2]:

$$\sum_{j=1}^{n} T_j + \frac{n}{b} - n\tau(\frac{1}{1 - e^{-b\tau}}) = 0$$

with respect to b. As above this needs to be done by a numerical method. Then our estimated transformed times V_j^* becomes:

$$V_j^* = \frac{e^{b^*T_j} - 1}{e^{b^*\tau} - 1} \qquad (\text{for} \quad j = 1, ..., n)$$

Chapter 8

Implementation NHPP Models

In this chapter we apply the LT (2006) method and Gibbs sampling in order to make exact statistical inference concerning NHPP models. We consider both goodness of fit testing and also exact confidence intervals for the unknown parameter b in our models when applying the LT (2006) method. This confidence interval could not be found using the Gibbs sampling. We then present the results of a power comparison between 5 of the different test statistics given in chapter 7. We also compare the LT (2006) method to the Gibbs Block algorithm, concerning how fast they converge. In the following we only consider the case of failure censoring, but all the reults are easily obtained in the case of time censoring.

8.1 Power Law Intensity

8.1.1 Simulating Samples

Assuming that the observed failure times $T = (T_1, ..., T_n)$ comes from a NHPP with power law intensity function, we are able to simulate samples $T_s = (T_1', ..., T_n')$ of the failure times T given the sufficient statistic S=$(\log T_n, \sum_{j=1}^{n-1} \log T_j)$, by the LT (2006) method. In figure (8.1a) we see a plot of the failure times $T = (T_1, ..., T_{10})$ of dataset a, in addition to 5 such conditional simulations of T. From the plot it seems that the observed failure times T of dataset a, are more "regularly" distributed througout the time interval than the simulated samples T_s.

8.1.2 Goodness of fit testing

It is now of interest to test if our two datasets are compatible with power law NHPP, by the LT (2006) method and by applying 5 of the test statistics given in the previous chapter. All the test statistics are functions of the V_j^*, which are expected to behave much similar to uniform variables. Figure (8.1b) gives a plot of $V^*(T)=(V_1^*, ..., V_{10}^*)$ for dataset a, in addition to 5 simulations of V^*. From the plot it seems that the $V_j^*(T)$ based on dataset a are more "regularly" distributed in the interval [0,1] than the simulated $V_j^*(U)$.

In chapter 6 it was demonstrated that in the power law case V_j^* was not dependent on the unknown parameter b and could be simulated by means of a random vector $U \sim$uniform[0,1]. Thus in the power law case we are able to simulate the unconditional distribution of the test statistics which are all functions of the V_j^*.

Dataset a

We start by applying Greenwoods two-sided test statistic to check if dataset a is consistent with power law NHPP. The simulated (unconditional) distribution of Greenwoods test statistic under the nullhypothesis that dataset a comes from a power law NHPP can be seen in figure (8.2).

We calculate g_{obs}=0.1263, which is seen to fall on the left tale of the distribution, and the resulting observed p-value is obtained by:

$$p_{obs} = 2 \cdot P_{H_0}(G \leq 0.1263) = 2 \cdot 0.02435 = 0.0480$$

Hence the Greenwood test statistic imply some evidence against the power law assumption for dataset a. This is the same result as obtained in [6].

It is now of interest to check if the other test statistics given in chapter 7 would imply the same results concerning dataset a. Hence 4 other test statistics are picked: Cramer-von Mises (W^2), Modified Kolmogorov Smirnov (D), Modified Kuiper (V) and Modified Watson U^2. The simulated (unconditional) distribution of these statistics under the nullhypothesis of power law NHPP are given in figure (8.3).

These are all one-sided tests and the nullhypothesis is rejected if:

$$p_{obs} = P_{H_0}(W(T) \geq w_{obs}) \leq \alpha$$

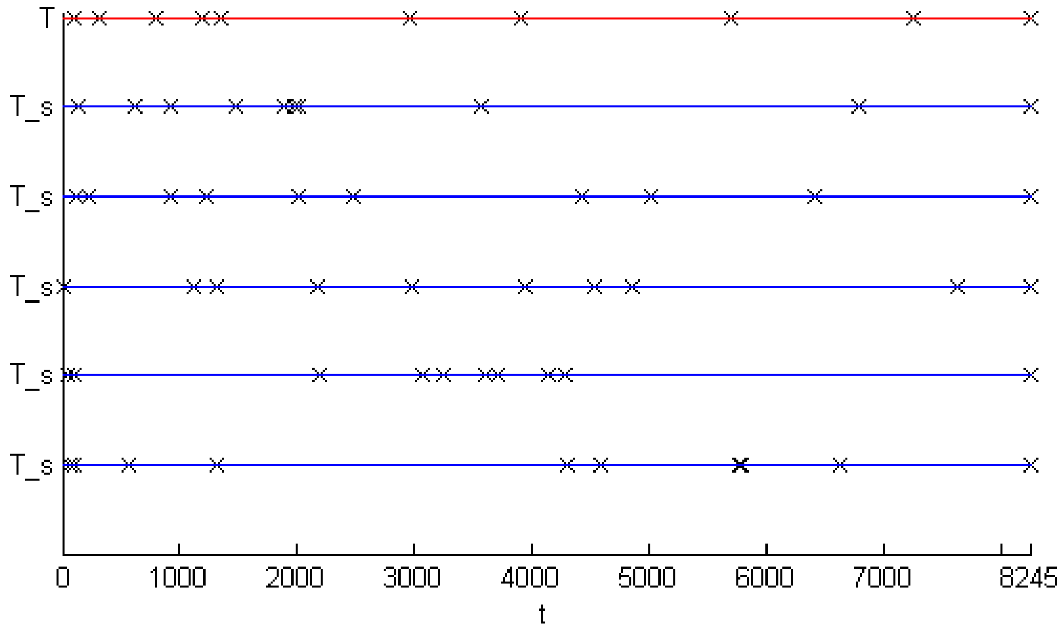
The value for each test statistic based on the observed failure times T is calculated, and the resulting p-values are given in table (8.1).

Table 8.1: Simulated (unconditional) p-values assuming dataset a is from power law NHPP

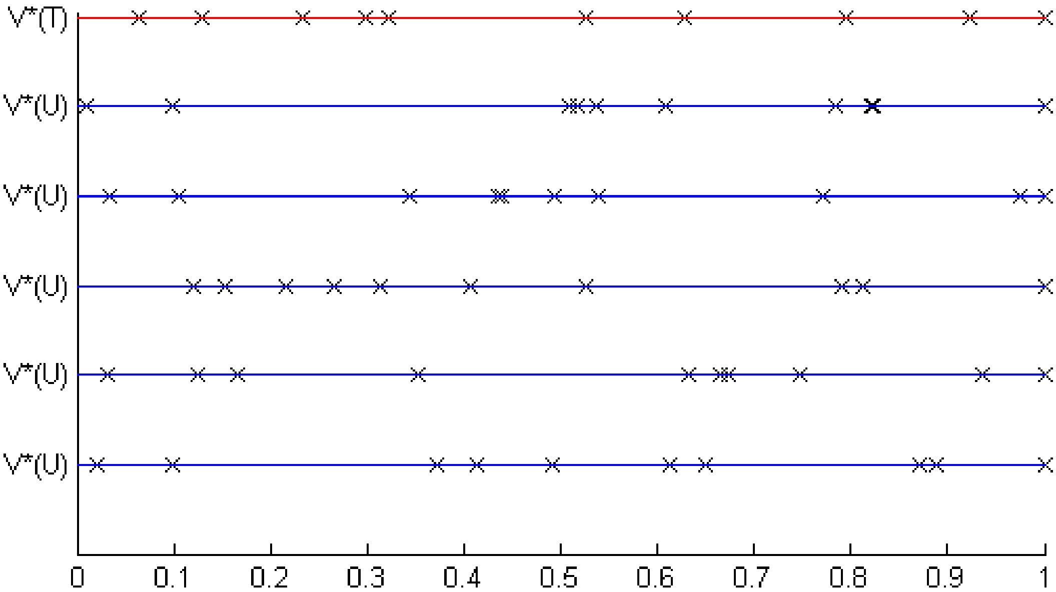
Test Statistic	W^2	D	V	U^2	G
p_{obs}	0.6621	0.5066	0.8069	0.8937	0.048

Neither of these 4 other test statistics imply any evidence against power law NHPP for dataset a, and the observed p-values are seen to deviate remarkebly from the Greenwood test statistic, which leads to the question of why this happens?

In order to answer this question we need to take a closer look at the test statistics.



(a) Displays the failure times $T = (T_1, ..., T_{10})$ for dataset a, in addition to 5 simulations, T_s, of T conditioned on the sufficient statistics $S = (\log T_n, \sum_{j=1}^{n-1} \log T_j)$ by the LT (2006) method.



