ISSN: 2321-7782 (Online)
Volume 1, Issue 7, December 2013
International Journal of Advance Research in Computer Science and Management Studies
Research Paper
Available online at: www.ijarcsms.com

Detecting Pareto Type II Software Reliability using SPRT

K.Sita Kumari[1]
Associate Professor
Department of IT
V.R.Siddhartha Engg Colleg
Vijayawada - India

Dr.R.Satya Prasad[2]
Associate Professor
Department of Computer Science
Acharya Nagarjuna University
Guntur - India

***Abstract:** As the volumes of data/software is getting increased in the internet day by day, there is a need for the people to have the tools/mechanism to assess the software for reliability as it takes more time to come to conclusion in classical hypothesis. By adopting Sequential Analysis of Statistical science it can be decided very quickly whether the software which is developed is reliable or unreliable. To implement this, Sequential Probability Ratio Test (SPRT) is applied for Pareto Type II model. In this paper, the performance of SPRT on time domain data is evaluated using Pareto Type II Model with ordered statistics. The results are analyzed for 4 different data sets and the parameters are estimated using Maximum Likelihood Estimation. The experimental results elucidated that the reliability of software can be assessed with less number of observations comparatively.*

***Keyword:** Pareto Type II, Mean Value Function, SPRT, NHPP, MLE, Decision lines, Software Reliability, Order Statistics.*

I. INTRODUCTION

In Classical Hypothesis testing, the number of test cases is fixed at the beginning of the testing. After the entire data is collected, the analysis is done and the conclusions are drawn. But where as in Sequential Analysis, each and every test case is analyzed soon after it has been collected and also compared with some threshold value and incorporates the new information obtained with the current test case. This permits one to draw conclusions during the data collection, so that the final conclusion may be made at much earlier stage. Wald's procedure is well suited if the data is collected sequentially.

Order statistics deals with properties and applications of ordered random variables and of functions of these variables [2]. The use of order statistics is significant when failures are frequent or inter failure time is less. Let X denote a continuous random variable with probability density function f(x) and cumulative distribution function F(x), and let (X1 , X2 , …, Xn) denote a random sample of size n drawn on X. The original sample observations may be unordered with respect to magnitude. A transformation is required to produce a corresponding ordered sample. Let (X(1) , X(2) , …, X(n)) denote the ordered random sample such that X(1) < X(2) < … < X(n); then (X(1), X(2), …, X(n)) are collectively known as the order statistics derived from the parent X. The various distributional characteristics can be known from Balakrishnan and Cohen [3]. The inter-failure time data is grouped into non overlapping successive sub groups of size 4 or 5 and add the failure times with in each sub group. The probability distribution of such a time lapse would be that of the rth ordered statistics in a subgroup of size r, which would be equal to r^{th} power of the distribution function of the original variable (m(t)). The order statistics is preferable when the failure data set is large. We implemented the Pareto Type II model for 4th order and 5th order statistics.

The software failure data is of two types which is very important for analyzing the software reliability. Time Domain Data records the occurrence times of the failures or the times of successive failures. Interval Domain Data assumes that the data are given for the cumulative number of detected errors in a given time interval [1]. The probability equation of the stochastic process representing the failure occurrences is given by a homogeneous Poisson process with the expression

$$P[\,N(t) = n] = \frac{e^{-bt}(bt)^n}{n!}$$

The Time domain approach requires more data collection efforts and highly accurate in parameter estimation than the interval domain approach. In this paper, we considered a popular SRGM known as Pareto Type II model with ordered statistics in assessing the reliability of a software product concluding whether to accept or reject the software which is developed. The parameters estimation for the proposed model is stated in section IV and the application of decision rules for the live data sets is described in Section V.

II. Wald's Sequential Test

The Sequential Probability Ratio Test (SPRT) was developed by Abraham Wald [5]. The SPRT procedure is used for quality control studies during the manufacturing of software products. The Sequential tests can be performed with fewer observations as compared to fixed sample size sets. In classical hypothesis testing is performed on large volumes of data which consumes large amount of time and it can be reduced to a large extent by implementing Sequential Probability Ratio Tests. For SPRT methodology the Stieber[4] suggested the following procedure to reject or accept a software system.

