International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-2 Issue-5, April 2013

A Sequential Probability Ratio Test in Assessing Software Quality Using LPETM

D. Haritha, R. Satya Prasad

Abstract— Rapid growth of software usage enforces us to assess the Software reliability, a critical task in the development of a software system. In this Paper a well known test procedure of statistical science called as Sequential Probability Ratio Test(SPRT) is adopted for Logarithmic Poisson Execution Time Model (LPETM) in assessing the reliability of a developed software. It requires considerably less number of observations when compared with the other existing testing procedures. The model is inspected by using live Data Sets.

Index Terms— Software reliability, SPRT, Maximum Likelihood Estimation, Software testing, Mean value function.

I. INTRODUCTION

Software failure data is essential in the analysis of software reliability applications. The failure data is of two types. Time Domain Approach records the failure occurrence times . where as interval domain approach deals with number of failures in a given time period. The Time domain approach requires more data collection efforts and highly accurate in parameter estimation than the interval domain approach. The probability equation of the stochastic process representing the failure occurrences is given by a homogeneous Poisson process with the expression

$$P\left[N(t)=n\right]=\frac{e^{-\lambda t}(\lambda t)^n}{n!} \qquad (1.1)$$

Stieber (1997) has identified that classical testing strategies results in misleading of reliability predictions. He pointed that statistical methods can be successfully applied to the failure data. In this paper we consider a popular SRGM known as Logarithmic Poison Execution Time Model in short LPETM in detecting unreliable software components in order to accept/reject a developed software. The parameter estimation for the given model is shown in section 4. Application of the decision rule to detect unreliable software components is given in Section 5.

Manuscript published on 30 April 2013.

*Correspondence Author(s)
Smt. D. Haritha, Assistant Professor, Department of ECM, KL University, Vaddeswaram, AP, India.
Dr. R. Satya Prasad, Assoc. Professor, Department of computer science, Acharya Nagarjuna University, Guntur, AP, India.

II. SEQUENTIAL PROBABILITY RATIO TEST FOR A POISSON PROCESS

A. Wald of Columbia university developed the Sequential Probability Ratio Test. Sequential Tests are well known as they require fewer observations on the average than fixed sample size tests. SPRTs are widely used for statistical quality control in manufacturing and military applications. An SPRT for homogeneous Poisson processes is described below. Let {N(t),t ≥0} be a homogeneous Poisson process with rate ‘λ’. Where N(t) is the number of failures up to time ‘ t’ and ‘ λ’ is the failure rate. To ensure failure free process let us consider two small numbers ‘α’ and ‘β’, where ‘α’ is the probability of falsely rejecting the system. That is rejecting the system even if $\lambda \le \lambda_0$. β is the probability of falsely accepting the system .That is accepting the system even if $\lambda \ge \lambda_1$. With specified choices of λ_0 and λ_1 such that $0 < \lambda_0 < \lambda_1$, the probability of finding N(t) failures in the time span (0,t) with λ_1,λ_0 as the

failure rates are respectively given by, $P_1=\dfrac{e^{-\lambda_1 t}\left[\lambda_1 t\right]^{N(t)}}{N(t)!}$

(2.1)

$$P_0=\frac{e^{-\lambda_0 t}\left[\lambda_0 t\right]^{N(t)}}{N(t)!} \qquad (2.2)$$

The ratio $\dfrac{P_1}{P_0}$ 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 $t_1<t_2<t_3<........<t_K$ and the corresponding realizations. $N(t_1),N(t_2),........N(t_K)$ of N(t). Simplification of $\dfrac{P_1}{P_0}$ gives

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

The decision rule of SPRT is to decide in favour of λ_1, in favour of λ_0 or to continue by observing the number of failures at a later time than 't' according as $\dfrac{P_1}{P_0}$ 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.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

We decide the given software product as unreliable, reliable or continue the test process with one more observation in failure data.

$$\frac{P_1}{P_0} \geq A \qquad (2.3)$$

$$\frac{P_1}{P_0} \leq B \qquad (2.4)$$

$$B < \frac{P_1}{P_0} < A \qquad (2.5)$$

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

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

Where ' α ' and ' β ' are the risk probabilities. A simplified version of the above decision processes is to reject the system as unreliable if N(t) falls for the first time above the line

$$N_U(t) = a.t + b_2 \qquad (2.6)$$

to accept the system to be reliable if N(t) falls for the first time below the line

$$N_L(t) = a.t - b_1 \qquad (2.7)$$

To continue the test with one more observation on (t, N(t)) as the random graph of [t, N(t)] is between the two linear boundaries given by equations (2.6) and (2.7) where

$$a = \frac{\lambda_1 - \lambda_0}{\log\left(\frac{\lambda_1}{\lambda_0}\right)} \qquad (2.8)$$

$$b_1 = \frac{\log\left[\frac{1-\alpha}{\beta}\right]}{\log\left(\frac{\lambda_1}{\lambda_0}\right)} \qquad (2.9)$$

$$b_2 = \frac{\log\left[\frac{1-\beta}{\alpha}\right]}{\log\left(\frac{\lambda_1}{\lambda_0}\right)} \qquad (2.10)$$

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

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

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

If λ_0 and λ_1 are chosen in this way, the slope of N_U (t) and N_L (t) equals λ. The other two ways of choosing λ_0 and λ_1 are from past projects (for a comparison of the projects) and from part of the data to compare the reliability of different functional areas.

