Chapter - IV

ESTIMATION OF PARAMETERS FOR A NHPP MODEL

§ 4.1 Introduction

In previous chapter we discussed inference related to the parameters of JM model. This chapter is devoted to the study of parameter estimation for a NHPP model. A stochastic model for which the software failure phenomenon based on a NHPP is suggested by Goel and Okumoto (1979). The difference between this model and JM model is that- for this model, the total number of bugs in the software is a random variable with mean a, while it is constant in the JM model. Secondly, for this model the times between failures are assumed dependent, while the JM model assumes independence. Like JM model, this model is also used widely.

In section (4.2) we give some preliminary results related to NHPP, which are useful for our further study. In this section a joint pdf of $\underline{S} = (S_1, S_2, \ldots, S_n)$ is obtained, where S_k denote the time to k^{th} failure. Also the distribution of residual number of faults and conditional reliability is obtained. Section (4.3) deals with MLE's of the parameters of GO79b model. In this section we obtain a necessary and sufficient condition for existence and uniqueness of MLE. Distribution of residual number of bugs and conditional reliability for GO79b model is obtained in the same section. In section (4.4) we study the modified GO79b model given by Hossian and Dahiya (1993). Since there is a positive probability, for no solution for ML equations inside the parameter space for GO79b model and due to the difficulty of improper pdf, Hossian and Dahiya modified GO79b model by introducing a control variable c in the pdf.Hossian and Dahiya claims that, the modified GO79b model eases the condition of existence of solution for ML equation and gives a better estimate for the parameters. Distribution of residual number of bugs and conditional reliability are also obtained for this model.

§ 4.2 Some results related with NHPP

Let a sequence of random variables $\{T_i\}_{i=1}^{\infty}$ represents a sequence of time between failures associated with the counting process defined in (2.6) of Chapter - II. Then T_i denote the time between $(i-1)^{th}$ and i^{th} failure. Define,

$$S_k = \sum_{i=1}^{k} T_i$$

Which denote the time to 'k' failures.

4.2.1 The joint density function of $\underline{S} = (S_1, S_2, \ldots, S_n)$.

Lemma (4.2.1) : The joint density function of $\underline{S} = (S_1, S_2, \ldots, S_n)$ is given by,

$$f_{\underline{S}}(\underline{s}) = \exp[-\mu(s_n)] \prod_{i=1}^{n} \mu'(s_i) \qquad (4.2.1)$$

where, $\mu(t) = E[M(t)]$ and $\mu'(t) = d\mu(t)/dt$.

Proof : We have,

$$P[S_1 > t] = P[T_1 > t]$$

$$= P[\text{ no failure before } t\ ]$$

Since the number of failures in interval (0,t] has a Poisson distribution with mean $\mu(t)$, we have,

$$P[S_1 > t] = \exp[-\mu(t)]$$

Therefore the pdf of S_1 is,

$$f_{S_1}(s_1) = \mu'(t) \exp[-\mu(t)] \qquad (4.2.2)$$

In order to obtain the joint pdf of (S_1, S_2), we proceed as follows :

Given $S_1 = s_1$, we have,

$$P[S_2 > (t+s_1) | S_1 = s_1] = P[\text{ no failure in } (s_1,\ s_1+t)\ ]$$

$$= P[\ M(s_1+t) - M(s_1)\ ]$$

$$= \exp[\ -\ \{\mu(s_1+t) - \mu(s_1)\ \}]$$

due to independent increment property of the Poisson process.

Thus the conditional density function of S_2 given S_1 is,

$$f_{S_2|S_1}(s_2) = \mu'(s_2)\ \exp[\ -\ \{\mu(s_2) - \mu(s_1)\ \}] \tag{4.2.3}$$

Hence, the joint pdf of (S_1, S_2) is,

$$f_{S_1,S_2}(s_1,s_2) = f_{S_1}(s_1)\ f_{S_2|S_1}(s_2)$$

$$= \Big\{\mu'(s_1)\ \exp[-\mu(s_1)]\Big\}\ \Big\{\mu'(s_2)\ \exp[-\{\mu(s_2)-\mu(s_1)\}]\Big\}$$

$$= \mu'(s_1)\ \ \mu'(s_2)\ \exp[-\mu(s_2)] \tag{4.2.4}$$

Repeating the above argument for $S_3,\ldots,S_n$; we obtain the joint density (4.2.1).

Hence the proof.

4.2.2 Distribution of residual number of bugs

Suppose that 'y' faults have been found by time t_o. Since the Poisson process $\{M(t),\ t\geq 0\}$ has independent increments, the conditional distribution of M(t) given $M(t_o)=y$ for $t > t_o$ is obtained as,

$$P[\ M(t)=x\ |\ M(t_o)=y\ ] = P[\ \{M(t) - M(t_o)\}= (x-y)\ ]$$

$$= \{\mu(t) - \mu(t_o)\}^{(x-y)}\ \frac{\exp[-(\mu(t)-\mu(t_o))]}{(x-y)!} \tag{4.2.5}$$

which is the distribution of additional failures during $(t_o, t]$.