(b) Displays $V^*(T) = (V_1^*, ..., V_{10}^*)$ for dataset a, in addition to 5 simulations $V^*(U)$, of $V^*(T)$, simulated by means of a random vector U ~uiform(0,1).

Figure 8.1

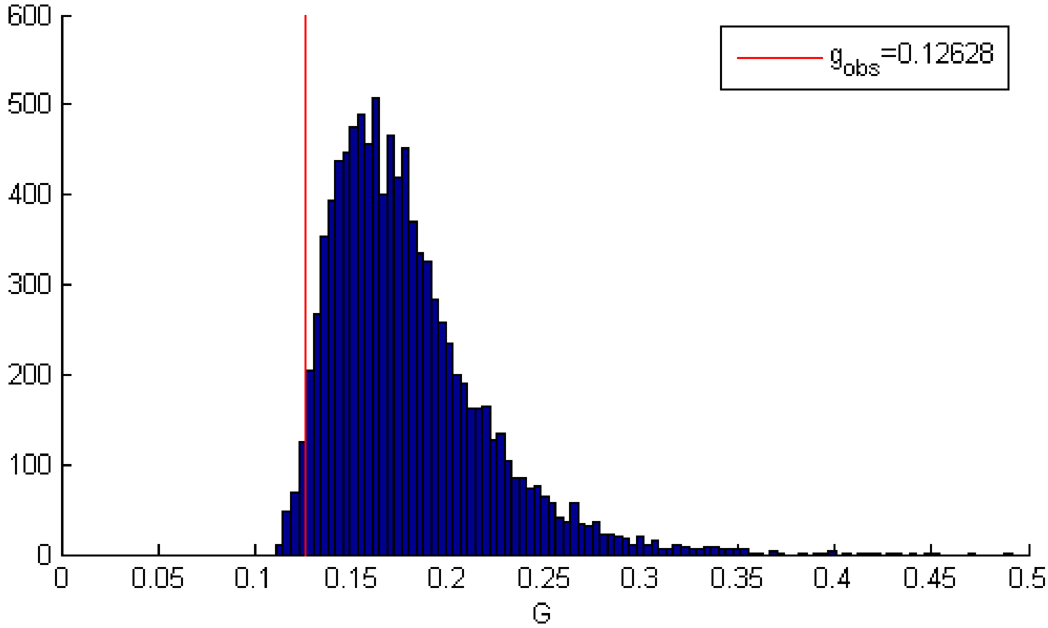


Figure 8.2: The simulated (unconditional) distribution of Greenwoods test statistic (G) under the nullhypothesis that dataset a comes from power law NHPP.

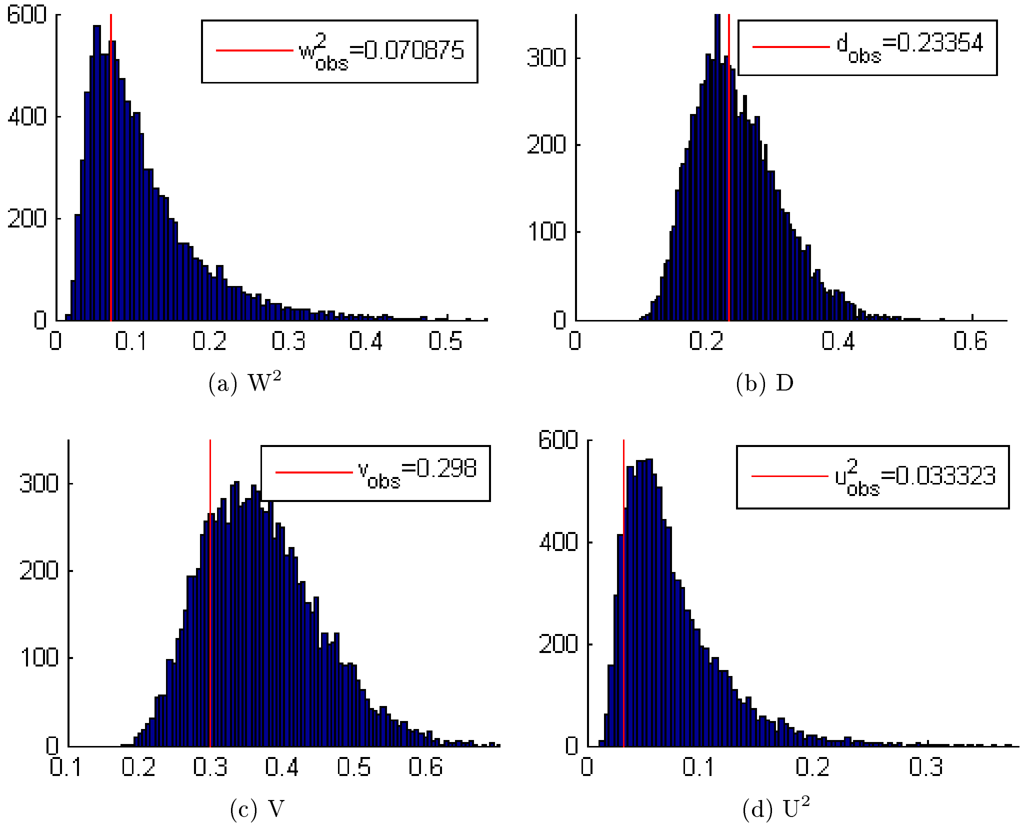
If the model under the nullhypothesis is correct the set $V^*(T) = (V_1^*, ..., V_{n-1})$ based on the observed failure times is expected to behave much similar to uniform [0,1] variables.

If we consider the Greenwood test statistic it is seen that if the V_j^* are too "regularly" distributed g_{obs} would be too small and hence lead to rejection. On the other hand if the V_j^* are too much "clumped" together in the interval [0,1], g_{obs} would be too large and also lead to rejection.

If we consider the 4 other test statistics, it is seen that these are all one-sided tests and the nullhypothesis is rejected only when the observed value of the test statistic is too large. This occurs if the V_j^* are too much "clumped" together as for the Greenwood statistic.

If we again consider the plot of V^* in figure (8.1b) it is seen that $V^*(T)$ based on the observed failure times T are much more "regularly" distributed throughout the interval [0,1] (almost equally spaced), than what is the case for the simulated V^*.

This explains why only Greenwood test statistic imply evidence against the nullhypothesis that dataset a comes from power law NHPP.



(a) W^2 (b) D

(c) V (d) U^2

Figure 8.3: The simulated (unconditional) distribution of the test statistics a) Cramer-von Mises (W^2) b) Modified Kolmogorov-Smirnov (D) c) Modified Kuiper (V) and d) Modified Watson (U^2), under the nullhypothesis of power law NHPP for dataset a.

Dataset USS Halfbeak

We now want to check if dataset d is consistent with power law NHPP. We follow the same procedure as given above for dataset a. The simulated (unconditional) distribution of the test statistics Greenwood (G), Cramer-von Mises (W^2), Modified Kolmogorov Smirnov (D), Modified Kuiper (V) and Modified Watson U^2, under the nullhypothesis that dataset d comes from power law NHPP are given in figure (8.4).

We calculate the the corresponding value w_{obs} for each of the 5 statistics and the observed p-values are given in table (8.2).

It is clear that all the test statistics implies strong evidence against the nullhypothesis that dataset d comes from a power law NHPP.

Table 8.2: Simulated (unconditional) p-values assuming dataset d is from power law NHPP

Test Statistic	W^2	D	V	U^2	G
p_{obs}	0.0001	0.0001	0.0005	0	0.01

8.1.3 Exact Confidence Intervals

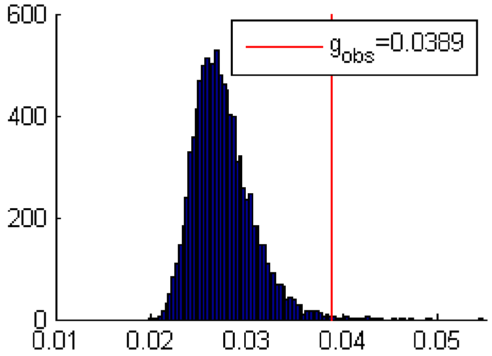
When applying the LT (2006) method to simulate samples of T given $S = s$ we estimate the parameter $b(U, s)$ for each sample. Hence if we order the m estimates $\hat{b}$ for b, $\tilde{b}_1 < ... < \tilde{b}_m$, then $(\tilde{b}_k, \tilde{b}_{m-k+1})$ is an exact 1-2k/(m+1) confidence interval for b [6]. In [6] it is seen that for the power law NHPP the above interval (for m$\rightarrow \infty$) 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.

