Research Paper
Impact Factor: 7.184
Peer Reviewed & Indexed Journal
www.iibarr.com

IJBARR
E- ISSN -2347-856X
ISSN -2348-0653

FAILURE DATA SETS- A STUDY

Dr. Sattaru Janardhana Rao
Team Lead (Academic Affairs), Andhra Pradesh Information Technology Academy, Vijayawada-520010.

Abstract
This paper lists several application data sets that, throughout the thesis they are used to implement and illustrate the use of Sequential Probability Ratio Test in software reliability using Half Logistic Software Reliability Growth Model.

Key Words: Data, Sets, Software.

Introduction:

Time domain data sets

Data Set #1: Data collected from (Xie *et al.*,2002).
DataSet#2:USNavalTacticalDataSystems(NTDS)

The Software data set was extracted from information about failures in the development of software for the real-time multi-computer complex of the US Naval Fleet Computer Programming Center of the US Naval Tactical Data Systems (NTDS) (Goel,1979). The software consists of 38 different project modules. The time horizon is divided into four phases: production phase, test phase, user phase, and subsequent test phase. The 26 software failures were found during the production phase.

DataSet#3:AT&TSystem T Project
The AT&T‟s System T is a network-management system developed by AT&T that receives data from telemetry events, such as alarms, facility-performance information, and diagnostic messages, and forwards them to operators for further action. The system has been tested and failure data has been collected (Ehrlich *et al.*, 1993). The following Table shows the failures and the inter-failure times (in CPU units).

DataSet#4:On-LineData Entry IBM Software Package
The data reported by Ohba (1984) are recorded from testing an on-line data entry software package developed at IBM. The following table shows the number of errors and the inter failure time.

DataSet#5:Lyu. **Data Set #6: S2**.

Table1.9.1.1:DataSet#1

F.No	TBF(h)	F.No	TBF(h)	F.No	TBF(h)
1	30.02	**11**	0.47	**21**	70.47
2	1.44	**12**	6.23	**22**	17.07
3	22.47	**13**	3.39	**23**	3.99
4	1.36	**14**	9.11	**24**	176.06
5	3.43	**15**	2.18	**25**	81.07
6	13.2	**16**	15.53	**26**	2.27
7	5.15	**17**	25.72	**27**	15.63
8	3.83	**18**	2.79	**28**	120.78
9	21	**19**	1.92	**29**	30.81
10	12.97	**20**	4.13	**30**	34.19

Table1.9.1.2:DataSet#2

F.No	TBF	F.No	TBF	F.No	TBF
1	9	**10**	7	**19**	6
2	12	**11**	1	**20**	1
3	11	**12**	6	**21**	11
4	4	**13**	1	**22**	33
5	7	**14**	9	**23**	7
6	2	**15**	4	**24**	91
7	5	**16**	1	**25**	2
8	8	**17**	3	**26**	1
9	5	**18**	3		

Table1.9.1.3:DataSet#3

F.No	IFT	F.No	IFT	F.No	IFT
1	10	**6**	12	**11**	19
2	9	**7**	18	**12**	30
3	13	**8**	15	**13**	32
4	11	**9**	22	**14**	25
5	15	**10**	25	**15**	40

Table1.9.1.4:DataSet#4

F.No	IFT	F.No	IFT	F.No	IFT
1	5.5	**9**	11.39	**17**	125.67
2	1.83	**10**	19.88	**18**	82.69
3	2.75	**11**	7.81	**19**	0.46
4	70.89	**12**	14.6	**20**	31.61
5	3.94	**13**	11.41	**21**	129.31
6	14.98	**14**	18.94	**22**	47.6
7	3.47	**15**	65.3		
8	9.96	**16**	0.04		

Table1.9.1.5:DataSet#5

F.No	IFT	F.No	IFT	F.No	IFT
1	0.5	**9**	1.4	**17**	3.2
2	1.2	**10**	3.5	**18**	2.5
3	2.8	**11**	3.4	**19**	2
4	2.7	**12**	1.2	**20**	4.5
5	2.8	**13**	0.9	**21**	3.5
6	3	**14**	1.7	**22**	5.2
7	1.8	**15**	1.4	**23**	7.2
8	0.9	**16**	2.7	**24**	10.7