Let $\bar{M}(t)$ be the number of errors remaining in the system at time t, then $\bar{M}(t) = \{M(\infty) - M(t)\}$. Thus substituting $t_o = t$ and $t=\infty$ in the equation (4.2.5) and since $\bar{M}(t)$ and M(t) are independent, we have,

$$P[\ \bar{M}(t)=x\ ] = P[\ \bar{M}(t)=x \mid M(t)=y\ ]$$

$$= \{\mu(\infty) - \mu(t)\}^{(x-y)} \frac{\exp[-(\mu(\infty)-\mu(t))]}{(x-y)!}$$

$$= \nu(t)^m \exp[-\nu(t)]/\ m! \qquad ;\text{for } m = 0,1,2,... \qquad (4.2.6)$$

where,

$$\nu(t) = E[\ \bar{M}(t)\ ]$$

$$= E[\ M(\infty)\ ] - E[\ M(t)\ ]$$

$$= \mu(\infty) - \mu(t) \qquad (4.2.7)$$

and m = (x-y).

The distribution of $\bar{M}(t)$ given in (4.2.6) is important in deciding the software system can be released or not. Such a decision can be based on the number of faults remaining in the software, because it plays an important role in software reliability.

4.2.3 Conditional reliability function

The conditional reliability function given that 'n' failures occurred at time S_n is obtained as,

$$R(t|S_n= s_n) = P[\ T>t \mid S_n= s_n\ ]$$

$$= P[\ \text{no failure in } (s_n,\ s_n+t)\ ]$$

Thus,

$$R(t|S_n = s_n) = P[\{M(s_n+t) - M(s_n)\} = 0]$$

Using equation (4.2.5) we have,

$$R(t|S_n = s_n) = \exp[-\{\mu(s_n+t) - \mu(s_n)\}] \qquad (4.2.8)$$

§ 4.3 Estimation of parameters for GO79b model

For GO79b model given in section (2.9), we have,

$$\mu(t) = a\ (1-\exp[-bt])$$

and

$$\mu'(t) = a\ b\ \exp[-bt]$$

From lemma (4.2.1), the likelihood function for given $\underline{s}=(s_1, s_2, \ldots, s_n)$ is,

$$L(a,b|\underline{s}) = \exp[-a\ (1 - \exp[-b\ s_n])] \prod_{i=1}^{n} \left\{ a\ b\ \exp[-\ b\ s_i] \right\}$$

Hence the log-likelihood is,

$$\mathbb{L}(a,b|\underline{s}) = \log L(a,b|\underline{s})$$

$$= \left\{-a\ (1 - \exp[-b\ s_n])\right\} + n\log(a) + n\log(b) - b\sum_{i=1}^{n} s_i \qquad (4.3.1)$$

The MLE's are the values of a and b that maximize equation (4.3.1). Differentiating equation (4.3.1) w.r.t. a and b separately and equating to zero we obtain the following ML equations.

$$n/a - (1 - \exp[- b\, s_n] = 0 \tag{4.3.2}$$

and

$$n/b - a\, s_n \exp[- b\, s_n] - \sum_{i=1}^{n} s_i = 0. \tag{4.3.3}$$

Substituting (4.3.2) in (4.3.3) we have,

$$\frac{n}{b} - \frac{n\, s_n \exp[-bs_n]}{(1-\exp[-bs_n])} - \sum_{i=1}^{n} s_i = 0$$

$$\frac{n}{b} - \frac{n\, s_n}{(\exp[\, bs_n] -1)} - \sum_{i=1}^{n} s_i = 0$$

$$\frac{1}{b} - \frac{s_n}{(\exp[\, bs_n] -1)} - \bar{s} = 0$$

$$\frac{\exp[bs_n] - 1 - bs_n}{b(\exp[\, bs_n] -1)} - \bar{s} = 0 \tag{4.3.4}$$

where $\bar{s} = \sum_{i=1}^{n} s_i \,/\, n$.

Now we define,

$$g(b) = \frac{\exp[bs_n] - 1 - bs_n}{b(\exp[\, bs_n] -1)} - \bar{s} \tag{4.3.5}$$

Now first we solve g(b) = 0 for 'b' and the same in equation (4.3.2) we get 'a'. In order to see the existence of b such that g(b) = 0 we study behavior of g(b) in the following lemma.

Lemma (4.3.1) : The function g(b) is a decreasing function of b.

Proof : Differentiating equation (4.3.5) w.r.t. b we have,

$$g'(b) = \frac{d}{db}\left[\, g(b)\right]$$

$$= \Big[\{b(\exp[bs_n]-1)(\exp[bs_n]s_n-s_n)\} - \{(\exp[b\ s_n]-1-b\ s_n)$$

$$[\{\exp[b\ s_n]-1\} + b\ s_n\exp[b\ s_n]\Big] / [b^2\ \{\exp[b\ s_n]-1\}^2]$$

$$= \Big[\{bs_n(\exp[bs_n]-1)^2\} - \Big((\exp[b\ s_n]-1)^2 + bs_n\ \exp[bs_n]$$

$$\{\exp[b\ s_n]-1\} - b\ s_n(\exp[b\ s_n]-1) - b^2 s_n^2\ \exp[bs_n]\ \Big)\Big]$$