Let $\{N(t), t \geq 0\}$ be a homogeneous Poisson process with rate 'λ'. N(t) = number of failures up to time ' t' and 'λ' is the failure rate (failures per unit time). For an instance a system is kept on test (suppose a software system, where we can conduct test based on a usage lineament and where we cannot correct mistakes) and its failure rate 'λ' will be assessed. The accurate value of 'λ' cannot be expected. But our desire is not to allow the system with a high probability if the failure rate is larger than λ1 for the provided data and is considered to admit it with a high probability, if it is smaller than λ0 $(0 < \lambda 0 < \lambda 1)$. There is always some risk to get the incorrect answers with statistical tests. So we have to specify two (small) numbers 'α' and 'β', where 'α' is the probability of falsely rejecting the system. That is, the system will be rejected even if $\lambda \leq \lambda 0$ at the "producer's" risk. β is the probability of falsely accepting the system, that the system is accepted even if $\lambda \geq \lambda 1$ which is at "consumer's" risk. With specified choices of λ0 and λ1 such that $0 < \lambda 0 < \lambda 1$, the probability of finding N(t) failures in the time span (0,t) with λ1,λ0 as the failure rates respectively and are given

$$P1 = \frac{e^{-\lambda_1 t}[\lambda_1 t]^{N(t)}}{N(t)!} \qquad 2.1$$

$$p_0 = \frac{e^{-\lambda_0 t}[\lambda_0 t]^{N(t)}}{N(t)!} \qquad 2.2$$

The ratio $\frac{p_1}{p_0}$ at any time 't' is considered as a measure of deciding the truth towards λ_0 or λ_1 given a sequence of time instants say t1<t2<…<tk and the corresponding realizations N(t1), N(t2), ---- N(tk). Simplification of $\frac{p_1}{p_0}$ gives

$$\frac{p_1}{p_0} = \exp(\lambda_0 - \lambda_1)\,t + \left[\frac{\lambda_1}{\lambda_0}\right]^{N(t)} \qquad 2.3$$

The decision rule of SPRT is to decide in favor of λ1 or λ0 or to continue by observing the number of failures until $\frac{P_1}{P_0}$ is greater than or equal to a constant, say A, less than or equal to a constant say B or in between the constants A and B. That is, we decide the given software product as unreliable, reliable or continue the test process with one more observation in failure data, according as

$$\frac{p_1}{p_0} \geq A \qquad 2.4$$

$$\frac{p_1}{p_0} \leq B \qquad 2.5$$

$$B < \frac{p_1}{p_0} < A \qquad 2.6$$

The approximate values of the constants A and B are taken as

$$A \cong \frac{1-\beta}{\alpha} \qquad B \cong \frac{\beta}{1-\alpha}$$

Where 'α' and 'β' are the risk probabilities as defined earlier. A simplified version of the above decision process can be illustrated as follows:

- Reject the system, if N(t) falls above the line NU(t) = a t + b2 for the first time and the software is unreliable
- Accept the system if N(t) falls below the line NL(t) = a t + b1 for the first time and the software is said to be reliable
- Continue the test with one or more observations on (t, N(t)) as the random graph of [t, N(t)] is between the two linear boundaries given by NU(t) and NL(t) where

$$a = \frac{\lambda_1 - \lambda_0}{\log\left[\frac{\lambda_1}{\lambda_0}\right]} \qquad 2.7$$

$$b_1 = \frac{\log\left[\frac{1-\alpha}{\beta}\right]}{\log\left[\frac{\lambda_1}{\lambda_0}\right]} \qquad 2.8$$

$$b_2 = \frac{\log\left[\frac{1-\beta}{\alpha}\right]}{\log\left[\frac{\lambda_1}{\lambda_0}\right]} \qquad 2.9$$

The parameters α, β, λ0, and λ1, can be chosen in several ways. One way suggested by Stieber (1997) is

$$\lambda_0 = \frac{\lambda.\log(q)}{q-1}, \lambda_1 = q\frac{\lambda.\log(q)}{q-1}$$

Where

$$q = \frac{\lambda_1}{\lambda_0}$$

If we opt λ0 and λ1 in this way, the slope of NU (t) and NL (t) equals λ. There are two other ways in selecting λ0 and λ1. They are selecting λ0 and λ1 from past projects for comparison of the projects and another one is from part of the data to compare the reliability of various functional areas (components).

