On Determining Sample Size and Testing Duration of Repairable System Test

Huairui Guo, PhD, ReliaSoft Corporation
Rong Pan, PhD, Arizona State University

Key Words: Repairable system; Maximum likelihood estimation; Reliability test plan; Fisher information matrix

SUMMARY & CONCLUSIONS

Reliability demonstration tests for non-repairable systems have been extensively discussed by many researchers. However, very few works have been done for repairable system tests. Demonstration tests for repairable systems can be time consuming and costly. Carefully planning sample size and test duration is very important. This paper develops a theoretical method, based on pivotal quantities and a confidence bound requirement for the reliability metrics of interest, to help test planners to determine the minimal sample sizes and test duration. A case study was given and the developed theoretical results were compared with simulation results. The comparison shows that the proposed method is accurate and efficient.

ACRONYMS

CI Confidence Interval
NHPP Non-homogeneous Poisson Process
MLE Maximum Likelihood Estimation

1 INTRODUCTION

Life test design for non-repairable parts has been extensively discussed in literature. For example, Lawless [1] developed a procedure for determining the required sample size and number of failures based on the required confidence intervals of model parameters, providing a table for Weibull and Extreme Value distributions. McKane, Escobar and Meeker [2] generalized Lawless's method and developed charts for sample size determination for log-location-scale distributions. In order to reduce the test time and the cost, life tests are usually conducted at higher stress levels than the normal usage condition. Many accelerated test plans for non-repairable systems have also been discussed in the recent years (see, *e.g.*, [3-5]). Recent research has made extensions of previous works, such as designing test plans for high usage rates [6] and for multiple objective functions [7].

However, very few papers have discussed test plans for repairable systems. Singh and Swaminathan [8] derived the exact confidence interval of system availability based on the assumption of exponential distributions for time-to-failure and time-to-repair. Usher and Taylor [9] showed the design of availability test plans that meet the stated levels of producer and consumer risks. The existing papers focus on a single system only and have not discussed test plans for a repairable system with certain failure intensity function. Due to the high cost and lengthy testing time that a typical system test consumes, it is necessary to develop a method for sample size determination for repairable system tests. With a proper sample size, the test results can be used to predict the number of failures in a given time period with greater confidence. Resources, such as budget, crew and spare parts, also can be properly allocated in advance.

In this paper, a method for planning a life test for repairable systems is proposed. A repairable system is assumed to undergo "minimal repair" at the occurrence of each failure so that the rate of system failure is not disturbed by the repair. This assumption is used by many NHPP (non-homogeneous Poisson process) models, such as the log linear failure intensity model by Cox and Lewis [10], Crow-AMSAA model [11], and bounded failure intensity models by Pulcini [12] and Attardi and Pulcini [13]. With the proposed method, a planner can determine the sample size and/or the total testing time so that the failure intensity function or the expected number of failures can be estimated with the required precision. This type of test can be applied at the final system validation stage, where the required system reliability needs to be demonstrated. The test results can also assist in allocating resources to maintain the system. For example, an automobile manufacturer may need to collect vehicle failure data with minimal effort in order to optimize its warranty policy. Therefore, it is important to design a test that will yield sufficient data and accurate results in the most time- and cost-effective way.

2 METHODOLOGY

In this section, based on the asymptotic properties of parameter estimation, a method for determining required sample size and test duration for repairable system tests is proposed. According to the required accuracy of predicted number of failures or model parameters, the sample size for the test can be properly determined. For convenience, the Crow-AMSAA model, which is a power law NHPP model, will be used in this section. The proposed method also can be extended to other NHPP models, such as the log-linear model and the bounded failure intensity model. The analytical solution for the power law model is provided and the results are compared with the simulation results provided in Section 3. The comparison shows that the proposed method is valid and will have a wide application for repairable system test planning.

