Burr Type III Quality Assessment using SPRT

N. Krishna Kumar, R. Satya Prasad, G. Sridevi

Abstract: The data or software system is obtaining exaggerated within the internet day by day, and a vivacious analysis is goes on towards the software system reliability. In this regard, there is a desire for the people to own the tools/mechanisms to observe whether or not the software system is reliable or not. Many strategies came into existence for assessing the software system reliability. A lot of time is used when classical approaches like hypothesis testing used for the reason that the conclusions solely are drawn when collecting vast amounts of knowledge. Adopting statistical mathematical strategies like sequential analysis will be applied to reach a decision quickly. We have a tendency to project to implement the test namely Sequential Probability Ratio Test (SPRT) based approach on Burr Type III model depends on interval domain data. For this, Maximum Likelihood Estimation (MLE) used to predict the parameters to be use SPRT on real time software system failure datasets borrowed from different software projects.

***Index Terms:** Burr type III distribution model, ML Estimation, Reliability of Software system, Sequential Probability Ratio Test.*

I. INTRODUCTION

Procedure of Wald's is particularly applicable for sequentially collected data [10]. There is a small divergence between sequential analysis and classical hypothesis. Assessment of software reliability needs effective tools and mechanisms. Based on the data collected conclusions are drawn in classical hypothesis testing. In this testing, first test cases number collected and fixed and then analysis is done. In sequential analysis, data collection by analyzing the individual test cases, compared with assumed some threshold level for integrating the information newly into the present test case provides conclusions to take final decisions at much earlier stage. So that at earlier stage only, decisions can be taken. This provides automatically saving money and human time. That is the benefit of sequential analysis.

The software reliability growth model applications could be tough and prediction of reliability will be misleading when classical testing strategies are used. Sequential analysis may be a technique of statistical inference and here variety of clarification needed as a result of the process is not determined previously of the experiment**.** At every stage, choice termination is based on the results from the observed data earlier. With average tiny fewer observations, sequentially testing procedure constructed to equally check the reliability is the key benefit of the test. The Ungrouped Data (time-domain data) and Grouped Data (interval-domain data) are the existing failure data types. The time-domain data records failures that occur by the side of individual times. The interval domain data within a fixed time period (within weeks) records count of variety of failures**.** The estimation of high accurate parameters related to the interval domain data requires more collection of data with existing software reliable models. The Failure occurrences random number is given by homogeneous Poisson process equation.

$$P[N(t) = n] = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \qquad (1)$$

Here we present a technique SPRT, which uses Maximum Likelihood Estimation (MLE) for detecting reliable software based on SPRT. There are two classes of Wald's SPRT procedure which are reliable or unreliable, pass or fail and certified or uncertified in which software can be under test and distinguished. SPRT can provide statistically optimal correct decision in a short period of time compared to all other tests with equivalent decision errors. Based on the estimated likelihood of the hypothesis, this procedure is working to detect fault based software systems. The Burr type III distribution model along with the principle of Stieber [1] is considered as a reliable software model to identify the reliable or unreliable software for accepting or rejecting the software developed.

II. WALD'S SEQUENTIAL POISSON PROCESS TEST

At Columbia University in the year 1943, the SPRT procedure was invented by Abraham Wald [10]. The SPRT method is used for software systems quality control, this procedure takes decisions between two simple hypotheses. We will use the SPRT procedure for the period of manufacturing of software products. To perform tests on fixed sample size sets with smaller number of observations that can be considered. The process for SPRT procedure for Homogeneous Poisson Process is discussed below. Consider a homogeneous Poisson process $\{N(t), t \geq 0\}$ with rate 'λ' with collection of failures N(t) and 'λ' which gives failure rate per unit of time interval known as failure rate up to the time t. For the time being system is kept on test to find 'λ'**.** When the failure rate is more than λ_1, we wish to reject a system and if less than $\boldsymbol{\lambda_0}$ accept the system. In statistical tests there is a more chance of getting wrong prediction of errors. For this 'α' and 'β' are the constants considered here indicates probability of falsely rejecting and accepting the system respectively.

That is if $\lambda \leq \lambda_0$ system rejected and is known as "producer's" risk similarly if $\lambda \leq \lambda_1$ then system is accepted known as "consumer's" risk.

