NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

A Simulation Study of Statistical Inference in Non-Homogeneous Poisson Processes with Emphasis on Frailty and Dynamic Behavior

by

Zeytu Gashaw Asfaw

PREPRINT
STATISTICS NO. 6/2013

NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
http://www.math.ntnu.no/preprint/statistics/2013/S6-2013.pdf
Zeytu Gashaw Asfaw has homepage: http://www.math.ntnu.no/∼zeytugas
E-mail: zeytugas@math.ntnu.no
Address: Department of Mathematical Sciences, NTNU, N-7491 Trondheim,
Norway.

A Simulation Study of Statistical Inference in Non-homogeneous Poisson Processes with Emphasis on Frailty and Dynamic Behavior

Zeytu Gashaw Asfaw
Department of Mathematical Sciences
Norwegian University of Science and Technology(NTNU)
P.O.Box 7491,Trondheim,Norway
zeytugas@math.ntnu.no

September 17, 2013

Abstract

A study of recurrent events for repairable systems is presented. The basic model is the nonhomogeneous Poisson process with power law intensity function. When several similar systems are under observation, the assumption that the corresponding processes are independent and identically distributed is often questionable. In practice there may be an unobserved heterogeneity among the systems. We consider two seemingly different approaches for analysis of such differences, namely by using frailities and by using dynamic models. The relation between the two approaches is investigated, both theoretically and in a simulation study. Detailed derivations of likelihood functions are provided, and maximum likelihood is used as the inference tool throughout the paper. A possible conclusion is that the two approaches are very similar, so that frailty models may be viewed as an alternative to dynamic models.

1 Notation

t	Failure time
S	Starting time
T	Ending time
τ_j	Ending time of observation for system j
n_j	Total number of failure per system
n	Total number of failure
m	Total number of system
$w(t)$	Failure rate (ROCOF)
$W(t)$	Cumulative failure rate (CROCOF)
$N(t)$	Number of failure in (0, t)
$E[N(t)]$	Expected number of failures in (0, t)
$Var[N(t)]$	Variance of number of failures in (0, t)
λ	Parameter of Power law model
β	Parameter of Power law model
δ	Parameter of Fraility model
γ	Parameter of LEYP model
$z_1(t)$	Hazard rate of T_1
$z_2(t)$	Hazard rate of T_2
$z_3(t)$	Hazard rate of T_3
$G_1(t)$	Hazard rate of T_1
$G_2(t)$	Hazard rate of T_2
$G_3(t)$	Hazard rate of T_3

2 Introduction

Survival analysis involves the modeling of time to event data. Classically, death or failure are considered as "events" in the survival literature, considering only single events, after which the individual or machine is dead or broken. More recently, many concepts in survival analysis have been modelled by counting process theory, which adds flexibility in that it allows modeling, for example recurrent events.

In the reliability literature, systems are generally classified into non-repairable and repairable. Non-repairable systems are those that do not get repaired when they fail. Thus, non-repairable system can fail only once, and a lifetime model such as the Weibull distribution provides the distribution of the time at which such a system fails. Most household products can be good examples of non-repairable systems.

On the other hand, repairable systems are those systems (machines, industrial plants, software, etc.) which, in the event of a failure, can be restored to satisfactory operation by any action, including parts replacements or changes to adjustable settings. A repairable system is often modeled by means of a counting process. But, to what extent can the system perform after being returned back to its regular operation? We may have that the system's performance is in the same state that the system had at the start of the operation, which means a renewal process or as good as new condition. Or, its performance may be in the same state as before the failure, which leads to a non-homogeneous Poisson process (NHPP), i.e. as bad as old condition.

NHPPs which is the main concern of this paper are useful due to their flexible assumption that events are occurring randomly in time at varying rates, instead of events being just as likely to occur in all intervals of equal size, which is the property of homogeneous Poisson processes (HPP).

There are three primary approaches to evaluating multivariate survival processes: Marginal models, Frailty models and Dynamic models (see Aalen et. al, 2008). We are focused on the last two due to the fact that marginal models, unlike fraility and dynamic models, focus on parts of the avaliable data instead of giving more realistic descriptions of the full data sets. For each of these model types we consider parametric modelling and inference. Although there is a fairly rich literature on the corresponding models and methods, all their features and particularly their interrelations are yet not fully understood nor fully investigated. The objective of the study is to perform such a study, by considering comparable models both theoretically and in a simulation study.

2.1 Definition and Properties of NHPPs

A counting process is a non-homogeneous (or non-stationary) Poisson process with rate function $w(t)$ for $t \geq 0$, if :

1. N(0)=0.
2. $Pr\{N(t+\Delta t)-N(t)=1\}=w(t)\Delta t+o(\Delta t)$, for all t where $\frac{O(\Delta t)}{\Delta t}\to 0$ as $\Delta t\to 0$ and $N(t)$ is the number of events occurring within $(0,t]$
3. $Pr\{N(t+\Delta t)-N(t)\geq 2\}=o(\Delta t)$, which means that the system will not experience more than one failure at the same time.

The NHPP is fully characterized by ROCOF(rate of occurrence of failure) and usually denoted by w(t). This function is also called the peril rate of the NHPP.

Its cumulative rate of the process is

$$W(t)=\int_0^t w(s)ds$$

(later called the CROCOF)
Then, the probability of seeing n events in the interval $(0,t]$ is

$$Pr[N(t)=n]=\frac{[W(t)]^n}{n!}e^{-W(t)}$$

for $n=0,1,2,...$

The mean number of failures in $(0,t]$ is therefore

$$E[N(t)]=W(t)$$

and its variance is

$$Var[N(t)]=W(t)$$

Likewise, the probability of seeing n events in the interval $(t,t+s]$ is

$$Pr[(N(t+s)-N(t))=n]=e^{-[W(t+s)-W(t)]}\frac{[W(t+s)-W(t)]^n}{n!}$$

What is said above is that $N(t+s)-N(t)$ is Poisson distributed with expected value $\int_t^{t+s}w(s)ds$ where $w(s)$ is the time dependent intensity function.

Another probabilistic property of NHPP which can help us to simulate the events of NHPP from that of HPP is stated as follows. If $t_1, t_2, ...$ are event times in a unit HPP, then $W^{-1}(t_1), W^{-1}(t_2), ...$ are event times in an NHPP with cumulative intensity function $W(t)$. Let us use the CROCOF of power law model to show how NHPP events are simulated from HPP. The CROCOF of power law is,

$$W(t) = \lambda t^{\beta}$$

Then, equate W(t) to the exponentially distributed random number having parameter one, $u \sim exp(1)$.

$$u = \lambda t^{\beta}$$
$$\Rightarrow t = [\frac{u}{\lambda}]^{\frac{1}{\beta}}$$

Graphically,

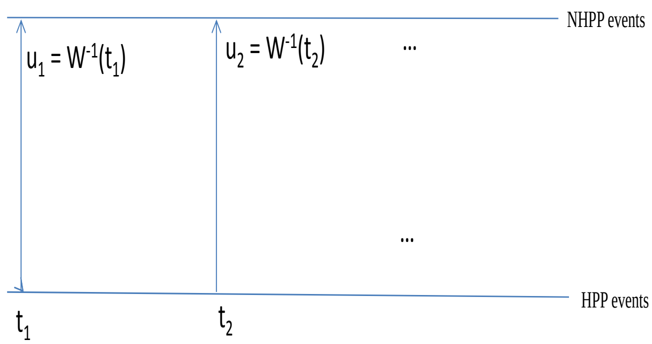

Figure 1: Simulation of NHPP events from HPP

The basic difference of NHPP from HPP is that the rate of occurrence of failures varies with time rather than being a constant. This implies that for an NHPP model the inter occurrence times are neither independent nor identically distributed. In line with this, frequently NHPP is used to model repairable systems that are subject to a minimal repair strategy, with negligible repair times. Minimal repair means that a failed system is restored just back to functioning state and the system continues as if nothing had happened.This implies that the likelihood of system failure is the same immediately before and after a failure.

2.2 Parametric Models of NHPPs

Several parametric models have been established to portray the ROCOF of an NHPP, but here we are concerned to the most celebrated process model which is the power law model. This model is favored for several reasons.The first reason for the popularity of this model is that it has a very practical foundation in terms of minimal repair. The second reason here is that if the time to first failure follows the Weibull distribution, then each succeeding failure is governed by the Power law model in the case of minimal repair. From the aforementioned discussion we can say the Power law model is an extension of the Weibull distribution.

In the power law model the ROCOF of the NHPP is defined as

$$w(t) = \lambda\beta t^{\beta-1} \quad \text{for} \;\; \lambda \geq 0, \;\; \beta \geq 0 \text{ and } \;\; t \geq 0$$

Its cumulative rate of ocurrence of failure (CROCOF) is

$$\begin{aligned} W(t) &= \int_0^t w(s)ds \\ &= \lambda t^{\beta} \end{aligned}$$

This intensity function was introduced in Crow (1972) as a stochastic model for the Duane reliability growth postulate. Moreover, it is referred to as a Weibull Poisson Process or the Power law Poisson Process.

The parameter β in the Power law model can give information about the system as follows:

If $0 < \beta < 1$, then the system is improving(happy).
If $\beta > 1$, then the system is deteriorating(sad).
If $\beta = 1$, then the model is reduces to an HPP.
The case $\beta = 2$ is seen to give a linearly increasing ROCOF.

2.3 Maximum Likelihood Estimation of Power law model

Suppose that the number of systems under study is m and the j^{th} system is observed continuously from time S_j to time T_j, $j = 1, 2, 3, ..., m$. During the period $[S_j, T_j]$, let n_j be the number of failures experienced by the j^{th} system and let $t_{i,j}$ be the age of this system at the i^{th} occurrence of failure, $i = 1, 2, ..., n_j$.

It is also possible that the system boundaries S_j and T_j may be observed failure times for the j^{th} system. If $t_{n_j j} = T_j$, then the data on the j^{th} system are said to be failure terminated and T_j is a random variable with n_j fixed. If $t_{n_j j} < T_j$, then the data on the j^{th} system are said to be time terminated with n_j a random variable. Suppose that data are available from m independent systems with the same intensity function $w(t)$ and system j is observed in the interval $[S_j, T_j]$, $j = 1, 2, ..., m$, and the system j recurrence times are denoted by $t_{1j}, t_{2j}, ..., t_{n_j j}$.

Then, the NHPP likelihood function is simply the product of the individual system likelihoods

$$L = \prod_{j=1}^{m} \left[\prod_{i=1}^{n_j} [w(t_{ij})]\, exp\left\{-[W(T_j) - W(S_j)]\right\}\right]$$

Due to the monotonicity characteristics of log transformation and for theoretical as well as technical reasons it is well recommended to work with the logarithm of the likelihood function or with the negative logarithm of it. Although the shape of these(likelihood and log-likelihood) functions are different, they have their maximum point at the same value.

Hence,

The log-likelihood function of NHPP is

$$\begin{aligned} l &= \log(L) \\ &= \sum_{j=1}^{m} \left[\sum_{i=1}^{n_j} [\log w(t_{ij})] - [W(T_j) - W(S_j)]\right]. \end{aligned}$$

The log-likelihood function of power law model with intensity function $w(t) = \lambda\beta t^{\beta-1}$ is,

$$
\begin{aligned}
l &= \log(L) \\
&= \sum_{j=1}^{m}\left[\sum_{i=1}^{n_j}\left[\log(\lambda\beta t_{ij}^{\beta-1})\right] - \left[\lambda T_j^{\beta} - \lambda S_j^{\beta}\right]\right] \\
&= \sum_{j=1}^{m}\left[\sum_{i=1}^{n_j}[\log\lambda + \log\beta + (\beta-1)\log t_{ij}] - \left[\lambda T_j^{\beta} - \lambda S_j^{\beta}\right]\right] \\
&= \sum_{j=1}^{m}\left[n_j\log\lambda + n_j\log\beta + (\beta-1)\sum_{i=1}^{n_j}\log t_{ij} - \lambda\left[T_j^{\beta} - S_j^{\beta}\right]\right] \\
&= \sum_{j=1}^{m} n_j\log\lambda + \sum_{j=1}^{m} n_j\log\beta + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} - \lambda\sum_{j=1}^{m}\left[T_j^{\beta} - S_j^{\beta}\right] \\
&= n\log\lambda + n\log\beta + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} - \lambda\sum_{j=1}^{m}\left[T_j^{\beta} - S_j^{\beta}\right] \\
&= \sum_{j=1}^{m} n_j\log\lambda + \sum_{j=1}^{m} n_j\log\beta + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} - \lambda\sum_{j=1}^{m}\left[T_j^{\beta} - S_j^{\beta}\right] \\
&= n\log\lambda + n\log\beta + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} - \lambda m T^{\beta}
\end{aligned}
$$

where $n = \sum_{j=1}^{m} n_j$

In the last line above we set $S_j = 0$ i.e. all systems have the same intial point which is zero and all $T_j = T$, where T is a constant number. The standard, analytical method of finding the MLEs is to take the first partial derivatives of the likelihood/log-likelihood function with respect to each parameter in the model and equate to zero.

Hence,

$$
\begin{aligned}
\frac{\partial l}{\partial\lambda} &= \frac{n}{\lambda} - mT^{\beta} \\
\hat{\lambda} &= \frac{n}{m[T^{\hat{\beta}}]}
\end{aligned}
$$

Similarly,

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} + \sum_{j=1}^{m} \sum_{i=1}^{n_j} \log t_{ij} - \lambda m T^{\beta} [\log(T)]$$

Setting this equal to zero and using the above $\hat{\lambda}$ we get

$$\hat{\beta} = \frac{n}{n \log(T) - \sum_{j=1}^{m} \sum_{i=1}^{n_j} \log t_{ij}}$$

This gives an explicit solution for $\hat{\beta}$ which can afterwards be substituted in the expression for $\hat{\lambda}$.

We might consider the Fisher information matrix for the computation of variances and covariances of the MLEs. Fisher information matrix is used to measure the amount of information that the observed data carries about the unknown parameters. The log-likelihood function is twice differentiable with respect to each parameter,and

$$\frac{\partial^2 l(\lambda, \beta)}{\partial \lambda^2} = \frac{-n}{\lambda^2}$$

Similarly for β parameter

$$\frac{\partial^2 l(\lambda, \beta)}{\partial \beta^2} = \frac{-n}{\beta^2} - \lambda m T^{\beta} (log(T))$$

Second mixed-partial derivatives of the log-likelihood,

$$\frac{\partial^2 l(\lambda, \beta)}{\partial \lambda \partial \beta} = -m T^{\beta} (log(T))$$

Thus, the Fisher information matrix is

$$I\left(\lambda, \beta\right) = \begin{bmatrix} \frac{n}{\lambda^2} & m T^{\beta} (log(T)) \\ \\ m T^{\beta} (log(T)) & \frac{n}{\beta^2} + \lambda m T^{\beta} (log(T)) \end{bmatrix}$$

For large sample size, maximum likelihood estimate have an approximate normal distribution centered on the true parameter and the variance, which is given by Fisher information matrix after substituting the maximum likelihood estimates for λ and β. Thus, asymptotically, maximum likelihood estimator is normally distributed.