Table1.9.1.6:DataSet#6

F.No	TBF(s)	F.No	TBF(s)	F.No	TBF(s)	F.No	TBF(s)
1	3.183	**15**	6.783	**29**	13.333	**43**	260
2	3.7	**16**	0.833	**30**	65.167	**44**	0
3	4.667	**17**	11	**31**	115	**45**	0
4	4.833	**18**	25.117	**32**	55	**46**	5
5	4.834	**19**	10.417	**33**	25.166	**47**	150.35
6	6.416	**20**	15.2	**34**	3.25	**48**	41.983
7	9.5	**21**	10.633	**35**	32.6	**49**	114.834
8	10.167	**22**	4.883	**36**	2.25	**50**	55.8
9	6.083	**23**	20.2	**37**	11.017	**51**	45.833
10	6.5	**24**	10.2	**38**	0.833	**52**	111.25
11	4.584	**25**	11.25	**39**	12.15	**53**	115.75
12	6	**26**	20.25	**40**	15	**54**	131.65
13	13.333	**27**	45.25	**41**	3		
14	20.167	**28**	59.184	**42**	70.417		

Time domain data sets for ordered statistics

Data Sets:The Real-time Control System Data

The data sets were listed in"DATA"directory Containing 45 industry project failure data sets in the Hand book of Software Reliability Engineering(Lyu,1996).

Table1.9.2.1:DataSet#7,SYS1 data

F.No	TBF(s)	F.No	TBF(s)	F.No	TBF(s)	F.No	TBF(s)
1	3	**35**	227	**69**	529	**103**	108
2	30	**36**	65	**70**	379	**104**	0
3	113	**37**	176	**71**	44	**105**	3110
4	81	**38**	58	**72**	129	**106**	1247
5	115	**39**	457	**73**	810	**107**	943
6	9	**40**	300	**74**	290	**108**	700
7	2	**41**	97	**75**	300	**109**	875
8	91	**42**	263	**76**	529	**110**	245
9	112	**43**	452	**77**	281	**111**	729
10	15	**44**	255	**78**	160	**112**	1897
11	138	**45**	197	**79**	828	**113**	447
12	50	**46**	193	**80**	1011	**114**	386
13	77	**47**	6	**81**	445	**115**	446
14	24	**48**	79	**82**	296	**116**	122
15	108	**49**	816	**83**	1755	**117**	990
16	88	**50**	1351	**84**	1064	**118**	948
17	670	**51**	148	**85**	1783	**119**	1082
18	120	**52**	21	**86**	860	**120**	22
19	26	**53**	233	**87**	983	**121**	75
20	114	**54**	134	**88**	707	**122**	482
21	325	**55**	357	**89**	33	**123**	5509

22	55	**56**	193	**90**	868	**124**	100
23	242	**57**	236	**91**	724	**125**	10
24	68	**58**	31	**92**	2323	**126**	1071
25	422	**59**	369	**93**	2930	**127**	371
26	180	**60**	748	**94**	1461	**128**	790
27	10	**61**	0	**95**	843	**129**	6150
28	1146	**62**	232	**96**	12	**130**	3321
29	600	**63**	330	**97**	261	**131**	1045
30	15	**64**	365	**98**	1800	**132**	648
31	36	**65**	1222	**99**	865	**133**	5485
32	4	**66**	543	**100**	1435	**134**	1160
33	0	**67**	10	**101**	30	**135**	1864
34	8	**68**	16	**102**	143	**136**	4116

Table1.9.2.2:DataSet#8,SYS2 data

F.No	TBF(ms)	F.No	TBF(ms)	F.No	TBF(ms)	F.No	TBF(ms)
1	479	**23**	437	**45**	460	**67**	1866
2	266	**24**	340	**46**	565	**68**	490
3	277	**25**	405	**47**	1119	**69**	1487
4	554	**26**	535	**48**	437	**70**	4322
5	1034	**27**	277	**49**	927	**71**	1418
6	249	**28**	363	**50**	4462	**72**	1023
7	693	**29**	522	**51**	714	**73**	5490
8	597	**30**	613	**52**	181	**74**	1520
9	117	**31**	277	**53**	1485	**75**	3281
10	170	**32**	1300	**54**	757	**76**	2716
11	117	**33**	821	**55**	3154	**77**	2175
12	1274	**34**	213	**56**	2115	**78**	3505
13	469	**35**	1620	**57**	884	**79**	725
14	1174	**36**	1601	**58**	2037	**80**	1963
15	693	**37**	298	**59**	1481	**81**	3979
16	1908	**38**	874	**60**	559	**82**	1090
17	135	**39**	618	**61**	490	**83**	245
18	277	**40**	2640	**62**	593	**84**	1194
19	596	**41**	5	**63**	1769	**85**	994
20	757	**42**	149	**64**	85	**86**	3902
21	437	**43**	1034	**65**	2836		
22	2230	**44**	2441	**66**	213		