By considering $0 < \lambda_0 < \lambda_1$, time span (0, t) with λ_1, λ_0, the failure rates N(t) of probabilities are estimated with the following equations (2) and (3).

Revised Manuscript Received on June 05, 2019

N.Krishna Kumar, Research Scholar, Department of Computer Science & Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India.

Dr R.Satya Prasad, Professor & Head, Department of Computer Science & Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India.

Dr G. Sridevi, Professor, Department of Computer Science & Engineering, MVR College of Engineering and Technology, Paritala, Andhra Pradesh, India.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

$$P_1 = \frac{e^{-\lambda_1 t}[\lambda_1 t]^{N(t)}}{N(t)!} \quad (2)$$

$$P_0 = \frac{e^{-\lambda_0 t}[\lambda_0 t]^{N(t)}}{N(t)!} \quad (3)$$

The ratio $\frac{P_1}{P_0}$ is used to find the truth in concern to the λ_0 or λ_1, at any point of time 't' say $t_1 < t_2 < \cdots < t_k$ along with realizations $N(t_1), N(t_2) \ldots N(t_k)$ of $N(t)$. The ratio $\frac{P_1}{P_0}$ after simplification is

$$\frac{P_1}{P_0} = exp(\lambda_0 - \lambda_1)\, t + \left[\frac{\lambda_1}{\lambda_0}\right]^{N(t)}$$

The decision rules of SPRT procedure is as follows:

i) Failures observation until is greater than or equal to a constant say A.

ii) Observing the failures numbers less than or greater than the constant say B.

iii) Or observing the failures in between the A and B which are constants.

Given software can be tested for its reliability or unreliability.

$$\frac{p_1}{p_0} \geq A \quad (4)$$

$$\frac{p_1}{p_0} \leq B \quad (5)$$

$$B < \frac{p_1}{p_0} < A \quad (6)$$

The rough constant values for A and B are considered as

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

As outlined earlier, the constants 'α ' and 'β ' are the risk probabilities. The above decision processes simply illustrated as follows:

- The software is unreliable and system rejected if failure rate is present above the line

$N_U(t) = at + b_2.$

- The software is reliable and system accepted if failure rate is lies below the line $N_L(t) = at - b_1$.
- With increased number of observations, repeate the test as the random graph of is between the two linear boundaries given by $N_U(t)$ and $N_L(t)$ and

$$a = \frac{\lambda_1 - \lambda_0}{\log\left[\frac{\lambda_1}{\lambda_0}\right]} \quad (7)$$

$$b_1 = \frac{log\left[\frac{1-\alpha}{\beta}\right]}{log\left[\frac{\lambda_1}{\lambda_0}\right]} \quad (8)$$

$$b_2 = \frac{log\left[\frac{1-\beta}{\alpha}\right]}{log\left[\frac{\lambda_1}{\lambda_0}\right]} \quad (9)$$

The parameters α, β, λ_0 and λ_1 can be chosen in several ways. One way suggested by Stieber (1997) [1] is

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

$$\lambda_1 = q\,\frac{\lambda \log q}{q-1} \quad \text{Where} \quad q = \frac{\lambda_1}{\lambda_0}$$

If λ_0 and λ_1 are chosen in this manner, the slope of $N_U(t)$ and $N_L(t)$ equals λ.

III. SPRT PROCEDURE FOR BURR TYPE III SRGM

From section II, $N(t) = \lambda(t)$ gives number failures average value in time 't' for the Poisson process. The Poisson process can be specified to determine the mean value function m(t) as

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

Based on $m(t)$, various Poisson processes exists. For the Burr type III distribution model [10], NHPP is considered and the mean value function is given as

$$m(t) = a\left[\left(1+t^{-c}\right)^{-b}\right], \quad t \geq 0$$

Then we written

Then we written

$$P_1 = \frac{e^{-m_1 t}[m_1 t]^{N(t)}}{N(t)!}$$

$$P_0 = \frac{e^{-m_0 t}[m_0 t]^{N(t)}}{N(t)!}$$

Where $m_1(t), m_0(t)$ is mean function value to find reliable and unreliable software's respectively, having 'a', 'b' and 'c' parameters. The two specifications of NHPP for 'b' are considered as b_0, b_1 where $(b_0 < b_1)$ and two specifications of c say c_0, c_1 where $(c_0 < c_1)$. The implementation procedure for SPRT is discussed below [2]. The system can be accepted and said to be reliable if

$$\frac{P_1}{P_0} \leq B$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$i.e., \quad \frac{e^{-m_1(t)}[m_1(t)]^{N(t)}}{e^{-m_0(t)}[m_0(t)]^{N(t)}} \leq B$$

$$i.e., N(t) \leq \frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \quad (10)$$