2.1 *Maximum Likelihood Estimation of a Power Law NHPP model*

For a power law NHPP model the expected number of failures and failure intensity function for a single repairable system are derived from the following equation:

$$\lambda(t) = \theta\beta t^{\beta-1} \tag{1}$$

where θ is a scale parameter (sometimes called the intrinsic failure rate) and β is a shape parameter. The expected number of failures at time *t* of a Poisson process is the same as the cumulative failure intensity value, which is:

$$E[N(t)] = \Lambda(t) = \int_0^t \theta\beta u^{\beta-1} du = \theta t^{\beta} \tag{2}$$

For a single repairable system, the likelihood function of observing *n* failures at times $t_1, t_2, ..., t_n$ with the censoring time at *T* is:

$$L(\theta,\beta \mid t_1,t_2,...,t_n,T) = \prod_{i=1}^{n}\left(\lambda(t_i)e^{-[\Lambda(t_i)-\Lambda(t_{i-1})]}\right)\times e^{-[\Lambda(T)-\Lambda(t_n)]} \tag{3a}$$

$$= \prod_{i=1}^{n}\lambda(t_i)\times e^{-\Lambda(T)} = \theta^n\beta^n e^{-\theta T^{\beta}}\prod_{i=1}^{n} t_i^{\beta-1}$$

Taking the logarithm of the likelihood function, we get the log-likelihood function:

$$\ln L(\theta,\beta \mid t_1,t_2,...,t_n,T) = n\ln(\theta\beta) - \theta T^{\beta} + \sum_{i=1}^{n}(\beta-1)\ln t_i \tag{3b}$$

Assume there are *m* identical repairable systems. These systems are tested until time *T* and the failure times of these systems are recorded as $\{t_{1,1},t_{1,2},...,t_{1,n_1}\},...,\{t_{m,1},t_{m,2},...t_{m,n_m}\}$, then the likelihood and log-likelihood functions for all the *m* systems are:

$$L(\theta,\beta \mid \{t_{j,1},t_{j,2},...,t_{j,n_j}\}_{j=1,...,m},T) = (\theta\beta)^{\sum_{j=1}^{m} n_j} e^{-\theta T^{\beta} m}\prod_{j=1}^{m}\prod_{i=1}^{n_j} t_{j,i}^{\beta-1} \tag{4a}$$

and:

$$\ln L(\theta,\beta \mid \{t_{j,1},t_{j,2},...,t_{j,n_j}\}_{j=1,...,m},T) \tag{4b}$$
$$= (\sum_{j=1}^{m} n_j)\ln(\theta\beta) - \theta T^{\beta} m + (\beta-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\ln t_{j,i}$$

In order to obtain the MLEs of θ and β, we take the first derivative of the log-likelihood function with respect to θ and β and set them to zero. It can be shown that [11]:

$$\hat{\theta} = \frac{\sum_{j=1}^{m} n_j}{T^{\hat{\beta}} m} \tag{5}$$

and:

$$\hat{\beta} = \frac{\sum_{j=1}^{m} n_j}{\sum_{j=1}^{m} n_j \ln T - \sum_{j=1}^{m}\sum_{i=1}^{n_j}\ln t_{j,i}} \tag{6}$$

2.2 *Confidence Intervals for θ, β and $N(t)$*

Normal-approximation confidence regions are usually used for ML estimates of model parameters and functions of model parameters. In this section, we will build the confidence interval for θ, β and $N(t)$. Similar work was done by Lawless [1] for non-repairable systems with Weibull and Extreme Value distributions, Nelson and Schmee [3] for lognormal distribution and McKane, Escobar and Meeker [2] for general log-location-scale distributions. The Fisher information matrix is used to construct the variance-covariance matrix for model parameters. By taking the expected value of the negative of second derivatives of the log-likelihood function, we obtain the expected Fisher information matrix as:

$$I = -E\begin{bmatrix} -\frac{\sum_{j=1}^{m} n_j}{\beta^2} - \theta m T^{\beta}\ln^2 T & -mT^{\beta}\ln T \\ -mT^{\beta}\ln T & \frac{\sum_{j=1}^{m} n_j}{\theta^2} \end{bmatrix} \tag{7}$$
$$= mT^{\beta}\begin{bmatrix} \frac{\theta}{\beta^2} + \theta\ln^2 T & \ln T \\ \ln T & \frac{1}{\theta} \end{bmatrix}$$