III. SEQUENTIAL TEST FOR SOFTWARE RELIABILITY GROWTH MODEL

We know that for any poisson process, the expected value of N(t) = λt called the average number of failures experienced in time 't'. Consider a Poisson process with a general function m(t) as its mean value function the probability equation of a such a process is

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

Depending on the forms of m(t) we get various Poisson processes called NHPP for our model the mean value function is

m(t)=a.log(1+bt)

We may write

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

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

where $m_1(t), m_0(t)$ are values of the mean value function at specified sets of its parameters indicating reliable software and unreliable software respectively. For instance the model we have been considering its m(t) function, contains a pair of parameters a, b with 'a' as a multiplier. Also a, b are positive. Let P_0, P_1 be values of the NHPP at two specifications of b say $b_0, b_1 (b_0 < b_1)$ respectively. It can be shown that for our models m(t) at b_1 is greater than that at b_0. Symbolically $m_0(t) < m_1(t)$.Then the SPRT procedure is as follows:

Accept the system to be reliable $\frac{P_1}{P_0} \leq B$

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

i.e., $$N(t) \leq \frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \qquad (3.1)$$

Decide the system to be unreliable and reject if $\frac{P_1}{P_0} \geq A$

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-2 Issue-5, April 2013

i.e., $$N(t) \geq \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \qquad (3.2)$$

Continue the test procedure as long as

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

Substituting the appropriate expressions of the mean value function m(t) of LPETM we get the decision rules and are given in followings lines
m(t)=a.log(1+bt)

Acceptance region:

$$N(t) \leq \frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left(e^{-b_0 t} - e^{-b_1 t}\right)}{\log\left(\frac{1-e^{-b_1 t}}{1-e^{-b_0 t}}\right)} \qquad (3.4)$$

Rejection region:

$$N(t) \geq \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left(e^{-b_0 t} - e^{-b_1 t}\right)}{\log\left(\frac{1-e^{-b_1 t}}{1-e^{-b_0 t}}\right)} \qquad (3.5)$$

Continuation region:

$$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left(e^{-b_0 t} - e^{-b_1 t}\right)}{\log\left(\frac{1-e^{-b_1 t}}{1-e^{-b_0 t}}\right)} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left(e^{-b_0 t} - e^{-b_1 t}\right)}{\log\left(\frac{1-e^{-b_1 t}}{1-e^{-b_0 t}}\right)} \qquad (3.6)$$

It may be noted that in the above model the decision rules are exclusively based on the strength of the sequential procedure (α,β) and the values of the mean value functions namely, $m_0(t)$, $m_1(t)$. If the mean value function is linear in 't' passing through origin, that is, m(t) = λt the decision rules become decision lines as described by Stieber (1997). In that sense equations (3.1), (3.2) , (3.3) can be regarded as generalizations to the decision procedure of Stieber (1997).The applications of these results for live software failure data are presented with analysis in Section 5.

IV. Parameter Estimation

Parameters can be estimated by applying a well known technique of Maximum Likelihood Estimation (MLE).The ML Estimation determine the parameters that maximize the probability (likelihood) of the sample data. The log likelihood function for Time domain data (pham, 2006) is given by:

$$\text{LLF} = \sum_{i=1}^{n} \log[\lambda(s_i)] - m(s_n) \qquad (3.7)$$

The parameters 'a' and 'b' are estimated using iterative Newton Raphson Method, which is given as

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

m(t)=a.log(1+bt) (3.8)

Where $\lambda(t) = \frac{\partial}{\partial t} m(t)$

$$\text{LLF} = \text{Log L} = \sum_{i=1}^{n} \log\left[\frac{ab}{1+bs_i}\right] - a.\log(1+bs_n)$$
(3.9)

The unknown parameters a & b of the given LPETM can be obtained as follows:
Firstly, take the partial derivative of eqn.(3.9) w.r.t 'a' and equate it to zero we get,

$$\sum_{i=1}^{n} \frac{\frac{\partial}{\partial a}\left[\frac{ab}{1+bs_i}\right]}{\left[\frac{ab}{1+bs_i}\right]} - \frac{\partial}{\partial a}(a.\log(1+bs_n)) = 0$$

$$a = \frac{n}{\log(1+bs_n)} \qquad (3.10)$$

The parameter 'b' is estimated by iterative Newton Raphson Method using $b_{n+1} = b_n - \frac{g(b_n)}{g'(b_n)}$. which is substituted in finding 'a'. where g(b) & g'(b) are expressed as follows.

$$g(b) = \frac{\partial \log L}{\partial b} = 0 \qquad g'(b) = \frac{\partial^2 \log L}{\partial b^2} = 0$$

$$g(b) = \sum_{i=1}^{n} \frac{1}{b} - \frac{s_n}{\log(1+bs_n)(1+bs_n)} - \frac{s_i}{1+bs_i} \qquad (3.11)$$

$$g'(b) = \sum_{i=1}^{n} \frac{s_n \log(1+bs_n).s_n + (1+bs_n)\frac{s_n}{\log(1+bs_n)}}{[\log(1+bs_n)(1+bs_n)]^2}$$

$$+ \frac{s_I^2}{(1+BS_I)^2} - \frac{1}{B^2} \qquad (3.12)$$

.