Dataset a

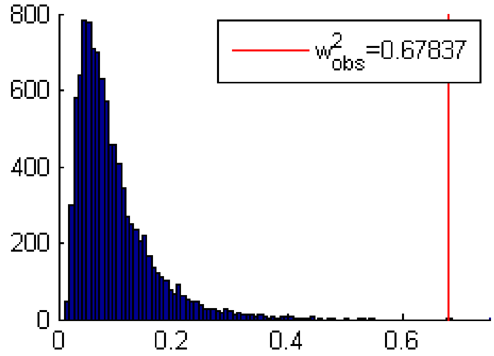
Figure (8.5a) shows the distribution of the estimated values $\hat{b}$ for dataset a under the assumption of power law intensity, applying LT (2006). The resulting 90% confidence interval for b is [0.2914, 0.8949]. This interval agrees with the one obtained in [6] calculated by the classical approach (as m$\rightarrow \infty$). Since the interval does not contain 1, this indicates reliability growth [6].

Dataset USS Halfbeak

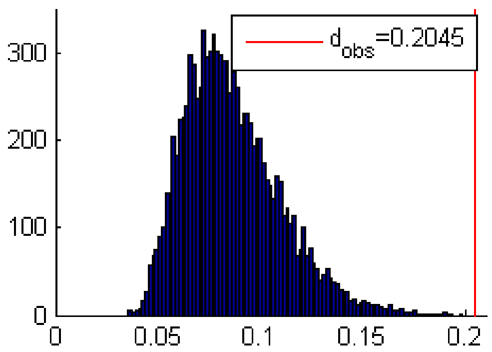
Figure (8.5b) shows the distribution of the estimated values $\hat{b}$ for dataset d, under the assumption of power law NHPP, applying LT (2006). The resulting 90% confidence interval for b is [2.2123, 3.2995]. This interval indicates reliability reduction.



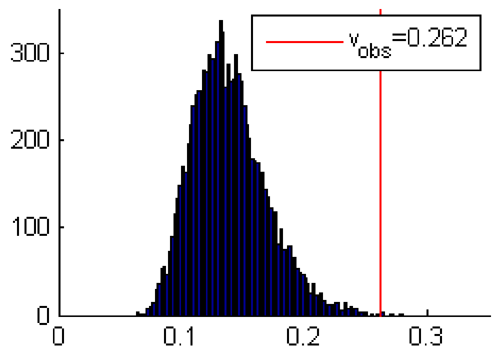
(a) G



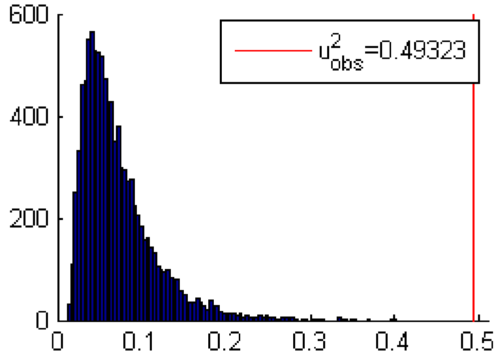
(b) W^2



(c) D

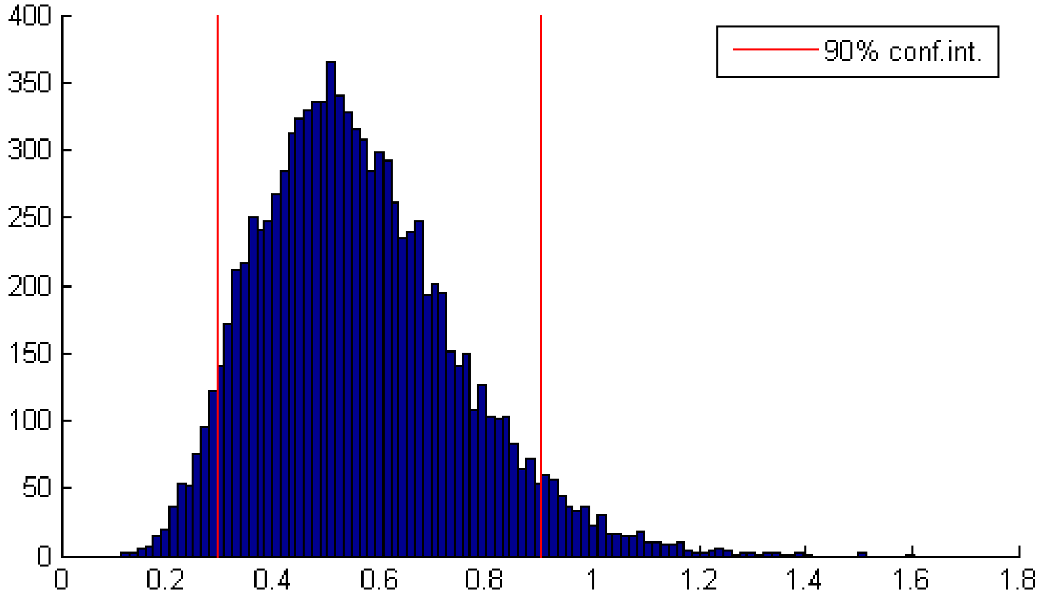


(d) V

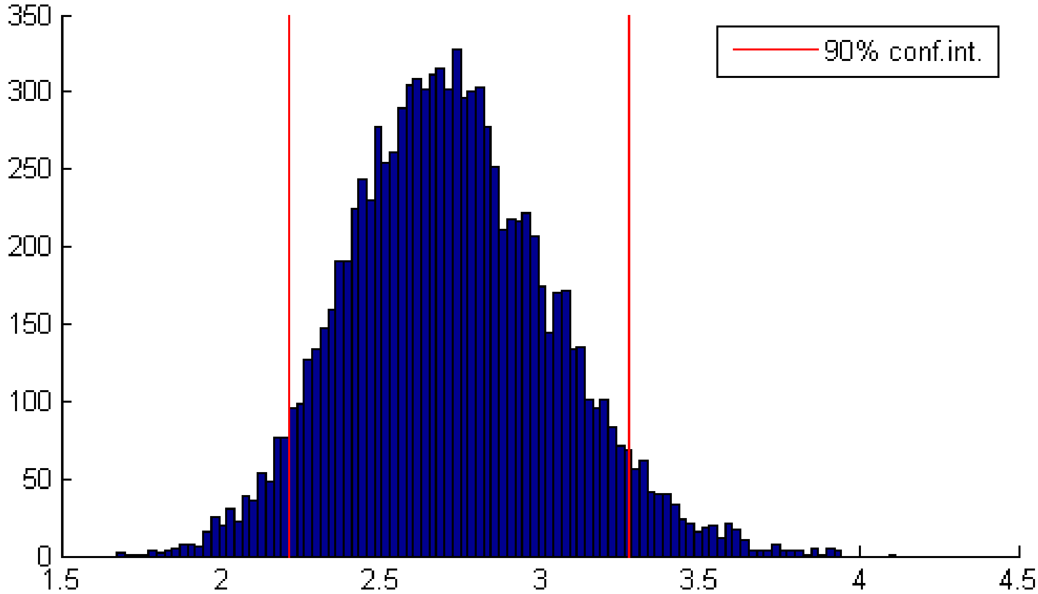


(e) U^2

Figure 8.4: The simulated (unconditional) distribution of the test statistics a) Greenwood, b) Cramer-von Mises c) Kolmogorov-Smirnov d) Modified Kuiper and e) Modified Watson, under the nullhypothesis of power law NHPP for dataset d.



(a) Distribution of $\hat{b}$ under the nullhypothesis of power law NHPP for dataset a, simulated by LT (2006).



(b) Distribution of $\hat{b}$ under the nullhypothesis of power law NHPP for dataset d, simulated by LT (2006).

Figure 8.5

8.1.4 Discussion of Results

The test statistics implies that both dataset a and d are not compatible with power law NHPP. For dataset a, it was only Greenwoods test statistic that indicated that the nullhypothesis should be rejected. For dataset d, all the test statistics lead to the conclusion that these data where not compatible with the nullhypothesis of power law NHPP. The estimated confidence intervals of the parameter b implies that dataset a has reliability growth, and dataset d has reliability reduction.

From the study of the datasets it is seen that it is important to understand the test statistics applied, and their characteristics. This agrees with a statement in [1] which deals with choosing test statistics. It is argued that engineering insight and understanding of the test statistics and their characteristics should provide as a basis for choosing the "correct" test statistic, or even construct "purpose built" tests for the specific data at hand.