Once $\hat{\lambda}$ and $\hat{\beta}$ have been estimated, the maximum likelihood estimate of the intensity function is given by:

$$w(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}, t > 0$$

and then we can draw failure intensity versus time.

3 Frailties in NHPP

3.1 Definition and Parametric Model

The notion of fraility provides a convenient way to introduce random effects, association and unobserved heterogeneity into models for survival variables. It may be considered as unmeasured risk factors, where the relevant covariates are not included in the model's specification and unknown to exist. This may otherwise be a problem in having inconsistent parameter estimates and wrong standard estimate values.

The term fraility itself was introduced by Vaupel et al. (1979) for univariate survival models, but was substantially promoted by applications to multivariate survival data from around 1980. Fraility models extend popular models such as the Cox model. Normally, survival analysis implicitly assumes a homogeneous population to be studied. In many applications, however, the study population can not be assumed to be homogeneous but must be considered as a heterogeneous sample.

Here we consider parametric models for NHPP, with a serious consideration of frailties (hidden heterogeneity) among systems. This is done in accordance with the definition of frailties in connection with the power law model. Recall that the CROCOF of power law model is

$$W(t) = \lambda t^{\beta} \text{ where } \lambda > 0,\ \beta > 0 \text{ and } t \geq 0$$

With a consideration of frailties this model can be written as $W(t) = a\lambda t^{\beta}$ where $\lambda > 0, \beta > 0, t \geq 0$ and a is a gamma distributed random number with mean 1 and variance δ. The idea is then that in the case of m systems, each system has its own value of a, i.e. $a_1, a_2, ..., a_m$, which are assumed to be independent and identically distributed with the distribution just given.

Although we have several potential frailty models to choose for the above "a's" we choose gamma frailties deliberately due to the following reason:There is no physical justification to prefer gamma frailties instead of the other but only in the line of computational and analytical aspect we prefer it. From a computational and analytical perspective, it fits very well to failure data because it is easy to derive the closed form expressions of unconditional survival, cumulative density and hazard function. This is due to the simplicity of the Laplace transform. The density of the two-parameter gamma distribution is given as

$$h_a(a) = \frac{a^{k-1}e^{-\frac{a}{\theta}}}{\theta^k \Gamma(k)}$$

where $a \geq 0$, k is shape parameter and θ is scale parameter.
Moreover, E(a)=kθ and Var(a)=k θ^2 But we want to have E(a)=1 and Var(a)=δ. Thus we have k=$\frac{1}{\delta}$ and θ=δ and density

$$h_a(a) = \frac{a^{\frac{1}{\delta}-1} exp^{\frac{-a}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}}$$

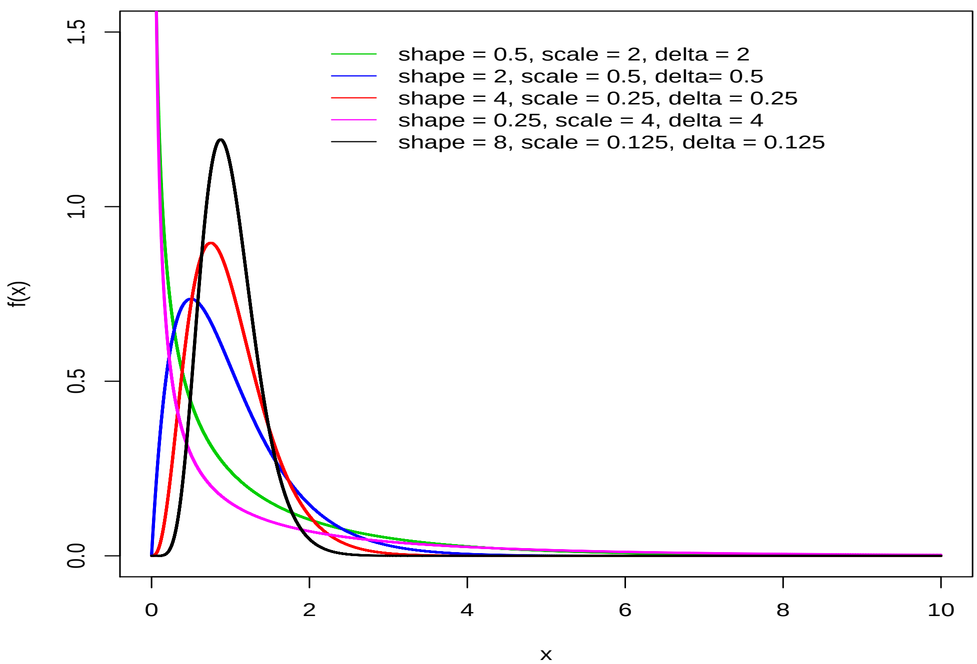

Figure 2: Graph of Gamma Densities with expected value 1

3.2 Maximum Likelihood Estimation of Power law model with gamma distributed frailty

We considered the likelihood function for m systems without consideration of frailty but in this subsection we are eager to see the change in parameter estimation with a consideration of frailty a. We use similar argument as before, but now with "δ" as an additional parameter and CROCOF should be multiplied with the frailty.

Individual system likelihood function is:

$$L_j(a_j) = \prod_{i=1}^{n_j} a_j w(t_{ij}) exp\left[-a_j\left[W(T_j) - W(S_j)\right]\right]$$

Since a_j is a random variable we should find the expected value of $L_j(a_j)$ with respect to the distribution of a_j. In our case the distribution of a_j is gamma with expected value 1 and its probability density function is

$$h(a_j) = \frac{a_j^{\frac{1}{\delta}-1} e^{\frac{-a_j}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}}$$

The expected value of $L_j(a_j)$ is

$$
\begin{aligned}
L_j &= E\left[L_j(a_j)\right] \\
&= \int L_j(a_j)h(a_j)da_j \\
&= \int \prod_{i=1}^{n_j} a_j w(t_{ij}) exp\left[-a_j\left[W(T_j) - W(S_j)\right]\right] h(a_j)da_j \\
&= \int \prod_{i=1}^{n_j} a_j w(t_{ij}) exp\left[-a_j\left[W(T_j) - W(S_j)\right]\right] \frac{a_j^{\frac{1}{\delta}-1} e^{\frac{-a_j}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} da_j \\
&= \int \prod_{i=1}^{n_j} a_j \left[\lambda\beta t_{ij}^{\beta-1}\right] exp\left[-a_j\left[\lambda T_j^{\beta} - \lambda S_j^{\beta})\right]\right] \frac{a_j^{\frac{1}{\delta}-1} e^{\frac{-a_j}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} da_j \\
&= \int_0^{\infty} a_j^{n_j}\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1} exp\left[-a_j\left[\lambda T_j^{\beta} - \lambda S_j^{\beta})\right]\right] \frac{a_j^{\frac{1}{\delta}-1} e^{\frac{-a_j}{\delta}}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}} da_j \\
&= \left[\frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}}\right] \int_0^{\infty} a_j^{n_j+\frac{1}{\delta}-1} exp\left[-a_j\left[\lambda T_j^{\beta} - \lambda S_j^{\beta} + \frac{1}{\delta}\right]\right] da_j
\end{aligned}
$$

Let $r_j = n_j + \frac{1}{\delta} - 1$ and $s_j = \lambda T_j^{\beta} - \lambda S_j^{\beta} + \frac{1}{\delta}$

Then the above equation can be written as:

$$
L_j = \left[\frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}}\right] \int_0^{\infty} a_j^{r_j} e^{-a_j s_j} da_j
$$

To have the integrand expression of the above integration let us substitute $v = a_j s_j$. After a certain mathematical operation we have the following expression:

$$
L_j = \left[\frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}}\right] \left[\frac{1}{s_j^{r_j+1}} \int_0^{\infty} v^{r_j} e^{-v} dv\right]
$$

But, $\frac{1}{s_j^{r_j+1}}\int_0^\infty v^{r_j}e^{-v}dv$ equals $\frac{1}{s_j^{r_j+1}}\Gamma(r_j+1)$ by using gamma function

Hence,

$$L_j = \frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}\Gamma(n_j+\frac{1}{\delta})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}\left[\lambda T_j^{\beta}-\lambda S_j^{\beta}+\frac{1}{\delta}\right]^{n_j+\frac{1}{\delta}}}$$

Although power law model can have a potential to model systems that start from any time t, we restrict to time zero as starting operation time of all systems i.e $S_j = 0\ \forall j = 1, 2, ..., m$ in this study. Then, the aforementioned individual likelihood function simplifies to

$$L_j = \frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}\Gamma(n_j+\frac{1}{\delta})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}\left[\lambda T_j^{\beta}+\frac{1}{\delta}\right]^{n_j+\frac{1}{\delta}}}$$

Thus, the total likelihood function is

$$L = \prod_{j=1}^{m} L_j$$

The log likelihood function is

$$
\begin{aligned}
l(\lambda,\beta,\delta) &= \log L \\
&= \log\left[\prod_{j=1}^{m} L_j\right] \\
&= \log\left[\prod_{j=1}^{m} \frac{\lambda^{n_j}\beta^{n_j}\left(\prod_{i=1}^{n_j} t_{ij}\right)^{\beta-1}\Gamma(n_j+\frac{1}{\delta})}{\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}\left[\lambda T_j^{\beta}+\frac{1}{\delta}\right]^{n_j+\frac{1}{\delta}}}\right] \\
&= \sum_{j=1}^{m}\left\{\log\left[\lambda^{n_j}\beta^{n_j}\left(\sum_{i=1}^{n_j} t_{ij}\right)^{\beta-1}\Gamma(n_j+\frac{1}{\delta})\right] - \log\left[\Gamma(\frac{1}{\delta})\delta^{\frac{1}{\delta}}\left[\lambda T_j^{\beta}+\frac{1}{\delta}\right]^{n_j+\frac{1}{\delta}}\right]\right\} \\
&= \sum_{j=1}^{m}\left\{n_j\log\lambda + n_j\log\beta + (\beta-1)\log\left(\sum_{i=1}^{n_j} t_{ij}\right) + \log\Gamma(n_j+\frac{1}{\delta})\right\} \\
&- \sum_{j=1}^{m}\left\{\log\Gamma(\frac{1}{\delta}) + (\frac{1}{\delta})\log\delta + \left[n_j+\frac{1}{\delta}\right]\log\left[\lambda T_j^{\beta}+\frac{1}{\delta}\right]\right\} \\
&= n\log\lambda + n\log\beta + (\beta-1)\sum_{j=1}^{m}\log\left(\sum_{i=1}^{n_j} t_{ij}\right) + \sum_{j=1}^{m}\log\Gamma(n_j+\frac{1}{\delta}) \\
&- \left[m\log\Gamma(\frac{1}{\delta}) + m\frac{1}{\delta}\log\delta + \sum_{j=1}^{m}\left[\left[n_j+\frac{1}{\delta}\right]\log\left[\lambda T_j^{\beta}+\frac{1}{\delta}\right]\right]\right]
\end{aligned}
$$

Hereafter, let all T_j=T and T=τ.

$$
\begin{aligned}
&= n\log\lambda + n\log\beta + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} + \sum_{j=1}^{m}\log\Gamma(n_j+\frac{1}{\delta}) \\
&- \left[m\log\Gamma(\frac{1}{\delta}) + m\frac{1}{\delta}\log\delta + \sum_{j=1}^{m}\left[\left[n_j+\frac{1}{\delta}\right]\log\left[\lambda T^{\beta}+\frac{1}{\delta}\right]\right]\right] \\
&= n\log\lambda + n\log\beta + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} + \sum_{j=1}^{m}\log\Gamma(n_j+\frac{1}{\delta}) \\
&- \left[m\log\Gamma(\frac{1}{\delta}) + m\frac{1}{\delta}\log\delta + \left[n+\frac{m}{\delta}\right]\log[\lambda T^{\beta}+\frac{1}{\delta}]\right]
\end{aligned}
$$

Partial derivative of $l\left(\lambda,\beta,\delta\right)$ with respect to λ is

$$\frac{\partial l\left(\lambda,\beta,\delta\right)}{\partial\lambda}=\frac{n}{\lambda}-\left[\frac{\tau^{\beta}}{\lambda\tau^{\beta}+\frac{1}{\delta}}\right]\left[n+\frac{m}{\delta}\right]$$

Then,

$$\frac{\partial l\left(\lambda,\beta,\delta\right)}{\partial\lambda}=0$$

$$\Rightarrow\frac{n}{\lambda}=\left[\frac{\tau^{\beta}}{\lambda\tau^{\beta}+\frac{1}{\delta}}\right]\left[n+\frac{m}{\delta}\right]$$

$$\Rightarrow\hat{\lambda}=\frac{n}{m\tau^{\hat{\beta}}}$$

Similarly, Partial derivative of $l\left(\lambda,\beta,\delta\right)$ with respect to β is

$$\frac{\partial l\left(\lambda,\beta,\delta\right)}{\partial\beta}=\frac{n}{\beta}+\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij}-\left[\frac{\lambda\tau^{\beta}\log(\tau))}{\lambda\tau^{\beta}+\frac{1}{\delta}}\right]\left[n+\frac{m}{\delta}\right]$$

Hence,

$$\frac{\partial l\left(\lambda,\beta,\delta\right)}{\partial\beta}=0$$

$$\Rightarrow\frac{n}{\beta}+\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij}-n\log\left(\tau\right)$$

$$\Rightarrow\hat{\beta}=\frac{n}{n\log(\tau)-\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij}}$$

Thus, $\hat{\lambda}$ and $\hat{\beta}$ are exactly the same as for the power law without frailities.