The system can be rejected and said to be unreliable if

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

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

$$i.e., N(t) \geq \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \quad (11)$$

Continue the test procedure if

$$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \quad (12)$$

The appropriate decision rules can be obtained by substituting of mean value function $m(t)$ and are given in followings lines.

Region of Acceptance:

$$N(t) \leq \frac{\log\left(\frac{\beta}{(1-\alpha)}\right) + a\left[\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{c_0}\right)^{-b_0}\right]}{\log\left[\frac{\left(1+t^{c_1}\right)^{-b_1}}{\left(1+t^{c_0}\right)^{-b_0}}\right]} \quad (13)$$

Region of Rejection:

$$N(t) \geq \frac{\log\left(\frac{\beta}{(1-\alpha)}\right) + a\left[\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{c_0}\right)^{-b_0}\right]}{\log\left[\frac{\left(1+t^{c_1}\right)^{-b_1}}{\left(1+t^{c_0}\right)^{-b_0}}\right]} \quad (14)$$

Region of Continuation:

$$\frac{\log\left(\frac{\beta}{(1-\alpha)}\right) + a\left[\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{c_0}\right)^{-b_0}\right]}{\log\left[\frac{\left(1+t^{c_1}\right)^{-b_1}}{\left(1+t^{c_0}\right)^{-b_0}}\right]} < N(t) < \frac{\log\left(\frac{\beta}{(1-\alpha)}\right) + a\left[\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{c_0}\right)^{-b_0}\right]}{\log\left[\frac{\left(1+t^{c_1}\right)^{-b_1}}{\left(1+t^{c_0}\right)^{-b_0}}\right]} \quad (15)$$

The sequential procedure (α, β) and mean value functions m_0 (t), m_1 (t) are the strengths of the decision rules made by the SPRT based burr type III Model.

IV. PARAMETER ESTIMATION

The parameters estimation is most important in prediction of reliability of the software. To estimate parameters here we used MLE estimation process for parameter assessment to take decision about the parameters which gives more probability of the specimen data, applicable for the grouped and ungrouped data. At present soft computing techniques or using genetic algorithms are used to estimate the parameters. The Burr type III distribution model whose mean function is given by [10]

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

Where maximum likelihood (ML) estimates the parameters 'a', 'b' and 'c'. The required likelihood function is given by [9]

$$LogL = \sum_{i=1}^{k}(n_i - n_{i-1})\log\left[m(t_i) - m(t_{i-1})\right] - m(t_k)$$

$$LogL = \sum_{i=1}^{k}(n_i - n_{i-1})\log\left\{a\left[\left(1+t_i^{-c}\right)^{-b}\right] - a\left[\left(1+t_{i-1}^{-c}\right)^{-b}\right]\right\}$$
$$-a\left[\left(1+t_k^{-c}\right)^{-b}\right]$$

Differentiate with respect to 'a' and equating to '0'.

$$\frac{\partial Log\ L}{\partial a} = 0$$

$$\therefore a = \sum_{i=1}^{k}(n_i - n_{i-1})\left(1+t_k^{-c}\right)^{b} \quad (16)$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The Newton Raphson Method used to get the parameter value 'b'.

$$\frac{\partial Log\,L}{\partial b} = g(b) = \sum_{i=1}^{k} (n_i - n_{i-1}) \left\{ \begin{bmatrix} -\log\left(1+t_i^{-1}\right) - \log(1+t_{i-1}^{-1}) + \\ \dfrac{(1+t_{i-1}^{-1})^b \log(1+t_{i-1}^{-1}) - (1+t_i^{-1})^b \log(1+t_i^{-1})}{(1+t_{i-1}^{-1})^b - (1+t_i^{-1})^b} \end{bmatrix} \atop -\log\left(\dfrac{1}{1+t_k^{-1}}\right) \right\} \quad (17)$$