III. Sequential Probability ratio test for Pareto Type II SRGM

From Section II, it is understood that the desired value of N (t) = λt which indicates the average number of failures experienced in time't' for the Poisson process. It also represents the mean value function for the Poisson Process. On the other hand, m (t) is considered as mean value function for Poisson Process in general and the probability equation of m (t) can be specified as

$$P[N(t) = Y] = \frac{[m(t)]^y}{y!} . e^{-m(t)}$$

For the Pareto Type II Distribution, the mean value function for NHPP is given by

$$m(t) = a\left[1 - \frac{c^b}{(t+c)^b}\right]$$

N(t) = λt which indicated the average number of failures experienced in time 't' for the Poisson process which is also called the mean value function.

Based on m(t), various Poisson Processes exists. For the Pareto Type II Distribution model, Non Homogeneous Poisson Process (NHPP) is considered and the mean value function is given as

$$m(t) = a(1 - \frac{c^b}{(t+c)^b})$$

It can be written as

$$p_1 = \frac{e^{-m_1(t)}[m_1(t)]^{N(t)}}{N(t)!} \qquad 3.1$$

$$p_0 = \frac{e^{-m_0(t)}[m_0(t)]^{N(t)}}{N(t)!} \qquad 3.2$$

Where m1(t), m0(t) represents the mean value function at stated parameters indicating reliable software and unreliable software respectively. The mean value function m(t) comprises the parameters 'a', 'b' and 'c'. The two specifications of NHPP for b are considered as b0 , b1 where (b0 < b1) and the two specifications of c are considered as c0 , c1 where (c0 < c1). For our proposed model, m(t) at b1 is said to be greater than b0 and m(t) at c1 is said to be greater than c0. The same can be denoted symbolically as m0(t) < m1(t). The details of the SPRT procedure is as follows

System is said to be accepted and be reliable if

$$\frac{p_1}{p_0} \le B$$

i.e.,

$$\frac{e^{-m_1(t)[m_1(t)]^{N(t)}}}{e^{-m_0(t)[m_0(t)]^{N(t)}}} \le B$$

i.e.,

$$N(t) \le \frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log\ m_1(t) - \log m_0(t)} \qquad 3.3$$

System is said to be rejected and be unreliable if

$$\frac{p_1}{p_0} \geq A$$

i.e.,

$$\frac{e^{-m_1(t)[m_1(t)]^{N(t)}}}{e^{-m_0(t)[m_0(t)]^{N(t)}}} \geq A$$

i.e.,

$$N(t) \geq \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log\ m_1(t) - \log m_0(t)}$$

3.4

Continue the test procedure as long as it satisfies the condition

$$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log\ m_1(t) - \log m_0(t)} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log\ m_1(t) - \log m_0(t)}$$

3.5

Substituting the m(t) of the Pareto model [6] in the above expressions, the decisions rules are obtained and are enumerated as follows

Acceptance region:

$$N(t) \leq \frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left[\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}} - \frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}}{1-\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}}}\right]}$$

3.6

Rejection region:

$$N(t) \geq \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left[\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}} - \frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}}{1-\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}}}\right]}$$

3.7

Continuation region :

$$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left[\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}} - \frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}}{1-\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}}}\right]} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left[\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}} - \frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{{c_1}^{b_1}}{(t+c_1)^{b_1}}}{1-\frac{{c_0}^{b_0}}{(t+c_0)^{b_0}}}\right]}$$

3.8

For the specified model, it may be noted that the decision rules are exclusively based on the strength of the sequential process (α, β) and the value of the mean value function namely, $m_0(t), m_1(t)$ The decision rules become decision lines as described by Stieber (1997) [4], if the mean value function is linear in 't' passing through origin that is $m(t) = \lambda t$.. The equations 3.6, 3.7 & 3.8 are considered as generalizations for the decision procedure given by Stieber (1997) [4][5].

