International Journal of Innovations in Engineering and Technology (IJIET)
http://dx.doi.org/10.21172/ijiet.83.015

An Analogy of Burr Type III and Pareto Type II using SPRT

CH.Smitha Chowdary
Scholar, Department of Computer Science,
Krishna University, Andhra Pradesh, INDIA-521001.

Dr.R.Satya Prasad
Associate Professor, Department of computer science,
Acharya Nagarjuna University, Andhra Pradesh, INDIA-522502.

Dr.R.Kiran Kumar
Asst. Professor & Head of Department of Computer Science,,
Krishna University, Andhra Pradesh, INDIA-521001.

Abstract- Software reliability is viewed as one of the key attributes of software quality, an important quality metric for any developed software. The aim is to statistically measure, ensure and predict the reliability of the software product and is accessed by the use of any statistical model, whose unknown parameters are estimated from the available software failure data. Researchers developed Sequential Probability Ratio Test (SPRT) to judge upon reliable/unreliable software based software reliability techniques. An analogy between Burr Type III and Pareto Type II with SPRT mechanism based on time domain data of Non Homogenous Poisson Process (NHPP) is presented here.

Keywords – Software Reliability, Sequential Probability Ratio Test, Maximum Likelihood Estimation, Burr Type-III, Pareto Type-II.

I. INTRODUCTION

Software reliability is defined as "probability of fault free operations provided by the software product under consideration over a specified period of time in a specified operational environment" [1], assessment of which needs effective tools and mechanisms. Statistical testing is a mechanism that represents a well-founded approach to the estimation of software reliability. Classical Hypothesis Testing of statistical testing is performed with volumes of data without analysis. The use of classical testing strategies in the application of software reliability growth models may be difficult and reliability predictions can be misleading. However statistical methods can be successfully applied to the failure data [2]. Sequential analysis or sequential hypothesis testing is statistical analysis where the specimen size is not determined at the earlier stage, but the data are assessed, while they are being collected and when the prominent results are noticed, the next sampling is stopped in accordance with a predefined stopping rule. In this way a conclusion may be acquired in an earlier stage than it would be with more classical hypothesis testing or assessment at consequently lower financial and/or human cost [3]. A best of sequential methods which is implemented to test statistically a hypothesis, is that a test procedure can be developed that requires on average a small number of observations which equally test the reliability of the procedure based on a predetermined number of observations [4] [5]. Stieber's observations are demonstrated by applying the well-known Sequential Probability Ratio Test (SPRT) of Wald [6] for a software failure data to detect unreliable software components and compare the reliability of different software versions which is a way towards managing the process of reliable software, instead of crafting unreliable software.

Analysis of software reliability requires software failure data. The failure data are of two types: the time-domain data and interval-domain data, where the former records the independent times at which the failure has occurred and the latter counts the number of failures occurring during a fixed time period. With the existing reliability models of software, time-domain data provides better accuracy in the estimation of parameters, but involves more data collection efforts [7]. The stochastic process probability equation that represents the failure occurrence is given by a homogeneous Poisson process with the expression.

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

This paper describes an analogy of Burr Type III and Pareto Type II for detecting reliable software based on the SPRT, using Maximum Likelihood Estimation (MLE) of parameter estimation. The Wald's SPRT procedure is a method that can be used to distinguish the software under test into one of the two categories like reliable/unreliable, pass/fail and certified/uncertified [8]. SPRT is the optimal statistical test that makes the correct decision in the shortest time among all tests that are subject to the same level of decision errors [9]. It is used to detect the fault based on the calculated likelihood of the hypotheses. We considered a popular software reliability growth model, Burr Type III for which the principle of Stieber [2] has been adopted, which helped in detecting whether the software is reliable or unreliable in order to accept or reject the developed software, later Burr Type III and Pareto Type II model results are compared in order make a decision on which model has better performance.

The theory of Burr Type III and Pareto Type II of Software Reliability Growth Models (SRGM's) and Sequential Test of SPRT for Burr type III and Pareto Type II Software Reliability Growth Model are illustrated in section II. Application of the decision rule to detect the unreliable software with reference to the Software Reliability Growth Model Burr Type III and Pareto Type II is depicted in section III. The conclusion is provided in section IV.