$$/\ [b^2\ \{\exp[b\ s_n]-1\}^2]$$

$$= \Big[\{bs_n(\exp[bs_n]-1)^2\} - \Big((\exp[b\ s_n]-1)^2 + bs_n\{\exp[bs_n]-1\}^2$$

$$- b^2 s_n^2\ \exp[bs_n]\ \Big)\Big]\ /\ [b^2\ \{\exp[b\ s_n]-1\}^2]$$

$$= \Big(2\ \exp[bs_n] - \exp[2bs_n] - 1 + b^2 s_n^2\ \exp[bs_n]\Big)$$

$$/\ [b^2\ \{\exp[b\ s_n]-1\}^2]$$

$$= \Big(2 - \exp[bs_n] - \exp[-bs_n] + b^2 s_n^2\ \Big)\ \exp[bs_n]$$

$$/\ [b^2\ \{\exp[b\ s_n]-1\}^2]$$

$$= \Big(- \{\exp[bs_n] + \exp[-bs_n]\} + \{2+b^2 s_n^2\}\Big)\ \exp[bs_n]$$

$$/\ [b^2\ \{\exp[b\ s_n]-1\}^2]$$

Using result 1(b) given in appendix A, we write the right side of above equation,

$$g'(b) \le \left(- \{2+b^2 s_n^2\} + \{2+b^2 s_n^2\} \right) \exp[b s_n] \,/\, [b^2 \{\exp[b\, s_n]-1\}^2]$$

$$= 0$$

Which shows that, the function g(b) is a decreasing function of b.

Theorem (4.3.1) : The necessary and sufficient condition for equations,

$$n/a - (1 - \exp[-\, b\, s_n\,] = 0 \qquad (4.3.2)$$

and

$$n/b - a\, s_n \exp[-\, b\, s_n\,] - \sum_{i=1}^{n} s_i = 0. \qquad (4.3.3)$$

to have finite positive root is

$$s_n > 2\, \bar{s} \qquad (4.3.6)$$

where $\bar{s} = \sum_{i=1}^{n} s_i \,/\, n$.

Proof : From equations (4.3.2) and (4.3.3) we have,

$$\frac{\exp[b s_n] - 1 - b s_n}{b(\exp[\, b s_n] - 1)} - \bar{s} = 0 \qquad (4.3.4)$$

It suffices to show that (4.3.6) is the necessary and sufficient condition for the existence of positive root of (4.3.4). The l.h.s. of equation (4.3.4) is the function g(b) defined in (4.3.5).

Now,

$$\lim_{b \to o} g(b) = \lim_{b \to o} \left[\frac{\exp(bs_n) - 1 - bs_n}{b(\exp(bs_n) - 1)} - \bar{s} \right]$$

Applying L'Hospital's rule we obtain,

$$\lim_{b \to o} g(b) = \lim_{b \to o} \left[\frac{\exp(bs_n) s_n - s_n}{bs_n \exp(bs_n) + (\exp(bs_n) - 1)} - \bar{s} \right]$$

Again applying L'Hospital's rule we have,

$$\lim_{b \to o} g(b) = \lim_{b \to o} \left[\frac{s_n^2 \exp(bs_n)}{bs_n^2 \exp(bs_n) + \exp(bs_n) s_n + \exp(bs_n) s_n} - \bar{s} \right]$$

$$= \lim_{b \to o} \left[\frac{s_n}{(bs_n + 2)} - \bar{s} \right]$$

$$= \frac{s_n}{2} - \bar{s} \qquad (4.3.7)$$

and

$$\lim_{b \to \infty} g(b) = \lim_{b \to \infty} \left[\frac{s_n}{(bs_n + 2)} - \bar{s} \right]$$

$$= (- \bar{s}) < 0 \qquad (4.3.8)$$

Since, g(b) is decreasing function in b, using (4.3.7) and (4.3.8) we conclude that, the equation (4.3.6) has a positive root if and only if

$$\lim_{b \to o} g(b) > 0$$

This implies that,

$s_n /2 > \bar{s}$ or $s_n > 2\bar{s}$.

Hence the proof.

Since g(b) is decreasing for all b≥0 [by lemma (4.3.1), g(b)=0 has unique solution if and only if $s_n > 2\bar{s}$. By substituting this root into equation (4.3.2) we obtain the unique solution for a. Hence the MLE's for a and b are unique if and only if the condition $s_n > 2\bar{s}$ is satisfied.

4.3.1 Distribution of $\bar{M}(t)$ for GO79b Model

For this model we have,

$\mu(t) = a\ (1-\exp[-bt])$

Using (4.2.6), we obtain the distribution of number of failures remained ; $\bar{M}(t)$ in the software as,

$P[\ \bar{M}(t)=x\ ] = \nu(t)^m \exp[-\nu(t)]/\ m!$; for m = 0,1,2,...

where,

$$\begin{aligned}\nu(t) &= E\ [\ \bar{M}(t)\ ] \\ &= E\ [\ M(\infty)\ ] - E\ [\ M(t)\ ] \\ &= \mu(\infty) - \mu(t) \\ &= a - a\ (1-\exp[-bt]) \\ &= a\ \exp[-bt] \qquad (4.3.9)\end{aligned}$$

and m = (x-y) ; x failures are remained and y failures are observed.