Likewise, using the digamma function ψ defined by

$$\psi(x)=\frac{d}{dx}\log\Gamma(x)=\frac{\Gamma^{'}(x)}{\Gamma(x)}$$

, the partial derivative of $l\left(\lambda,\beta,\delta\right)$ with respect to δ is

$$\begin{aligned}
\frac{\partial l\left(\lambda,\beta,\delta\right)}{\partial\delta} &= -\frac{1}{\delta^2}\sum_{j=1}^{m}\psi(n_j+\frac{1}{\delta})+\frac{m}{\delta^2}\psi(\frac{1}{\delta})-m\left[-\frac{1}{\delta^2}\log(\delta)+\frac{1}{\delta^2}\right]\\
&-\left[-\frac{m}{\delta^2}\log[\lambda\tau^\beta+\frac{1}{\delta}]-\frac{1}{\delta^2}\left[\frac{n+\frac{m}{\delta}}{\lambda\tau^\beta+\frac{1}{\delta}}\right]\right]\\
&= -\frac{1}{\delta^2}\sum_{j=1}^{m}\psi(n_j+\frac{1}{\delta})+\frac{m}{\delta^2}\psi(\frac{1}{\delta})+\frac{m}{\delta^2}\log(\delta)\\
&-\frac{m}{\delta^2}+\frac{m}{\delta^2}\log[\lambda\tau^\beta+\frac{1}{\delta}]+\frac{1}{\delta^2}\left[\frac{n+\frac{m}{\delta}}{\lambda\tau^\beta+\frac{1}{\delta}}\right]\\
&= \left[\frac{-1}{\delta^2}\right]\left\{\sum_{j=1}^{m}\psi(n_j+\frac{1}{\delta})-m\psi(\frac{1}{\delta})-m\log(\delta)+m\right\}\\
&-\left[\frac{-1}{\delta^2}\right]\left\{m\log\left[\lambda\tau^\beta+\frac{1}{\delta}\right]-\frac{n+\frac{m}{\delta}}{\lambda\tau^\beta+\frac{1}{\delta}}\right\}
\end{aligned}$$

It might be difficult to have the explicit solution of this expression by equating to zero so that using an iterative procedure is recommendable. Therefore, a function of δ will be utilized in The Newton-Raphson Method.

Since we have the explicit formula for λ and β estimate, which is independent of δ, $\frac{\partial l(\lambda,\beta,\delta)}{\partial\delta}$ is a function of δ only and we can denote it by, $f(\delta)$

$$f(\delta)=\left[\frac{-1}{\delta^2}\right]\left\{\sum_{j=1}^{m}\psi(n_j+\frac{1}{\delta})-m\psi(\frac{1}{\delta})-m\log(\delta)+m-m\log\left[\lambda\tau^\beta+\frac{1}{\delta}\right]-\frac{n+\frac{m}{\delta}}{\lambda\tau^\beta+\frac{1}{\delta}}\right\}$$

Note that in Matlab, psi means digamma function and psi(x) computes the digamma function of x. Similarly, psi(k,x) computes the polygamma function of x, which is the k^{th} derivative of the digamma function at x, denoted by $\psi^k(k,x)$.

4 Dynamic Models: Extending the NHPP

4.1 Introduction

In the previous section we concentrated on modeling recurrent events for repairable systems by non-homogeneous Poisson processes with and without frailty. But, it might be difficult to quantify the effect of the repair by an amount proportional to the current intensity of the processes. Moreover, the number of repair actions up to the current time may have a heavier impact on failure intensity than aging. Due to this fact, in the following we are interested in the dynamic aspect of repairable systems to make a comparison between them.

4.2 Maximum Likelihood Estimation in dynamic model

Here we are considering maximum likelihood estimation of a dynamic model. An intensity process that depends on previous repair actions is termed as conditional intensity. The LEYP model (Linear Extension of Yule Process) (Babykina and Couallier, 2009; Le Gat, 2013) assumes that the conditional intensity evolves as

$$E[dN_j(t)|N_{t_-}] = w_j(t)dt,$$

where $N_j(t)$ counts the number of events for process j, and the history N_{t_-} contains information on (fixed and time-dependent) covariates as well as censoring and observed events in all counting processes prior to time t. Here we look at the situation where

$$w_j(t) = [1 + \gamma N_j(t)]\lambda\beta t^{\beta-1}$$

We suppose that the data concerns m systems with a consideration of these systems in a calender time interval [S,T] where S and T are the starting and the ending time of observation.

The likelihood function for the j^{th} process may be expressed as

$$L_j(\theta) = \left[\prod_{i=1}^{n_j} w_j\left(t_{ij}\right)\right] exp\left[-W_j(\tau_j)\right] \qquad (*)$$

To write the explicit form of this likelihood function we should define the ROCOF $w_j\,(t_{ij})$ to be $(1+\gamma N_j(t_{ij}))w_0(t_{ij})$ where $N_j(t_{ij})$ is the number of previous observed failures for process j and $w_0(t_{ij})$ could be power law model i.e $w_0(t_{ij}) = \lambda\beta t_{ij}^{\beta-1}$.

Next we show how to obtain W(t) for the LEYP model.

In fact we can use any positive number as a starting time of a single recurrent event process but for simplicity we consider t=0 as initial point. Let $0 \leq T_1 < T_2 < ...$ denote the event times, where T_k and T_{k+1} are the time of the k^{th} and $(k+1)^{th}$ events, in respective order. In counting processes $[N(t), 0 \leq t]$ records the cumulative number of events generated by the process but while we look in depth on the processes, the counting processes can be written as $N(t) = \sum_{k=1}^{\infty} I(T_k \leq t)$ counting the number of events occuring over the time interval [0, t].

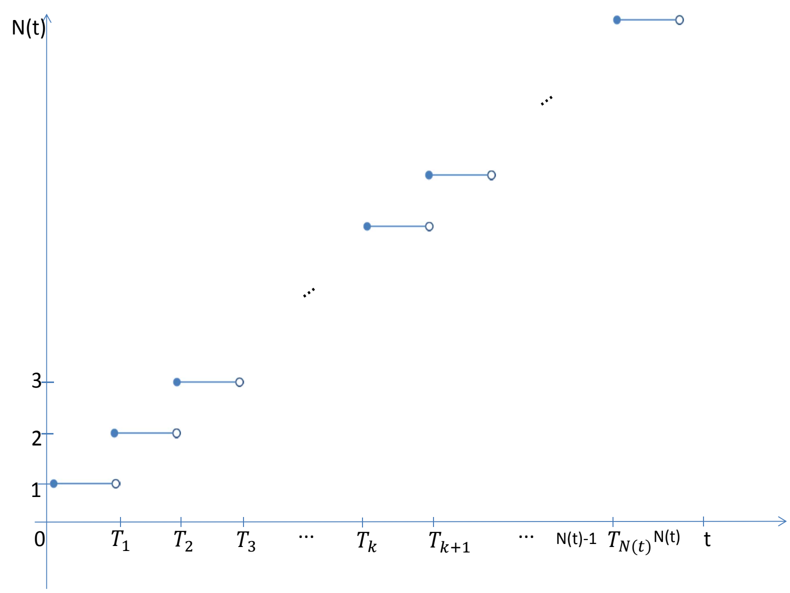

Figure 3: Counting processes representation of data on recurrent events

Now W(t) for LEYP model can derived as follows, being a function of $T_1, T_2, ..., T_{N(t)}$

$$
\begin{aligned}
W(t) &= \int_0^t w(u)du \\
&= \int_0^t (1+\gamma N(u))w_0(u)du \\
&= \sum_{k=0}^{N(t)-1} \int_{T_k}^{T_{k+1}} (1+\gamma k))w_0(u)du + \int_{T_{N(t)}}^{t} (1+\gamma N(u)))w_0(u)du \\
&= \sum_{k=0}^{N(t)-1} (1+\gamma k)(W_0(T_{k+1}) - W_0(T_k)) + (1+\gamma N(t))(W_0(t) - W_0(T_{N(t)})) \\
&= \sum_{k=0}^{N(t)-1} (W_0(T_{k+1}) - W_0(T_k)) + (W_0(t) - W_0(T_{N(t)})) \\
&+ \gamma \sum_{k=0}^{N(t)-1} k(W_0(T_{k+1}) - W_0(T_k)) + \gamma N(t)(W_0(t) - W_0(T_{N(t)}))
\end{aligned}
$$

By manipulating through all k values i.e $k = 0, 1, 2, ..., N(t) - 1$ we arrive at

$$W(t) = W_0(t) + \gamma \left[N(t)W_0(t) - \sum_{k=1}^{N(t)} W_0(T_k) \right]$$

Hence, from (*) on page 18,

$$L_j(\theta) = \left[\prod_{i=1}^{n_j} (1+\gamma N_j(t_{ij}))w_0(t_{ij}) \right] exp\left[-\left[W_j(\tau_j) \right] \right]$$

Substituting the ROCOF of power law model on the above equation is

$$L_j(\theta) = \left[\prod_{i=1}^{n_j} (1+\gamma N_j(t_{ij})\lambda\beta t_{ij}^{\beta-1}) \right] exp\left[- \left[\lambda\tau_j^\beta + \gamma \left[N(\tau_j)\lambda\tau_j^\beta - \sum_{i=1}^{n_j} \lambda t_{ij}^\beta \right] \right] \right]$$

Thus, the total likelihood function is

$$L = \prod_{j=1}^{m} L_j(\theta)$$

The log likelihood function is

$$\begin{aligned}
l &= \log L \\
&= \log \left[\prod_{j=1}^{m} L_j(\theta) \right] \\
&= \log \left[\prod_{j=1}^{m} \left\{ \left[\prod_{i=1}^{n_j} (1 + \gamma N_j(t_{ij}) \lambda \beta t_{ij}^{\beta-1}) \right] exp \left[- \left[\lambda \tau_j^{\beta} + \gamma \left[n_j \lambda \tau_j^{\beta} - \sum_{i=1}^{n_j} \lambda t_{ij}^{\beta} \right] \right] \right] \right\} \right] \\
&= \sum_{j=1}^{m} \left\{ \left[\sum_{i=1}^{n_j} \log(1 + \gamma N_j(t_{ij}) \lambda \beta t_{ij}^{\beta-1}) \right] - \left[\lambda \tau_j^{\beta} + \gamma \left[n_j \lambda \tau_j^{\beta} - \sum_{i=1}^{n_j} \lambda t_{ij}^{\beta} \right] \right] \right\} \\
&= \sum_{j=1}^{m} \left[\sum_{i=1}^{n_j} [\log(1 + \gamma N_j(t_{ij})) + \log \lambda + \log \beta + (\beta - 1) \log(t_{ij})] \right] \\
&- \sum_{j=1}^{m} \left[\lambda \tau_j^{\beta} + \gamma \left[n_j \lambda \tau_j^{\beta} - \sum_{i=1}^{n_j} \lambda t_{ij}^{\beta} \right] \right] \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n_j} \log \left(1 + \gamma N_j(t_{ij})\right) + \sum_{j=1}^{m} n_j \log \lambda + \sum_{j=1}^{m} n_j \log \beta + (\beta - 1) \sum_{j=1}^{m} \sum_{i=1}^{n_j} \log \left(t_{ij}\right) \\
&- \lambda \sum_{j=1}^{m} \left[\tau_j^{\beta} + \gamma \left[n_j \tau_j^{\beta} - \sum_{i=1}^{n_j} t_{ij}^{\beta} \right] \right] \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n_j} \log \left(1 + \gamma (i - 1)\right) + n \log \lambda + n \log \beta + (\beta - 1) \sum_{j=1}^{m} \sum_{i=1}^{n_j} \log \left(t_{ij}\right) \\
&- \lambda \sum_{j=1}^{m} \left[\tau^{\beta} + \gamma \left[n_j \tau^{\beta} - \sum_{i=1}^{n_j} t_{ij}^{\beta} \right] \right]
\end{aligned}$$

Note that $N_j(t_{ij}) = i - 1$ since before the event at t_{ij} we had $i - 1$ events.

Recall property of gamma function:

$$\Gamma[k+1] = k\Gamma[k]$$

So,

$$\begin{aligned}
\Gamma[\frac{1}{\gamma}+n_j] &= \Gamma[\frac{1}{\gamma}+n_j-1+1] \\
&= [\frac{1}{\gamma}+n_j-1]\Gamma[\frac{1}{\gamma}+n_j-1] \\
&= [\frac{1}{\gamma}+n_j-1][\frac{1}{\gamma}+n_j-2]\Gamma[\frac{1}{\gamma}+n_j-2] \\
&= [\frac{1}{\gamma}+n_j-1][\frac{1}{\gamma}+n_j-2][\frac{1}{\gamma}+n_j-3]\Gamma[\frac{1}{\gamma}+n_j-3] \\
&. \\
&. \\
&. \\
&= \frac{1}{\gamma^{n_j}}\,[[1+\gamma(n_j-1)][1+\gamma(n_j-2)][1+\gamma(n_j-3)]\cdots[1+\gamma]]\,\Gamma[\frac{1}{\gamma}]
\end{aligned}$$

Hence,

$$\log[\Gamma[\frac{1}{\gamma}+n_j]] = -n_j\log(\gamma) + \sum_{i=1}^{n_j}\log[1+\gamma(i-1)] + \log[\Gamma(\frac{1}{\gamma})]$$

Thus, for all systems,

$$\sum_{j=1}^{m}\log[\Gamma[\frac{1}{\gamma}+n_j]] = -\sum_{j=1}^{m}n_j\log(\gamma) + \sum_{j=1}^{m}\sum_{i=1}^{n_j}\log[1+\gamma(i-1)] + \sum_{j=1}^{m}\log[\Gamma(\frac{1}{\gamma})]$$

$$\Rightarrow \sum_{j=1}^{m}\sum_{i=1}^{n_j}\log[1+\gamma(i-1)] = \sum_{j=1}^{m}\log[\Gamma[\frac{1}{\gamma}+n_j]] + n\log(\gamma) - m\log[\Gamma(\frac{1}{\gamma})]$$

Hence, complete likelihood function for dynamic model:

$$l = n\log(\lambda) + n\log(\beta) + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log\left(t_{ij}\right) + \sum_{j=1}^{m}\log\left[\Gamma(\frac{1}{\gamma}+n_j)\right]$$
$$- m\log\left[\Gamma\left(\frac{1}{\gamma}\right)\right] + n\log(\gamma) + \gamma\lambda\sum_{j=1}^{m}\sum_{i=1}^{n_j}(t_{ij}^{\beta}) - \lambda\tau^{\beta}\left[m+n\gamma\right]$$

Partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to λ

$$\frac{\partial l\left(\lambda,\beta,\gamma\right)}{\partial\lambda} = \frac{n}{\lambda} + \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}(t_{ij}^{\beta}) - \tau^{\beta}\left[m+n\gamma\right]$$

Second partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to λ

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\lambda^2} = \frac{-n}{\lambda^2}$$

Mixed partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to λ and then β

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\lambda\partial\beta} = \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}(t_{ij}^{\beta}\log(t_{ij})) - \tau^{\beta}\log(\tau)\left[m+n\gamma\right]$$

Mixed partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to λ and then γ

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\lambda\partial\gamma} = \sum_{j=1}^{m}\sum_{i=1}^{n_j}(t_{ij}^{\beta}) - n\tau^{\beta}$$

Partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to β

$$\frac{\partial l\left(\lambda,\beta,\gamma\right)}{\partial\beta} = \frac{n}{\beta} + \sum_{j=1}^{m}\sum_{i=1}^{n_j}\log(t_{ij}) + \lambda\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}(t_{ij}^{\beta}\log(t_{ij})) - \lambda\tau^{\beta}\log(\tau)\left[m+n\gamma\right]$$

Second partial derivative of $l(\lambda, \beta, \gamma)$ with respect to β

$$\frac{\partial^2 l(\lambda, \beta, \gamma)}{\partial \beta^2} = \frac{-n}{\beta^2} + \lambda\gamma \sum_{j=1}^{m} \sum_{i=1}^{n_j} (t_{ij}^{\beta} [\log(t_{ij})]^2) - \lambda\tau^{\beta} [\log(\tau)]^2 [m + n\gamma]$$

Mixed partial derivative of $l(\lambda, \beta, \gamma)$ with respect to β and then λ

$$\frac{\partial^2 l(\lambda, \beta, \gamma)}{\partial \beta \partial \lambda} = \gamma \sum_{j=1}^{m} \sum_{i=1}^{n_j} (t_{ij}^{\beta} \log(t_{ij})) - \tau^{\beta} \log(\tau) [m + n\gamma]$$

which is the same expression as

$$\frac{\partial^2 l(\lambda, \beta, \gamma)}{\partial \lambda \partial \beta}$$

Mixed partial derivative of $l(\lambda, \beta, \gamma)$ with respect to β and then γ

$$\frac{\partial^2 l(\lambda, \beta, \gamma)}{\partial \beta \partial \gamma} = \lambda \sum_{j=1}^{m} \sum_{i=1}^{n_j} (t_{ij}^{\beta} \log(t_{ij})) - n\lambda\tau^{\beta} \log(\tau)$$

In mathematics, the trigamma function, denoted $\psi_1(x)$,is the second of the polygamma functions, and is defined as

$$\begin{aligned} \psi_1(x) &= \frac{d^2}{dx^2} \log \Gamma(x) \\ &= \frac{d}{dx} \psi(x) \end{aligned}$$

where $\psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function, which is the logarithmic derivative of the gamma function. As stated before, in matlab,digamma function at x, $\psi(x)$, is psi(x).

