



Half Logistic Software Reliability Growth Model

Dr.R.Satya Prasad *
Dept. of Computer Science Engineering
Acharya Nagarjuna University
profersp@gmail.com

Dr.R.R.L. Kantam
Department of Statistics
Acharya Nagarjuna University
kantam_rrl@rediffmail.com

Abstract: A software reliability growth model based on non-homogeneous Poisson process (NHPP) is derived with its mean value function satisfying conditions similar to those of Goel and Okumoto (1979) and Musa and Okumoto (1984). The model turned out to be the (folded) half logistic distribution of Balakrishnan (1985). Reliability performance measures of the model are presented. Its parameters are estimated by maximum likelihood and approximate maximum likelihood methods. Asymptotic dispersion matrix, interval estimation also are derived. The results are illustrated with software failure data. The performance of the proposed model is evaluated in relation that of Goel and Okumoto (1979)

Keyword: Count process, mean value function, intensity function, software reliability, maximum likelihood estimation.

I. INTRODUCTION

Software reliability is the probability of failure free operation of a computer program in a specified environment for a specified time. A failure is a departure of a program operation from program requirements. A software reliability model provides a general form in terms of a random process describing failures for characterizing software reliability or a related quantity as a function of experienced failures over time.

In this backdrop Musa and Okumoto (1984) suggested a model for software reliability measurement and named it as logarithmic Poisson execution time model (LPETM). Motivated by their approach we develop a new software reliability model in this paper and present its performance in measuring software reliability. The development of our model, the reliability measurement characteristics and estimation procedures associated with it are given in Sections 2, 3, 4 respectively. Illustration of the results and the evaluation of the proposed model with a live data are presented in section 5.

II. THE PROPOSED MODEL

Let $[N(t), t \geq 0], m(t), \lambda(t)$ be the counting process, mean value function and intensity function of a software failure phenomenon. LPETM was derived with the following assumptions:

Assumption (i): There is no failure observed at time $t = 0$. That is, $N(0) = 0$ with probability one.

Assumption (ii): The failure intensity $\lambda(t)$ will decrease exponentially with the expected number of failures experienced up to time 't' so that the relation between $m(t)$ and $\lambda(t)$ is

$$\lambda(t) = \lambda_0 \cdot e^{-\theta m(t)} \quad (1)$$

Where λ_0 and θ are the initial failure intensity and the rate of reduction in the normalized failure intensity per failure respectively.

Assumption (iii): For a small interval Δt the probabilities of one, more than one failure during $[t, t+\Delta t]$ are $\lambda(t) \cdot \Delta t + O(\Delta t)$ and $O(\Delta t)$, respectively, where $\frac{O(\Delta t)}{\Delta t} \rightarrow 0$ as $\Delta t \rightarrow 0$. Note that the probability of no failure during $[t, t+\Delta t]$ is given by $1 - \lambda(t)\Delta t + O(\Delta t)$. Using assumptions (i) and (ii) the mean value function and intensity function are derived to be

$$m(t) = \frac{1}{\theta} \log(\lambda_0 \theta t + 1) \quad (2)$$

$$\lambda(t) = \frac{\lambda_0}{(\lambda_0 \theta t + 1)} \quad (3)$$

It may be noted that $m(t) \rightarrow \infty$ as $t \rightarrow \infty$. Hence this model is also called 'infinite failures' model. Using assumptions (i) and (iii) the probability distribution of the stochastic process $N(t)$ is given by

$$P[N(t) = y] = \frac{[m(t)]^y}{y!} \cdot e^{-m(t)} \quad (4)$$

Where $m(t)$ is given by equation (2).