4.3.2 Conditional reliability function

Using (4.2.8) we obtain the conditional reliability function as,

$$R(t|S_n = s_n) = \exp[-\{\mu(s_n+t) - \mu(s_n)\}]$$

$$= \exp\left[-\{ a\ (1-\exp[-b(s_n+t)]) - a\ (1-\exp[-bs_n])\right]$$

$$= \exp\left[-\{ -\exp[-b(s_n+t)] + \exp[-bs_n]\}\ \right] \qquad (4.3.10)$$

In table (4.1), we report Maximum Likelihood Estimate for a and b. Also we report estimate for residual number of bugs for GO79b model. These estimates are computed for the data which collected from an automization project at the Dutch Aerospace Laboratory (NLR).

§ 4.4 Estimation of parameters for modified GO79b model

In the above section we have seen that, with positive probability, there is no solution for ML equations inside the parameter space. Because of this difficulty and the difficulty of improper pdf {time to first failure}, Hossian and Dahiya (1993) have suggested a modified NHPP model by introducing a control variable ´c´ into GO79b model. They denote this modified model as HD/G-O. The HD/G-O model does not get rid of aforesaid problems completely. But eases the condition of existence of solution for ML equations of GO79b model, reduces the probability mass at infinity and gives a better estimates of the model parameters.

Table 4.1

MLE & residual number of failures for GO79b model

Sr. No.	t_i	$s_k = \sum_{i=1}^{k} t_i$	$\bar{s}_k = \sum_{i=1}^{k} s_i$	$\hat{a}$	$\hat{b}$ (10^{-3})	$\hat{\nu}(s_k)$
1	880					
2	3430					
3	2860					
4	11760					
5	4750	23680	54970	14.3041	1.8163	9
6	240					
7	2300					
8	8570					
9	4620					
10	1060	40470	219780	*	*	*
11	3820					
12	14800					
13	1770					
14	24270					
15	4800	89930	559080	23.1321	1.1625	8
16	470					
17	40					
18	10170					
19	1120					
20	980	102710	1044970	*	*	*
21	24300					
22	17500					
23	4450					
24	4860					
25	640	154460	1773730	64.5329	0.3172	40
26	3990					
27	26840					
28	2270					
29	200					
30	39180	256940	2719730	35.4919	0.7262	5
31	14910					
32	14670					
33	16310					
34	38410					
35	1120	312360	4114530	44.5232	0.4938	10
36	30560					
37	6210					
38	120					
39	20210					
40	26400	395860	5921150	50.4479	0.3978	10

In the table (4.1), * indicates that the condition for existence of MLE is not satisfied and hence MLE does not exist for a, b and residual number of bugs.

4.4.1 Development of HD/G-O model

In the GO79b model the pdf {time to first failure} is,

$$f(t) = \mu'(t) \exp[-\mu(t)] \qquad \{\text{from } (4.2.2)\}$$

$$= (a\, b\, e^{-bt}) \exp[-a(1 - e^{-bt})]$$

$$= (a\, b\, e^{-bt}) \exp[a\, e^{-bt}]/ e^{a} \qquad (4.4.1)$$

Now consider,

$$\int_{o}^{\infty} f(t)\, dt = \int_{o}^{\infty} \left[(a\, b\, e^{-bt}) \exp[a\, e^{-bt}]/ e^{a}\right] dt$$

Substituting $(a\, e^{-bt}) = y$ **we get,**

$$\int_{o}^{\infty} f(t)\, dt = \int_{a}^{o} \left[\exp[y]/ e^{a} \right] (-dy)$$

$$= \int_{o}^{a} [\exp[y]/ e^{a}]\, dy$$

$$= [e^{a} - 1]/ e^{a}$$

Which implies that f(t) is improper pdf. The corresponding proper pdf is,

$$f_1(t) = f(t) \,/\, \left([e^a - 1]/\, e^a\right)$$

$$= (a\, b\, e^{-bt})\, \exp[a\, e^{-bt}]/\, [e^a - 1] \qquad (4.4.2)$$

for a, b, t > 0

This led us to believe that with-

$$f_1(t) = (a\, b\, e^{-bt})\, \exp[a\, e^{-bt}]/\, [e^a - c] \qquad (4.4.3)$$

for a, b, t > 0 and 0 ≤ c ≤ 1.

we might do better than the GO79b model. The corresponding mean value function of this model is,

$$\mu(t) = \log\{(e^a - c)/\, (\exp[a\, e^{-bt}] - c)\} \qquad (4.4.4)$$

for a, b, t > 0 and 0 ≤ c ≤ 1.

The model (4.4.3), when c=0, is GO79b model and when c=1, the corresponding pdf is proper as given in (4.4.2). In this kind of model we anticipate that $\mu(\infty)$ is finite. But when c=1, $m(\infty)=\infty$; giving rise to a new problem in determining the mean total number of failures in the system. So we try to search for a $\{c: 0 \le c \le 1\}$ that gives a better (in some sense) estimated mean number of time failures in the system than the GO79b model estimate. Therefore we modify the condition on c as

$$o \le c < 1 \qquad (4.4.5)$$

From (4.4.4) we have, $\mu(0) = 0$ and

$\mu(\infty) = \log\{(e^{a} - c)/ (1 - c)\}$

$\mu(\infty)$ is the mean failures to be detected eventually and finite.