IV. PARAMETER ESTIMATION FOR PARETO TYPE II MODEL WITH ORDERED STATISTIC

Parameter estimation is of primary importance in software reliability prediction. Once the analytical solution for m (t) is known for a given model, parameter estimation is achieved by applying a well-known technique of Maximum Likelihood Estimation (MLE) [1]. The main idea behind maximum likelihood parameter assessment is to decide the parameters that maximize the probability (likelihood) of the specimen data. In the other words, MLE methods are versatile and applicable to most models and to different types of data. The mean value function of Pareto type II model is given by

$$m(t) = a\left(1 - \frac{c^b}{(t+c)^b}\right)$$

4.1

Where a, b and c which maximize L are called as maximum likelihood estimators (MLEs). The parameters a, b, c are estimated with Maximum Likelihood (ML) estimation

To derive m (t) for r^{th} ordered statistics, m (t) to the power r has to be considered

$$m(t) = \left(a\left(1 - \frac{c^b}{(t+c)^b}\right)\right)^r$$

4.2

Differentiating m (t) with respect to t

$$m'(t) = r\left[a - \frac{ac^b}{(t+c)^b}\right]^{r-1} * \left[\frac{abc^b}{(t+c)^{b+1}}\right]$$

4.3

The required likelihood function is given by

$$L = e^{-m(t)} * \prod_{i=1}^{n} m'(t_i)$$

4.4

$$L = e^{-\left[a - \frac{ac^b}{(t+c)^b}\right]^r} * \prod_{i=1}^{n} r\left[a - \frac{ac^b}{(t_i+c)^b}\right]^{r-1}\left[\frac{abc^b}{(t_i+c)^{b+1}}\right]$$

4.5

The log likelihood equation is used to estimate the unknown parameters a, b and c and is given by

$$\log L = -a^r\left[1 - \frac{c^b}{(t+c)^b}\right]^r + \sum_{i=1}^{n} \log r + \sum_{i=1}^{n} (r-1)\log\left[a - \frac{ac^b}{(t+c)^b}\right] + \sum_{i=1}^{n}[(\log a + \log b + b\log c) - (b+1)\log(t_i + c)]$$

4.6

The parameters ‘a’, ‘b’ and ‘c’ are computed by considering the following equations

$$\frac{\partial \text{Log L}}{\partial a} = 0,\ \frac{\partial \text{Log L}}{\partial b} = 0,\ \frac{\partial^2 \text{Log L}}{\partial b^2} = 0,\ \frac{\partial \text{Log L}}{\partial c} = 0,\ \frac{\partial^2 \text{Log L}}{\partial c^2} = 0$$

Differentiating with respect to a, b and c and equating to zero the following equations are obtained

Differentiating with respect to ‘a’

$$\frac{\partial \text{Log L}}{\partial a} = -ra^{r-1}\left[1-\frac{c^b}{(t+c)^b}\right]^r + 0 + \sum_{i=1}^{n}(r-1)\frac{1}{\left(a-\frac{ac^b}{(t+c)^b}\right)}\left[1-\frac{c^b}{(t+c)^b}\right] + \frac{n}{a}$$

4.7

$$a^r = n * \left[\frac{(t+c)^b}{(t+c)^b - c^b}\right]^r$$

4.8

Differentiating with respect to ‘b’

$$\frac{\partial \text{Log L}}{\partial b} = a^r r\left(1-\left(\frac{c}{t+c}\right)^b\right)^{r-1}\left(\frac{c}{t+c}\right)^b \log\left(\frac{c}{t+c}\right) + \sum_{i=1}^{n}[(r-1) * \frac{(t+c)^b}{(t+c)^b-c^b} * -\left(\frac{c}{t+c}\right)^b \log\left(\frac{c}{t+c}\right)] + \sum_{i=1}^{n}[\frac{1}{b} + \log c - \log(t_i + c$$

4.9

Taking c=1,

$$= nr\left[\frac{1}{(t+1)^b - 1}\right]\log\left(\frac{1}{t+1}\right) + \sum_{i=1}^{n}(r-1)\log(t_i+1)\frac{1}{[(t_i+1)^b-1]} + \frac{n}{b} - \sum_{i=1}^{n}\log(t_i+1)$$

$$g(b) = \frac{nr}{(t+1)^b-1}\log\left(\frac{1}{t+1}\right) + \sum_{i=1}^{n}(r-1)\log(t_i+1)\frac{1}{[(t_i+1)^b-1]} + \frac{n}{b} - \sum_{i=1}^{n}\log(t_i+1)$$

4.10

$$g'(b) = nr\log\left(\frac{1}{t+1}\right)(-1) * \frac{(t+1)^b\log(t+1)}{[(t+1)^b-1]^2} - \sum_{i=1}^{n}(r-1)\log(t_i+1)(t_i+1)^b\log(t_i+1)[(t_i+1)^b-1]^{-2} - \frac{n}{b^2}$$