On lines of Goel and Okumoto (1979), let us specify that the mean value function $m(t)$ is finite valued, non decreasing, non negative and bounded with the boundary conditions

$$m(t) = \begin{cases} 0, & t = 0 \\ a, & t \rightarrow \infty \end{cases}$$

Here 'a' represents the expected number of software failures eventually detected. If $\lambda(t)$ is the corresponding intensity function, the basic assumption of LPETM says that $\lambda(t)$ is a decreasing function of $m(t)$ as a result of repair action following early failures. We model this assumption as a relation between $m(t)$ and $\lambda(t)$ so that one can be expressed in terms of the other. Our proposed relation is

$$\lambda(t) = \frac{b}{2a} [a^2 - m^2(t)]$$

Where 'b' is a positive constant, serving the purpose of constant of proportional fall in $\lambda(t)$. This relation indicates a decreasing trend for $\lambda(t)$ with increase in $m(t)$ - a characteristic similar to that of LPETM which requires exponential decrease of $\lambda(t)$ with increase in $m(t)$, whereas our proposed model requires a quadratic decrease of $\lambda(t)$ with $m(t)$. According to our proposition we get the following differential equation

$$\frac{dm(t)}{dt} = \frac{b}{2a} [a^2 - m^2(t)]$$

Whose solution is

$$m(t) = \frac{a(1 - e^{-bt})}{(1 + e^{-bt})}$$

(5)

We propose an NHPP with its mean value function given in equation (5). Its intensity function is

$$\lambda(t) = \frac{2abe^{-bt}}{(1 + e^{-bt})^2}$$

(6)

Our proposed model is derived on lines of Goel and Okumoto (1979) specifying a relation between $m(t)$ and $\lambda(t)$ as motivated from LPETM. Our model turns out to be the probability model of half logistic distribution of Balakrishnan (1985). This model is paid a considerable attention as a reliability model recently by many authors and some of those published works are Kantam *et al.*, (1994), Kantam and Dharmarao (1994), Kantam and Rosaiah (1998), Kantam *et al.*, (2000), Kantam and Srinivasarao (2004) and the references therein. In short, we abbreviate this model as HLSRGM.

III. RELIABILITY PERFORMANCE MEASURES

Using the mean value function $m(t)$ for HLSRGM, the probability distribution of the counting process $N(t)$ is given by

$$P[N(t) = y] = \frac{[m(t)]^y}{y!} e^{-m(t)}, y = 0, 1, 2, \dots (7)$$

Where $m(t)$ is given by Equation (5). Here $N(t)$ is random number of failures experienced by time 't'. Suppose that y_e failures have been observed during $[0, t_e]$. Then the conditional distribution of $N(t)$ given $N(t_e) = y_e$ for $t > t_e$ is the distribution of number of failures during $[t_e, t]$. That is,

$$P[N(t) = y / N(t_e) = y_e] = \frac{[m(t) - m(t_e)]^{y - y_e}}{(y - y_e)!} \times e^{-[m(t) - m(t_e)]} \tag{8}$$

Let X_k be the time between $K-1$ and K^{th} failures of the software product. And S_k be the time up to the K^{th} failure. The probability that X_k exceeds a real number 'x' given that the total time up to the $(K-1)^{\text{th}}$ failure is equal to 's' is

$$R_{X_k} / S_{k-1}(x/s) = e^{-[m(s+x) - m(s)]}$$

(9)

Equation (9) is called the software reliability. If we take the negative of derivative of Equation (9) with respect to x, we get the conditional density as

$$f(X_k / S_{k-1}) = \lambda(s+x) e^{-[m(s+x) - m(s)]}$$

(10)

The hazard rate is given by the ratio of density to reliability which becomes

$$h(X_k / S_{k-1}) = \lambda(s+x) \tag{11}$$

In equations (9), (10) and (11) the expressions for $m(t)$ and $\lambda(t)$ are given by the respective quantities of HLSRGM.

IV. ESTIMATION BASED ON INTER FAILURE TIMES

We may recall that the mean value function and intensity function of HLSRGM are given by

$$m(t) = \frac{a(1 - e^{-bt})}{(1 + e^{-bt})}, a > 0, b > 0, t \geq 0$$

(12)

$$\lambda(t) = \frac{2abe^{-bt}}{(1 + e^{-bt})^2} \tag{13}$$

The constants 'a', 'b' which appear in the mean value function and hence in NHPP, in intensity function (error detection rate) and various other expressions are called parameters of the model. In order to have an assessment of the software reliability 'a', 'b' are to be known or they are to be estimated from a software failure data. Suppose we have 'n' time instants at which the first, second, third..., n^{th} failures of software are experienced. In other words if S_k is the total time to the k^{th} failure, s_k is an observation of random variable S_k and 'n' such failures are successively recorded, the

joint probability of such failure time realizations $s_1, s_2, s_3, \dots, s_n$ is

$$L = e^{-m(s_n)} \prod_{k=1}^n \lambda(s_k) \tag{14}$$

The function given in equation (14) is called the likelihood function of the given failure data. Values of 'a', 'b' that would maximize L are called maximum likelihood estimators (MLEs) and the method is called maximum likelihood (ML) method of estimation. Accordingly 'a', 'b' would be solutions of the equations

$$\frac{\partial \log L}{\partial a} = 0, \frac{\partial \log L}{\partial b} = 0.$$