So, here after we will use thus facts in first and second derivative of $l\left(\lambda,\beta,\gamma\right)$ and $l\left(\beta,\gamma\right)$ with respect to γ

Partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to γ

$$\frac{\partial l\left(\lambda,\beta,\gamma\right)}{\partial\gamma}=-\frac{1}{\gamma^2}\sum_{j=1}^{m}\psi(\frac{1}{\gamma}+n_j)+\frac{m}{\gamma^2}\psi(\frac{1}{\gamma})+\frac{n}{\gamma}+\lambda\left[\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}-n\tau^{\beta}\right]$$

Second partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to γ

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\gamma^2}=\frac{1}{\gamma^4}\sum_{j=1}^{m}\left[\psi_1(\frac{1}{\gamma}+n_j)+2\gamma\psi(\frac{1}{\gamma}+n_j)\right]-\frac{m}{\gamma^4}\left[\psi_1(\frac{1}{\gamma})+2\gamma\psi(\frac{1}{\gamma})\right]-\frac{n}{\gamma^2}$$

Mixed partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to γ and then λ

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\gamma\partial\lambda}=\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left[t_{ij}^{\beta}\right]-n\tau^{\beta}$$

which is the same expression as

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\lambda\partial\gamma}$$

Mixed partial derivative of $l\left(\lambda,\beta,\gamma\right)$ with respect to γ and then β

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\gamma\partial\beta}=\lambda\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left[t_{ij}^{\beta}\log\left[t_{ij}\right]\right]-n\lambda\tau^{\beta}\left[\log(\tau)\right]$$

which is the same expression as

$$\frac{\partial^2 l\left(\lambda,\beta,\gamma\right)}{\partial\beta\partial\gamma}$$

Since it is difficult to get the explicit solution of λ, β and γ the Newton-Raphson method will be used.The Newton-Raphson method converges relatively fast for most functions regardless of the initial value even if difficult to set the best initial value. The steps that we should follow in the aforementioned method:
Step1: Set initial value for all the three parameter λ_0,β_0 and γ_0.
Step2: Iterative formula:

$$\begin{bmatrix} \lambda_{i+1} \\ \beta_{i+1} \\ \gamma_{i+1} \end{bmatrix} = \begin{bmatrix} \lambda_i \\ \beta_i \\ \gamma_i \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\lambda^2} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\lambda\partial\beta} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\lambda\partial\gamma} \\ \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\beta\partial\lambda} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\beta^2} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\beta\partial\gamma} \\ \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\gamma\partial\lambda} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\gamma\partial\beta} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\gamma^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial l(\lambda,\beta,\gamma)}{\partial\lambda} \\ \frac{\partial l(\lambda,\beta,\gamma)}{\partial\beta} \\ \frac{\partial l(\lambda,\beta,\gamma)}{\partial\gamma} \end{bmatrix}$$

where the matrix H= $\begin{bmatrix} \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\lambda^2} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\lambda\partial\beta} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\lambda\partial\gamma} \\ \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\beta\partial\lambda} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\beta^2} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\beta\partial\gamma} \\ \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\gamma\partial\lambda} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\gamma\partial\beta} & \frac{\partial^2 l(\lambda,\beta,\gamma)}{\partial\gamma^2} \end{bmatrix}$
is called Hessian Matrix.

It might be easier to find solution from profile likelihood so

$$\frac{\partial l\left(\lambda,\beta,\gamma\right)}{\partial\lambda} = 0$$

$$\Rightarrow \lambda = \frac{n}{\tau^\beta\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^\beta}$$

put in to complete likelihood function and then the profile likelihood function is:

$$l = n\log(n) + n\log(\beta) - n\log\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right] + \left[\beta-1\right]\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log\left[t_{ij}\right]$$

$$+\sum_{j=1}^{m}\log\left[\Gamma\left[\frac{1}{\gamma}+n_j\right]\right] - m\log\left[\Gamma[\frac{1}{\gamma}]\right] + n\log(\gamma) - n$$

Partial derivative of $l\left(\beta,\gamma\right)$ with respect to β

$$\frac{\partial l\left(\beta,\gamma\right)}{\partial\beta} = \frac{n}{\beta} + \sum_{j=1}^{m}\sum_{i=1}^{n_j}\log(t_{ij}) - \frac{n\left[\tau^{\beta}\log(\tau)\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log(t_{ij})t_{ij}^{\beta}\right]}{\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}}$$

Second partial derivative of $l\left(\beta,\gamma\right)$ with respect to β

$$\begin{aligned}\frac{\partial^2 l\left(\beta,\gamma\right)}{\partial\beta^2} &= \frac{-n}{\beta^2} \\ &- \frac{n\left[\tau^{\beta}\left[\log(\tau)\right]^2\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left[\log(t_{ij})\right]^2 t_{ij}^{\beta}\right]\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]}{\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]^2} \\ &+ \frac{n\left[\tau^{\beta}\log(\tau)\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\log(t_{ij})\right]^2}{\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]^2}\end{aligned}$$

Mixed partial derivative of $l\left(\beta,\gamma\right)$ with respect to β and then γ

$$\begin{aligned}\frac{\partial^2 l\left(\beta,\gamma\right)}{\partial\beta\partial\gamma} &= -\frac{n\left[n\tau^{\beta}\left[\log(\tau)\right] - \sum_{j=1}^{m}\sum_{i=1}^{n_j}\left[\log(t_{ij})\right] t_{ij}^{\beta}\right]\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]}{\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]^2} \\ &+ \frac{n\left[\tau^{\beta}\log(\tau)\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\log(t_{ij})\right]\left[n\tau^{\beta} - \sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]}{\left[\tau^{\beta}\left[m+n\gamma\right] - \gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j} t_{ij}^{\beta}\right]^2}\end{aligned}$$

Partial derivative of $l\left(\beta,\gamma\right)$ with respect to γ

$$\frac{\partial l\left(\beta,\gamma\right)}{\partial\gamma}=-\frac{n\left[n\tau^{\beta}-\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]}{\left[\tau^{\beta}\left[m+n\gamma\right]-\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]}$$
$$-\frac{1}{\gamma^2}\sum_{j=1}^{m}\psi(\frac{1}{\gamma}+n_j)+\frac{m}{\gamma^2}\psi(\frac{1}{\gamma})+\frac{n}{\gamma}$$

Second partial derivative of $l\left(\beta,\gamma\right)$ with respect to γ

$$\frac{\partial^2 l\left(\beta,\gamma\right)}{\partial\gamma^2}=\frac{n\left[n\tau^{\beta}-\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]^2}{\left[\tau^{\beta}\left[m+n\gamma\right]-\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]^2}$$
$$+\frac{1}{\gamma^4}\sum_{j=1}^{m}\left[\psi_1(\frac{1}{\gamma}+n_j)+2\gamma\psi(\frac{1}{\gamma}+n_j)\right]-\frac{m}{\gamma^4}\left[\psi_1(\frac{1}{\gamma})+2\gamma\psi(\frac{1}{\gamma})\right]-\frac{n}{\gamma^2}$$

Mixed partial derivative of $l\left(\beta,\gamma\right)$ with respect to γ and then β

$$\frac{\partial^2 l\left(\beta,\gamma\right)}{\partial\gamma\partial\beta}=-\frac{n\left[n\tau^{\beta}\left[\log(\tau)\right]-\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left[\log(t_{ij})\right]t_{ij}^{\beta}\right]\left[\tau^{\beta}\left[m+n\gamma\right]-\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]}{\left[\tau^{\beta}\left[m+n\gamma\right]-\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]^2}$$
$$+\frac{n\left[\tau^{\beta}\log(\tau)\left[m+n\gamma\right]-\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\log(t_{ij})\right]\left[n\tau^{\beta}-\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]}{\left[\tau^{\beta}\left[m+n\gamma\right]-\gamma\sum_{j=1}^{m}\sum_{i=1}^{n_j}t_{ij}^{\beta}\right]^2}$$

which is the same expression as

$$\frac{\partial^2 l(\beta,\gamma)}{\partial\beta\partial\gamma}$$

Step in Newton Raphson's method for profile likelihood function is: Step1: Set initial value for all the three parameter β_0 and γ_0.

Step2: Iterative formula:

$$\begin{bmatrix} \beta_{i+1} \\ \gamma_{i+1} \end{bmatrix} = \begin{bmatrix} \beta_i \\ \gamma_i \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 l(\beta,\gamma)}{\partial \beta^2} & \frac{\partial^2 l(\beta,\gamma)}{\partial \beta \partial \gamma} \\ \\ \frac{\partial^2 l(\beta,\gamma)}{\partial \gamma \partial \beta} & \frac{\partial^2 l(\beta,\gamma)}{\partial \gamma^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial l(\beta,\gamma)}{\partial \beta} \\ \\ \frac{\partial l(\beta,\gamma)}{\partial \gamma} \end{bmatrix}$$

where the matrix H= $\begin{bmatrix} \frac{\partial^2 l(\beta,\gamma)}{\partial \beta^2} & \frac{\partial^2 l(\beta,\gamma)}{\partial \beta \gamma} \\ \\ \frac{\partial^2 l(\beta,\gamma)}{\partial \gamma \partial \beta} & \frac{\partial^2 l(\beta,\gamma)}{\partial \gamma^2} \end{bmatrix}$

is called Hessian Matrix.

5 Interrelation between dynamic behaviour and frailty model for Poisson processes

It is generally agreed that fraility represents an unmeasured risk factor that eventually leads to wrong conclusions if not taken into account. Moreover, there is even a misunderstanding in the concept itself. That is, it is hard to differentiate between static and dynamic fraility.

So, the first and the critical point is a confirmation of whether there is fraility or not. Second, is this fraility static or dynamic? Sometimes the current frail may depend on the past. Thus, we are keenly interested to see the interrelation between static/fixed fraility for each individual and dynamic/stochastic processes that change over time (Aalen et al., 2008).

The idea of intensity functions and counting processes are vital for modelling and statistical analysis of recurrent events. The event intensity function gives the instantaneous probability of an event occuring at t, conditional on the process history. The intensity is defined formally as

$$\lambda(t|H(t)) = lim_{\Delta t \downarrow 0} \frac{Pr[\Delta N(t)=1|H(t)]}{\Delta t}$$

where $H(t) = [N(s) : 0 \leq s < t]$ denote the history of the process at time t (see e.g. Cook and Lawless, 2006).

Just referenced to (Aalen et al., 2008), in this study, individual intensity given the fraility variable a_j is

$$\lambda(t) = aw(t)$$

where $w(t) = \lambda\beta t^{\beta-1}$ is a common baseline intensity and fixed function, that is, independent of the past, while a_j, $j = 1, 2, ..., m$, are independent identically distributed random variables give in the multiplicative factor that determines the risk of an individual. Here we considered a_j to be gamma distributed with scale parameter δ and shape parameter $\frac{1}{\delta}$.

Hence, the conditional intensity of the fraility model,

$$\lambda(t) = w(t)\tfrac{\frac{1}{\delta}+N(t-)}{\delta+A(t)}$$

where $A(t) = \int_0^t w(u)du \equiv \lambda t^{\beta}$

Thus

$$\lambda(t) = \left[\frac{w(t)}{\delta[\delta + A(t)]}\right][1 + \delta N(t-)]$$

$$= \left[\frac{\lambda\beta t^{\beta-1}}{\delta[\delta + \lambda t^{\beta}]}\right][1 + \delta N(t-)]$$

Recall LEYP model:

$$\lambda(t) = w^{*}(t)[1 + \gamma N(t-)]$$

Hence,

$$Fraility \Leftrightarrow LEYP\text{: if } w^{*}(t) = \tfrac{w(t)}{\delta[\delta+A(t)]}$$

This bi-implication shows us that frailty models may alternatively be viewed as dynamic models. Hereafter, we are interested in confirming this theoretical observation by a simulation study.

6 Simulate m Systems with dynamic ROCOF

As mentioned before we suppose that the data concerns on m systems with a consideration of these systems in a calender time interval [S,T] where S and T are the starting and the ending time of observation. We might look the system at the start of the operation i.e S=0 and up to T=10.

We can use ordinary power law model $w(t) = \lambda\beta t^{\beta-1}$ to simulate failure time $T_1 = S_1$; to simulate T_2 in the interval $[S_1, \infty)$ we might use the intensity $w(t) = (1+\gamma)\lambda\beta t^{\beta-1}$; to simulate T_3 from the interval $[S_2, \infty)$ where $S_2 = S_1 + T_2$, we can consider the intensity $w(t) = (1+2\gamma)\lambda\beta t^{\beta-1}$ and so on.