Also by various plots of the data (simulations of T, V^*, confidence intervals etc.) such as demonstrated might lead to a better understanding of the observed failure times T, and better statistical inference.

8.2 Log Linear Intensity

Having concluded that the datasets are not consistent with power law NHPP we want to test if they are consistent with log linear NHPP. In order to compute conditional p-values (by simulation) we have applied both the LT (2006) method and Gibbs sampling in this section.

8.2.1 Simulating Samples

Assuming the observed failure times $T = (T_1, ..., T_n)$ are NHPP with log linear intensity we are able to simulate samples $T_s = (T'_1, ..., T'_n)$ of the failure times T given the sufficient statistic S=$(T_n, \sum_{j=1}^{n-1} T_j)$, by the LT (2006) method. Figure (8.6a) shows 5 such samples, together with the observed failure times $T = (T_1, ..., T_{10})$ for dataset a. From the plot it is again seen that dataset a looks more "regularly" distributed throughout the time interval, than the simulated samples T_s.

8.2.2 Goodness of fit testing

It is now of interest to check if our two datasets are compatible with the log linear NHPP assumption. As we have seen all the test statistics are functions of the V_j^*, which are claimed to behave much similar to the uniform[0,1] distribution. Figure (8.6b) gives a plot of $V^*(T)$ for dataset a, in addition to $V^*(T_s)$ for 5 simulated samples T_s of T given

$S = (T_n, \sum_{j=1}^{n-1} T_j)$, applying the LT (2006) method. The plot indicates that the V_j^* based on dataset a are more "regularly" distributed on the interval [0,1] than the simulated V_j^*, but less significant than in the power law case.

Dataset a

We start by considering Greenwoods test statistic, and calculate w_{obs}=0.1466. In order to simulate (by LT 2006) the conditional p-value:

$$p_{obs} = P_{H_0}(W \leq 0.1466 | S = s) = \frac{E[\phi(T_s)|\frac{\pi(b)}{\partial_b \tau(U,b)}|_{b=\hat{b}(U,s_2)}]}{E[|\frac{\pi(b)}{\partial_b \tau(U,b)}|_{b=\hat{b}(U,s_2)}]}$$

we have to consider the weights determined by the choice of the arbitrary function $\pi(b)$. We try two different choices for the function $\pi(b)$:

Suggested by [8] :

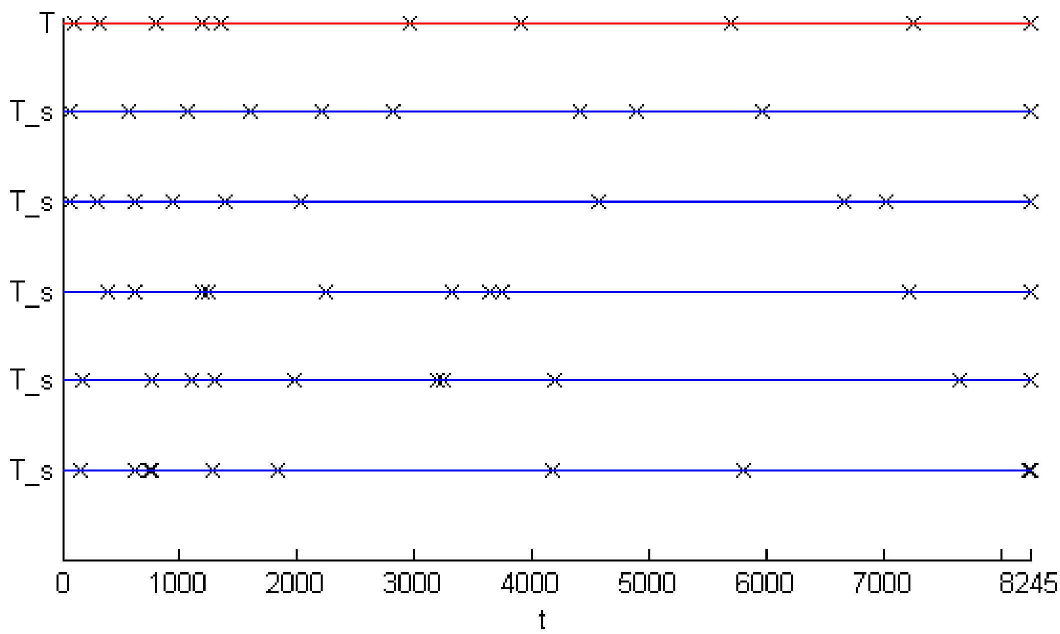
We choose the function π(b)=$\frac{1}{|b|}$, and compute (by simulation) the conditional p-value 5 times with 10 000 samples for each simulation. This resulted in the p-values [0.7834, 0.2134, 0.6230, 0.5840, 0.7472.] Hence the convergence seems to be very slow. By evalutaing the weights by a plot it is seen that most of them are close to 1, but there are a few weights that are very large. The reason is that when the estimated value $\hat{b}$ approaches 0 the corresponding weight will become very large and hence dominate the resulting conditional p-value. This explains the large variations in the simulated conditional p-value obtained above.

This problem could be avoided by putting a restriction on the weights. Hence we only accept a sample if the corresponding weight is less than 10. We run 5 new simulations with this new restriction imposed, again with 10 000 samples for each simulation. The corresponding p-values were [0.4703, 0.4747, 0.4579, 0.4706, 0.4570] with number of discardes samples [151,174,163,167,154]. Figure (8.7a) displays the distribution of the weights with this new restriction imposed.

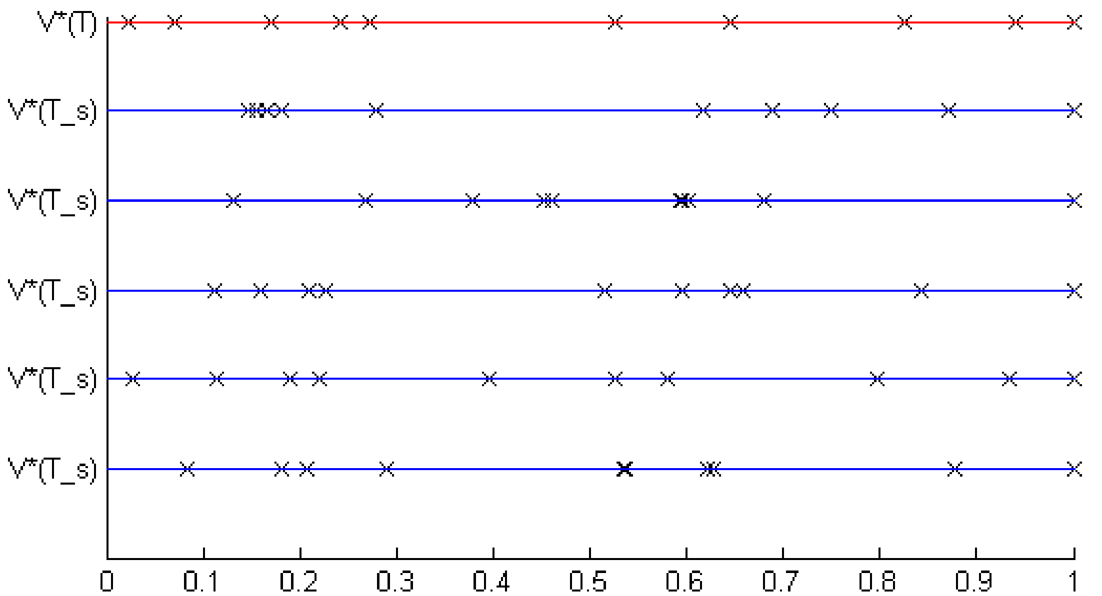
Jeffrey's Prior

We now choose the function π(b)=$\sqrt{\frac{1}{b^2} + \frac{1}{2-(e^b+e^{-b})}}$, which is known as Jeffrey's prior. We simulate the conditional p-value 5 times with 10 000 samples for each simulation. This resulted in the p-values [0.4611,0.4714, 0.4671, 0.4803, 0.4655]. Figure (8.7b) displays the distribution of the weights by using Jeffrey's prior as the choice for the arbitrary function $\pi(b)$. We see that there are a few weights that are slightly larger than the other, and we try to impose the restriction that only weights$<$0.6 are counted for.

We simulate the conditional p-value 5 new times with 10 000 samples for each simulation.This resulted in the p-values [0.4835, 0.4724, 0.4694, 0.4803, 0.4776] with number of discarded samples [167, 174, 157, 164, 165]. Hence this does not seem to improve the convergence, and the restriction does not seem necessary for this choice of $\pi(b)$.



(a) Displays the failure times $T = (T_1, ..., T_{10})$ for dataset a, in addition to 5 simulations, T_s, of T conditioned on the sufficient statistics $S = (T_n, \sum_{j=1}^{n-1} T_j)$ by the LT (2006) method.