Substituting the expressions for m(t), λ(t) given by equations (12) and (13) in equation (14), taking logarithms, differentiating with respect to 'a', 'b' and equating to zero, after some joint simplification we get

$$a = n \left[\frac{(1 + e^{-bs_n})}{(1 - e^{-bs_n})} \right] \tag{15}$$

$$g(b) = \sum_{k=1}^n s_k \frac{n}{b} - 2 \sum_{k=1}^{n-1} \frac{s_k e^{-ts_k}}{(1 + e^{-ts_k})} - \frac{2s_n e^{-ts_n}}{(1 + e^{-ts_n})} \left[1 - \frac{n}{1 - e^{-ts_n}} \right] = 0 \tag{16}$$

MLE of 'b' is an iterative solution of equation (16) which when substituted in equation (4.4) gives MLE of 'a'. In order to get the asymptotic variances and co-variance of the MLEs of 'a', 'b' we need the elements of the information matrix which can be obtained through the following second order partial derivatives.

$$\frac{\partial^2 \log L}{\partial a^2} = \frac{-n}{a^2} \tag{17}$$

$$\frac{\partial^2 \log L}{\partial a \partial b} = \frac{2s_n e^{-2bs_n}}{(1 + e^{-bs_n})^2} \tag{18}$$

$$\frac{\partial^2 \log L}{\partial b^2} = \frac{-n}{b^2} - 2 \sum_{k=1}^n s_k^2 e^{-ts_k} \left[\frac{2e^{-ts_k} - (1 + e^{-ts_k})}{(1 + e^{-ts_k})^3} \right] + 2 \sum_{k=1}^n s_k^2 e^{-ts_k} \left[\frac{s_k e^{-ts_k} - (1 + e^{-ts_k})}{(1 + e^{-ts_k})^2} \right] \tag{19}$$

Expected values of negatives of the above three derivatives would be the following information matrix

$$E \begin{bmatrix} -\frac{\partial^2 \log L}{\partial a^2} & -\frac{\partial^2 \log L}{\partial a \partial b} \\ -\frac{\partial^2 \log L}{\partial a \partial b} & -\frac{\partial^2 \log L}{\partial b^2} \end{bmatrix}$$

Inverse of the above matrix is the asymptotic variance covariance matrix of the MLEs of 'a', 'b'. Generally the

above partial derivatives evaluated at the MLEs of 'a', 'b' are used to get consistent estimator of the asymptotic variance covariance matrix.

In order to overcome the numerical iterative way of solving the log likelihood equations and to get analytical estimators rather than iterative, some approximations in estimating the equations can be adopted from Kantam and Dharmarao (1994), Kantam and Sriram (2001) and the references therein. We use two such approximations here to get modified MLEs of 'a' and 'b'. Equation (4.5) can be written as

$$b \sum_{k=1}^n s_k - n - 2 \sum_{k=1}^n \frac{z_k e^{-z_k}}{(1 + e^{-z_k})} + \frac{2n z_n e^{-z_n}}{(1 - e^{-2z_n})} = 0 \tag{20}$$

Let us approximate the following expressions in the L.H.S of equation (20) by linear functions in the neighborhoods of the corresponding variables.

$$\frac{z_k e^{-z_k}}{1 + e^{-z_k}} = \alpha_k + \beta_k \cdot z_k, k=1, 2, \dots, n \tag{21}$$

$$\frac{z_n e^{-z_n}}{1 - e^{-2z_n}} = \gamma_n + \delta_n \cdot z_n \tag{22}$$

Where the intercepts and slopes β_k, δ_n in equations (21) and (22) are to be suitably found. With such values, equations (21) and (22) when used in equation (4.9) would give an approximate MLE for 'b' as

$$\hat{b} = \frac{n + 2 \sum_{k=1}^n \alpha_k - 2n \gamma_n}{\sum_{k=1}^n s_k - 2 \sum_{k=1}^n \beta_k \cdot s_k + 2n \delta_n s_n} \tag{23}$$