4.11

Differentiating with respect to c

$$\frac{\partial \text{Log L}}{\partial c} = -nr\left[\frac{(t+c)^b}{(t+c)^b-c^b}\right]^r\left[\frac{(t+c)^b-c^b}{(t+c)^b}\right]^{r-1}\left[-b\left(\frac{c}{t+c}\right)^{b-1} * \frac{t}{(t+c)^2}\right] + \sum_{i=1}^{n}(r-1) * \frac{(t+c)^b}{(t+c)^b-c^b}\left[-b\left(\frac{c}{t+c}\right)^{b-1} * \frac{t}{(t+c)^2}\right] + \frac{nb}{c} - \sum_{i=1}^{n}\left(\frac{b+1}{t_i+c}\right)$$

`

4.12

Taking b=1,

$$= \frac{nrct}{(t+c)^1-c^1} * \frac{1}{c} * \frac{1}{t+c} - \sum_{i=1}^{n} t * \frac{(r-1)c}{(t+c)^1-c^1} * \frac{1}{c} * \frac{1}{t+c} + \frac{n}{c} - \sum_{i=1}^{n}\frac{2}{t_i+c}$$

$$g(c) = \frac{nr}{t+c} - \sum_{i=1}^{n}\frac{(r-1)}{(t_i+c)} + \frac{n}{c} - \sum_{i=1}^{n}\left(\frac{2}{t_i+c}\right)$$

4.13

$$g'(c) = \frac{-nr}{(t+c)^2} + \sum_{i=1}^{n}\frac{(r-1)}{(t_i+c)^2} - \frac{n}{c^2} + \sum_{i=1}^{n}\frac{2}{(t_i+c)^2}$$

4.14

The values of the parameters 'b' and 'c' in the above equations are obtained using Newton Raphson Method. The point estimates of the parameters yields 'b'and 'c' after solving the above equations simultaneously. The equations need to be solved iteratively to obtain the values of b and c and by substituting them in the log likelihood equation of 'a' which provides the analytical solution for the MLE of 'a'.

V. SPRT Analysis

In this Section, the SPRT procedure is applied on four different data sets for 4th ordered and 5th ordered statistics referred by Musa [9] and Michael.R.Lyu [8] and the decisions are evaluated based on the mean value function. The specifications for b0, b1 and c0, c1 are chosen based on the parameter estimates b and c such that b0 < b < b1 and c0 < c < c1. The estimates for the four different data sets are specified in the following table.

TABLE I

Estimates of a, b, c & Specifications of b0, b1, c0, c1

Data Set	Order	a	b	c	b_0	b_1	c_0	c_1
CSR2, Michael R.Lyu, 1996a	4	32.007075	0.999752	4.890103	0.649752	1.349752	4.640103	5.140103
	5	25.007721	1.000377	4.883801	0.65038	1.35038	4.6338	5.1338
SYS2, Michael R.Lyu, 1996a	4	21.000000	4.748504	5.051261	4.395804	5.098504	4.801261	5.301261
	5	17.000000	4.312880	5.054862	3.96288	4.66288	4.80486	5.30486
SYS3, Michael R.Lyu, 1996a	4	51.044234	1.000120	3.496916	0.65012	1.35012	3.246916	3.746916
	5	41.043687	1.000142	3.469722	0.65014	1.350142	3.219722	3.719722
Musa, 1975	4	26.026781	1.000276	3.961974	0.500276	1.500276	2.461974	5.461974
	5	27.00713	1.000108	4.470552	0.50011	1.50011	2.97055	5.97055

Using the specification b0, b1 and c0, c1 , the mean value functions m0(t) and m1(t) are computed for each 't'. Later the decisions are made based on the decision rules specified by the equations 3.6, 3.7 & 3.8 for the data sets. At each't' of the data set , the strengths (α , β) are considered as (0.05, 0.2). [6][7]. SPRT procedure is applied on four different data sets and the relevant calculations are given in the Table 2 [8][9].