(17

$$\frac{\partial^2 LogL}{\partial b^2} = g^{'}(b) = \sum_{i=1}^{k} (n_i - n_{i-1}) \left[\frac{(1+t_i^{-1})^b (1+t_{i-1}^{-1})^b \log\left(\dfrac{1+t_i^{-1}}{1+t_{i-1}^{-1}}\right) \log\left(\dfrac{1+t_{i-1}^{-1}}{1+t_i^{-1}}\right)}{\left[(1+t_{i-1}^{-1})^b - (1+t_i^{-1})^b\right]^2} \right] \quad (18)$$

The iterative Newton Raphson Method is used to obtain the parameter 'c'.

$$g(c) = \frac{\partial LogL}{\partial c} = 0$$

$$\frac{\partial LogL}{\partial c} = g(c) = \sum_{i=1}^{k} (n_i - n_{i-1}) \left[-\log\left(\frac{1}{t_i}\right)\left(\frac{t_i^{-c}}{1+t_i^{-c}}\right) - \log\left(\frac{1}{t_i - 1}\right)\left(\frac{t_{i-1}^{c}}{1+t_{i-1}^{-c}}\right) + \left(\frac{t_{i-1}^{c} \log\left(\dfrac{1}{t_{i-1}}\right) - t_i^{-c} \log\left(\dfrac{1}{t_i}\right)}{\left(t_{i-1}^{-c} - t_i^{-c}\right)} \right) \right]$$
$$-\sum_{i=1}^{k} (n_i - n_{i-1}) \left(\frac{t_k^{-c}}{1+t_k^{-c}}\right) \log\left(\frac{1}{t_k}\right) \quad (19)$$

$$\frac{\partial^2 Log\,L}{\partial c^2} = g^{'}(c) = \sum_{i=1}^{k} (n_i - n_{i-1}) - \log\left(\frac{1}{t_i}\right)^2 \left(\frac{t_i^{-c}}{(1+t_i^{-c})^2}\right) - \log\left(\frac{1}{t_{i-1}}\right)^2 \frac{t_{i-1}^{-c}}{(1+t_{i-1}^{-c})^2} +$$
$$\frac{t_i^{-c} t_{i-1}^{-c}}{(t_{i-1}^{-c} - t_i^{-c})^2} \left\{ \log\left(\frac{t_{i-1}}{t_i}\right) \log\left(\frac{t_i}{t_{i-1}}\right) \right\} + \sum_{i=1}^{k} (n_i - n_{i-1}) \log\left(\frac{1}{t_k}\right)^2 \frac{t_k^{-c}}{(1+t_k^{-c})^2} \quad (20)$$

V. SPRT ANALYSIS ON REAL DATASETS

The SPRT analysis is carried out on six real time datasets which are different taken from from Pham (2005) [3] and wood (1996) [7]. The b0, b1 and c0, c1 specifications are taken based on the parameters b and c estimation such that $b_0 < b < b_1$ *and* $c_0 < c < c_1$. These six different datasets estimations are presented in Table 1. The $m_0(t)$ and $m_1(t)$ (20) values for every t are computed using specifications b_0, b_1 and c_0, c_1. By the equations 10, 11 and 12 for the datasets decisions are made by the use of decision rules. The strengths (α, β) with values (0.05, 0.2) are taken to apply the procedure SPRT on six distinct datasets and the needed calculations are given in Table 2.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table 1. Parameter Estimations

Dataset's	Estimation of the parameter of 'a'	Estimation of the parameter of 'b'	Value of b_0	Value of b_1	Estimation of the parameter 'c'	Value of c_0	Value of c_1
Phase 1	5.306901	15.901524	15.401524	16.401524	0.748596	0.248596	1.248596
Phase 2	41.590454	0.978993	0.478993	1.478993	1.083119	0.583119	1.583119
Release 1	48.185326	8.123505	7.623505	8.623505	0.883849	0.383849	1.383849
Release 2	54.32094	9.180969	8.680969	9.680969	0.872395	0.372395	1.372395
Release 3	29.529605	8.548219	8.048219	9.048219	1.039893	0.539893	1.539893
Release 4	21.402067	8.849993	8.349993	9.349993	0.87423	0.37423	1.37423