We suggest two methods to get the slopes and intercepts in the R.H.S of equations (21) and (22)

Method I

Let $F(z) = \frac{1 - e^{-z}}{1 + e^{-z}}$

$$p_i = \frac{i}{n+1}, i=1, 2, \dots, n$$

$$F(u'_i) = p_i - \sqrt{\frac{p_i(1-p_i)}{n}}$$

$$F(u''_i) = p_i + \sqrt{\frac{p_i(1-p_i)}{n}}$$

Given a natural number 'n' we can get the values of u'_i and u''_i by inverting the above equations through the function F(z). If G(.) , H(.) are the symbols for the L.H.S of equations (21) and (22) we get

$$\beta_k = \frac{G(u''_k) - G(u'_k)}{u''_k - u'_k}, k=1, 2, \dots, n$$

$$\alpha_k = G(u'_k) - \beta_k u'_k, k=1, 2, \dots, n$$

$$\delta_n = \frac{H(u_n'') - H(u_n')}{u_n'' - u_n'}$$

$$\gamma_n = H(u_n') - \delta_n \cdot u_n'$$

It can be seen that the evaluation of $\alpha_k, \beta_k, \delta_n, \gamma_n$ is based on only a specified natural number 'n' and can be computed free from any data. Given the data observations and sample size using these values along with the sample data in equation (4.12) we get an approximate MLE of 'b'. Equation (4.4) gives approximate MLE of 'a'.

Method II.

In this method we consider the Taylor series expansion up to the first derivative of the functions G (.), H (.) around $u_i, (i = 1, 2, \dots, n)$ and u_n respectively where u_i is the solution of the following equations

$$\frac{1 - e^{-u_i}}{1 + e^{-u_i}} = p_i, i = 1, 2, \dots, n.$$

Accordingly the expressions for the slopes and intercepts of equation (4.10) and (4.11) are given by

$$\beta_k = G'(u_i), \alpha_k = G(u_i) - \beta_k u_i, i = 1, 2, \dots, n.$$

$$\delta_n = H'(u_n), \gamma_n = H(u_n) - \delta_n \cdot u_n.$$

where $G'(.), H'(.)$ are the derivatives of $G(.), H(.)$ respectively. These are

$$G'(z) = \frac{e^{-z}(1 + e^{-z} - z)}{(1 + e^{-z})^2}$$

$$H'(z) = \frac{e^{-z}(1 - z - e^{-2z} - z \cdot e^{-2z})}{(1 - e^{-2z})^2}$$

As mentioned in method I, here also the slopes and intercepts can be calculated free from magnitudes of data observations. These values when used in equation (4.9) would give another approximate MLE of 'b' and subsequently another approximate MLE of the parameter 'a' from equation (4.4). Thus given software failure data in the form of inter failure times of a finite number of failures in terms of X_k or S_k we can get estimates of the parameters 'a', 'b', of HLSRGM by exact ML method as iterative solution of equation (4.5) and an approximate MLE by two methods of approximation as described above. In the case of approximations, the basic principle is, some expressions of the estimating equations are approximated by linear functions. The larger the size of the sample the closer the approximation. Hence the exactness of the approximation becomes finer for large values of 'n'. Therefore, approximate MLEs and exact MLEs differ little, when the number of experienced software failures is large. As supported by Bhattacharya (1985), we can consider approximated expressions of log L for the elements of the information matrix in order to get the asymptotic

variances and covariance of the estimators of a, b. Hence

$$\frac{\partial^2 \text{Log}L}{\partial a^2} = \frac{-n}{a^2} \tag{24}$$

$$\frac{\partial^2 \text{Log}L}{\partial a \partial b} = \frac{2be^{-bs_n}}{(1 + e^{-bs_n})^2} \tag{25}$$

$$\frac{\partial^2 \text{Log}L}{\partial b^2} = \frac{-n}{b^2} - 2 \sum_{k=1}^n \frac{s_k b e^{-ts_k}}{(1 + e^{-ts_k})^2} + \frac{2s_n a b e^{-ts_n}}{(1 + e^{-ts_n})^3} (1 - e^{-ts_n}) \tag{26}$$