How to simulate T_1? To generate the failure time T_1 we can see the following procedures

Step 1: Let us take $z_1(t) = \lambda\beta t^{\beta-1}$ the hazard rate of T_1. Its survival function $G_1(t) = P(T_1 > t) = e^{-\int_0^t z_1(u)du}$. By integrating the intensity function the survival function is $G_1(t) = e^{-\lambda t^\beta}$

Step 2: Draw a random variable $u_1 \sim u[0,1]$ and equate to the survival function $G_1(t) = e^{-\lambda t^\beta}$

Thus,

$$
\begin{aligned}
e^{-\lambda t^\beta} &= u_1 \\
\Rightarrow -\log u_1 &= \lambda T_1^\beta \\
\Rightarrow S_1 = T_1 &= \left(-\frac{\log u_1}{\lambda}\right)^{1/\beta}
\end{aligned}
$$

How to simulate T_2?

Claim: T_2 has hazard rate $z_2(t) = (1+\gamma)\lambda\beta(S_1+t)^{\beta-1}$, which is conditional on S_1. The survival function of T_2 conditional on S_1 is

$$
\begin{aligned}
G_2(t) &= P(T_2 > t) \\
&= e^{-\int_0^t z_2(u)du}
\end{aligned}
$$

$$= e^{-\int_0^t (1+\gamma)\lambda\beta(S_1+u)^{\beta-1}du}$$

$$= e^{-(1+\gamma)\lambda\beta\int_0^t (S_1+u)^{\beta-1}du}$$

$$= e^{-(1+\gamma)\lambda\beta\int_{S_1}^{S_1+t} x^{\beta-1}dx}$$

$$= e^{-(1+\gamma)\lambda\left[(S_1+t)^{\beta}-S_1^{\beta}\right]}$$

Let us draw a random variable $u_2 \sim u[0,1]$ and equate to the survival function $u_2 = e^{-(1+\gamma)\lambda\left[(S_1+T_2)^{\beta}-S_1^{\beta}\right]}$

This implies that

$$u_2 = e^{-(1+\gamma)\lambda\left[(S_1+T_2)^{\beta}-S_1^{\beta}\right]}$$

$$\Rightarrow \log u_2 = -(1+\gamma)\lambda\left[(S_1+T_2)^{\beta}-S_1^{\beta}\right]$$

$$\Rightarrow -\frac{\log u_2}{(1+\gamma)\lambda} = (S_1+T_2)^{\beta}-S_1^{\beta}$$

$$\Rightarrow S_1+T_2 = \left(S_1^{\beta} - \frac{\log u_2}{(1+\gamma)\,\lambda}\right)^{1/\beta}$$

$$\Rightarrow T_2 = \left(S_1^{\beta} - \frac{\log u_2}{(1+\gamma)\lambda}\right)^{1/\beta} - S_1$$

Thus,

$$S_2 = \left(S_1^{\beta} - \frac{\log u_2}{(1+\gamma)\lambda}\right)^{1/\beta}$$

How to simulate T_3?
Claim: T_3 has hazard rate $z_3(t) = (1+2\gamma)\lambda\beta(S_2+t)^{\beta-1}$, which is conditional on S_2. The survival function of T_3 is

$$
\begin{aligned}
G_3(t) &= P(T_3 > t) \\
&= e^{-\int_0^t z_3(u)du} \\
&= e^{-\int_0^t (1+2\gamma)\lambda\beta(S_2+u)^{\beta-1}du} \\
&= e^{-(1+2\gamma)\lambda\beta\int_0^t (S_2+u)^{\beta-1}du} \\
&= e^{-(1+2\gamma)\lambda\beta\int_{S_2}^{S_2+t} x^{\beta-1}dx} \\
&= e^{-(1+2\gamma)\lambda\left[(S_2+t)^{\beta}-S_2^{\beta}\right]}
\end{aligned}
$$

Let us draw a random variable $u_3 \sim u[0,1]$ and equate to the survival function $u_3 = e^{-(1+\gamma)\lambda\left[(S_2+T_2)^{\beta}-S_2^{\beta}\right]}$

This implies that

$$
\begin{aligned}
u_3 &= e^{-(1+2\gamma)\lambda\left[(S_2+T_3)^{\beta}-S_2^{\beta}\right]} \\
\Rightarrow \log u_3 &= -(1+2\gamma)\lambda\left[(S_2+T_3)^{\beta}-S_2^{\beta}\right] \\
\Rightarrow -\frac{\log u_3}{(1+2\gamma)\lambda} &= (S_2+T_3)^{\beta}-S_2^{\beta} \\
\Rightarrow S_2+T_3 &= \left(S_2^{\beta}-\frac{\log u_3}{(1+2\gamma)\,\lambda}\right)^{1/\beta} \\
\Rightarrow T_3 &= \left(S_2^{\beta}-\frac{\log u_3}{(1+2\gamma)\lambda}\right)^{1/\beta} - S_2
\end{aligned}
$$

Thus,

$$
S_3 = \left(S_2^{\beta}-\frac{\log u_3}{(1+2\gamma)\lambda}\right)^{1/\beta}
$$

Similarly,

$$S_4 = \left(S_3^\beta - \frac{\log u_4}{(1+3\gamma)\lambda} \right)^{1/\beta}$$

,

$$S_5 = \left(S_4^\beta - \frac{\log u_5}{(1+4\gamma)\lambda} \right)^{1/\beta}$$

and so on. Graphically,

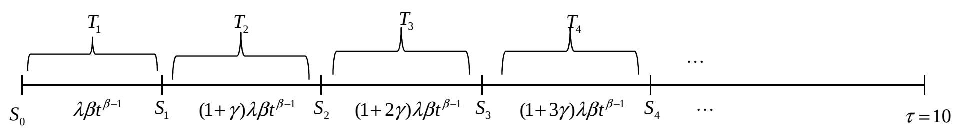

Figure 4: Relation between the interoccurrence times (T_i), caleander time S_i and intensity $w(t) = (1 + i\gamma)\lambda\beta t^{\beta-1}$ where $i = 0, 1, ...$

If $\gamma = 0$, then LEYP process is an ordinary power law process.

$$S_1 = T_1 = \left(-\frac{\log u_1}{\lambda} \right)^{1/\beta}$$

$$S_2 = \left(S_1^\beta - \frac{\log u_2}{\lambda} \right)^{1/\beta}$$

$$S_3 = \left(S_2^\beta - \frac{\log u_3}{\lambda} \right)^{1/\beta}$$

and so on but stop while we are passing τ. So we can simulate ordinary power law process from LEYP model by making $\gamma = 0$.

We may describe it graphically as:

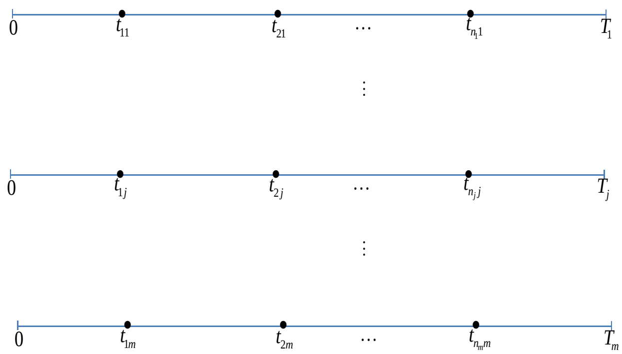

Figure 5: Observation of failure times of m systems.

Similarly, we can simulate power law with frailty from LEYP model. For process ♯1, draw an "a_1" from gamma distribution with expected value 1 and variance δ. Then the failure times are simulated as

$$S_1^{(1)} = T_1^{(1)} = \left(-\frac{\log u_1}{\lambda \cdot a_1}\right)^{1/\beta}$$

$$S_2^{(1)} = \left(S_1^{(1)\beta} - \frac{\log u_2}{\lambda \cdot a_1}\right)^{1/\beta}$$

$$S_3^{(1)} = \left(S_2^{(1)\beta} - \frac{\log u_3}{\lambda \cdot a_1}\right)^{1/\beta}$$

and so on.

For process ♯2, draw an "a_2" from gamma distribution.

$$S_1^{(2)} = T_1^{(2)} = \left(-\frac{\log u_1}{\lambda \cdot a_2}\right)^{1/\beta}$$

$$S_2^{(2)} = \left(S_1^{(2)\beta} - \frac{\log u_2}{\lambda \cdot a_2}\right)^{1/\beta}$$

$$S_3^{(2)} = \left(S_2^{(2)\beta} - \frac{\log u_3}{\lambda \cdot a_2}\right)^{1/\beta}$$

and so on. Graphically, provided $\beta = \frac{3}{2}$,

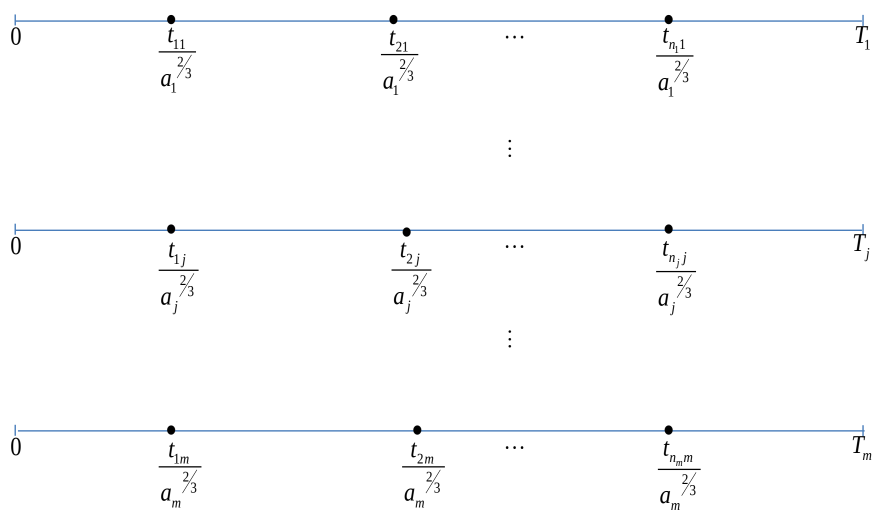

Figure 6: Observation of failure times of m system with frailties a

7 Maximum Likelihood Estimation

Although the method of maximum likelihood is an efficient method once we have an explicit likelihood function,it is a routine procedure for obtaining estimators for unknown parameters from a set of data. It's estimate for θ is a value of θ which maximize the likelihood function over the parameter space. It is a single parameter value which is most likely in light of what have been observed.

Definition:

1. The likelihood function is the joint probability (density) function of observable random variables but it is viewed as the function of the parameters given the realized random variables.

   Mathematically,let $x_1, x_2, ..., x_n$ be a random sample of size n from the discrete pdf $p_X(x;\theta)$. The likelihood function, L(θ), is the product of the pdf evaluated at the n $x_i^{'}s$. That is,

   $$L(\theta) = \Pi_{i=1}^{n} p_X(x_i;\theta)$$

   Similarly,if $x_1, x_2, ..., x_n$ be a random sample of size n from a continuous pdf, $f_X(x;\theta)$. The likelihood function can be written as

   $$L(\theta) = \Pi_{i=1}^{n} f_X(x_i;\theta)$$

   where θ is an unknown parameter in both cases. Moreover, let θ_l is the value of the parameter such that $L(\theta_l) \geq L(\theta)$ for all possible values of θ. Then θ_l is maximum likelihood estimate(MLE) θ.

2. The function $l(\theta) = lnL(\theta)$ is the log likelihood function of $x_1, x_2, ..., x_n$.

3. The function $S(\theta) = \frac{\partial}{\partial x} l(\theta)$ is the score function of $x_1, x_2, ..., x_n$.

4. The function $I(\theta) = -\frac{\partial^2}{\partial x} l(\theta)$ is the information matrix of $x_1, x_2, ..., x_n$.

The Fisher information matrix is used to calculate the covariance matrices associated with maximum-likelihood estimates so that we can easily estimate the standard deviation of estimates.

8 Preliminary Analysis

The main objective of the preliminary analysis is to give a simple overview about simulation of a single process observed on the time interval [0,10].

8.1 Ordinary power law model

This is simulation of a single process observed on the interval [0,10], where parameter values are $\lambda = 2$ and $\beta = 1.5$.

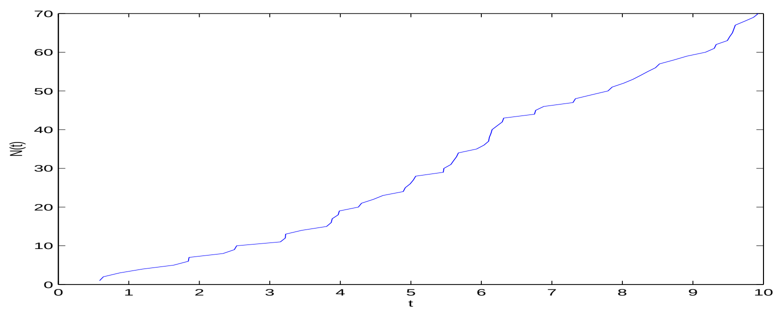

Figure 7: Random failures time t Vs number of failure N(t)

8.2 Fraility

These are simulations of single processes observed on the interval [0,10] for the same $\lambda = 2$ and $\beta = 1.5$ but varying δ values