(b) Displays $V^*(T) = (V_1^*, ..., V_{10}^*)$ for dataset a, in addition to 5 simulations $V^*(T_s)$, of $V^*(T)$, applying the LT (2006) method.

Figure 8.6

From the two plots in figure (8.7) it is seen that the Jeffrey's prior results in more even distributed weights and seems to be the better choice of the two functions considered.

Figure (8.8) displays the conditional distribution of Greenwoods test statistic assuming dataset a is log linear NHPP,using Jeffrey's Prior. The conditional p-value is 0.4599 and this does not imply any evidence against the nullhypothesis that dataset a is consistent with a log linear NHPP. In [6] the conditional p-value:

$$p_{obs} = 2 \cdot P_{H_0}(W \leq 0.1466|S = s) = 2 \cdot 0.217 = 0.434$$

which is seen to be a little bit off the p-value obtained here for dataset a. The choice of $\pi(b)$ in [6] could be the reason for this small deviation between the results. The number of simulations could be another reason. As it is seen from the calculations above the conditional p-value has small variations using 10 000 samples. More samples would increase the accuracy, and decrease the variation between each simulation.

It is now of interest to check if the other test statistics given in chapter 7 would imply the same results concerning the assumption that dataset a comes from a log linear NHPP. Again the 4 statistics Cramer-von Mises (W^2), Modified Kolmogorov-Smirnov (D), Modified Kuiper (V) and Modified Watson (U^2) are considered. The conditional p-values are now simulated by the alternative approach of **Gibbs sampling**, applying the Gibbs Block algorithm, given in chapter 6.

The resulting p-values are given in table (8.3).

Hence neither of these test statistics would reject the nullhypothesis that dataset a is consistent with log linear NHPP, either.

Table 8.3: Conditional p-values assuming dataset a is from log linear NHPP, simulated by Gibbs Block algorithm.

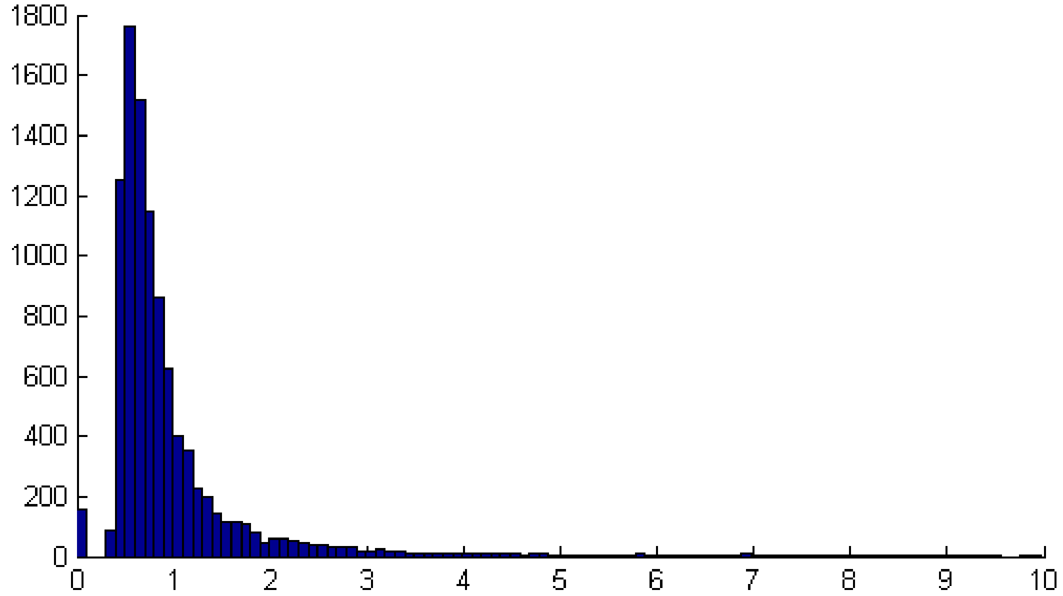
Test Statistic	W^2	D	V	U^2
p_{obs}	0.4255	0.2824	0.6521	0.7063

Datasett USS Halfbeak

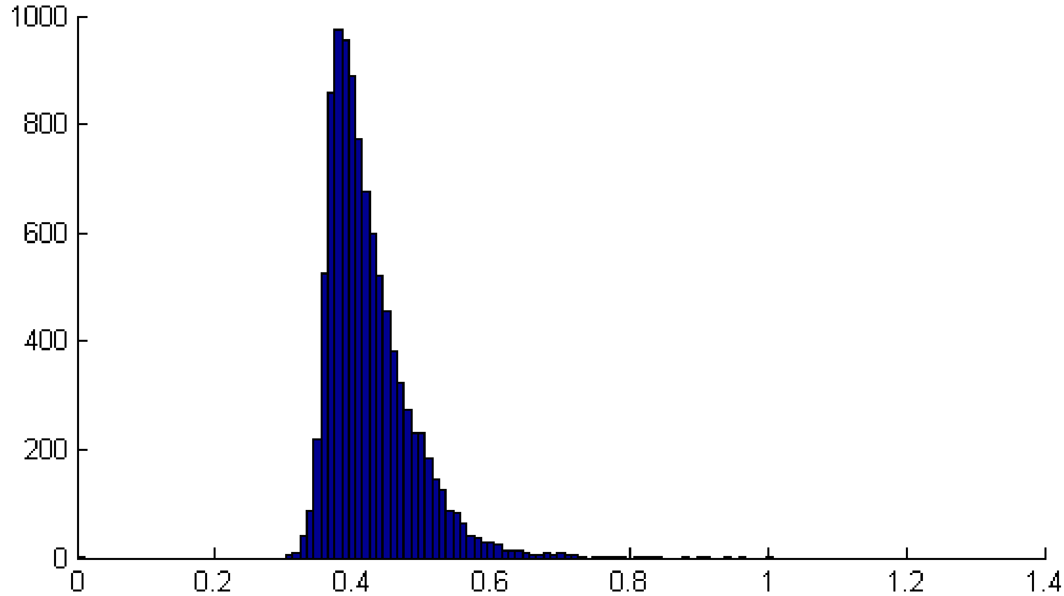
We now want to test if dataset d is consistent with a log linear NHPP. Again the test statistics Greenwood (G), Cramer-von Mises (W^2), Modified Kolmogorov-Smirnov (D), Modified Kuiper (V) and Modified Watson (U^2) are chosen.

The conditional distribution of the Greenwood test statistic is simulated by the LT (2006) method, while the distribution of the 4 other test statistics are simulated by the Gibbs Block algorithm. The resulting obtained p-values are given in table 8.4.

From the conditional p-values it is seen that all the test statistics imply evidence that dataset d is not consistent with log linear NHPP. It is worth to notice that the p-values for the statistics Greenwood (G) and Modified Kolmogorov Smirnov (D) has less significant p-values than the other statistics.



(a) Distribution of weights for $\pi(b) = \frac{1}{|b|}$ for 10 000 samples with restriction that only weights<10 are counted for.



(b) Distribution of weights by applying Jeffrey's prior for 10 000 samples

Figure 8.7

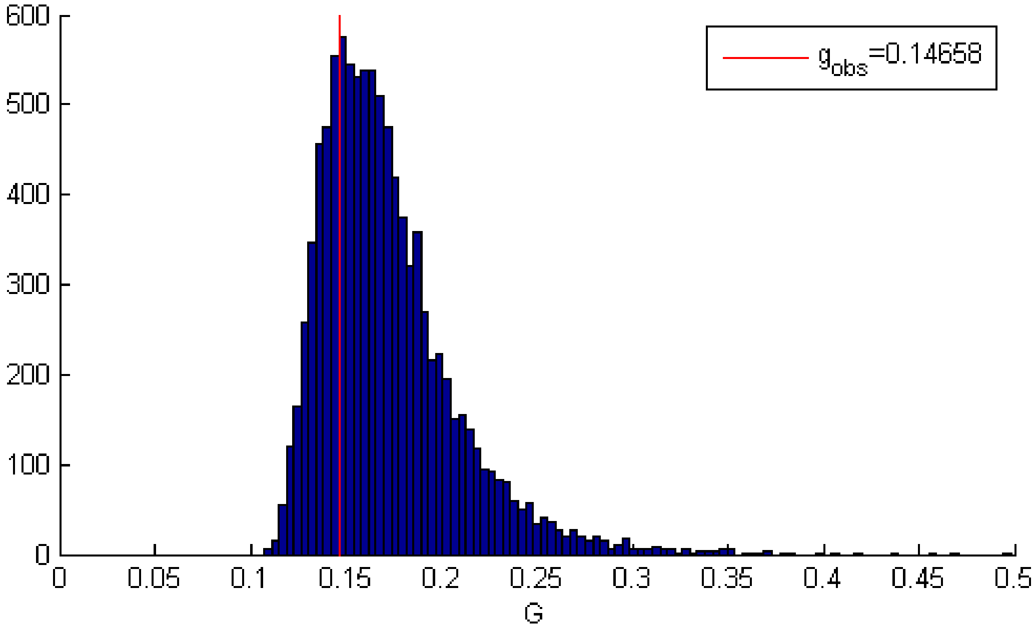


Figure 8.8: The simulated (LT 2006) conditional distribution of Greenwoods test statistics (G) under the nullhypothesis that dataset a comes from log linear NHPP, using Jeffreys prior.