V. SPRT Analysis

In this section we evaluate the decision rules based on the given mean value function for two different data sets. Based on the estimated value of the parameter 'b', we have chosen the specifications of b_0, b_1 that are to be equidistant such that $b_0 < b < b_1$. The choices are given in the following table.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 5.1: Specifications of b0, b1

Data Set	Esti-mate of a	Esti-mate of b	b_0	b_1
1	175.124	0.000253	0.000223	0.000283
2	192.015	0.00058	0.00029	0.00087

Using the selected b_0,b_1 and $m_0(t)$,$m_1(t)$ we have calculated the decision rules given by Equations 3.4, 3.5, sequentially at each 't' of the data sets taking the strength (α, β) as (0.05,0.05). These are presented for the model in Tables 5.2.

Table 5.2: SPRT Analysis for LPETM

Data Set	T	N(t)	Acceptance Region(≤)	Rejection Region(≥)	Decision
1	30.02	1	-9.3669	15.4402	Rejection
	1.44	2	-12.212	12.5052	
	22.47	3	-10.116	14.6669	
	1.36	4	-12.220	12.4969	
	3.43	5	-12.014	12.7102	
	13.2	6	-11.039	13.7154	
	5.15	7	-11.842	12.8873	
	3.83	8	-11.974	12.7514	
	21	9	-10.262	14.5161	
	12.97	10	-11.062	13.6917	
	0.47	11	-12.309	12.4052	
	6.23	12	-11.734	12.9985	
	3.39	13	-11.974	12.7514	
2	9	1	-0.5917	4.78132	Acceptance
	21	2	2.17436	----	

VI. CONCLUSION

The above table shows of 2 data sets using SPRT. We are succeeded in proving that sequential tests will give an early conclusion about the reliability or unreliability of a developed software product as par with the existing reliability models.

REFERENCES

1. GOEL, A.L and OKUMOTO, K. (1979). "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.
2. MUSA, J.D., and OKUMOTO, K. (1984). "A Logorithmic Poisson Execution Time Model For Software Reliability Measurement", Proceeding Seventh International Conference on Software Engineering, Orlando, 230-238.
3. PHAM, H.(2005). "A Generalized Logistic Software Reliability Growth Model", OPSEARCH, Vol.42, No.4, 322-331.
4. Pham. H., 2006. "System software reliability", Springer.
5. STIEBER, H.A.(1997). "Statistical Quality Control: How To Detect Unreliable Software Components", Proceedings the 8th International Symposium on Software Reliability Engineering, 8-12.
6. WALD (1947)."Sequential Analysis", Wiley,New York.
7. WOOD, A.(1996)."Predicting Software Reliability",IEEE Computer, 2253-2264.
8. R.Satya Prasad and G. Krishna Mohan.(2011)."Detection Of Reliable Software Using SPRT On Time Domain Data",International Journal of Computer Science, Engineering and Applications, Vol.1, No.4, pp.92-99.
9. R. Satya Prasad, N. Supriya and G. Krishna Mohan (2011)."Detection Of Reliable Software Using SPRT" International Journal of Advanced Computer Science and Applications Vol.2, No: 8, pp.60-63.
10. R. Satya Prasad and D. Haritha (2011). " Discovery of Reliable Software using GOM on Interval Domain Data" , International Journal of Computer Applications Volume 32– No.5, pp.7-12.
11. R. Satya Prasad and D. Haritha (2011). "Detection of Reliable Software using HLSRGM" ,International Journal of Computer Information Systems pp.49-53.
12. R. Satya Prasad and D. Haritha (2012). "Assessing Reliable Software using SPRT based on LPETM" , Interna tional Journal of Computer Applications Volume 47– No:19, pp.6-11.

AUTHOR PROFILE

Smt. D. Haritha Received her B.Tech Degree in Computer Science & Engineering from JNTUH and M.Tech in CSE from ANU, and pursuing Ph.D from ANU. She is working as an Assistant Professor in the department of ECM, KL University. Her research interests include Software Engineering, Computer Networks, Embedded Systems.

Dr. R. Satya Prasad received Ph.D degree in Computer Science in the faculty of Engineering in 2007 from ANU. He received gold medal from ANU for his outstanding performance in Masters Degree. He is currently working as Associate Professor in the department of Computer science , ANU. His current research is focused on Software Engineering.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)