The approximated variance-covariance matrix of the β and θ estimators is:

$$\Sigma = \begin{bmatrix} \mathrm{var}(\hat{\beta}) & \mathrm{cov}(\hat{\beta},\hat{\theta}) \\ \mathrm{cov}(\hat{\beta},\hat{\theta}) & \mathrm{var}(\hat{\theta}) \end{bmatrix} = I^{-1} \tag{8}$$
$$= \left(\frac{\beta^2}{mT^{\beta}}\right)\begin{bmatrix} \frac{1}{\theta} & -\ln T \\ -\ln T & \frac{\theta}{\beta^2} + \theta\ln^2 T \end{bmatrix}$$

Therefore, the large sample approximations of the standard errors of β and θ estimators are:

$$std(\hat{\beta}) = \frac{\beta}{\sqrt{mT^{\beta}\theta}} \tag{9}$$

and:

$$std(\hat{\theta}) = \sqrt{\frac{\theta}{mT^{\beta}}(1+\beta^2\ln^2 T)} \tag{10}$$

Usually, the model parameter estimators, especially for the parameters that can only have positive values, are assumed lognormally distributed [1, 3, 14]. In the simulation part of Section 3, we will see that the estimates of β indeed can be fitted very well by a lognormal distribution, while the estimates of θ slightly deviate from a lognormal distribution. In this analysis, following the tradition, we will assume that both estimators of β and θ are lognormally distributed. The $(1-\alpha)\%$ confidence interval of β is:

$$[\hat{\beta}e^{-z_{\alpha/2}std(\hat{\beta})/\hat{\beta}}, \hat{\beta}e^{z_{\alpha/2}std(\hat{\beta})/\hat{\beta}}] \tag{11}$$

The confidence interval of θ is:

$$[\hat{\theta}e^{-z_{\alpha/2}std(\hat{\theta})/\hat{\theta}}, \hat{\theta}e^{z_{\alpha/2}std(\hat{\theta})/\hat{\theta}}] \tag{12}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile of a standard normal distribution.

Since the expected number of total failures is a function of β and θ, we may derive the variance of the expected number of failures at a given time *t* for a single repairable system. The estimated number of failures expected for a single repairable system is $\hat{\Lambda}(t) = \hat{E}[N(t)] = \hat{\theta} t^{\hat{\beta}}$ and its variance is:

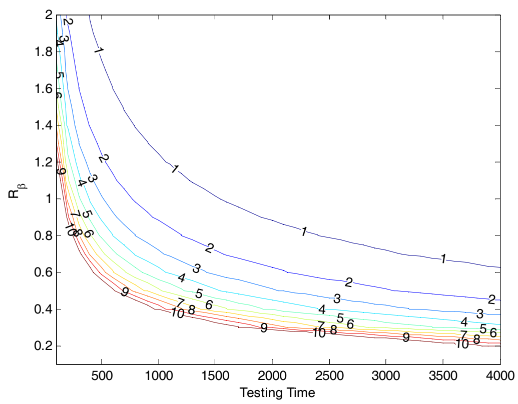
$$\begin{aligned} Var(\hat{\Lambda}(t)) &\approx (t^{\beta})^2 Var(\hat{\theta}) + (\theta t^{\beta} \ln(t))^2 Var(\hat{\beta}) \\ &\quad + 2(t^{\beta})(\theta t^{\beta} \ln(t)) Cov(\hat{\theta}, \hat{\beta}) \\ &= t^{2\beta} \frac{\theta(1+\beta^2 \ln^2 T)}{mT^{\beta}} + \theta^2 t^{2\beta} \ln^2(t) \frac{\beta^2}{mT^{\beta}\theta} \\ &\quad - 2t^{2\beta}\theta \ln(t) \frac{\beta^2 \ln T}{mT^{\beta}} \\ &= \frac{t^{2\beta}\theta}{mT^{\beta}} [1 + \beta^2 (\ln T - \ln t)^2] \end{aligned} \tag{13}$$

The corresponding confidence interval of $\Lambda(t)$ can be easily obtained as $[\hat{\Lambda}(t) e^{-z_{\alpha/2} std(\hat{\Lambda}(t))/\hat{\Lambda}(t)}, \hat{\Lambda}(t) e^{z_{\alpha/2} std(\hat{\Lambda}(t))/\hat{\Lambda}(t)}]$.