TABLE II

SPRT Analysis for 4th and 5th order statistics

Data set	T	N(t)	R.H.S of equation (3.3) Acceptance Region(≤)	R.H.S of equation (3.4) Rejection Region (≥)	Decision
CSR2, Michael R.Lyu, 1996a 4th order	1557	1	-1.521980456	60.40735943	
	1639	2	-2.363664708	61.19302918	
	1973	3	-5.602221781	64.19955817	
	2183	4	-7.505588675	65.95636072	
	2714	5	-11.95893562	70.04422893	

	3455	6	-17.51898409	75.11493547	Continuous
	5045	7	-27.74940129	84.38301639	
	5087	8	-27.99628871	84.60596339	
	5222	9	-28.78314626	85.31633798	
	5608	10	-30.97939179	87.29771755	
	6599	11	-36.29872995	92.08922517	
	7042	12	-38.54832807	94.11288595	
	7565	13	-41.11601258	96.42099482	
	7612	14	-41.34240155	96.62441697	
	8496	15	-45.47844618	100.338781	
	9356	16	-49.3022912	103.7696182	
	10662	17	-54.79021776	108.6890393	
	12523	18	-62.06280593	115.2015831	
	13656	19	-66.22896226	118.9295173	
	24480	20	-99.5717941	148.7164213	
	26136	21	-103.9571962	152.6296836	
	31174	22	-116.5080636	163.825532	
	34077	23	-123.2820175	169.8661884	
	35422	24	-126.3226574	172.5772865	
	37476	25	-130.8570432	176.6198361	
	39336	26	-134.857272	180.185794	
	47688	27	-151.7528155	195.2438059	
	50119	28	-156.3869831	199.3731567	
	58707	29	-171.9191153	213.2112304	
	69259	30	-189.5210589	228.8901434	
	78723	31	-204.1975401	241.961117	
	88694	32	-218.7346684	254.9064607	
CSR2, Michael R.Lyu, 1996a 5th Order	1579	1	-7.561305554	52.84022539	Continuous
	1738	2	-9.125044688	54.28592842	
	2030	3	-11.82847923	56.77376905	
	2714	4	-17.50705839	61.9637791	
	3491	5	-23.16671186	67.1030995	
	5054	6	-32.92437338	75.91492717	
	5222	7	-33.87844045	76.7741518	
	5608	8	-36.01484645	78.69698421	
	6602	9	-41.20397052	83.36134978	
	7233	10	-44.29930882	86.14023478	
	7603	11	-46.05241999	87.7131417	
	8496	12	-50.11722526	91.35773735	
	9632	13	-54.99351438	95.72612833	
	11629	14	-62.91915039	102.8191299	
	12793	15	-67.22823924	106.6725145	
	24480	16	-102.7292559	138.3688006	
	26809	17	-108.6901104	143.6850387	
	31869	18	-120.8088953	154.490045	
	35386	19	-128.6738583	161.5004524	

	37476	20	-133.1636927	165.5018759	
	47320	21	-152.8016038	182.9995211	
	49620	22	-157.0866688	186.8168374	
	58648	23	-173.0169432	201.0063531	
	69259	24	-190.246013	216.3500391	
	78785	25	-204.6200376	229.1494057	
SYS2, Michael R.Lyu, 1996a 4th Order	1576	1	-6442026.289	5731564.611	Continuous
	4149	2	-72085685.61	64135279.97	
	5827	3	-168312027.3	149748664.8	
	10071	4	-660009377.1	587215929.1	
	11,836	5	-987913656.4	878955122.1	
	15280	6	-1869716572	1663502587	
	16572	7	-2289983140	2037417284	
	16860	8	-2390691963	2127018770	
	23827	9	-5672327868	5046717814	
	29257	10	-9473258650	8428437873	
	32886	11	-12687098419	11287818124	
	35467	12	-15322798642	13632822762	
	41151	13	-22213406419	19763454409	
	48662	14	-33768970062	30044536494	
	53623	15	-43037470210	38290798967	
	56483	16	-49003893338	43599175765	
	61888	17	-61572191482	54781296246	
	70138	18	-84170079399	74886827035	
	83146	19	-1.28752E+11	1.14552E+11	
	91514	20	-1.63611E+11	1.45566E+11	
	98022	21	-1.94244E+11	1.72821E+11	
SYS2, Michael R.Lyu, 1996a 5th order	2610	1	-396372600	352656065.4	Continuous
	4436	2	-1904736682	1694660266	
	8163	3	-11589772426	10311517959	
	11836	4	-34830986458	30989421399	
	15,685	5	-80197251654	71352168796	
	17995	6	-1.20476E+11	1.07189E+11	
	22226	7	-2.25215E+11	2.00376E+11	
	29257	8	-5.08423E+11	4.52349E+11	
	33346	9	-7.49067E+11	6.66451E+11	
	40856	10	-1.36735E+12	1.21655E+12	
	47147	11	-2.08971E+12	1.85924E+12	
	54223	12	-3.16093E+12	2.8123E+12	
	59996	13	-4.26841E+12	3.79764E+12	
	68374	14	-6.28787E+12	5.59437E+12	
	81107	15	-1.04114E+13	9.26308E+12	
	92191	16	-1.52549E+13	1.35724E+13	
	99693	17	-1.91728E+13	1.70582E+13	