Lemma (4.4.1) : The conditional cdf, $\bar{F}(x|t)$ corresponding to HD/GO79b model is DFR.

Proof : From pdf given in (4.4.3) we have,

$$F(x) = \int_{o}^{x} f_{1}(t)\, dt$$

Thus,

$$F(x) = \int_{o}^{x} \left\{ (a\ b\ e^{-bt}) \exp[a\ e^{-bt}]/ [e^{a}-c] \right\} dt$$

Putting $(a\ e^{-bt}) = y$ we have,

$$F(x) = \int_{a}^{ae^{-bx}} \left\{ e^{y}/ [e^{a}-c] \right\} (-dy)$$

$$= \left\{ e^{a} - \exp[a\ e^{-bx}] \right\} / [e^{a}-c] \qquad ; \text{for } x \geq 0 \qquad (4.4.6)$$

Hence,

$$\bar{F}(x) = 1 - F(x)$$

$$= \left\{ \exp[a\ e^{-bx}] -c \right\} / [e^{a}-c] \qquad ; \text{for } x \geq \qquad (4.4.7)$$

Therefore the corresponding conditional reliability of a unit age t is,

$\bar{F}(x|t) = \bar{F}(x+t)/\bar{F}(t)$

$$= \left(\exp[a\ e^{-b(x+t)}] - c \right) / \left(\exp[a\ e^{-bt}] - c \right) \qquad (4.4.8)$$

To show that $\bar{F}(x|t)$ is increasing in t, we show that, $(d/dt)[\log\bar{F}(x|t)]$ is positive for all $x > 0$.

We have from (4.4.8)

$$\log[\bar{F}(x|t)] = \log\left(\exp[a\ e^{-b(x+t)}] - c\right) - \log\left(\exp[a\ e^{-bt}] - c\right)$$

Thus,

$$\frac{d}{dt}\left(\bar{F}(x|t)\right) = \left(-ab\ \exp[a\ e^{-b(x+t)}]\ e^{-b(x+t)}\right) / \left(\exp[a\ e^{-b(x+t)}] - c\right)$$

$$+ \left(ab\ \exp[a\ e^{-bt}]\ e^{-bt}\right) / \left(\exp[a\ e^{-bt}] - c\right)$$

Let $u = e^{-bx}$ and $v = e^{-bt}$; then we have, $uv = e^{-b(x+t)}$ and hence,

$$\frac{d}{dt}\left(\bar{F}(x|t)\right) = \left(-ab\ \exp[a\ uv]\ uv\right) / \left(\exp[a\ uv] - c\right)$$

$$+ \left(ab\ \exp[a\ v]\ v\right) / \left(\exp[a\ v] - c\right)$$

$$\frac{d}{dt}\left(\bar{F}(x|t)\right) = abv \left\{ \left(-\exp[a\ uv]\ u\right) / \left(\exp[a\ uv] - c\right) + \left(\exp[a\ v]\right) / \left(\exp[a\ v] - c\right) \right\}$$

We observe that, for given c [$0 \le c < 1$],

$$\left(\exp[a\ uv]\ u \right) / \left(\exp[a\ uv] - c \right)$$

is decreasing function of u.

As $0 \le u \le 1$ for all $x \ge 0$, we can write,

$$\frac{d}{dt}\left(\bar{F}(x|t)\right) \ge abv \left\{ \left(- \exp[a\ v] \right) / \left(\exp[a\ v] - c \right) + \left(\exp[a\ v] \right) / \left(\exp[a\ v] - c \right) \right\}$$

$$= 0$$

Hence the proof.

4.4.2 Estimation of parameters

For HD/G-O model we have,

$$\mu(t) = \log\{(e^{a} - c)/ (\exp[a\ e^{-bt}] - c)\}$$

$$= \log\{(e^{a} - c)\} - \log\{(\exp[a\ e^{-bt}] - c)\}$$

Therefore,

$$\mu'(t) = \frac{d}{dt}\left[\mu(t)\right]$$

$$= \frac{\exp[ae^{-bt}]\ a\ e^{-bt}(-b)}{(\exp[ae^{-bt}] - c)}$$

$$= \frac{(ab\ \exp[ae^{-bt} - bt])}{(\exp[ae^{-bt}] - c)}$$

Let $\Phi(a,b,t) = a\, e^{-bt}$ then,

$\mu(t) = \log\{(e^{a} - c)\} - \log\{(\exp[\Phi(a,b,t)] - c)\}$ and

$$\mu'(t) = \frac{(ab\, \exp[\Phi(a,b,t)] - bt)}{(\exp[\Phi(a,b,t)] - c)}$$

Thus from lemma (4.2.1), the likelihood function for given $\underline{s} = (s_1, s_2, \ldots, s_n)$ is,

$$L(a,b,c|\underline{s}) = \exp[-\mu(s_n)] \prod_{i=1}^{n} \left\{ \mu'(s_i) \right\}$$

$$= \frac{(\exp[\Phi(a,b,s_n)] - c)}{(e^{a} - c)} \prod_{i=1}^{n} \left[\frac{\exp[\Phi(a,b,s_i) - bs_i]}{(\exp[\Phi(a,b,s_i)] - c)}\, ab \right]$$