II. SEQUENTIAL TEST FOR BURR TYPE III AND PARETO TYPE II

A. Burr Type III and Pareto Type II SRGM's.

An SRGM is a mathematical relationship between time span of testing or using the software and the cumulative number of faults that are detected and repaired. Burr Type III and Pareto Type II are SRGM's that analyze the reliability of software systems using the Time Domain Data. Non Homogeneous Poisson Process (NHPP) is one of the important classes of SRGM used due to its mathematical traceability and wide applicability. Mean value function for the model parameters of Burr Type III and Pareto Type II are determined using maximum likelihood function through which software reliability can be estimated [10][11].

B. Sequential Test for Software Reliability Growth Models.

For any Poisson process, the expected value of N(t) = λ(t) called the average number of failures experienced in time 't', is also called the mean value function of the Poisson process. In other way if we consider the general function of a Poisson process (not necessarily linear) m(t) as its mean value function then probability equation of a such a process is[12]

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

Depending on the forms of m(t) we get various Poisson processes called NHPP[13], for the Burr Type III model and Pareto Type II model. The respective mean value functions are given as [14][15]

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

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

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

$m_1(t)$, $m_0(t)$ here represent the mean value function for the stated parameters depicting reliable software and unreliable software. The mean value function m(t) consists 'a', 'b' and 'c' parameters which are estimated through Maximum Likelihood Estimation (MLE). Two specifications of NHPP for b here, say b_0, b_1 with the condition $b_0 < b_1$ and two specifications of c here, say c_0, c_1 with the condition $c_0 < c_1$ are considered. For the models considered, m(t) at b_1 is said to be greater than b_0 and m(t) at c_1 is said to be greater than c_0 and this is denoted symbolically as $m_0(t) < m_1(t)$. The implementation of SPRT procedure for the models is illustrated below. [14]

System is said to be accepted as it is reliable if

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

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

i.e $$N(t) \le \frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \qquad (2)$$

System is said to be rejected as it is unreliable if

$$\frac{P_1}{P_0} \ge A$$

i.e., $$N(t) \ge \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \qquad (3)$$

Continue the test procedure as long as it is between

i.e., $$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + m_1(t) - m_0(t)}{\log m_1(t) - \log m_0(t)} \qquad (4)$$

Substituting the mean value function of the Burr Type III in appropriate expressions of sequential test, we get the respective decision rules that are given in the following lines [14].

Acceptance Region of Burr Type III:

i.e.,
$$N(t) \le \frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left(\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{-c_0}\right)^{-b_0}\right)}{\log\left[\frac{\left(1+t^{-c_1}\right)^{-b_1}}{\left(1+t^{-c_0}\right)^{-b_0}}\right]} \quad (5)$$

Rejection Region of Burr Type III:

i.e.,
$$N(t) \ge \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left(\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{-c_0}\right)^{-b_0}\right)}{\log\left[\frac{\left(1+t^{-c_1}\right)^{-b_1}}{\left(1+t^{-c_0}\right)^{-b_0}}\right]} \quad (6)$$

Continuation Region of Burr Type III:

i.e.,
$$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left(\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{-c_0}\right)^{-b_0}\right)}{\log\left[\frac{\left(1+t^{-c_1}\right)^{-b_1}}{\left(1+t^{-c_0}\right)^{-b_0}}\right]} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left(\left(1+t^{-c_1}\right)^{-b_1} - \left(1+t^{-c_0}\right)^{-b_0}\right)}{\log\left[\frac{\left(1+t^{-c_1}\right)^{-b_1}}{\left(1+t^{-c_0}\right)^{-b_0}}\right]} \quad (7)$$

Substituting the mean value function of the Pareto Type II in appropriate expressions of sequential test, we get the respective decision rules that are given in the following lines [15].

Acceptance Region of Pareto Type II:

i.e.,
$$N(t) \le \frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left[\frac{c_0^{\,b_0}}{\left(t+c_0\right)^{b_0}} - \frac{c_1^{\,b_1}}{\left(t+c_1\right)^{b_1}}\right]}{\log\left[\frac{1-\frac{c_1^{\,b_1}}{\left(t+c_1\right)^{b_1}}}{1-\frac{c_0^{\,b_0}}{\left(t+c_0\right)^{b_0}}}\right]} \quad (8)$$

Rejection Region of Pareto Type II:

i.e.,
$$N(t) \leq \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left[\frac{c_0^{\,b_0}}{(t+c_0)^{b_0}} - \frac{c_1^{\,b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{c_1^{\,b_1}}{(t+c_1)^{b_1}}}{1-\frac{c_0^{\,b_0}}{(t+c_0)^{b_0}}}\right]} \quad (9)$$