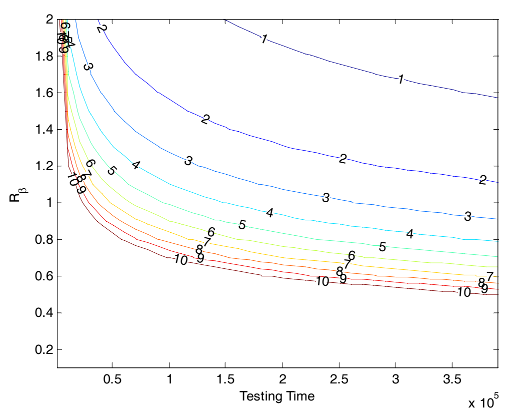
2.3 *Sample Size and Test Duration Determination*

As discussed at the beginning of this paper, the purpose of a test is to get the accurate estimation of model parameters or the function of model parameters for a repairable system. To achieve the required precision of an estimation, for example, where the expected log ratio of the upper and lower bounds in the confidence interval of β (or equivalently, where the expected difference of the upper and lower bounds of the $\ln \beta$), is less than R_{β}, we have:

$$2 z_{\alpha/2} / \sqrt{m T^{\beta} \theta} \le R_{\beta} \tag{14}$$



(a)



(b)

Similarly, for the precision requirement of θ, we have:

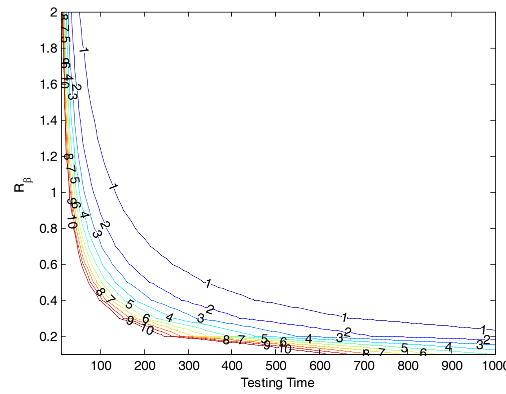
$$2 z_{\alpha/2} \sqrt{\frac{1}{m T^{\beta} \theta} (1 + \beta^2 \ln^2 T)} \le R_{\theta} \tag{15}$$

The expected number of failures at time T for m repairable systems is $E[N_m(T)] = E[\sum_{j=1}^{m} n_j] = m\theta T^{\beta}$. From (13) and (14), one can see that the larger the sample size, the higher the precision of the estimation of model parameters β and θ will be.

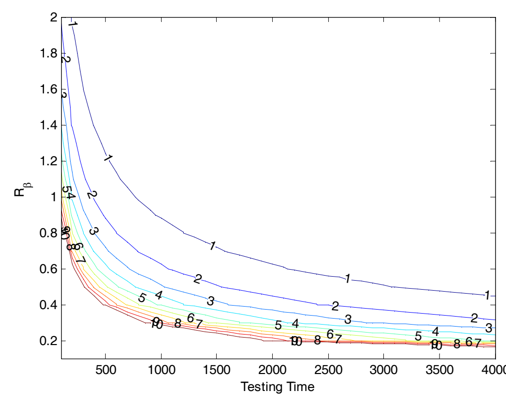
If the test plan criterion is in terms of the precision of the expected number of failures $\Lambda(t)$ at time t, we can find the similar relationship of test sample size, m, and test duration, T, with the required precision, $R_{\Lambda(t)}$. It is:

$$2 z_{\alpha/2} \sqrt{\frac{1}{m T^{\beta} \theta} (1 + \beta^2 (\ln T - \ln t)^2)} \le R_{\Lambda(t)} \tag{16}$$

From Eqns. 14-16, the contour plots of the minimum number of systems and the allowable test duration for a given



(c)



(d)

Figure 1- Contour plots of minimal sample size versus testing time and precision requirement of model parameter β.