If we compare these p-values compared to the p-values under the assumption that dataset d is from a power law NHPP, it is clear that the log linear assumption seems more reasonable. This agrees with results given in [2], where the USS Halfbeak data is considered by various plots and likelihood ratio tests based on asymptotic chi-square distributions, to determine whether the power law or log linear law fits the data better.

Table 8.4: Conditional p-values assuming dataset d is from log linear NHPP, Greenwood simulated by LT (2006), while the 4 others are simulated by Gibbs Block algorithm.

Test Statistic	W^2	D	V	U^2	G
p_{obs}	0.0087	0.0457	0.0034	0.0012	0.0404

8.2.3 Exact Confidence Intervals

When applying the LT (2006) method to simulate samples of T given $S = s$ we estimate the parameter $b(U, s)$ for each sample. Hence if order the m estimates $\hat{b}$ for b, $\tilde{b}_1 < ... < \tilde{b}_m$, then $(\tilde{b}_k, \tilde{b}_{m-k+1})$ is an exact 1-2k/(m+1) confidence interval for b [6].

Dataset a

Figure (8.9a) shows the distribution of the estimated values $\hat{b}$ for dataset a under the assumption of log linear intensity. The resulting 90% confidence interval for b is $[-5.58 \cdot 10^{-4},\ -3.99 \cdot 10^{-5}]$ which agrees with the interval obtained by [6]. This also indicates realiability growth for dataset a.

Dataset USS Halfbeak

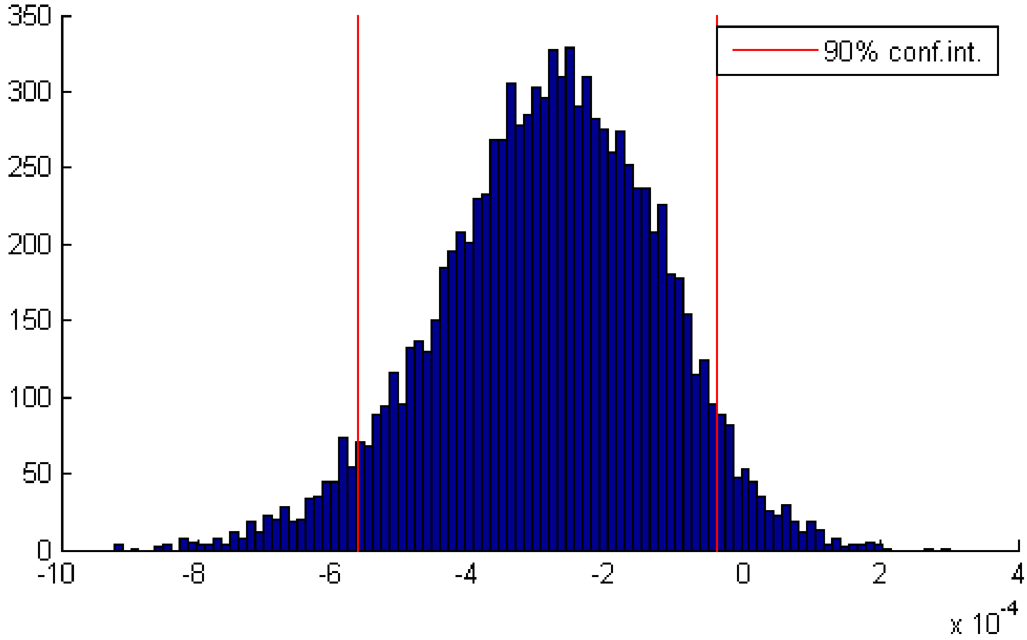
Figure (8.9b) shows the distribution of the estimated values $\hat{b}$ for dataset d, under the assumption of log linear NHPP. The resulting 90% confidence interval for b is $[1.121 \cdot 10^{-4},\ 1.832 \cdot 10^{-4}]$. This interval indicates reliability reduction.

8.2.4 Discussion of Results

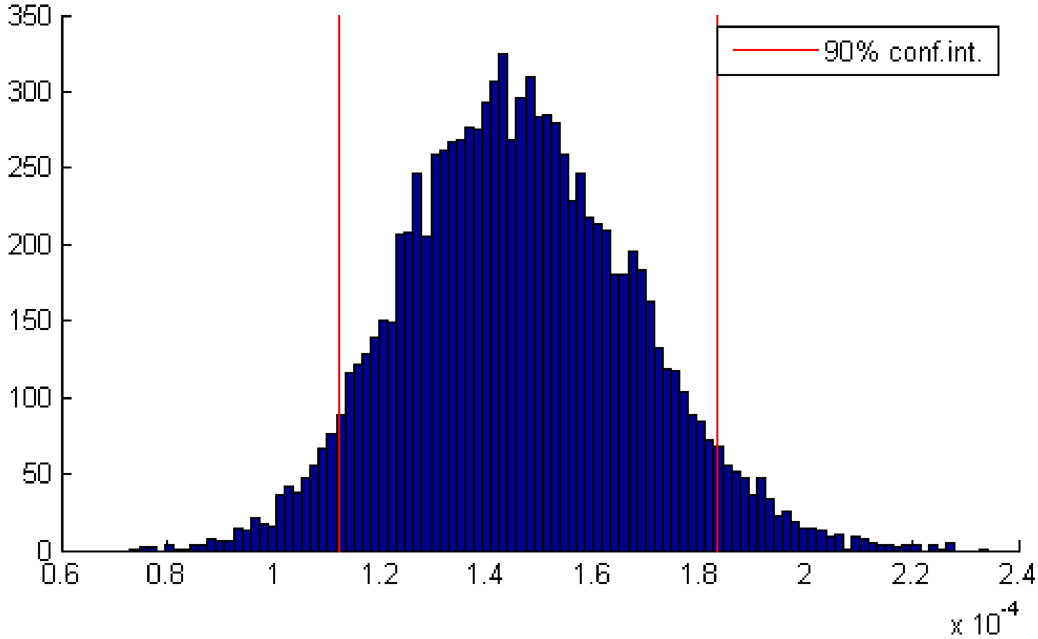
It is seen that the test statistics does not imply any evidence against the nullhypothesis that dataset a is consistent with log linear NHPP. For dataset d, all the test statistics imply evidence against the nullhypothesis that dataset d is consistent with log linear NHPP. But the p-values are less significant than for the assumption that dataset d is power law NHPP. This result is consistent with a discussion in [2].

When the p-values are obtained by simulating the distribution of the test statistics by the LT (2006) method, we need to choose the arbitrary function $\pi(b)$, such that the weigths for each sample could be calculated. We considered two choices for this function. It was demonstrated that Jeffrey's prior seemed to be the better choice of these two, due to more even distribution of the weights and less fluctuation in the calculated observed p-value for each simulation.

An exact conficence interval for the unknown parameter b was obtained for both datasets. These indicated reliability growth of dataset a, and reliability reduction for the USS Halfbeak dataset.



(a) Distribution of $\hat{b}$ under the nullhypothesis of log linear NHPP for dataset a, applying LT (2006).



(b) Distribution of $\hat{b}$ under the nullhypothesis of log linear NHPP for dataset dapplying LT (2006)

Figure 8.9

8.3 Power Comparison

In this section we present new results of a power comparison of 5 of the test statistics presented in the previous chapter. The statistics considered are Cramer-von Mises (W^2), Modified Kolmogorov-Smirnov (D), Modified Kuiper (V), Modified Watson (U^2) and Greenwood (G).

The nullhypothesis is that a set of observed failure times $T = (T_1, ..., T_n)$ are compatible with log linear NHPP. Now we simulate the observed failure times T from a power law NHPP, with intensity function $\lambda(t) = abt^{b-1}$ and check if the nullhypothesis is rejected, by applying the 5 different test statistics given above.