SYS3, michael R.Lyu, 1996a 4th order	89	1	33.24077253	60.68919885	Accept
SYS3,Michael R.Lyu, 1996a 5th order	93	1	29.00322299	44.15318842	Accept
MUSA(1975) 4th order	227	1	9.855679613	37.73859391	Accept
MUSA(1975) 5th order	342	1	9.318302926	40.28438295	Accept

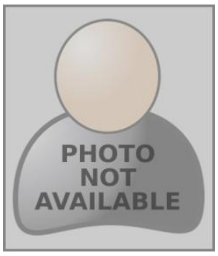
VI. CONCLUSION

The analysis for the data sets MUSA (1975), SYS3, SYS2 and CSR2 with SPRT is shown in the Table2. It is observed that we are able to come up with an early conclusion about the reliability or unreliability of a software product as compared with the existing models. The results of the table exemplifies that the model has given a decision of acceptance at first iteration itself for the data sets SYS3 and MUSA and continue for SYS2 and CSR2. Therefore, we conclude that applying SPRT on the data sets, the proposed Pareto Type II model – an order statistic approach is performing well in arriving at a decision.

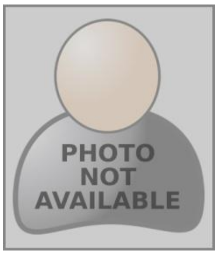
References

1. Pham. H., 2006. “System software reliability”, Springer.
2. Arak M. Mathai ;Order Statistics from a Logistic Dstribution and Applications to Survival and Reliability Analysis;IEEE Transactions on Reliability, vol.52, No.2; 2003
3. Balakrishnan.N., Clifford Cohen; Order Statistics and Inference; Academic Press inc.;1991.
4. Stieber, H.A. “Statistical Quality Control: How To Detect Unreliable Software Components”, Proceedings the 8th International Symposium on Software Reliability Engineering, 8-12. 1997.
5. Wald. A., “Sequential Analysis”, John Wiley and Son, Inc, New York. 1947
6. R Satya Prasad, N Geetha Rani,“Pareto Type II SRGM : Application of SPRT based on Interval Domain Data”, International Journal of Computer Information Systems, Vol. 5, No. 5, 2012 pp. 1-6, 2012.
7. R Satya Prasad, N Geetha Rani,“Pareto type II software reliability growth model”, International Journal of Software Engineering, Vol. 2, Issue 4, 2011. pp. 81-86
8. Michael. R. Lyu, “The hand book of software reliability engineering”, McGrawHill & IEEE Computer Society press.
9. Musa, J.D., Iannino, A., Okumoto, k., 1987. “Software Reliability: Measurement Prediction Application”. McGraw-Hill, New York

AUTHOR(S) PROFILE



Mrs. K.Sita Kumarai, received M.Sc degree from Acharya Nagarjuna University, Guntur and M.Tech degree from Dr.MGR university, Chennai. She is currently pursuing her Ph.D from Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. Presently she is working as an Associate Professor in the department of Information Technology, V.R.Siddhartha Engineering College, Kanuru, Vijayawada.



Dr. R. Satya Prasad, received Ph.D.degree incomputer science in the faculty of Engineering in 2007 from Acharya Nagarjuna University, Andhra Pradesh. He received gold medal from Acharya Nagarjuna University for his outstanding performance in master’s degree. He is currently working as Associate Professor in the department of Computer Science &Engg., Acharya Nagarjuna University. His current research is focused on Software engineering. He has published several papers in National & International Journals