Hence the log-likelihood, for given c such that $0 \le c < 1$, is,

$$\mathbb{L}(a,b|c,\underline{s}) = \log L(a,b|c,\underline{s})$$

$$= n \log(a) + n \log(b) - \log(e^{a} - c) + \sum_{i=1}^{n} \Phi(a,b,s_i)$$

$$- b \sum_{i=1}^{n} s_i - \sum_{i=1}^{n-1} \log\{\exp[-\Phi(a,b,s_i)] - c\} \qquad (4.4.9)$$

The MLE's are the values of a and b that maximize equation (4.4.1). Differentiating equation (4.4.1) w.r.t. a and b separately and equating to zero we obtain the following ML equations. Here, $\frac{d}{da}\left[\mathbb{L}(a,b|c,\underline{s})\right] = 0$ implies that,

$$\frac{n}{a} - \frac{e^{a}}{(e^{a}-c)} + \sum_{i=1}^{n} \exp[-bs_i] - \sum_{i=1}^{n-1}\left[\frac{\exp[\Phi(a,b,s_i)]\ \exp[-bs_i]}{(\exp[\Phi(a,b,s_i)] - c)}\right] = 0$$

$$\frac{n}{a} - \frac{e^{a}}{(e^{a}-c)} + \sum_{i=1}^{n} \Phi(a,b,s_i)/a - \sum_{i=1}^{n-1}\left[\frac{\exp[\Phi(a,b,s_i)]\ \exp[-bs_i]}{(\exp[\Phi(a,b,s_i)] - c)}\right] = 0$$

$$\frac{n}{a} - \frac{e^{a}}{(e^{a}-c)} + \Phi(a,b,s_n)/a$$

$$+ \sum_{i=1}^{n-1}\left[\frac{(\exp[\Phi(a,b,s_i)] - c) - \exp[\Phi(a,b,s_i)]}{(\exp[\Phi(a,b,s_i)] - c)}\right]\Phi(a,b,s_i)/a = 0$$

$$\frac{n}{a} - \frac{1}{(1-ce^{-a})} + \Phi(a,b,s_n)/a - \sum_{i=1}^{n-1}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c)}\right] = 0$$

(4.4.10)

and $\frac{d}{db}\left[\mathbb{L}(a,b|c,\underline{s})\right] = 0$ implies that,

$$\frac{n}{b} - \sum_{i=1}^{n} \Phi(a,b,s_i)\ s_i - \sum_{i=1}^{n} s_i + \sum_{i=1}^{n-1}\left[\frac{\exp[\Phi(a,b,s_i)]\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c)}\right]s_i = 0$$

$$\frac{n}{b} - s_n\ \Phi(a,b,s_n) - \sum_{i=1}^{n} s_i$$

$$- \sum_{i=1}^{n-1}\left[1 - \frac{\exp[\Phi(a,b,s_i)]}{(\exp[\Phi(a,b,s_i)] - c)}\right]s_i\Phi(a,b,s_i) = 0$$

$$\frac{n}{b} - s_n\ \Phi(a,b,s_n) - \sum_{i=1}^{n} s_i + \sum_{i=1}^{n-1}\left[\frac{c}{(\exp[\Phi(a,b,s_i)] - c)}\right]s_i\Phi(a,b,s_i) = 0$$

(4.4.11)

Lemma (4.4.2) : Let

$$f_b(a) = \frac{n}{a} - \frac{1}{(1-ce^{-a})} + \Phi(a,b,s_n)/a - \sum_{i=1}^{n-1}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c\ )}\right] \tag{4.4.12}$$

then for given b, the function $f_b(a)$ is decreasing function of a.

Proof : Differentiating equation (4.4.12) w.r.t. a, we get,

$$f_b'(a) = \frac{d}{db}\left[\ f_b'(a)\right]$$

$$= -\frac{n}{a^2} + \frac{c\ e^{-a}}{(1-ce^{-a})^2} - \sum_{i=1}^{n-1}\frac{d}{da}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c\ )}\right]$$

Thus,

$$f_b'(a) = -\frac{n}{a^2} + \frac{c\ e^{a}}{(e^{a}-c\ )^2} - \sum_{i=1}^{n-1}\frac{d}{da}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c\ )}\right]$$

Here,

$$\frac{d}{da}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c\ )}\right] = \frac{d}{da}\left[\frac{c\ \exp[-bs_i]}{(\exp[\Phi(a,b,s_i)] - c\ )}\right]$$

$$= \left[-\frac{c\ \exp[-2bs_i]\ \exp[\Phi(a,b,s_i)]}{(\exp[\Phi(a,b,s_i)] - c)^2}\right]$$

Therefore,

$$f_b'(a) = -\frac{n}{a^2} + \frac{c\ e^{a}}{(e^{a}-c\ )^2} + \sum_{i=1}^{n-1}\left[\frac{c\ \exp[-2bs_i]\ \exp[\Phi(a,b,s_i)]}{(\exp[\Phi(a,b,s_i)] - c)^2}\right]$$