The values of the parameters for the curves in Fig. 1 are:

(a) Constant failure intensity, $\beta = 1$ $\theta = 0.01$;

(b) Decreasing failure intensity function, $\beta = 0.5$ $\theta = 0.01$ (note: the scale of time axis is enlarged);

(c) Increasing failure intensity function, $\beta = 1.5$ $\theta = 0.01$ (note: the scale of time axis is shrunk);

(d) Constant failure intensity and larger intrinsic failure rate, $\beta = 1$ $\theta = 0.02$.

For all plots, $\alpha = 0.05$ and $z_{\alpha/2} = 1.96$

precision can be drawn. In order to get these plots, the pivotal values of β and θ should be provided. This contour plot can be used by test planners to design an efficient test for a repairable system. Figure 1 provides the contour plots based on the precision requirement of model parameter β. Similar graphs can be plotted for the precision of θ or $\Lambda(t)$.

If, instead of defining the required precision of each parameter directly, we define the ratios of each parameter to its estimator's standard error as $\hat{SN}_{\beta} = \beta / std(\hat{\beta})$ and $\hat{SN}_{\theta} = \theta / std(\hat{\theta})$, we obtain:

$$\hat{SN}_{\beta} = \sqrt{mT^{\beta}\theta} \tag{17}$$

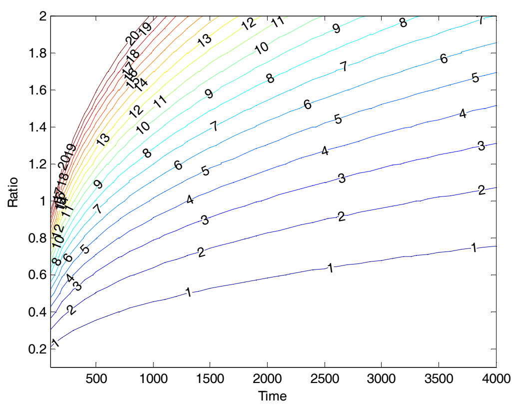
and

$$\hat{SN}_{\theta} = \frac{\sqrt{mT^{\beta}\theta}}{\sqrt{1+\beta^2 \ln^2 T}} \tag{18}$$

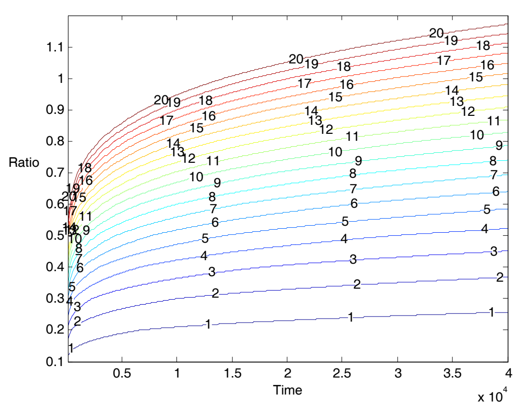
We can define that these ratios must be larger than a specific value, *r*. Since the denominator of the latter equation is always larger than 1, to have both ratios larger than *r* is equivalent to having:

$$\frac{mT^{\beta}}{1+\beta^2 \ln^2 T} \geq \frac{r^2}{\theta} \tag{19}$$

The minimum requirement of sample size and testing time is given in the above formula.



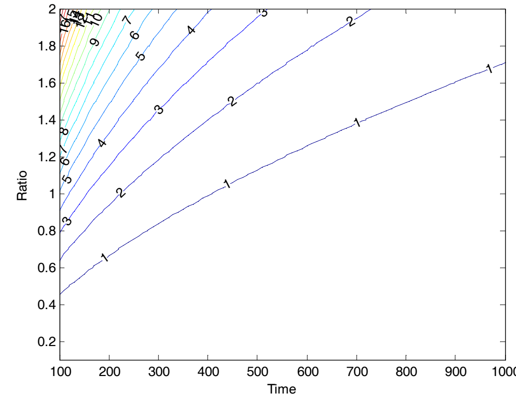
(a)