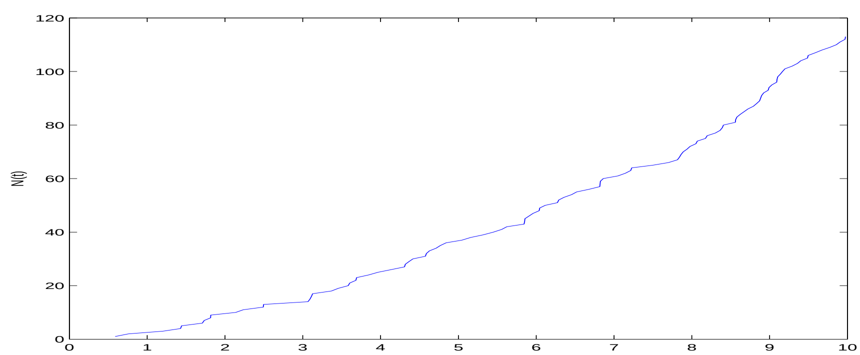

Figure 8: Random failures time t Vs number of failure N(t), δ=0.2

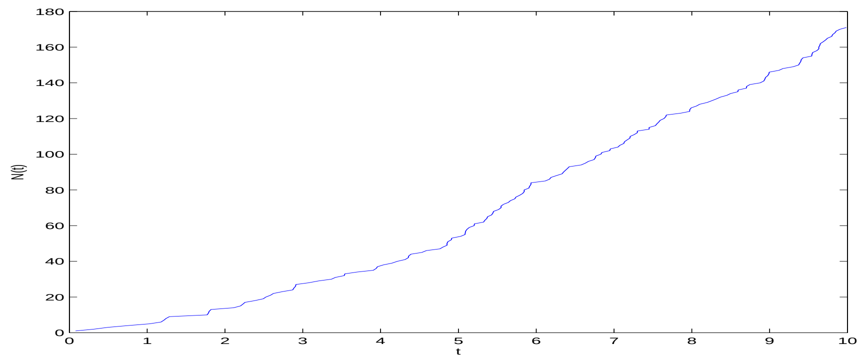

Figure 9: Random failures time t Vs number of failure N(t), δ=0.4

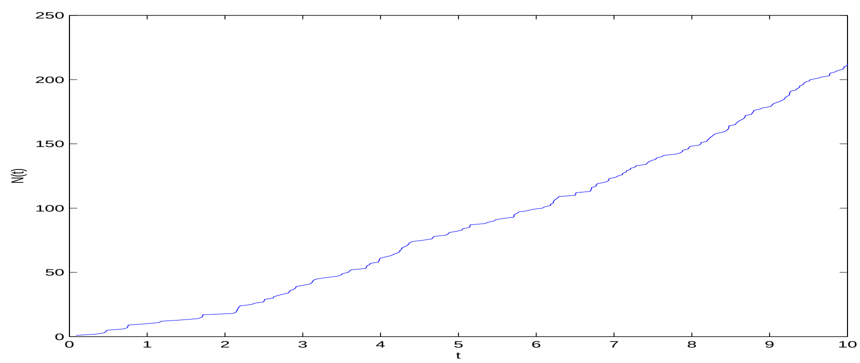

Figure 10: Random failures time t Vs number of failure N(t), δ=0.6

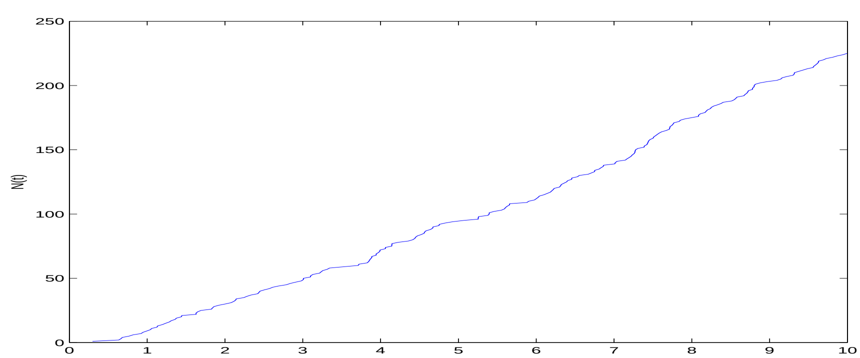

Figure 11: Random failures time t Vs number of failure N(t), δ=0.8

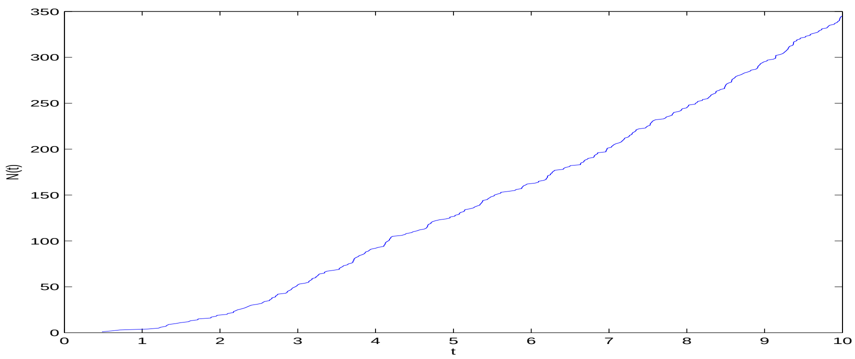

Figure 12: Random failures time t Vs number of failure N(t), δ=1

8.3 Dynamic Behaviour

These are simulations of single processes observed on the interval [0,10] for the same $\lambda = 2$ and $\beta = 1.5$ but varying γ values

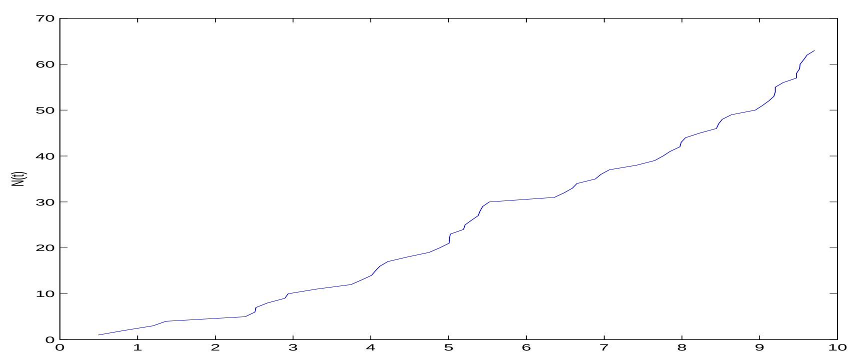

Figure 13: Random failures time t Vs number of failure N(t), γ=0.001

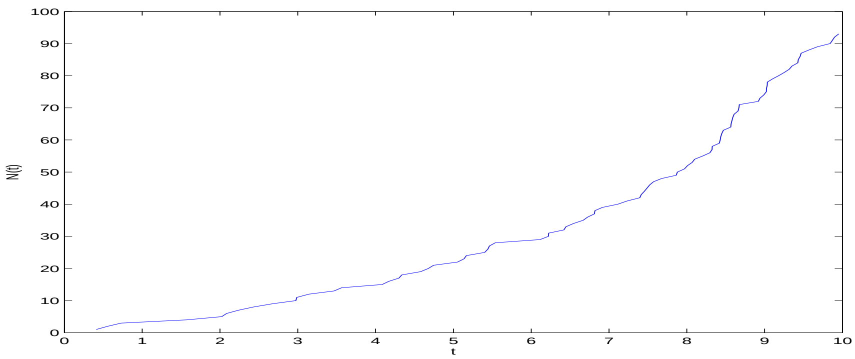

Figure 14: Random failures time t Vs number of failure N(t), γ=0.01

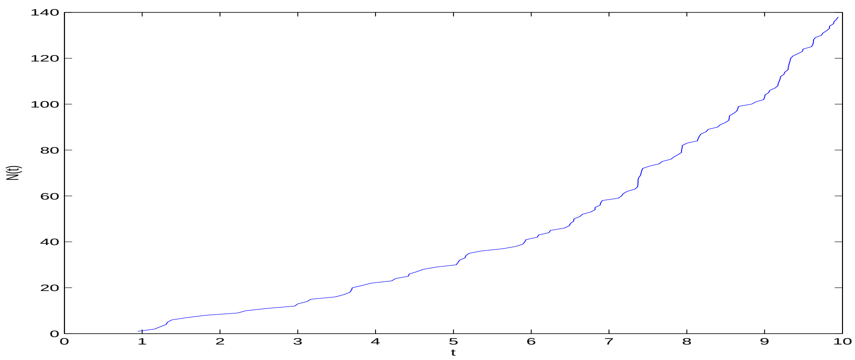

Figure 15: Random failures time t Vs number of failure N(t), γ=0.02

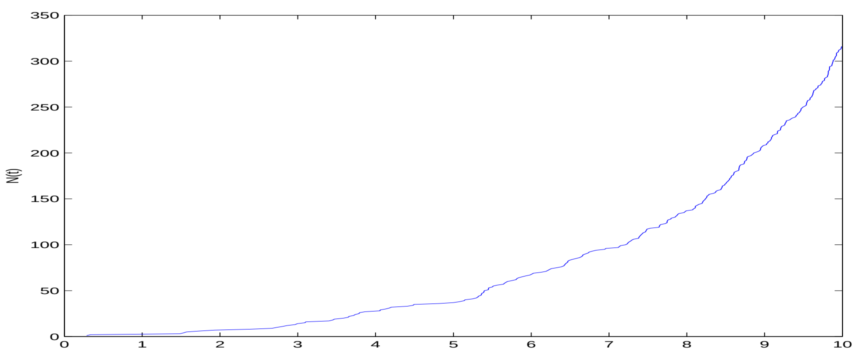

Figure 16: Random failures time t Vs number of failure N(t), γ=0.04

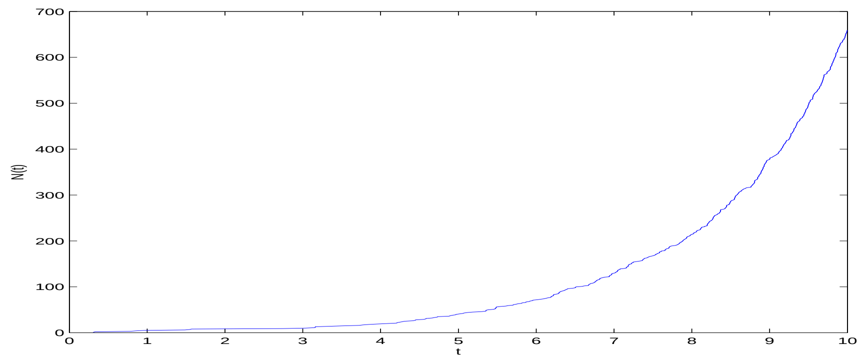

Figure 17: Random failures time t Vs number of failure N(t), γ=0.06

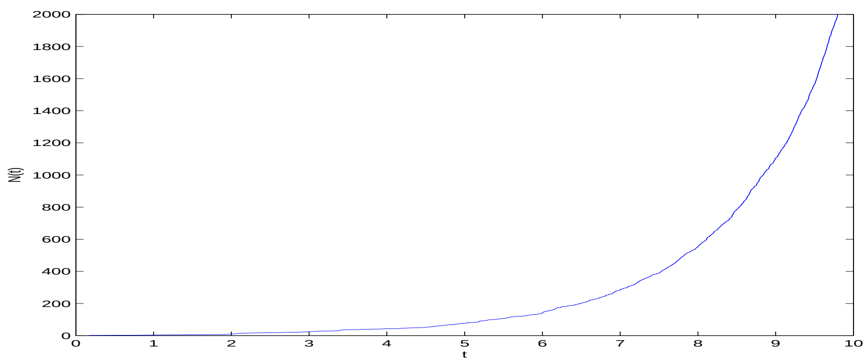

Figure 18: Random failures time t Vs number of failure N(t),γ=0.08

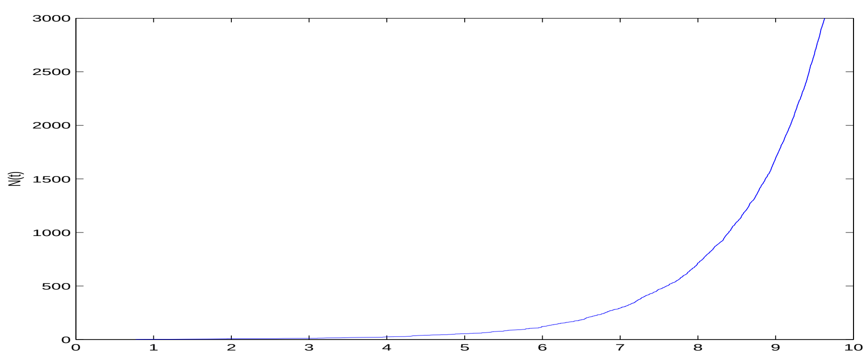

Figure 19: Random failures time t Vs number of failure N(t), γ=0.1

9 Simulation Study

9.1 Power Law Model

9.1.1 Maximum Likelihood Estimate

Maximum likelihood estimates of ordinary power law model for single simulation with a given value m=20, λ=2, β=1.5. The ML estimates are $\hat{\lambda}$=1.9926 and $\hat{\beta}$=1.4999. This resulted in the Fisher information

$$I\left(\hat{\lambda},\hat{\beta}\right)=\begin{bmatrix} 0.0407 & -0.0082 \\ -0.0082 & 0.0018 \end{bmatrix}$$

From the Fisher information matrix we can further derive the standard deviation of $\hat{\lambda}$ and $\hat{\beta}$ to be 0.2018 and 0.0423 in respective order. Its maximum likelihood,

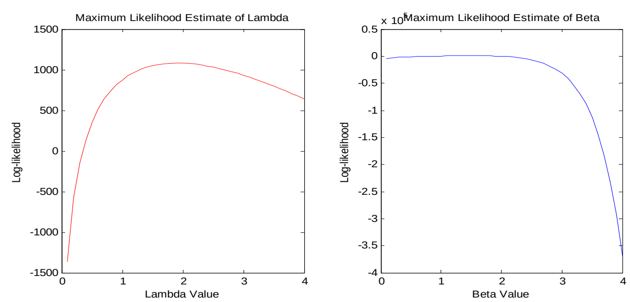

Figure 20: Maximum likelihood estimates of ordinary power law model

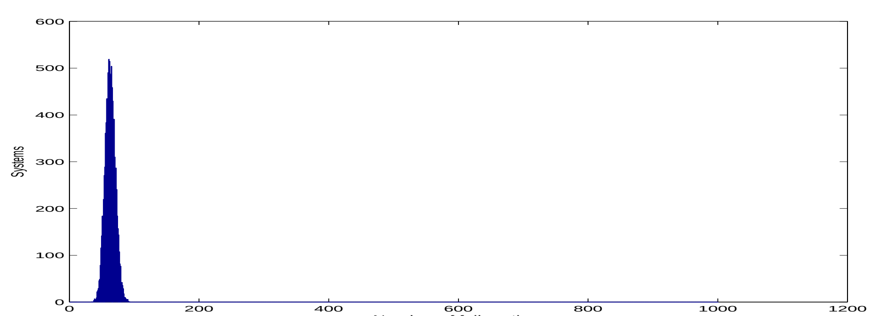

Figure 21: Histogram of Number of failure Vs Systems; λ=2, β=1.5, 10000 data sets and m=20 systems per data sets

Hereafter, average (Ave.)and standared deviation (St.D) are denoted as follows:

For estimate [$\hat{\lambda}$, $\hat{\beta}$, $\hat{\delta}$ and $\hat{\gamma}$]: Average (Ave.) is the sum of all estimates divided by number of data sets and standard deviation (St.D) is the average distance between the estimates and the mean(Average of estimates).

For number of failures [n]: Average (Ave.) is the sum of all number of failure per system divided by the product of number of system per data set ,and number of total data sets. Its standard deviation (St.D) is the square root of the quadratic distance between the number of failures per process and the mean(Average number of failures).