Table 2. SPRT analysis results

Dataset's	T	N(t)	Region of Acceptance (<=)	Region of Rejection (>=)	Decision either Accepted/Rejected
Phase 1	1	1	-1.867224	3.322366	Rejected
	2	1	-0.024145	0.043672	
Phase 2	1	3	-8.840581	-6.517145	Rejected
Release 1	1	16	-0.867249	1.367936	Rejected
Release 2	1	13	0.014984	0.192914	Rejected
Release 3	1	6	0.002566	0.277293	Rejected
Release 4	1	1	-0.042226	0.209797	Rejected

VI. CONCLUSION

The SPRT methodology for the proposed Burr type III SRGM was applied on 6 real time datasets. At an early stage only, our model is accomplishing well to reach a decision. The results exemplifies that the Burr type III distribution model at different intervals of time can provide rejection decision for all the data sets used**.** Therefore, we have a tendency to might conclude that, we are able to come to an early decision of reliable or unreliable of software system by applying SPRT procedure on any type of data set.

REFERENCES

1. Stieber, H.A.(1997). "Statistical Quality Control: How To Detect Unreliable Software Components", Proceedings the 8th International Symposium on Software Reliability Engineering, 8-12.
2. Murali Mohan, K. V.; Satya Prasad, R.; Sridevi, G.. Detection of Burr Type XII Reliable Software using Sequential Process Ratio Test. **Indian Journal of Science and Technology**, [S.l.], jul. 2015. ISSN 0974 -5645.
3. Pham. H., "System software reliability", Springer. 2006.
4. John D.Musa; "Software Quality and Reliability Basics"; AT&T Bell Laboratories. CH 2468-7/87/0000/014,1987 IEEE.
5. K.B.Misra."Handbook of Performability Engineering". Springer. 2008.
6. WOOD, A. "Predicting Software Reliability*"*, IEEE Computer, 2253-2264. 1996.
7. Satya Prasad R, Krishna kumar N and Sridevi Gutta, " Burr Type III software reliability growth model with Interval Domain Data", International Journal Research , Vol.5, issue. 20, sep 2018, pp.404-410, ISSN: 2348-6848.
8. Mohan D. Pant and Todd C. Headrick, "A Method for Simulating Burr Type III and Type XII Distributions through -Moments and -Correlations," ISRN Applied Mathematics, vol. 2013.
9. Smitha, C.H. & Satya Prasad, R & Kiran Kumar, R. (2016). Software reliability using SPRT: Burr type III process model. International Journal of Electrical and Computer Engineering (IJECE). 6. 3060-3067. 10.11591/ijece.v6i6.11511.
10. Wald A. Sequential Analysis. New Impression edition. New York: John Wiley and Son, Inc; 1947 Sep 30.

AUTHORS PROFILE

Mr N.Krishna Kumar has completed his graduation from Sri Venkateswara University and Post graduation from Anna University in the year 2005. He is currently pursing Ph.D in the Department of Computer Science and Engineering, Acharya Nagarjuna University, Andhra Pradesh, India. He is working as Associate Professor in Krishna Chaitanya Institute of Science & Technology, Nellore.

Dr R.Satya Prasad received Ph.D. degree in computer science in the faculty of Engineering in 2007 from

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Acharya Nagarjuna University, Guntur, Andhra Pradesh. He have a satisfactory consistent academic track of record and received Gold medal from Acharya Nagarjuna University for his outstanding performance in Master's degree. He is currently working as Professor & Head in the Department of Computer Science & engineering, Acharya Nagarjuna University. He has occupied various academic responsibilities like practical examiner, project adjudicator, external member of Board of Examiners for various universities and colleges in and around in Andhra Pradesh. His current research is focused on Software Engineering. He has published nearly 200 research papers in National & International Journals.

Dr G. Sridevi received Ph.D. degree in Computer Science & Engineering in 2015 from Acharya Nagarjuna University. She is currently working as a Professor in the Department of Computer Science & Engineering, MVR College of Engineering and Technology, Paritala, Andhra Pradesh. She has published 50 research papers in National & International journals. Her research interests lies in Data Mining and Software Engineering.

we written as