This could be done for different values of the parameters a and b in the power law model, and a comparison for various sample sizes are also possible. This study can provide as a guide in selecting test statistic and sample size.

For each set of simulated failure times $T = (T_1, ..., T_n)$ from the power law NHPP, 10 000 samples T_s conditioned on $S = (T_n, \sum_{j=1}^{n-1} T_j)$ are simulated, under the assumption that T comes from the log linear NHPP, applying the Gibbs Block algorithm given in chapter 6. These samples are used to determine the conditional observed p-value:

$$p_{obs} = P_{H_0}(W(T) \geq w_{obs} | S = s)$$

This is done for 1 000 different sets of simulated failure times T, and the frequencies of $p_{obs} \leq \alpha$=0.05 is counted. This process was repeated 3 times and the average values of the frequencies (power) were obtained. This is done for different values of the parameter b in the power law model, while the parameter a=1 is held constant.

In [10, 11] the "opposite" study is performed. The main difference is that the nullhypothesis is that the observed failure times are power law NHPP and the power comparison study is performed under the alternative hypothesis of log linear NHPP.

Table (8.5) shows the result of the study when the parameter b=0.1, and for different sample sizes n. It is seen that all the test statistics has high a high power for $n \geq 10$ (W^2),$n \geq 12$ (D, V, U^2) and $n \geq 15$ (G). The Cramer-von Mises statistic has the highest power for $n \geq 10$. All the test statistics have rather low power for sample size n=5, but the Greenwood test is seen to have the highest power for this sample size. For $n \geq 10$ the Greenwood test is seen have the lowest power.

Table (8.6) displays the results of the study when the parameter b=0.3. It is seen that the power for all the statistics are remarkebly lower in this situation. The Cramer-von Mises statistic is still seen to have the highest power for $n \geq 10$. But larger sample sizes are required in order for the tests to have an acceptable power, than was needed for b=0.1 above.

Table (8.7) gives the results when the parameter b=0.5. Again the power for all the test statistics is lower than for the two situations given above. The Cramer-von Mises, Modified Kolmogorov-Smirnov, Modified Kuiper and Modified Watson still has an acceptable power for sample size n=100, but for smaller sample sizes the power is low for all the statistics. The Greenwood statistic is again seen to have the lowest power.

Table 8.5: Power comparison simualating samples from power law NHPP with intensity function $\lambda(t) = abt^{b-1}$, b=0.1, a=1

n	W^2	D	V	U^2	G
5	0.38	0.24	0.10	0.14	0.50
10	0.84	0.77	0.65	0.65	0.67
12	0.91	0.85	0.80	0.81	0.70
15	0.97	0.94	0.91	0.90	0.79
20	1	0.99	0.98	0.99	0.90
30	1	1	1	1	0.98
50	1	1	1	1	1
100	1	1	1	1	1

Table 8.6: Power comparison simualating samples from power law NHPP with intensity function $\lambda(t) = abt^{b-1}$, b=0.3, a=1

n	W^2	D	V	U^2	G
5	0.17	0.13	0.08	0.09	0.20
10	0.54	0.42	0.28	0.26	0.26
12	0.56	0.53	0.35	0.38	0.30
15	0.75	0.62	0.48	0.50	0.32
20	0.86	0.77	0.64	0.67	0.39
30	0.95	0.94	0.85	0.89	0.47
50	0.99	0.98	0.98	0.99	0.64
100	1	1	1	1	0.85

Table 8.7: Power comparison simualating samples from power law NHPP with intensity function $\lambda(t) = abt^{b-1}$, b=0.5, a=1

n	W^2	D	V	U^2	G
5	0.08	0.07	0.05	0.04	0.08
10	0.18	0.15	0.09	0.09	0.11
12	0.20	0.18	0.12	0.10	0.11
15	0.25	0.21	0.16	0.15	0.12
20	0.33	0.29	0.22	0.23	0.13
30	0.47	0.43	0.33	0.30	0.15
50	0.66	0.63	0.50	0.54	0.15
100	0.89	0.88	0.82	0.87	0.20

Table (8.8) displays the results for b=1.5. It is seen that neither statistics have an acceptable power for any sample size n between 5 and 100, in this situation.

Table 8.8: Power comparison simualating samples from power law NHPP with intensity function $\lambda(t) = abt^{b-1}$, b=1.5, a=1

n	W^2	D	V	U^2	G
5	0.07	0.07	0.06	0.05	0.05
10	0.07	0.07	0.06	0.07	0.05
12	0.08	0.07	0.06	0.06	0.04
15	0.09	0.08	0.07	0.07	0.05
20	0.09	0.09	0.09	0.08	0.04
30	0.09	0.09	0.09	0.08	0.05
50	0.12	0.10	0.09	0.09	0.04
100	0.15	0.13	0.12	0.14	0.04

8.3.1 Discussion of results

From the results given above it is seen that the test statistics have an acceptable power for various parameter values b. Another, but expected characteristic, is that the power increases with larger sample size n.

For b=[0.1, 0.3] all the statistics have a high power of rejecting the simulated failure times T under the nullhypothesis of log linear NHPP, with Cramer-von Mises being the strongest and requiring the lowest sample size ($n \geq$10).

For b=[0.5, 1.5] neither of the tests have an acceptable power for $n \leq$50, and trying with values of b in the range of [2, 8] reveals that the power remains at this low level for all the statistics.

These results are in agreement with the ones obtained in [10, 11] which demonstrates that the power test has an acceptable power for b in the range of [-2.5, -2.0], in the log linear NHPP (remember that this is the "opposite" situation with regards to the hypothesis).

This is reasonable since that for some parameter values b the log linear intensity ($\lambda(t) = e^{a+bt}$) function can be well approximated by a power law intensity ($\lambda(t) = abt^{b-1}$) function [10].

8.4 Convergence Comparison

Suppose $T = (T_1, ..., T_n)$ is a set of observed failure times. Assuming that T is consistent with log linear NHPP, we have seen that this could be tested by choosing a test statistic

$W(T)$ and simulating the conditional p-value:

$$p_{obs} = P_{H_0}(W(T) \geq w_{obs}|S = s) \tag{8.1}$$

and reject the nullhypothesis if $p_{obs} \leq \alpha$ where α is a predetermined level of significance. In this paper we have presented two different approaches of how this conditional p-value could be be simulated. These are the LT (2006) method, which is the main focus in this paper in addition to Gibbs sampling which was introduced in chapter 6.

In this section we compare the two methods for the situation described above concerning how fast they converge to the "correct" conditional p-value.

We have applied two measures in order check the speed of convergence. These are

$$e = y_0 - \frac{1}{n}\sum_{j=1}^{m} p_j$$

and

$$\text{MSE} = \frac{1}{n}\sum_{j=1}^{m}(p_j - y_0)^2$$

where y_0 is the "correct" value of p_{obs}, and p_j is the conditional p-value obtained for simulation number j. In the following tests the number of simulations m=10.

In this paper we have tested the convergence applying dataset a, and the test statistic $W(T)$ is chosen to be Greenwoods test statistic (G), given in chapter 7. Remember that this is a two-sided test statistic, and the conditional p-value given in equation (8.1) is found by $p_{obs} = 2 \cdot \{\min(p_{obs}, 1 - p_{obs})\}$.

The "correct" conditional p-value is set to y_0=0.4673. This is found by applying the LT (2006) method. I simulated 2 million samples under the nullhypothesis that dataset a was log linear NHPP given the sufficient statistic $S = (T_n, \sum_{j=1}^{n-1} T_j)$ and calculated the conditional p-value p_{obs}. This procedure was repeated 3 times, and the "correct" conditional p-value y_0 was set to be the average of these 3 simulations.

When applying the LT (2006) method Jeffreys prior is chosen as the arbitrary function $\pi(b)$ in the weighting scheme. When simulating by Gibbs sampling this is done by Gibbs Block algorithm given in chapter 6, which is seen to be the fastest of the 3 Gibbs algorithms.

In the first test the two methods were compared with regards to how well they performed in 1 minute. It was seen that the LT (2006) method could simulate 15 000 samples for each of the m=10 conditional p-values, resulting in the corresponding measures $e = 10^{-3}$ and MSE=10^{-5}. The Gibbs Block algorithm could simulate 1 100 000 samples for each of the 10 conditional p-values resulting in the corresponding measures $e = 10^{-4}$ and MSE=10^{-6}. These results are given in table (8.9).

The second test was applied to see how long the two methods would need in order to obtain an accuracy of MSE=10^{-5} and MSE=10^{-6}. The LT (2006) method needed approximately 35 seconds to obtain accuracy of MSE=10^{-5}, and approximately 315 seconds to obtain accuracy of MSE=10^{-6}. The Gibbs Block algorithm needed approximately 3.5 seconds to obtain accuracy of MSE=10^{-5}, and approximately 36 seconds to obtain accuracy of MSE=10^{-6}. The results of this test is given in table (8.10).