Data	m	True Value		n		Estimates			
10000	20	λ	β	Average	St.D	$\hat{\lambda}$		$\hat{\beta}$	
						Average	St.D	Average	St.D
		2	1.5	63.2896	7.9675	2.0034	0.2042	1.5012	0.0424
			1	19.9312	4.4597	2.0000	0.2495	1.0030	0.0504
			0.75	11.2457	3.3483	2.0015	0.2619	0.7526	0.0497

Table 1: Power law data and power law estimates

Data	m	True Value			n		Estimates			
10000	20	λ	β	δ	Average	St.D	$\hat{\lambda}$		$\hat{\beta}$	
							Average	St.D	Average	St.D
		2	1.5	0.2	62.5152	29.2594	1.9850	0.3041	1.5036	0.0488
				0.4	63.0461	40.7118	1.9918	0.3623	1.5023	0.0486
				0.6	62.6755	49.3998	1.9892	0.4078	1.5020	0.0485
				0.8	63.1969	56.7187	1.9962	0.4543	1.5018	0.0485
				1	61..5949	63.0386	1.9845	0.4952	1.5019	0.0494
			1	0.2	19.9349	9.9328	2.0050	0.3188	1.0024	0.0508
				0.4	19.9346	13.2303	1.9942	0.3754	1.0036	0.0504
				0.6	20.0902	16.4537	1.9949	0.4270	1.0033	0.0512
				0.8	20.0922	18.5500	1.9966	0.4700	1.0049	0.0520
				1	19.9214	20.3878	1.9906	0.5143	1.0085	0.0533
			0.75	0.2	11.3481	6.0628	1.9478	0.3118	0.7671	0.0469
				0.4	11.2743	7.7920	1.9453	0.3678	0.7680	0.0462
				0.6	11.5032	9.3936	1.9335	0.4249	0.7712	0.0479
				0.8	11.3861	10.5682	1.9247	0.4760	0.7745	0.0499
				1	11.5880	11.9995	1.9198	0.5246	0.7799	0.0523

Table 2: Fraility data and power law estimates

Table Summary
Case 1: $\beta > 1$
As δ increases: average number of failures per system are nearly constant but the standard deviation (St.D) increases; average of λ estimates are nearly constant but the standard deviation (St.D) increases; average and standared deviation (St.D)of β estimates are fairly constant.

Case 2: $\beta = 1$
As δ increase: similar to case 1.

Case 3: $\beta < 1$
As δ increase: average number of failures per system are very nearly constant but the standard deviation (St.D) increases; average of λ estimates are slightly decrease but standard deviation of λ estimates are increases; average and standared deviation (St.D)of β estimates are increases.

Data	m	True Value			n		Estimates			
10000	20	λ	β	γ	Average	St.D	$\hat{\lambda}$		$\hat{\beta}$	
							Average	St.D	Average	St.D
		2	1.5	0.001	65.0923	8.2762	1.9566	0.2013	1.5256	0.0428
				0.01	88.1634	12.6667	1.5068	0.1652	1.7700	0.0455
				0.02	126.9190	21.5008	0.9970	0.1171	2.1080	0.0483
				0.04	289.1689	59.6044	0.2704	0.0344	3.0315	0.0516
				0.06	726.9944	180.2163	0.0356	0.0045	4.3109	0.0495
			1	0.001	20.2509	4.5139	2.0015	0.2537	1.0073	0.0512
				0.01	22.1547	5.1818	1.9607	0.2539	1.0558	0.0517
				0.02	24.5958	6.0606	1.9192	0.2549	1.1105	0.0532
				0.04	30.8249	8.4493	1.7955	0.2505	1.2356	0.0551
				0.06	38.8179	11.3669	1.6233	0.2399	1.3812	0.0575
			0.75	0.001	11.2690	3.3855	1.9966	0.2703	0.7562	0.0515
				0.01	11.8937	3.5962	2.0140	0.2696	0.7743	0.0507
				0.02	12.6623	4.0317	2.0249	0.2805	0.7978	0.0523
				0.04	14.2409	4.7588	2.0406	0.2861	0.8462	0.0529
				0.06	16.0941	5.5937	2.0576	0.3010	0.8961	0.0550

Table 3: Dynamic data and power law estimates

Table Summary
Case 1: $\beta > 1$
As γ increase: average and standard deviation (St.D) of number of failure per system are highly increases; averages and standard deviation (St.D) of λ estimates are highly decrease; average of β estimates increases but the standard deviations (St.D) fairly constant.

Case 2: $\beta = 1$
As γ increase: average and standard deviation (St.D) of number of failure per system are increase; average of λ estimates are decrease and its standard deviation (St.D) estimates are fairly constant; average of β estimates are increase but the standard deviation (St.D) fairly constant.

Case 3: $\beta < 1$
As γ increase: average and standard deviation (St.D) of number of failure per system are slowly increases; average of λ estimates are fairly constant but the standard deviations (St.D) slowly increases; average of β estimates are slowly increase but standard deviation (St.D) fairly constant.

9.2 A gamma multiple(frailty)Power Law Model

9.2.1 Maximum Likelihood Estimate

Here we are keenly interested to estimate parameters λ, β and δ.

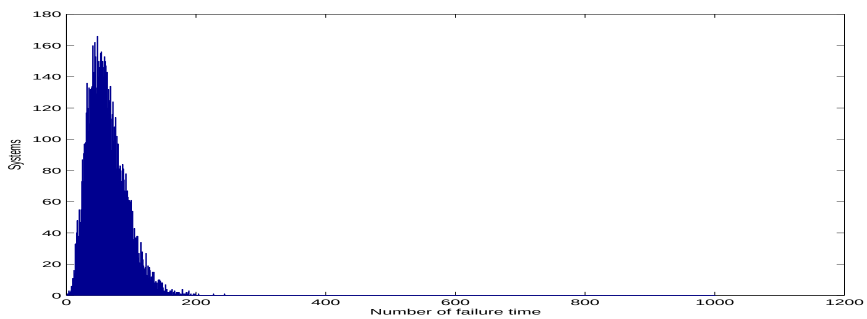

Figure 22: Histogram of Number of failure Vs Systems; λ=2, β=1.5, δ=0.2,10000 data sets and m=20 systems per data sets

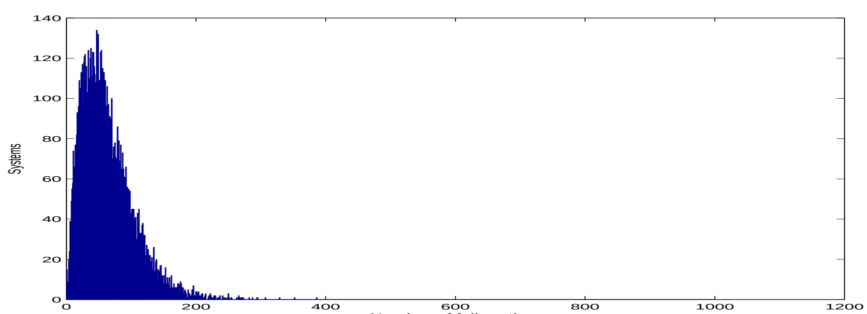

Figure 23: Histogram of Number of failure Vs Systems; λ=2, β=1.5, δ=0.4,10000 data sets and m=20 systems per data sets

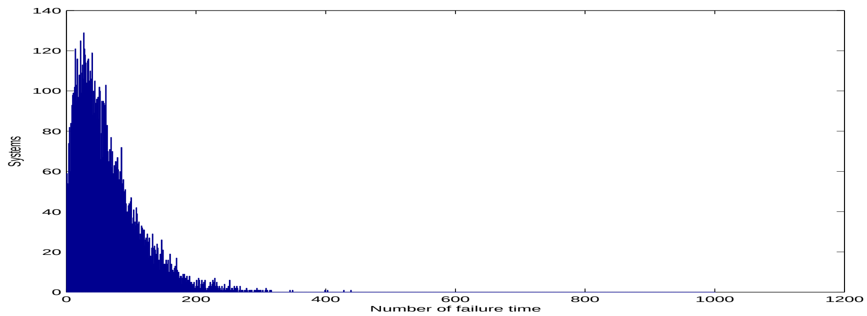

Figure 24: Histogram of Number of failure Vs Systems; λ=2, β=1.5, δ=0.6,10000 data sets and m=20 systems per data sets

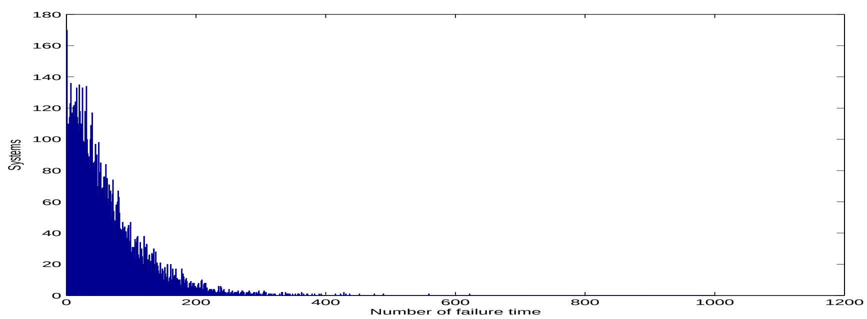

Figure 25: Histogram of Number of failure Vs Systems; λ=2, β=1.5, δ=0.8,10000 data sets and m=20 systems per data sets

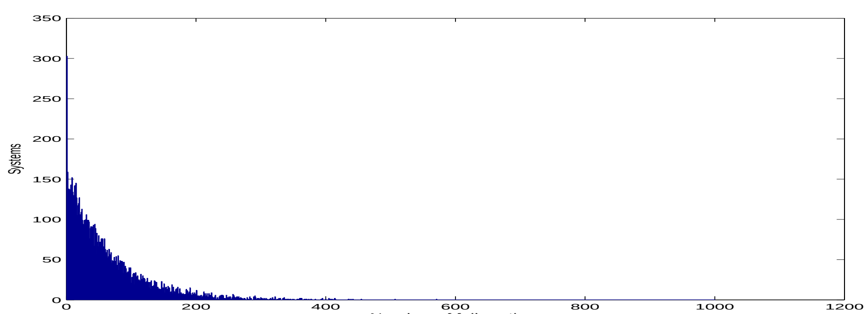

Figure 26: Histogram of Number of failure Vs Systems; λ=2, β=1.5, δ=1,10000 data sets and m=20 systems per data sets

Data	m	True Value			n		Estimates					
10000	20	λ	β	δ	Ave.	St.D	$\hat{\lambda}$		$\hat{\beta}$		$\hat{\delta}$	
							Ave.	St.D	Ave.	St.D	Ave.	St.D
		2	1.5	0.0001	63.2373	7.8956	2.0062	0.2024	1.5005	0.0423	0.002	0.0037
				0.01	63.4082	10.2008	2.0054	0.2071	1.5009	0.0422	0.0112	0.0097
				0.1	63.0648	21.4421	2.0075	0.2456	1.5005	0.0421	0.0942	0.0352
				0.2	63.2006	29.6606	2.0042	0.2853	1.5013	0.0425	0.1897	0.0651
				0.4	63.0469	40.1756	2.0033	0.3450	1.5009	0.0426	0.3795	0.1208
				0.6	63.3880	49.3983	2.0063	0.4033	1.5014	0.0431	0.5718	0.1750
				0.8	63.0149	56.1749	2.0050	0.4452	1.5013	0.0430	0.7511	0.2223
				1	63.4131	64.1655	2.0033	0.5023	1.5021	0.0432	0.9309	0.2610
			1	0.001	19.9770	4.6594	2.0057	0.2559	1.0030	0.0511	0.0071	0.0137
				0.2	19.9885	9.9238	1.9987	0.3231	1.0023	0.0509	0.1879	0.0774
				0.4	20.2335	13.4745	2.0052	0.3802	1.0026	0.0508	0.3742	0.1305
				0.6	19.8679	15.8587	1.9968	0.4285	1.0043	0.0514	0.5520	0.1769
				0.8	20.1814	18.2127	1.9923	0.4691	1.0051	0.0518	0.7158	0.2107
				1	20.4585	20.8251	1.9843	0.5149	1.0083	0.0524	0.8679	0.2421
			0.75	0.001	11.2590	3.3240	2.0083	0.2626	0.7521	0.0513	0.0123	0.0231
				0.2	11.8200	6.8717	2.0535	0.3142	0.7502	0.0493	0.1888	0.0924
				0.4	11.2779	7.8143	1.9923	0.3934	0.7557	0.0519	0.3620	0.1381
				0.6	11.1454	9.2048	1.9846	0.4330	0.7572	0.0522	0.5234	0.1763
				0.8	11.1906	10.4320	1.9702	0.4873	0.7624	0.0539	0.6685	0.2054
				1	11.1918	11.4700	1.959	0.5285	0.7663	0.0552	0.7917	0.2274

Table 4: Fraility data and fraility estimates

Table Summary
Case 1: $\beta > 1$
As δ increase: average number of failure per system are constant but its standard deviation (St.D) increases; average of λ estimates are fairly constant but the standard deviation (St.D) slowly increase: average and standard deviation (St.D) of β estimates are fairly constant; average and standard deviation (St.D) of δ estimates are increase.

Case 2: $\beta = 1$
As δ increase: similar as **Case 1**

Case 3: $\beta < 1$
As δ increase: average number of failure per system are fairly constant but the standard deviation (St.D) increases; average of λ estimates are decrease but the standard deviation (St.D) increases: average of β estimates are increase but the standard deviations (St.D) constant; average and standard deviation (St.D) of δ are increase.