Estimated variances and covariance of \hat{a} and \hat{b} are obtained from the elements of the inverse of the matrix evaluated at \hat{a}, \hat{b} .

$$\begin{bmatrix} \frac{-\partial^2 \text{Log}L}{\partial a^2} & \frac{-\partial^2 \text{Log}L}{\partial a \partial b} \\ \frac{-\partial^2 \text{Log}L}{\partial a \partial b} & \frac{-\partial^2 \text{Log}L}{\partial b^2} \end{bmatrix} \tag{27}$$

which can be used to get the estimated asymptotic variance of estimate of any reliability characteristic like mean value function, intensity function, reliability function etc., In all these cases a software reliability characteristic is a parametric function say $g(a, b)$. Estimated asymptotic variance of estimate of $g(a, b)$ is given by

$$\text{Est.Asy.var.} \\ g(\hat{a}, \hat{b}) = \left(\frac{\partial g}{\partial a}\right)^2 \text{Est.Asy.Var}(\hat{a}) + \left(\frac{\partial g}{\partial b}\right)^2 \text{Est.Asy.Var}(\hat{b}) + 2 \frac{\partial g}{\partial a} \cdot \frac{\partial g}{\partial b} \text{cov}(\hat{a}, \hat{b}) \tag{28}$$

With the help of asymptotic optimum properties of MLEs, we can get the $(1 - \alpha)\%$ equitable confidence intervals for 'a', 'b' or any parametric function of 'a', 'b' using the elements of inverse of the matrix (27) or the R.H.S of equation (28) evaluated at \hat{a}, \hat{b} . That is,

$$\hat{a} \pm z_{1-\alpha/2} \text{Est.Asy.Var}(\hat{a})$$

$$\hat{b} \pm z_{1-\alpha/2} \text{Est.Asy.Var}(\hat{b})$$

NTDS DATA

$$g(\hat{a}, \hat{b}) \pm z_{1-\alpha/2} \text{Est.Asy.Var} \left[g(\hat{a}, \hat{b}) \right]$$

Where $z_{1-\alpha/2}$ is the $(1 - \alpha / 2)^{th}$ percentile of standard normal distribution.

$$\text{Using the result namely } 2 \log \left[\frac{L(\hat{a}, \hat{b})}{L(a, b)} \right] \text{ follows a}$$

chi-square distribution with 2 degrees of freedom, $100(1 - \alpha)\%$ equitailed confidence interval for the parametric function $L(a, b)$ is

$$L(\hat{a}, \hat{b})e^{-\frac{u^*}{2}} < L(a, b) < L(\hat{a}, \hat{b})e^{-\frac{L^*}{2}} \quad (29)$$

where L^*, U^* are respectively $\frac{\alpha}{2}, 1 - \frac{\alpha}{2}$ percentiles

of a chi-square distribution with 2 degrees of freedom. The procedure of interval estimation described above can be adopted by using the exact MLEs of 'a', 'b' and the corresponding estimated variances, covariance of the MLEs with the help of equations (15) through (19).

V. ILLUSTRATION AND EVALUATION OF THE MODEL

We apply the results to the software failure data of Naval Tactical Data system (NTDS) borrowed from Jelinski and Moranda (1972), given in the following table.

Table I: Software failure data

Phase	Error No (check out)	Time between errors x_k days	Cumulative time $S_n = \sum x_k$ days
Production phase	1	9	9
	2	12	21
	3	11	32
	4	4	36
	5	7	43
	6	2	45
	7	5	50
	8	8	58
	9	5	63
	10	7	70
	11	1	71
	12	6	77
	13	1	78
	14	9	87
	15	4	91
	16	1	92
	17	3	95
	18	3	98
	19	6	104
	20	1	105
	21	11	116
	22	33	149
	23	7	156
	24	91	247
	25	2	249
	26	1	250
Test phase	27	87	337
	28	47	384
	29	12	396
	30	9	405
	31	135	540
user phase	32	258	798
Test phase	33	16	814
	34	35	849