Table1.9.2.3:DataSet#9,SYS3 data

F.No	TBF(s)	F.No	TBF(s)	F.No	TBF(s)	F.No	TBF(s)
1	39	**53**	4	**105**	9	**157**	108
2	10	**54**	23	**106**	26	**158**	38
3	4	**55**	9	**107**	62	**159**	86
4	36	**56**	13	**108**	239	**160**	7

5	4	**57**	165	**109**	13	**161**	22
6	5	**58**	14	**110**	4	**162**	80
7	4	**59**	22	**111**	85	**163**	239
8	91	**60**	41	**112**	85	**164**	3
9	49	**61**	12	**113**	240	**165**	39
10	1	**62**	138	**114**	178	**166**	63
11	25	**63**	95	**115**	34	**167**	152
12	1	**64**	49	**116**	102	**168**	63
13	4	**65**	62	**117**	9	**169**	80
14	30	**66**	2	**118**	146	**170**	245
15	42	**67**	35	**119**	59	**171**	196
16	9	**68**	89	**120**	48	**172**	46
17	49	**69**	90	**121**	25	**173**	152
18	44	**70**	69	**122**	25	**174**	102
19	32	**71**	22	**123**	111	**175**	9
20	3	**72**	15	**124**	5	**176**	228
21	78	**73**	19	**125**	31	**177**	220
22	1	**74**	42	**126**	51	**178**	208
23	30	**75**	14	**127**	6	**179**	78
24	205	**76**	11	**128**	193	**180**	3
25	5	**77**	41	**129**	27	**181**	83
26	129	**78**	210	**130**	25	**182**	6
27	103	**79**	16	**131**	96	**183**	212
28	224	**80**	30	**132**	26	**184**	91
29	186	**81**	37	**133**	30	**185**	3
30	53	**82**	66	**134**	30	**186**	10
31	14	**83**	9	**135**	17	**187**	172
32	9	**84**	16	**136**	320	**188**	21
33	2	**85**	14	**137**	78	**189**	173
34	10	**86**	24	**138**	39	**190**	371
35	1	**87**	12	**139**	13	**191**	40
36	34	**88**	159	**140**	13	**192**	48
37	170	**89**	89	**141**	19	**193**	126
38	129	**90**	118	**142**	128	**194**	90
39	4	**91**	29	**143**	34	**195**	149
40	4	**92**	21	**144**	84	**196**	30
41	35	**93**	18	**145**	40	**197**	317
42	5	**94**	2	**146**	177	**198**	500
43	5	**95**	114	**147**	349	**199**	673
44	22	**96**	37	**148**	274	**200**	432
45	36	**97**	46	**149**	82	**201**	66
46	35	**98**	17	**150**	58	**202**	168
47	121	**99**	1	**151**	31	**203**	66
48	23	**100**	150	**152**	114	**204**	66
49	33	**101**	382	**153**	39	**205**	128

50	48	**102**	160	**154**	88	**206**	49
51	32	**103**	66	**155**	84	**207**	332
52	21	**104**	206	**156**	232		

Table1.9.2.4:DataSet#10,DS2 Data

FNO	TBF (hrs)	FNO	TBF (hrs)	FNO	TBF (hrs)	FNO	TBF (hrs)
1	12	**29**	7	**57**	0	**85**	6
2	3	**30**	1	**58**	1	**86**	18
3	13	**31**	0	**59**	2	**87**	0
4	11	**32**	7	**60**	5	**88**	0
5	14	**33**	10	**61**	5	**89**	0
6	0	**34**	6	**62**	4	**90**	6
7	7	**35**	4	**63**	0	**91**	0
8	0	**36**	1	**64**	3	**92**	12
9	0	**37**	8	**65**	2	**93**	14
10	3	**38**	2	**66**	6	**94**	1
11	5	**39**	10	**67**	6	**95**	20
12	0	**40**	15	**68**	3	**96**	21
13	14	**41**	1	**69**	15	**97**	3
14	9	**42**	5	**70**	0	**98**	5
15	6	**43**	0	**71**	4	**99**	41
16	0	**44**	2	**72**	1	**100**	1
17	5	**45**	1	**73**	5	**101**	14
18	1	**46**	0	**74**	3	**102**	14
19	0	**47**	1	**75**	11	**103**	57
20	1	**48**	4	**76**	2	**104**	0
21	1	**49**	8	**77**	20	**105**	11
22	4	**50**	2	**78**	20	**106**	73
23	0	**51**	4	**79**	3	**107**	6
24	4	**52**	3	**80**	1	**108**	35
25	12	**53**	3	**81**	12	**109**	14
26	1	**54**	3	**82**	5	**110**	211
27	5	**55**	2	**83**	1	**111**	70
28	27	**56**	3	**84**	0		