Hence,

$$f_b'(a) = \left[\frac{c\,e^{a}}{(e^{a}-c)^2} - \frac{1}{a^2} \right]$$

$$+ \sum_{i=1}^{n-1} \left[\frac{c\,\exp[-2bs_i]\,\exp[\Phi(a,b,s_i)]}{(\exp[\Phi(a,b,s_i)]-c)^2} - \frac{1}{a^2} \right]$$

$$= \left[\frac{a^2 c\,e^{a} - (e^{a}-c)^2}{(e^{a}-c)^2\,a^2} \right]$$

$$+ \sum_{i=1}^{n-1} \left[\frac{a^2 c\,\exp[-2bs_i]\,\exp[\Phi(a,b,s_i)] - (\exp[\Phi(a,b,s_i)]-c)^2}{(\exp[\Phi(a,b,s_i)]-c)^2} \right]$$

(4.4.13)

Now the expression

$$ca^2e^a - (e^a-c)^2 = ca^2e^a - e^{2a} + 2ce^a - c^2$$

is increasing in c $(0 \le c < 1)$, therefore,

$$ca^2e^a - (e^a-c)^2 = ca^2e^a - e^{2a} + 2ce^a - c^2$$

$$\le a^2e^a - e^{2a} + 2e^a - 1$$

$$= [a^2 - e^a + 2 - e^{-a}]\, e^a$$

$$= [(2+a^2) - (e^a + e^{-a})]\, e^a$$

Using result 1(b) given in appendix A, we have,

$$ca^2e^a - (e^a-c)^2 \le [(e^a - e^{-a}) - (e^a + e^{-a})]\, e^a$$

Which is equivalent to,

$$ca^2 e^a - (e^a - c)^2 \le [\ -2\ e^{-a}]\ e^a$$

$$= -2 \ < 0$$

Similarly we can show that, the numerator of second term in expression (4.4.13) is negative. Thus we conclude that $f_b(a)$ defined above is a decreasing function

Lemma (4.4.3) : The upper and lower bounds of a is the solution of $f_b(a)=0$ are,

$$0 < a < n/(1-\exp[-bs_n]) \qquad (4.4.14)$$

Proof : In the above lemma we have seen that $f_b(a)$ is decreasing function of a (a>0).

Also we observe that, the term in the expression of $f_b(a)$,

$$\frac{1}{(1-ce^{-a})} - \sum_{i=1}^{n-1}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c\ )}\right]$$

is decreasing in a and

$$\lim_{a\to\infty}\left[\ \frac{1}{(1-ce^{-a})} - \sum_{i=1}^{n-1}\left[\frac{(c/a)\ \Phi(a,b,s_i)}{(\exp[\Phi(a,b,s_i)] - c\ )}\right]\right] = 1$$

Therefore,

$$f_b(a) \le \frac{n}{a} + \Phi(a,b,s_n)/a - 1 \qquad (4.4.15)$$

By (4.4.15) we have,

$$f_b(a) \le 0 \text{ if } \left[\frac{n}{a} + \Phi(a,b,s_n)/a - 1 \right] \le 0$$

Which implies that,

$$a \ge n/(1-\exp[-bs_n]) \tag{4.4.16}$$

and

$$f_b(a) \ge 0 \text{ as } a \longrightarrow 0. \tag{4.4.17}$$

From equations (4.4.16), (4.4.17) and lemma (4.4.1) we conclude that for given b, the root of a of $f_b(a)=0$ lies in the interval $(0,\ n(1-\exp[-bs_n])^{-1})$.

Theorem (4.4.1) : The sufficient condition for ML equations (4.4.10) and (4.4.11) to have finite root is $s_n > 2\,\bar{s}$ where

$$\bar{s} = \sum_{i=1}^{n} s_i \,/\, n\,.$$

Proof : Define,

$$g_a(b) = \frac{n}{b} - s_n\,\Phi(a,b,s_n) - \sum_{i=1}^{n} s_i + \sum_{i=1}^{n-1} \left[\frac{c}{(\exp[\Phi(a,b,s_i)]-c)} \right] s_i \Phi(a,b,s_i)$$

We need to determine the sufficient condition for $f_b(a)$ and $g_a(b)$ to have finite zeros. Here we observe that , $g_a(b)$ is decreasing function in a. To get the sufficient condition for the existence of finite zeros, we need to determine $\inf_a\{g_a(b)\}$ and

$\sup_a\{g_a(b)\}$.

$$\inf_a g_a(b) \geq \frac{n}{b} - n\, s_n \exp[-bs_n]/ (1-\exp[-bs_n]) - \sum_{i=1}^{n} s_i + \sum_{i=1}^{n-1}\left[\frac{c}{(\exp[\theta_i(b)] - c)}\right] s_i \theta_i(b)$$

$$= \frac{n}{b} - n\, s_n/(\exp[bs_n]-1) - \sum_{i=1}^{n} s_i + \sum_{i=1}^{n-1}\left[\frac{c}{(\exp[\theta_i(b)]-c)}\right] s_i \theta_i(b)$$

$$= n\left[\frac{\exp[bs_n]-1-bs_n}{b(\exp[bs_n]-1)}\right] - \sum_{i=1}^{n} s_i + \sum_{i=1}^{n-1}\left[\frac{c}{(\exp[\theta_i(b)]-c)}\right] s_i \theta_i(b)$$

where,

$$\theta_i(b) = \frac{n \exp[-bs_n]}{(1-\exp[-bs_n])}$$

and it tends to ∞ as b tends to zero.