Data	m	True V.		n		Estimates					
10000	20	λ	β	Ave.	St.D	$\hat{\lambda}$		$\hat{\beta}$		$\hat{\delta}$	
						Ave.	St.D	Ave.	St.D	Ave.	St.D
		2	1.5	63.1451	8.0541	2.0059	0.2057	1.5008	0.0430	0.0021	0.0036
			1	20.0288	4.4658	2.0051	0.2526	1.0019	0.0509	0.0064	0.0119
			0.75	11.2920	3.3512	1.9992	0.2653	0.7535	0.0507	0.0116	0.0215

Table 5: Power law data and fraility estimates

Data	m	True Value			n		Estimates					
1000	20	λ	β	γ	Ave.	St.D	$\hat{\lambda}$		$\hat{\beta}$		$\hat{\delta}$	
							Ave.	St.D	Ave.	St.D	Ave.	St.D
		2	1.5	0.00001	63.3176	8.0255	2.0030	0.2019	1.5017	0.0423	0.0020	0.0038
				0.001	65.2978	8.2883	1.9562	0.1972	1.5254	0.0425	0.0026	0.0042
				0.01	88.4093	12.8219	1.5084	0.1668	1.7694	0.0459	0.0111	0.0080
				0.02	127.0301	21.0328	0.9979	0.1159	2.1078	0.0480	0.0229	0.0119
				0.04	287.3850	61.2510	0.2698	0.0349	3.0325	0.0516	0.0387	0.0133
				0.06	719.8430	176.3796	0.0354	0.0044	4.3141	0.0479	0.0569	0.0194
			1	0.00001	19.9410	4.3837	2.0018	0.2592	1.0019	0.0513	0.0062	0.0114
				0.001	20.4110	4.6004	1.9959	0.2521	1.0097	0.0509	0.0069	0.0118
				0.01	21.9510	5.0593	1.9584	0.2518	1.0558	0.0514	0.0135	0.0174
				0.02	24.5330	6.0879	1.9214	0.2596	1.1113	0.0537	0.0238	0.0221
				0.04	30.8490	8.4581	1.8012	0.2496	1.2348	0.0551	0.0474	0.0296
				0.06	37.9590	11.3169	1.6259	0.2481	1.3816	0.0589	0.0682	0.0334
				0.08	50.1180	16.0797	1.4173	0.2131	1.5465	0.0587	0.0763	0.0305
				0.1	64.0240	22.7574	1.1807	0.1888	1.7359	0.0616	0.0943	0.0361
				0.2	269.6050	120.1785	0.2495	0.0397	3.0342	0.0576	0.1902	0.0602
			0.75	0.00001	11.2950	3.3210	2.0022	0.2644	0.7536	0.0490	0.0117	0.0215
				0.001	11.3640	3.5082	2.0046	0.2663	0.7550	0.0513	0.0126	0.0229
				0.01	11.8860	3.5045	2.0044	0.2715	0.7769	0.0520	0.0164	0.0253
				0.02	12.5950	4.1497	2.0310	0.2757	0.7970	0.0511	0.0261	0.0320
				0.04	14.3780	4.8219	2.0294	0.2925	0.8478	0.0541	0.0441	0.0410
				0.06	16.0900	5.8230	2.0615	0.3080	0.8961	0.0544	0.0668	0.0461
				0.08	18.5330	7.1101	2.0409	0.3171	0.9561	0.0574	0.0942	0.0534
				0.2	40.9700	19.5547	1.6950	0.2931	1.4002	0.0611	0.1937	0.0709
				0.4	219.4870	136.0986	0.5884	0.1085	2.5791	0.0532	0.3823	0.1157

Table 6: Dynamic data and fraility estimates

Table Summary
Case 1: $\beta > 1$
As γ increase: average and standard deviation (St.D) of number of failures per system are highly increases; average and standard deviation (St.D) of λ estimates are highly decreases; average of β estimates are increases but the standard deviation (St.D) fairly constant; average and standard deviation

(St.D) of δ estimates are increases.

Case 2: $\beta = 1$
As γ increase: average and standard deviation (St.D) of failures per system are increases; average and the standard deviation (St.D) decreases; average of β estimates are increases but the standard deviations (St.D) fairly constant; average and standard deviation (St.D) of δ estimates are increases.

Case 3: $\beta < 1$
As γ increase: average and standard deviation (St.D) of failure per system are increases; average and standard deviation (St.D) of λ estimates are decreases; average of β estimates increase but the standard deviations (St.D) fairly constant; average and standard deviation (St.D) of δ estimates are increases.

9.3 A dynamic view of Power Law Model

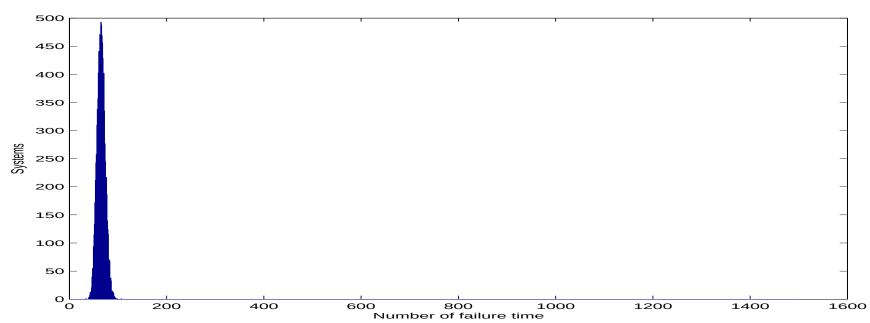

Figure 27: Histogram of Number of failure Vs Systems; λ=2, β=1.5,γ=0.001,10000 data sets and m=20 systems per data sets

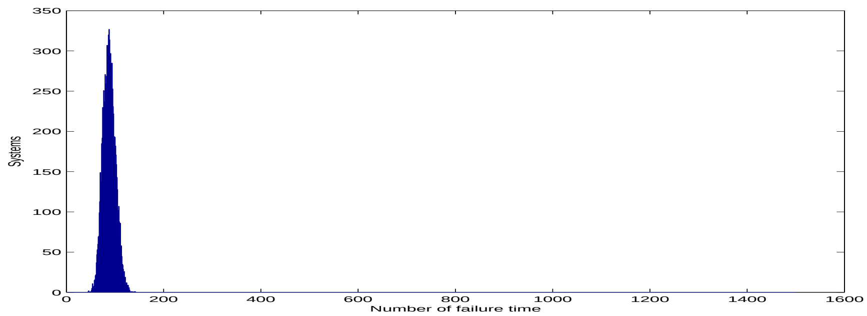

Figure 28: Histogram of Number of failure Vs Systems; λ=2, β=1.5,γ=0.01,10000 data sets and m=20 systems per data sets

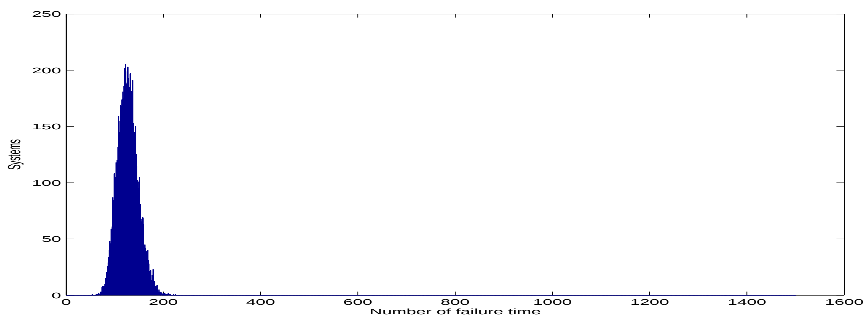

Figure 29: Histogram of Number of failure Vs Systems; λ=2, β=1.5,γ=0.02,10000 data sets and m=20 systems per data sets

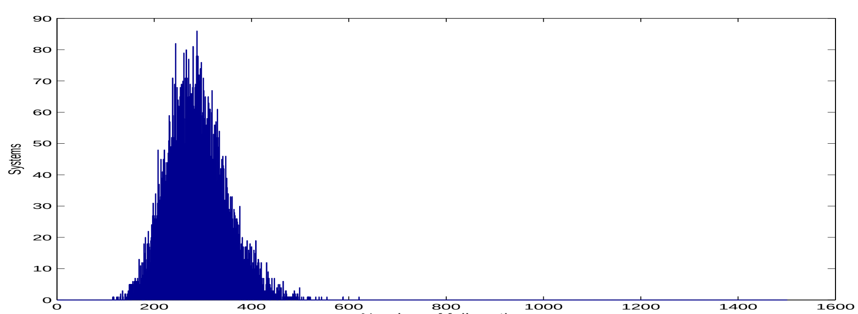

Figure 30: Histogram of Number of failure Vs Systems; λ=2, β=1.5,γ=0.04,10000 data sets and m=20 systems per data sets

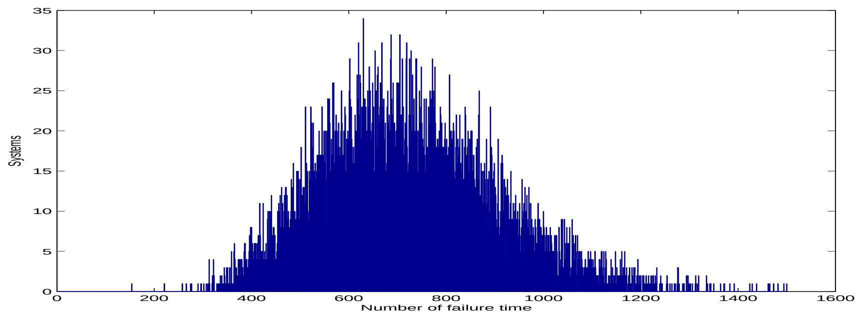

Figure 31: Histogram of Number of failure Vs Systems; λ=2, β=1.5,γ=0.06,10000 data sets and m=20 systems per data sets

10 Conclusion

In this paper, interrelation between fraility and dynamic models have been investigated. We have considered parameter δ as measure of fraility and γ as measure of dynamic behaviour. Moreover, these, parameters are considered as the main focus of the study and to see the difference from the baseline model. We were forced to use a smaller γ than δ to have a reasonable number of failures for dynamic behaviour. Unlike δ, as γ increases, a decrease in λ is seen, but the converse for β due to the fact that higher number of failures happen in system. Both features dynamic behaviour and frailty have great impact on analyses, avoiding wrong conclusions occurring if they are not taken into account. We have considered dynamic and fraility data sets and estimate their parameters by the fraility likelihood function. Often the true γ value and δ estimates are close to being equal. Hence, we can say that fraility models may be viewed as an alternative to dynamic models.

11 Acknowledgement

I am deeply grateful to Professor Bo Henry Lindqvist for his excellent guidance, constructive comments and encouragement throughout the study. His warmth for my topic and tremendous expertise is very much appreciated.

12 Literature and references

1. Aalen, O. , Borgan, O. and Gjessing, H. *Survival and Event History Analysis: A Process Point of View.* Springer, 2008.

2. Aalen, O. , Fosen, J. , Weedon-Fekjr, H., Borgan, O. and Husebye, E. , "Dynamic Analysis of Multivariate Failure Time Data", Biometrics, Vol. 60, No. 3 (Sep., 2004), pp. 764-773.

3. Babykina, G. and Couallier, V.,"Modelling Recurrent Events for Repairable Systems Under Worse Than Old Assumption", 2009.

4. Bhalotra, S. and Soest, A. V. , "Birth Spacing, Fertility and Neonatal Mortality in India: Dynamics, Frailty and Fecundity", Journal of Econometrics 143(2), pp. 274-290, April, 2008.

5. Borgan, O. , Fiaccone, R. L. , Henderson, R. , Barreto, M. L. , "Dynamic Analysis of Recurrent Event Data with Missing Observations, with Application to Infant Diarrhoea in Brazil", Scandinavian Journal of Statistics, Vol. 34, No. 1 (March 2007), pp. 53-69.

6. Coetzee, J. L. , "The role of NHPP models in the practical analysis of maintenance failure data", Reliability Engineering and System Safety 56 (1997) 161-168.

7. Cook, R. J. and Lawless, J. F. *The Statistical Analysis of Recurrent Events.* Springer, 2006.

8. Crow, L. H., "Estimation Procedures of the Duane Model", Proceedings US AMSAA Reliability Growth Symposium, aberdeen Proving Ground, MD, September 1972.

9. Crow, L. H.,"Evaluating the reliability of repairable systems", Proceedings Annual Reliability and Maintainability Symposium, IEEE, pp. 275-279, 1990.

10. Crow, L. H. , "New International Standards on reliability growth", Reliability and Maintainability Symposium. Annual Proceedings, 1991.

11. Crow, L. H., "Planning a Reliability Growth Program Utilizing Historical Data", Reliability and Maintainability Symposium, January, 2011.

12. Crow, L. H. and Shimi, I. N. , "Maximum Likelihood Estimation of Life-Distributions from Renewal Testing", The AAnnals of Mathematical Statistics, Vol. 43, No.6 (December, 1972), pp. 1827-1838.

13. Duane, J. H., "Learning curve approach tp reliability monitoring", IEEE Transactions on Aerospace 2, pp. 563-566,1964.

14. Elgmati, E. A. , "Dynamic Models versus Frailty Models for Recurrent Event Data", World Academy of Science, Engineering and Technology, issue: 0076, p.517-521, April 2013.

15. Engelhardt, M. and Lee, J. B. *Statistical Analysis of a Weibull process with left censored data*, 1992 kluwer academic publishers. Printed in the Netherlands.

16. Gamiz, M. L. , Kulasekera, K. B. , Limnios, N. , Lindqvist, B. H. *Applied Nonparametric Statistics in Reliability.* Springer, February, 2011.

17. Guilley, E. , Ghisletta, P. , Armi, F. , Berchtold, A. , Lalive C. , d'Epinay, Michel J. and Ribaupierre, A. , "Dynamics of Fraility and ADL Dependence in a Five-Year Longitudinal Study of Octogenarians, Research on Aging", originally published online 31 January 2008, DOI: 10.1177/0164027507312115.

18. Kalbfleisch, J. D. and Prentice, R. L. *The Statistical Analysis of Failure Time Data*, Second edition, 2002.

19. Larsen, R. J. and Marx, M. L. *An Introduction to Mathematical Statistics and Its applications.* 2012, fifth edition.

20. Lawless, J. F. , "Regression Methods for Poisson Process Data. Journal of the American Statistical Association", Vol.82,No.399(Sep.,1987), pp.808-815.

21. Lee, E. T. and Wang, J. W. *Statistical methods for survival data analysis*, Third edition, 2003.

22. Le Gat, Y., "Extending the Yule process to model recurrent pipe failures in water supply networks", Urban Water Journal (2013), DOI: 10.1080/1573062X.2013.783088

23. Meeker, W. Q. and Escobar, L. A. *Statistical Methods for Reliability Data*. July, 24, 1998.

24. Puts, M. T. E. , Lips, P. , Deeg, D. J. H. , "Static and dynamic measures of fraility predicted decline in performance-based and self-reported physical functioning", Journal of Clinical Epidemiology,vol 58, pp-1188-1198, 2005.

25. Rannestad, B., "Exact Statistical Inference in Non-homogeneous Poisson Processes, based on Simulation", Norwegian University of Science and Technology (NTNU), Department of Mathematical Sciences. July, 2007.

26. Rausand, M. and Hyland, A. *System Reliability Theory Models, statistical Methods, and Applications*, Second Edition, 2004.

27. Ruggeri, F. R. and Antonio P., "On the reliability of repairable systems: Methods and Applications", CNR-IMATI, Via A. M. Ampere 56, I-20131 Milano.

28. Ruggeri, F. and Sivaganesan, S. , "On modeling change points in Non-homogeneous Poisson processes", Statistical inference for stochastic processes (2005) 8:311-329 DOI 10.1007/s11203-005-6100-y.

29. Steenbergen, M. R. , "Maximum likelihood programming in R", January 2006.

30. Vaupel, J. , Manton K. and Stallard, E. , "The impact of heterogeneity in individual fraility on the dynamics of mortality", Demography, vol. 16, pp. 439-454, 1979.

31. Yan, J. , "Analysis of Multivariate Survival Data", Journal of the American Statistical Association, 100:469, 354-355, DOI: 10.1198/jasa.2005.s10, 2005.