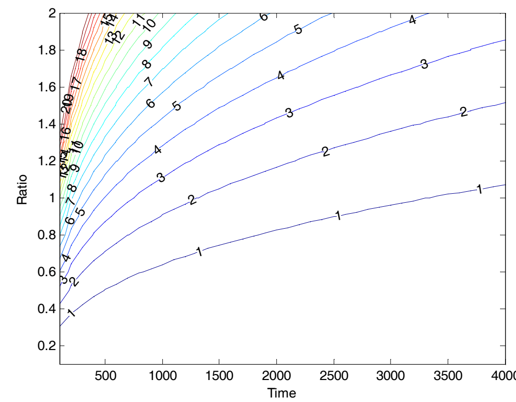
(b)

3 CASE STUDY AND SIMULATION COMPARISON

3.1 *Case Study of Sample Size Determination*



(c)



(d)

Figure 2- Contour plots of minimal sample size versus testing time and precision requirement, r.

The values of the parameters for the curves in Fig. 2 are:

(a) Constant failure intensity, $\beta = 1\ \theta = 0.01$;

(b) Decreasing failure intensity function, $\beta = 0.5\ \theta = 0.01$;

(c) Increasing failure intensity function, $\beta = 1.5$;

(d) Constant failure intensity and larger intrinsic failure rate, $\beta = 1\ \ \theta = 0.02$.

For all plots, $\alpha = 0.05$ *and* $z_{\alpha/2} = 1.96$

An engineer is required to design a test to estimate the expected number of failures for a system. From the test results, the engineer needs to predict the number of failures at 3000 operation cycles. It is required that the log ratio of 90% upper confidence bound to the lower confidence bound of the predicted number of failures be less than 1.2. Suppose that the power law NHPP model is an appropriate model to describe the failure process. From the historical data, it is known that if 10 units are tested for 200 cycles each, 5 failures are expected and if 10 units are tested for 400 cycles each, 12 failures are expected. Because of the constraints of the test, the allowable test duration for the ongoing test is 400 cycles, and the minimal sample size for the test must be decided.

From the above information, the expected number of failures can be obtained from θt^{β}, and

$$\theta \times 200^{\beta} = 0.5 \text{ and } \theta \times 400^{\beta} = 1.2$$

Solving for β and θ, we get the pivotal value of $\beta = 1.263$ and $\theta = 0.00062$. Using (16), we get

$$2z_{\alpha/2}\sqrt{\frac{1}{mT^{\beta}\theta}(1+\beta^2(\ln T - \ln t)^2)} \le R_{N(t)}$$

$$\Rightarrow m \ge 46.5$$

The engineer needs at least 47 units in the test to meet the precision requirement. Therefore, he decides to use 50 units. Using 50 units, the expected standard deviation of β, θ and $N(3000)$ will be:

$$std(\hat{\beta}) = \frac{\beta}{\sqrt{mT^{\beta}\theta}} = 0.1631$$

$$std(\hat{\theta}) = \sqrt{\frac{\theta}{mT^{\beta}}(1+\beta^2 \ln^2 T)} = 0.00061$$

$$std[\hat{N}(t=3000)] = t^{\beta}\sqrt{\frac{\theta}{mT^{\beta}}[1+\beta^2(\ln T - \ln t)^2]} = 5.39$$

Using (16), the expected bound ratio for expected number of failures at time 3000 is

$$2z_{\alpha/2}\sqrt{\frac{1}{mT^{\beta}\theta}(1+\beta^2(\ln T - \ln t)^2)} = 1.158$$

which is less than the required ratio value of 1.2.

3.2 *Comparison with the Simulation Results*

In order to verify the accuracy of the theoretical results, a simulation study is conducted. There are a total of 5000 simulation runs. In each simulation run, 50 systems are tested for 400 cycles. The simulated test results are used to estimate model parameters of β and θ. From the obtained 5000 βs and θs, the mean and the standard deviation of β and θ can be calculated. Similarly, the mean and standard deviation of $N(3000)$ also can be calculated from the simulation results. For easy comparison, the standard deviations calculated from the simulation results and from the Fisher information matrix are given in the following table.