$$\lim_{b \to 0} \inf_a g_a(b) \geq n \lim_{b \to 0}\left[\frac{\exp[bs_n]-1-bs_n}{b(\exp[bs_n]-1)}\right] - \sum_{i=1}^{n} s_i + \lim_{b \to 0} i \sum_{i=1}^{n-1}\left[\frac{c}{(\exp[\theta_i(b)]-c)}\right] s_i \theta_i(b)$$

Since,

$$\lim_{b \to 0}\left[\frac{\exp[bs_n]-1-bs_n}{b(\exp[bs_n]-1)}\right] = s_n/2 \qquad \{..\text{from eq.}(4.3.2)\}$$

and applying L'Hospital's rule,

$$\lim_{b \to 0} \sum_{i=1}^{n-1} \left[\frac{c}{(\exp[\theta_i(b)]-c)}\right] s_i \theta_i(b)$$

$$= \lim_{b \to 0} \sum_{i=1}^{n-1} \left[\frac{\theta_i(b)}{\exp[\theta_i(b)]\ \theta_i(b)}\right] s_i$$

$$= \lim_{b \to 0} \sum_{i=1}^{n-1} \left[\frac{1}{\exp[\theta_i(b)]}\right] s_i$$

$$= 0 \quad \{\text{due to } \theta_i(b) \to \infty \text{ as } b \to 0\}.$$

$$\lim_{b \to 0} \inf_a g_a(b) \geq n\, s_n/2 - \sum_{i=1}^{n} s_i \tag{4.4.18}$$

Again,

$$\sup_a g_a(b) \leq n/b - \sum_{i=1}^{n} s_i$$

and

$$\lim_{b \to \infty} \sup_a g_a(b) \leq \left(- \sum_{i=1}^{n} s_i\right) \tag{4.4.19}$$

From (4.4.18) and (4.4.19) it is clear that, we will have finite zeros if,

$$\lim_{b \to 0} \inf_a g_a(b) > 0$$

which implies that,

$$n\, s_n/2 - \sum_{i=1}^{n} s_i > 0 \quad \text{or} \quad s_n > 2\, \bar{s}.$$

Therefore, the ML equations have finite roots if $s_n > 2\, \bar{s}$. Hence the proof.

How to control the regularity variable c remains an open challenge. One possibility is to look at the sum of square of deviations of the observed and estimated values. In HD/GO79b model, as c increases from 0 to 1, the sum of square of deviation decreases. Therefore the method suggested by Hossian and Dahiya to control c is to use the mean square deviation (MSD) :

$$MSD = \sum_{i=1}^{n} (s_k - \hat{s}_k)^2 \qquad (4.4.20)$$

The estimated c is the minimum for which this MSD is almost unchanged for any further increase in c.

4.4.3 Distribution of $\bar{M}(t)$

For this model we have,

$$\mu(t) = \log\{(e^a - c)\} - \log\{(\exp[\Phi(a,b,t)] - c)\}$$

Using (4.2.6), we obtain the distribution of number of failures remained ; $\bar{M}(t)$ in the software as,

$$P[\ \bar{M}(t)=x\ ] = \nu(t)^m \exp[-\nu(t)]/\ m! \qquad ;\text{for } m = 0,1,2,...$$

where,

$$\begin{aligned}\nu(t) &= E\ [\ \bar{M}(t)\ ]\\ &= E\ [\ M(\infty)\ ] - E\ [\ M(t)\ ]\\ &= \mu(\infty) - \mu(t)\\ &= \log\{(e^a - c)/(1- c)\} - \log\{(e^a - c)/(\exp[\Phi(a,b,t)] - c)\}\\ &= \log\{(\exp[\Phi(a,b,t)] - c)/(1 - c)\} \qquad (4.4.21)\end{aligned}$$

and $m = (x-y)$; x failures are remained and y failures are observed.

4.4.4 Conditional reliability function

Using (4.2.8) we obtain the conditional reliability function as,

$$R(t|S_n = s_n) = \exp[-\{\mu(s_n+t) - \mu(s_n)\}]$$

$$= \exp\Big[-\Big\{\log\{(e^{\alpha} - c)\} - \log\{(\exp[\Phi(a,b,s_n+t)] - c)\} - \log\{(e^{\alpha} - c)\} + \log\{(\exp[\Phi(a,b,s_n)] - c)\}\Big\}\Big]$$

$$= \exp\Big[\log\{(\exp[\Phi(a,b,s_n+t)] - c)\} - \log\{(\exp[\Phi(a,b,s_n)] - c)\}\Big]$$

$$= \exp\Big[\log\{(\exp[\Phi(a,b,s_n+t)]-c)/(\exp[\Phi(a,b,s_n)]-c)\}\Big]$$

$$= \{(\exp[\Phi(a,b,s_n+t)]-c)/(\exp[\Phi(a,b,s_n)]-c)\} \quad (4.4.22)$$

4.5 Scope for Further Research

In practice, it is sometimes impossible to achieve "*Perfect*" debugging procedure. Whenever a fault is encountered, software engineer tries to remove the same from the software. Therefore we expect from our modeling strategies that, the failure rate should