Continuation Region of Pareto Type II:

i.e.,
$$\frac{\log\left(\frac{\beta}{1-\alpha}\right) + a\left[\frac{c_0^{\,b_0}}{(t+c_0)^{b_0}} - \frac{c_1^{\,b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{c_1^{\,b_1}}{(t+c_1)^{b_1}}}{1-\frac{c_0^{\,b_0}}{(t+c_0)^{b_0}}}\right]} < N(t) < \frac{\log\left(\frac{1-\beta}{\alpha}\right) + a\left[\frac{c_0^{\,b_0}}{(t+c_0)^{b_0}} - \frac{c_1^{\,b_1}}{(t+c_1)^{b_1}}\right]}{\log\left[\frac{1-\frac{c_1^{\,b_1}}{(t+c_1)^{b_1}}}{1-\frac{c_0^{\,b_0}}{(t+c_0)^{b_0}}}\right]} \quad (10)$$

For the models adopted, the observation states that the decision rules are solely based on the potency of the sequential procedure (α, β) and the values for the mean value functions namely $m_0(t)$, $m_1(t)$. As per the description of Stieber, the decision rules considered will become decision lines if the mean value function is linear in passing through origin, that is $m(t) = \lambda t$. The equations (2) and (3) are considered as generalizations for the decision procedure of Stieber based on which the SPRT procedure is applied to live software failure data sets and the results that were analyzed are illustrated in Section III.

III. SPRT Analysis of Live Data Sets

In this section, the SPRT methodology is applied on five different data sets that are borrowed from pham [16], lyu [17] and sonata [18] software services. The decisions are evaluated based on the considered mean value function. Based on the estimates of the parameters 'b' and 'c' in each mean value function, we have chosen the specifications of $b_0 = b - \delta$, $b_1 = b - \delta$ and $c_0 = c - \delta$, $c_1 = c - \delta$, and apply SPRT such that $b_0 < b < b_1$ and $c_0 < c < c_1$. The estimates are given in the Table 1 for Burr Type III and Table 2 for Pareto Type II.

Table -1 Burr Type III Estimates of a, b, c & specifications of b_0, b_1, c_0, c_1

Data sets	**Number of samples**	**Estimated Parameters**			**b_0**	**b_1**	**c_0**	**c_1**
		a	**b**	**c**				
NTDS	**26**	**34.465706**	**1.763647**	**1.810222**	**1.16365**	**2.36365**	**1.21022**	**2.41022**
AT&T	**22**	**26.839829**	**1.65869**	**1**	**1.05869**	**2.25869**	**0.4**	**1.6**
SONATA	**30**	**79.831359**	**6.74281**	**0.60244**	**6.14281**	**7.34281**	**0.00244**	**1.20244**

XIE	**30**	**33.310426**	**2.270095**	**1.371974**	**1.6701**	**2.8701**	**0.77197**	**1.97197**
IBM	**15**	**20.624785**	**1.71163**	**1.447815**	**1.11163**	**2.31163**	**0.84782**	**2.04782**

Table -2 Pareto Type II Estimates of a, b, c & specifications of b_0, b_1, c_0, c_1

Data set	**Number of samples**	**Estimated Parameter**						
		a	**b**	**c**	**b_0**	**b_1**	**c_0**	**c_1**
NTDS	**26**	**55.01871**	**0.998899**	**278.6101**	**0.648899**	**1.348899**	**273**	**283**
AT&T	**22**	**34.636895**	**0.999859**	**390.5077**	**0.649859**	**1.349859**	**385**	**395**
SONATA	**30**	**338.720949**	**0.999958**	**18851.72**	**0.649958**	**1.349958**	**18846**	**18856**
XIE	**30**	**44.834192**	**0.999825**	**365.1518**	**0.649825**	**1.349825**	**360**	**370**
IBM	**15**	**36.914829**	**0.99953**	**432.1917**	**0.64953**	**1.34953**	**427**	**437**

Using the specification of b_0, b_1, and c_0, c_1 from Table 1 and Table 2 the mean value functions $m_0(t)$ and $m_1(t)$ are computed for each 't' of the models. Later the decisions are made based on the decision rules specified by the equations (5), (6), (7) of Burr Type III and (8), (9), (10) of Pareto Type II for the data sets. At each 't' of the data set, the strengths of (α, β) are considered as (0.3,0.3) for Burr Type III and refer [11] for Pareto Type II. SPRT procedure is applied on five different data sets and the necessary calculations are given in the Table 3 for Burr Type III and Table 4 for Pareto Type II [11].