	Theta	Std (theta)	Beta	Std (beta)	N(t)	Std (N(t))
Theoretical Result	0.00062	0.00061	1.263	0.1631	15.3	5.39
Simulation Result	0.00085	0.00085	1.282	0.1660	16.9	6.75

Table 1- Theoretical and Simulation Results

In the process of getting theoretical results, we assume that β, θ and $N(t)$ follow lognormal distribution. In order to verify this assumption, we plot the lognormal probability plot for β, θ and $N(t)$.

From the probability plot, it can be seen that β is indeed lognormally distributed, while θ and $N(t)$ are slightly off from lognormal distribution.

Using the mean and standard deviation of $N(t)$. That are obtained from the simulation, and assuming $N(t)$ is lognormally distributed, we can get the upper and lower bound ratio, which is

$$2z_{\alpha/2}\frac{std[\hat{N}(t)]}{\hat{N}(t)} = 1.29$$

- This is close to the required bound ratio 1.2. The difference is caused by the lognormal approximation in (13) and the simulation itself.

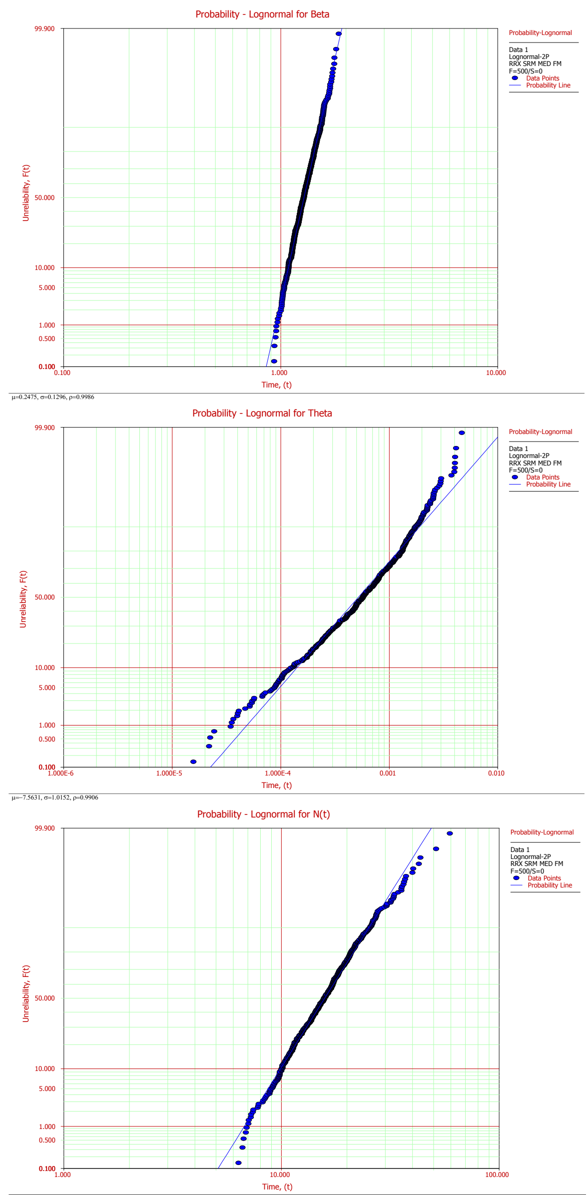


Figure 3 – Probability Plots for the Estimated Parameters

4 CONCLUSIONS

In this paper we discussed the reliability test plan for a repairable system, where the power law NHPP model is used for its failure intensity function. Confidence intervals for model parameters and for the expected number of failures at a given time are derived. They are used to determine the minimal sample size and testing time for a repairable system to meet the precision requirement of estimated system reliability. In general, the higher the precision required, the larger the number of samples and the longer the testing time that are needed. A case study was given and the developed theoretical results were compared with simulation results. The comparison shows that the proposed method is accurate and efficient.

5 ACKNOWLEDGEMENT

This research is partially supported by the NSF Grant CMMI-0654417.