We confirm the suitability of our model to the data by a test of goodness of fit known as Q-Q plot. We have taken the first 26 observations for the Q-Q plot. The correlation

coefficient between sample and population quantiles is 0.97 for our HLSRGM also indicating that the model fits well for the data. Solving equations (4.4) and (4.5) by Newton-Raphson(N-R) method for the NTDS data, (Program listing for N-R method is given), the iterative solutions for MLEs of 'b', 'a' are $\hat{b} = 0.011827, \hat{a} = 29$. Using equations (4.6), (4.7), (4.8) we get the estimated asymptotic variance covariance matrix of the MLEs of 'a', 'b' as

$$\begin{bmatrix} \hat{a}, \hat{b} \end{bmatrix} = \begin{bmatrix} 32.346894 & -0.000018747 \\ -0.000018747 & 0.000000475 \end{bmatrix}$$

Using equations of the approximate MLEs namely (2.4.10), (2.4.11), (2.4.12) we get the values of approximate MLEs of 'a', 'b' as follows:

	\hat{a}	\hat{b}
Method I:	27.59124139	0.014067485
Method II:	28.31665782	0.012618822

The asymptotic dispersion matrix \hat{a}, \hat{b} by Method I is

$$\begin{bmatrix} \hat{a}_1, \hat{b}_1 \end{bmatrix} = \begin{bmatrix} 29.27987 & 0.000000176 \\ 0.000000176 & 0.000007611 \end{bmatrix}$$

The asymptotic dispersion matrix \hat{a}, \hat{b} by Method II is

$$\begin{bmatrix} \hat{a}_2, \hat{b}_2 \end{bmatrix} = \begin{bmatrix} 30.83973515 & 0.000000187 \\ 0.000000187 & 0.000006124 \end{bmatrix}$$

Various quantities of interest can be obtained by substituting the estimates of 'a', 'b' in the appropriate equations of section 4.

A. Evaluation of the Model

Now we make an attempt to evaluate the performance of our model in relation to the GOM by two criteria namely Mean square Error (MSE) and Akaike's Information criterion (AIC)(Akaike,1974), defined as follows:

$$MSE = \sum_{i=1}^n \frac{[m(t_i) - \hat{m}(t_i)]^2}{n - k}$$

$$AIC = -2 \log L \left(\frac{\hat{a}, \hat{b}}{s_1, s_2, \dots, s_n} \right) + 2k$$

where $\hat{m}(t_i)$ is the estimated value of m(t) at estimates of 'a', 'b'; K is the number of parameters estimated (here K=2). For the NTDS data the values of MSE, AIC for our model by the three methods of estimation and for GOM by exact ML method of estimation are as follows:

Table II: NTDS data

HLSRGM				GOM (ML)
	ML	MML1	MML2	
MSE:	3.9999	3.868	3.8264	5.3995
AIC:	168.60	169.07	168.78	169.38

We find that MSE, AIC of our model are smaller than those of GOM. Regarding methods of estimation within our model we find that MMLE based on method-II of our model gives a least value of MSE. Hence method-II is preferable.

As mentioned earlier one of the basic assumptions common to LPETM and HLSRGM is that the failure intensity decreases non linearly with the expected number of failures experienced. Accordingly the specific relations between mean value function and intensity function of LPETM, HLSRGM are

LPETM: $\lambda(t) = \lambda_0 \cdot e^{-\theta m(t)}$ (30)

HLSRGM: $\lambda(t) = \frac{b}{2a} [a^2 - m^2(t)]$ (31)
 The solutions of these two relations when solved for m (t) (using the fact $m'(t) = \lambda(t)$) would be

LPETM: $m(t) = \frac{1}{\theta} \log(\lambda_0 \theta t + 1)$ (32)

HLSRGM: $m(t) = \frac{a(1 - e^{-bt})}{(1 + e^{-bt})}$ (33d)

The limiting values of m(t) as $t \rightarrow \infty$ in the above two equations are ∞, a respectively. Though the theoretical basis of these two models is similar, the asymptotic behavior of the mean value functions is different. We know that the mean value function of GOM is given by

GOM: $m(t) = a(1 - e^{-bt})$ (34)

And its limiting value is 'a'. Thus HLSRGM has a similarity with GOM with respect to the mean value function in the limit, and another similarity with LPETM in developing the mean value function as presented in equations (32), (33). We therefore, thought of comparing the relative suitability of these three models for a live data. Reparameterizing and rewriting the mean value functions of the three models we can get the following equations.