Table -3 SPRT analysis for 5 Data Sets with Burr Type III

Data Set	**T**	**N(t)**	**R.H.S. of equation(5) Acceptance region (≤)**	**R.H.S. of equation(6) Rejection region (≥)**	**Decision**
NTDS	**9**	**1**	**22.169838**	**2.7909024**	**ACCEPT**
AT&T	**5.5**	**1**	**3.7988452**	**2.8430065**	**ACCEPT**
SONATA	**52.5**	**1**	**16.809918**	**2.2387207**	**ACCEPT**
XIE	**30.02**	**1**	**3.488346**	**2.2740614**	**ACCEPT**
IBM	**10**	**1**	**4.0612657**	**1.6704087**	**ACCEPT**

Table -4 SPRT analysis for 5 Data Sets with Pareto Type II

Data Set	**T**	**N(t)**	**R.H.S. of equation(8) Acceptance region (≤)**	**R.H.S. of equation(9) Rejection region (≥)**	**Decision**
NTDS	**9**	**1**	**-0.625494**	**5.689154**	**Accept**
	21	**2**	**1.370656**	**7.804387**	
	32	**3**	**3.06055**	**9.601969**	**Continuous**
AT & T	**5.5**	**1**	**-1.762828**	**4.416667**	
	7.33	**2**	**-1.617226**	**4.575441**	
	10.08	**3**	**-1.401039**	**4.811382**	

	80.97	**4**	**3.265448**	**9.971477**	
	84.91	**5**	**3.482139**	**10.214789**	
	99.89	**6**	**4.272655**	**11.105808**	
	103.36	**7**	**4.448588**	**11.304863**	
	113.32	**8**	**4.939423**	**11.861747**	
	124.71	**9**	**5.476276**	**12.473563**	
	144.59	**10**	**6.355699**	**13.48242**	
	152.4	**11**	**6.682622**	**13.85972**	
	167	**12**	**7.267859**	**14.538445**	
	178.41	**13**	**7.703148**	**15.046188**	
	197.35	**14**	**8.386277**	**15.848453**	
	262.65	**15**	**10.420503**	**18.283354**	
	262.69	**16**	**10.421617**	**18.284709**	
	388.36	**17**	**13.319395**	**21.91564**	
	471.05	**18**	**14.731926**	**23.787736**	
	471.51	**19**	**14.738957**	**23.797297**	
	503.12	**20**	**15.203084**	**24.432892**	
	632.43	**21**	**16.773601**	**26.68379**	
	680.03	**22**	**17.245247**	**27.398149**	
	52.5	**1**	**-1.235305**	**4.701932**	
	105	**2**	**-0.342398**	**5.602728**	
	131.25	**3**	**0.10216**	**6.051227**	
	183.75	**4**	**0.987512**	**6.944455**	
	201.25	**5**	**1.281519**	**7.241086**	
	306.25	**6**	**3.034022**	**9.009312**	
	411.25	**7**	**4.766978**	**10.757959**	
	432.25	**8**	**5.111252**	**11.105367**	
	467.25	**9**	**5.68334**	**11.682677**	
	502.25	**10**	**6.253312**	**12.257866**	
	554.75	**11**	**7.104327**	**13.1167**	
	607.25	**12**	**7.950644**	**13.970829**	
	712.25	**13**	**9.629343**	**15.665127**	
	747.25	**14**	**10.184821**	**16.225799**	
	799.75	**15**	**11.014244**	**17.063005**	
	852.25	**16**	**11.839148**	**17.895684**	
	887.25	**17**	**12.386591**	**18.448306**	
	939.75	**18**	**13.204044**	**19.273521**	
	1044.75	**19**	**14.825719**	**20.910698**	
	1149.75	**20**	**16.429991**	**22.53044**	
	1254.75	**21**	**18.017137**	**24.133026**	
	1359.75	**22**	**19.587431**	**25.718731**	
	1412.25	**23**	**20.366342**	**26.505336**	
SONATA	**1464.75**	**24**	**21.14114**	**27.28782**	**Continuous**