REFERENCES

1. J. F. Lawless, "Confidence Interval Estimation for the Weibull and Extreme Value Distributions," *Technometrics*, vol. 20, 1978, pp 355-364.
2. S. W. McKane, L. A. Escobar and W. Q. Meeker, "Sample Size and Number of Failure Requirements for Demonstration Tests with Log-Location-Scale Distributions and Failure Censoring," *Technometrics*, vol. 47, 2005, pp 182-190.
3. W. Nelson and W. Meeker, "Theory for Optimum Accelerated Censored Life Tests for Weibull and Extreme Value Distributions," *Tehnometrics*, vol. 20, 1978, pp 171-177.
4. W. Meeker, "A Comparison of Accelerated Life Test Plans for Weibull and Lognormal distribution and Type I Censored Data," *Technometrics*, vol. 26, 1984, pp. 157-171.
5. F. G. Pascual and G Montepiedra, **"**Lognormal and Weibull accelerated life test plans under distribution misspecification," *IEEE Transactions on Reliability*, vol. 54, 2005, pp. 43-52,
6. G. Yang, "Accelerated Life Tests at Higher Usage Rates," *IEEE Transactions on Reliability*, vol. 54, 2005, pp. 53-57.
7. L. Tang and K. Xu, "A Multiple Objective Framework for Planning Accelerated Life Tests," *IEEE Transactions on Reliability,* vol. 54, 2005, pp. 58-63.
8. V. P. Singh and S. Swaminathan, "Sample Size for System Availability," *Proceedings Annual Reliability and Maintainability Symposium*, 2002, pp. 51-55.
9. J. Usher and G. Taylor, "Availability Demonstration Testing," *Quality and Reliability Engineering International*, vol. 22, 2006, pp. 473-479.
10. D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*. London: Methuen, 1966.
11. L. H. Crow, "Reliability Analysis for Complex Repairable Systems," In: F. Proschan and R.J. Serfling, Editors, *Reliability and Biometry, SIAM,* 1974, pp 379-410.
12. G. Pulcini, "A Bounded Intensity Process for the Reliability of Repairable Equipment", *Journal of Quality Technology*, vol. 33, 2001, pp 480-492.
13. L. Attardi and G Pulcini, "A New Model for Repairable Systems with Bounded Failure Intensity", *IEEE Transactions on Reliability,* vol. 54, 2005, pp. 72 – 82.
14. W. Nelson and J. Schmee, "Inference for (Log) Normal Life Distributions from Small Singly Censored Samples and BLUE," *Technometrics*, vol. 21, 1979, pp 43–54.

BIOGRAPHIES

Huairui Guo, Ph.D.
Director, Theoretical Development
ReliaSoft Corporation
1450 S. Eastside Loop
Tucson, Arizona 85710, USA

email: Harry.Guo@ReliaSoft.com

Dr. Huairui Guo is the Director of Theoretical Development Department at ReliaSoft Corporation. He received his Ph.D in Systems and Industrial Engineering from the University of Arizona. His dissertation focuses on Process Modeling, Diagnosis and Control for complex manufacturing processes. He also received his M.E. in Manufacturing (2002) from the National University of Singapore and M.S. in Reliability and Quality Engineering from the University of Arizona (2004). He has published papers in the areas of Quality Engineering including SPC, ANOVA and DOE and Reliability Engineering. His current research interests include repairable system modeling, accelerated life/degradation testing, warranty data analysis and robust optimization. He is a member of IEEE, IIE and SRE.

Rong Pan, Ph.D.
Assistant Professor
Department of Industrial Engineering
Arizona State University
Tempe, AZ 85287, USA

email: Rong.Pan@asu.edu

Dr. Rong Pan is an Assistant Professor of Industrial Engineering at Arizona State University. He received his Ph.D in Industrial Engineering from Penn State University in 2002. His research interests include reliability engineering, multivariate statistical quality control, time series analysis and control. He is currently leading a research project of modeling and analysis of profiled reliability testing, which is funded by the National Science Foundation. He is a senior member of ASQ and a member of INFORMS, IIE and SRE