Interval domain datasets

Table1.9.3.1:DataSet#11,Release2 data

Week Index	Fault	Week Index	Fault	Week Index	Fault	Week Index	Fault
1	13	**6**	8	**11**	6	**16**	2
2	5	**7**	13	**12**	5	**17**	3
3	8	**8**	14	**13**	4	**18**	1
4	8	**9**	9	**14**	6	**19**	2
5	6	**10**	5	**15**	2		

Table1.9.3.2:DataSet#12,Release3 data

Week Index	Fault	Week Index	Fault	Week Index	Fault	Week Index	Fault
1	6	**4**	7	**7**	8	**10**	2
2	3	**5**	8	**8**	6	**11**	1
3	4	**6**	12	**9**	3	**12**	1

Parameter estimation is of primary importance in softwarere liability prediction.Once the analytical solution for$m(t)$ is known for a given model,

parameter estimation is achieved by applying a technique of Maximum Likelihood Estimate (MLE). The idea behind maximum likelihood parameter estimation is to determine the parameters that maximize the probability (likelihood) of the sample data. The method of maximum likelihood is considered to be more robust and yields estimators with good statistical properties.Assuming that the data are given for the cumulative number of detected errors y_i in a given time-interval $(0,t_i)$

where i = 1,2, …, n. and $0 < t_1 < t_2 < \ldots < t_n$then the log likelihood function(LLF)takes on the following form.

Likelihood function by using λ(t)is:$L = e^{-m(t)} \prod_{i=1}^{n} \lambda(t_i)$

The logarithmic likelihood function for interval domain data (Pham, 2006) is given by:

$LogL = \sum (y_i - y_{i-1}) . \log\left[\lfloor m(t_i) - m(t_{i-1}) \rfloor \right] - m(t_n)$,,^“and,,^“are Maximum Likelyhood Estimates(MLEs)of parameters and the values canbe computed using iterative method for the given cumulative time between failures data. Using,,a“and ,,b“values wecan compute$m(t)$. Now the control limits are calculated by equating the cumulative distribution function to the standard values 0.00135, 0.99865, and 0.5. These limits are converted to $m(t_U)$, $m(t_C)$ and $m(t_L)$ form.They are used To find whether thes oftware process is in control or not by placing the points in failure control chart. A point below the control limit $m(t_L)$ indicates an alarming signal.A point above the control limit $m(t_U)$ indicates better quality.If the points are falling within the control limits it indicates the software process is in stable.

The respective brief contents of these four problems are given in the “introduction”. The numerical calculations and subsequent tables are provided at appropriate places in the respective chapters. The reprints of some of our findings in published form are appended towards the end of the thesis. List of references arranged alphabetically is also provided towards the end of thethesis.

References

1. Arnold B.C., And Balakrisha N.,(1989), “Relations, Bounds And approximations for order statistics”,Lecturenotes in Statisticsno-53,Springer-Verlag.
2. Balakrishna N., And Cohen A.C.,(1991), “ Order Statistics And Inference: Estimation Methods”, Academic Press.
3. David,H.A.Andnagaraja,H.N.,(2003).“Orderstatistics”, 3rdedition,Johnwiley& Sons.27-32
4. Kantam R.R.L., And Sriram B., (2001). “Variable Control Chartsbased Ongammadistribution”, Iapqr Transactions 26(2), 63-78.
5. Kantam,R.R.L.And Dharmarao, V., (1994).“HalfLogistic Distribution An imporivement overm.L.Estimtor”, Proceedings of 11annual Conferenceof Sds, 39-44.
6. Pham. H., (2006). “System Software Reliability”, Springer.
7. Xie,M.,Goh,T.N.Andranjan.P.,(2002).“Someeffective Control Chart Procedures For Reliability Monitoring”,Reliability engineering and system safety,77,143- 150.