GOM: $-\log\left(1 - \frac{m(t)}{a}\right) = bt$

(35)

LPETM: $e^{\frac{m(t)}{a}} - 1 = bt$

(36)

HLSRGM: $-\log\left[\frac{a - m(t)}{a + m(t)}\right] = bt$ (37)

Wherein equation (36), 'a' is the notation for $1/\theta$ and 'b' is the notation for $\lambda_0 \cdot \theta$ of equation (32). If the expressions on the LHS of the equations (35) through (37) are considered as dependent on 't', these indicate that each LHS is a linear function of 't' with slope 'b'. Hence if observations on t, m(t) in a live data are available and the value of 'a' for that data is known or assigned by a specific procedure the correlation between each [LHS, 't'] indicates the strength of closeness between the data and the respective model. Because, in a way 'a' is the limiting value of m(t), a reasonable substitute for 'a' can be the maximum number of experienced faults by a given software in the given maximum time. For the NTDS data we may take 'a' as 34 so that the

values of the correlation coefficient between the following pairs of variables,

GOM: $\left[t, -\log\left[1 - \frac{m(t)}{\hat{a}}\right]\right]$

LPETM: $\left[t, e^{\frac{m(t)}{\hat{a}}} - 1 \right]$

HLSRGM: $\left[t, -\log\left[\frac{\hat{a} - m(t)}{\hat{a} + m(t)}\right]\right]$

are -0.3954, 0.88, 0.9537 respectively where $\hat{a} = 34$; t, m(t) are columns 3 and 1 of NTDS DATA table. This shows that our HLSRGM is the best suitable model for the NTDS DATA with respect to explaining the reparameterized linear relations given in Equations (5.6), (5.7), and (5.8).

VI. REFERENCES

- [1]. HAKAIKE, "A new look at statistical model identification", IEEE Transactions on Automatic Control, 19, pp. 716-723,1974.
- [2]. N.Balakrishnan, "Order statistics from the Half Logistic distribution", J.Statist.Comput.Simul, 20, pp.287-309,1985.
- [3]. A.L.Goel and K.Okumoto, "A time dependent error detection rate model for Software Reliability and other performance measures",IEEE Transactions on Reliability, vol.R-28,pp.206-211,1979.
- [4]. Z.Jelinski and P.Moranda. "Software Reliability Research", In Statistical Computer Performance Evaluation, W. Freiberger, Ed.New York,: Academic, pp.465-484,1972.
- [5]. R.R.L.Kantam and V.Dharmarao, "Half Logistic Distribution-an improvement over M.l.estimator", proceedings of 11 annual conference of sds, pp.39-44,1994.
- [6]. R.R.L.Kantam and K.Rosaiah,"Half Logistic Distribution in acceptance sampling based on life tests", iapqr transactions vol.23, no.2, pp.117-125,1998.
- [7]. R.R.L.Kantam, K.Rosaiah and G.V.S.R.Anjaneyulu, "Estimation Of Reliability multi component stress-strength model:half logistic distribution", iapqr transactions ,vol.25no.2, pp.43-52,2000.
- [8]. R.R.L.Kantam and G.Srinivasarao, "A note on savings in experimental time under type ii censoring", economic quality control, vol.19, no.1, 91-95,2004.
- [9]. R.R.L.Kantam and B.Sriram, "Variable Control Charts based on gamma distribution", iapqr transactions 26(2), pp.63-78,2001.
- [10].R.R.L.Kantam A.Vasudevarao and V.L.Narasimham,"Linear Unbiased Estimation of scale parameter by absolute values of order statistics in logistic distribution", journal of statistical research, vol.28, no.1&2, pp.199-21,1994.
- [11]. J.D.Musa and K.Okumoto,"A Logarithmic Poisson Execution time model for software reliability measurement", proceeding seventh international conference on software engineering, orlando, pp.230-238,1984.