Table 8.9: Performance of LT (2006) and Gibbs Block algorithm in 1 minute

Simulation Method	LT (2006)	Gibbs Block alg.
e	10^{-3}	10^{-4}
MSE	10^{-5}	10^{-6}

Table 8.10: Performance of LT (2006) and Gibbs Block algorithm by the measure MSE

Simulation Method	LT (2006)	Gibbs Block alg.
MSE= 10^{-5}	≈35 s.	≈3.5 s.
MSE= 10^{-6}	≈315 s.	≈36 s.

8.4.1 Discussion of results

It is seen that the Gibbs block algorithm converges faster to the "correct" conditional p-value, under the assumption that the observed failure times T are log linear NHPP.

In the LT (2006) method the parameter $\hat{b}$ needs to found by solving the equation (given in chapter 6):

$$g(b) = \sum_{j=1}^{n-1} \log(1 + U_j(e^{bT_n} - 1)) - bs_2 = 0$$

with respect to b, in order to simulate a sample T_s. This needs to be done by a numerical technique and for the results given above the solution was obtained by the repeated bisection method.

When applying the bisection method a tolerance criterion needs to be set. Concerning the results given above the tolerance was such that the bisection method was terminated when tol=10^{-5} or the number of bisections exceeded 50 if $b > 0$ and 500 if $b<0$. This parameter estimation is time consuming and most certain the main reason why the LT (2006) converges slower than the Gibbs Block algorithm where such estimation is not necessary. Hence, there could be a potential for improving the LT (2006) method by performing this parameter estimation by faster methods. Also the tolerance limit in the bisection method could be less strict, but this might detoriate the convergence of the method.

There is also a need to calculate the weights in the LT (2006) method which is also time consuming when performing a large number of simulations.

Notice :

There is not performed a comparison of LT (2006) and Gibbs Block algorithm under the nullhypothesis of power law NHPP. All the test statistics considered here were independent of the unknown parameters $\theta = (a, b)$ in this model and could be simulated by means of a random vector U ~uniform[0,1].

Also if one considers a general test statistic $W \equiv W(T)$ the LT (2006) method is not expected to be any slower with respect to the convergence than by Gibbs sampling. In order to simulate a sample T_s applying the LT (2006) method in this situation, the parameter $\hat{b}$ is found by a single calculation of the equation (given in chapter 6):

$$\hat{b} = \frac{\sum_{j=1}^{n-1} \log U_j}{s_2 - (n-1) \log T_n}$$

(The Gibbs sampling is different for the power law NHPP than log linear NHPP but this is not considered here.)

Chapter 9

Concluding Remarks

We have presented a general method for Monte Carlo computation of conditional expectations of the form $E[\phi(T)|S=s]$ given a sufficient statistic S. The method is referred to as LT (2006).

The method was adjusted to the parametrizations power law and log linear law of the NHPP model, such that exact statistical inference could be made in these models. This included goodness-of-fit testing in addition to simulating exact confidence intervals of unknown parameters in our model.

In the power law NHPP model a certain pivotal condition was satisfied and the conditional expectation given above could be found by direct sampling by a simple parameter adjustment of the original statistical model.

Other characteristics for the power law parametrization was that the expectation of the test statistics considered in chapter 7, were seen to be independent of the unknown parameters $\theta=(a,b)$ in this model and could be simulated unconditionally by a random vector U ~uniform[0,1].

In the log linear NHPP model the pivotal condition was not satisfied, and in order to obtain the correct conditional expectation of the form given above a weighted sampling scheme was required. The convergence of the method was seen to be dependent on the choice of an arbitrarily chosen function $\pi(b)$ in the weighting scheme. Two different choices for this function was considered. Although Jeffrey's prior seemed to be the better choice of these two, the function suggested in [8] could very well be applied if one imposed a restriction on the weights.

There is a large potential for studying different test statistics and their applicability, for various situations. In [1] the choice of a suitable test statistic for a set of observed failure times T is considered, and in particular under the assumption of power law NHPP. It is argued that engineering experience should provide as a basis for choosing the "correct" statistic, and a single statistic can not detect all deviations from the NHPP model by itself. Constructing ”purpose built” statistics would be the best option in certain situations, which requires a good theoretical and practical understanding.

A power comparison was presented for 5 of the test statistics considered, under the nullhypothesis that a set of observed failure times $T = (T_1, ..., T_n)$ comes from a log linear NHPP and the alternative hypothesis of power law NHPP. This is a new result which extends the work done in [10, 11]. For some parameter values the power was high for all the statistics, with Cramer-von Mises being the strongest for sample sizes $n \geq 10$. The power increased as a function of the sample size n. However the most interesting results are for small sample sizes since the method of exact testing could provide accurate results when the use of existing asymptotic methods could not be justified.

A speed of convergence comparison of the method LT (2006) and the alternative approach of Gibbs sampling is given. This is performed under the assumption that the failure times T are consistent with log linear NHPP, and it was seen that the Gibbs sampling converged faster to the ”correct” conditional p-value. In the LT (2006) method a parameter estimation of the unknown parameter b is required for each simulated sample T_s, which is the main reason that Gibbs sampling is seen to be the fastest. This estimation needs to be done by a numerical method, and we have applied the repeated bisection. Hence there could be a potential for improving the speed of this estimation by different numerical techniques.

Considering further work it could be of interest to adjust the LT (2006) method to more general processes such as the modulated power law with intensity function given by $\lambda(t) = abt^{b-1}e^{\beta t}$ [2].

In addition the method LT (2006) could be adjusted to other than the NHPP model which is considered here.

Bibliography

[1] R.D Baker. New Tests of the Power Law Process. *Teknometrics*, 38, 1996.

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

[3] The University of Arizona Department of Mathematics. Sufficient Statistics. http://math.arizona.edu/ tgk/466/sufficient.pdf, 2007.

[4] Steinar Engen and Magnar Lillegård. Stochastic Simulations Conditioned on Sufficient Statistics. *Biometrika*, 84, 1997.

[5] Martin Emil Iden. Bruk av Betinget Parametrisk Bootstrapping i Analyse av Pålitelighetsdata. Master's thesis, NTNU, 1998.

[6] Bo Henry Lindqvist. Exact Statistical Inference for Nonhomogeneous Poisson Processes. Technical report, Department of Mathematical Sciences, NTNU, 2000.

[7] Bo Henry Lindqvist and Gunnar Taraldsen. Monte Carlo Conditioning on a Sufficient Statistic. *Biometrika*, 92, 2005.

[8] Bo Henry Lindqvist and Gunnar Taraldsen. Conditional Monte Carlo based on Sufficient Statistics with Applications. Festschrift Doksum, 2006.

[9] Bo Henry Lindqvist, Gunnar Taraldsen, Steinar Engen, and Magnar Lillegård. A Counterexample to a Claim about Stochastic Simulations. *Biometrika*, 90, 2003.

[10] Won J. Park and Yoon G. Kim. Goodness-of-Fit Tests for the Power-Law Process. *IEEE Transactions On Reliability*, 41, 1992.

[11] Won J. Park and Munsup Seoh. More goodness-of-Fit Tests for the Power-Law Process. *IEEE Transactions On Reliability*, 43, 1994.

[12] Marvin Rausand and Arnljot Høyland. *System Reliability Theory*. Wiley, 2 edition, 2004.

[13] Sheldon M. Ross. *Introduction to Probability Models*. Academic Press, 8 edition, 2003.

[14] Walpole, Myers, Myers, and Ye. *Probability & Statistics*. Prentice Hall, 7 edition, 2002.

Appendix A

Datasets

I have worked with two different datasets in the implementation. Dataset a is from [6] with n=10 observed failure times. The USS Halfbeak (also referred to as dataset d) is from [2], with n=71 observed failure times. Both datasets are failure censored.

Dataset a [6]

[103 315 801 1183 1345 2957 3909 5702 7261 8245]

USS Halfbeak Data [2]

[1382 2990 4124 6827 7472 7567 8845 9450 9794 10848 11993 12300 15413 16497 17352 17632 18122 19067 19172 19229 19360 19686 19940 19944 20121 20132 20431 20525 21057 21061 21309 21310 21378 21391 21456 21461 21603 21658 21688 21750 21815 21820 21822 21888 21930 21943 21946 22181 22311 22634 22635 22669 22691 22846 22947 23149 23305 23491 23526 23774 23791 23882 24006 24286 25000 25010 25048 25268 25400 25500 25518]