	1517.25	**25**	**21.911857**	**28.066217**	
	1569.75	**26**	**22.678526**	**28.840557**	
	1674.75	**27**	**24.199845**	**30.377198**	
	1727.25	**28**	**24.954558**	**31.139561**	
	1779.75	**29**	**25.705348**	**31.897993**	
	1832.25	**30**	**26.452245**	**32.652526**	
	30.02	**1**	**0.956814**	**7.340516**	
	31.46	**2**	**1.096422**	**7.490942**	
XIE	**53.93**	**3**	**3.145993**	**9.707533**	**Accept**
	10	**1**	**-1.425715**	**4.755247**	
	19	**2**	**-0.759069**	**5.480016**	
	32	**3**	**0.156309**	**6.478616**	
	43	**4**	**0.890075**	**7.282146**	
	58	**5**	**1.835309**	**8.321581**	
	70	**6**	**2.549062**	**9.109948**	
	88	**7**	**3.555208**	**10.226823**	
	103	**8**	**4.339626**	**11.102466**	
	125	**9**	**5.410192**	**12.305183**	
	150	**10**	**6.523917**	**13.566826**	
	169	**11**	**7.305376**	**14.459184**	
	199	**12**	**8.438666**	**15.765056**	
	231	**13**	**9.52864**	**17.035912**	
	256	**14**	**10.3058**	**17.952224**	
IBM	**296**	**15**	**11.431953**	**19.297328**	**Continuous**

IV.CONCLUSION

The SPRT methodology for the software reliability growth model of Burr type III and Pareto Type II applied for the software failure data sets are considered and an analogous study is made on them. From the observation we are able to come to a conclusion in a very less time regarding the reliability or unreliability of a software product and the results obtained from the datasets of Burr Type III exemplify that the model has given a decision of acceptance for all the data sets at very first instance of the data and has given an early decision in comparison with the Pareto Type II software reliability model.

REFERENCE

[1] Marinos, Swapna S. Gokhale Peter N., and Kishor S. Trivedi. "Important Milestones in Software Reliability Modeling." In Proceedings of Software Engineering and Know ledge Engineering (SEKE' 96), Lake Tahoe, NV, pp.345-352.1996.

[2] STIEBER, H.A.(1997). "Statistical Quality Control: How To Detect Unreliable Software Components", Proceedings the 8th International Symposium on Software Reliability Engineering, 8-12

[3] Dr. R. Satya Prasad,N Geetha Rani, Dr. R. R. L. Kantam, "Detection of Reliable Software Using SPRT & Pareto Type II SRGM", IJCST Vol. 3, Issue 4, Oct - Dec 2012,

[4] Card D., (1994), "Statistical Process Control for Software", IEEE Software, May, 95-97 Detection of Reliable Software Using SPRT & Pareto Type II SRGM.

[5] John D.Musa; "Software Quality and Reliability Basics"; AT&T Bell Laboratories. CH 2468-7/87/0000/014,1987 IEEE.

[6] Wald A. Sequential Analysis. New Impression edition. New York: John Wiley and Son, Inc; 1947 Sep 30.

[7] K.B.Misra."Handbook of Performability Engineering". Springer. 2008.

[8] Reckase MD. A Procedure for decision making using tailored testing. In: Weiss DJ, editor. New horizons in testing: Latent trait theory and computerized adaptive testing. New York: Academic Press; 1983. p. 237–54.

[9] V.Goutham, R.Satya Prasad "An SPRT Procedure for an Ungrouped Data using MMLE Approach", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727,Volume 14, Issue 6 (Sep. - Oct. 2013), PP 37-42.

[10] Ch.Smitha Chowdary, Dr R.Satya Prasad, K.Sobhana, "Burr Type III Software Reliability Growth Model," IOSR-JCE Volume17, Issue 1, Jan-Feb 2015.

[11] Dr.R.Satya Prasad, N.Geetha Rani & Prof.R.R.L. Kantam, "Pareto Type II Based Software Reliability Growth Model", International Journal of Software Engineering (IJSE), Volume (2) : Issue (4) , 2011.

[12] Goel, A.L., Okumoto, K., "Time- dependent error-detection rate model for software reliability and other performance measures". IEEE Trans. Reliab. R-28, 206-211,1979

[13] Satya Prasad R. Half Logistic Software Reliability Growth Model [PhD Thesis]. India: ANU; 2007.

[14] CH.Smitha, Dr.R.Satya Prasad, Dr.R.Kiran Kumar, "Burr Type III Process Model with SPRT for Software Reliability", International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2763, Issue 12, Volume 3 (December 2016)

[15] Dr. R. Satya Prasad, N Geetha Rani, Dr. R. R. L. Kantam, "Detection of Reliable Software Using SPRT & Pareto Type II SRGM", IJCST Vol. 3, Issue 4, Oct - Dec 2012.

[16] Pham H. System Software Reliability. Springer; 2006.

[17] Lyu MR. The Hand book of software Reliability engineering. McGrawHill and IEEE Computer Society Press; 1995. ISBN: 9-07-039400-8.

[18] Ashoka M. Data set. Bangalore: Sonata Software Limited; 2010.