Complex System Maintainability Verification under Mixture Distribution

Zhen-ya WU* and Jian-ping HAO
Shijiazhuang Campus of Army Engineering University, Shijiazhuang, 050003, China
*Corresponding author

Keywords: Complex system, Maintainability verification, Mean time to repair, Mixture distribution, Hypothesis testing.

Abstract. Maintainability verification is an important work in system development. However, two problems arise when the maintenance time of complex system conforms to a mixture distribution. The first problem is the occurrence of assuming log-normal distribution instead of mixture distribution and the second one is the limited sample size for maintainability verification. Regarding the problems mentioned above, firstly, the maintenance sample generation process of complex system is analyzed and the mixture distribution model of maintenance time is proposed. Then, the changes in two types of test risks are analyzed when assuming log-normal distribution instead of mixture distribution. Finally, a method to perform maintainability verification using similar maintenance task’s samples is proposed. A case study using the maintenance time data of satellite communication interface system shows the efficiency of the proposed method.

Introduction

The goal of maintainability verification is to validate whether the maintainability metric (e.g., MTTR) satisfies consumer’s requirement. According to MIL-STD-471A, the test method for verification of MTTR are under the assumption that maintenance time can be adequately described by a log-normal distribution[1]. If no specific assumption concerns the distribution of maintenance time, the verification for MTTR can still be performed as the sample mean is approximately normal for large sample size by the central limit theorem.

In most cases, the log-normal distribution is the best model to be used to describe maintenance time distribution[2,3], while in special cases, the multi-modal distribution, which is also called mixture distribution, is sometimes a good representation for data samples with more than one factor of time dependency[4]. A random variable has a mixture distribution if its probability density function can be expressed by a convex combination of some probability density functions[5]. One of the first analyses involving the use of mixture distribution was undertaken by Karl Pearson which fitted a mixture of two normal probability density functions[6]. Since 1980s, many monographs studied the finite mixture distribution and its application mainly in clustering analysis[7-9].

During MTTR verification, it is usually difficult to obtain enough samples to perform goodness of fit analysis. Thus, the occurrence of assuming log-normal distribution for maintenance time when in fact it has mixture distribution might happen, especially for complex system. However, there are two problems still not completely clarified when the maintenance time of specified item has a mixture distribution:

(1) During the operational test and evaluation stage, it is difficult to obtain enough samples to check whether the maintenance time of the specified item conforms to a mixture distribution. The consequences when assuming log-normal distribution instead of mixture distribution remain to be researched.

(2) In a complex system, some maintenance tasks may show much similarities regarding repair actions, which makes their maintenance time probability density functions vary slightly. Thus, with the bearing of certain degrees of error, whether it is feasible to combine the samples from similar maintenance tasks in order to meet the sample size requirement is still a challenging problem.

The rest of the paper is organized as follows. In section 2, a brief description of the sample generation process for complex system is given and the mixture distribution model of maintenance
time is proposed. In section 3, the error in producer’s and consumer’s risks when assuming log-normal distribution instead of mixture distribution and when combining samples from similar maintenance tasks are analyzed. Section 4 gives a case study using the maintenance records of satellite communication interface system to validate the proposed method. Finally, section 5 gives the conclusions of paper.

**Mixture Distribution Model of Complex System Maintenance Time**

**Basic Definitions of Mixture Distribution**

Suppose that a random variable or vector, \( X \), its distribution can be represented by a probability density function of the form

\[
p(x) = \lambda_1 f_1(x) + \cdots + \lambda_g f_g(x)
\]

where

\[
\lambda_j > 0, \quad \lambda_1 + \cdots + \lambda_g = 1
\]

and

\[
f_j(\cdot) \geq 0 \quad \int f_j(x) dx = 1 \quad j = 1, \ldots, g
\]

In such a case, \( X \) is said to have a finite mixture distribution and that \( p(\cdot) \), defined by Eq. (1), is a finite mixture density function.

The parameters \( \lambda_1, \ldots, \lambda_g \) denote the mixing weights and \( f_1(\cdot), \ldots, f_g(\cdot) \) denote the component densities of the mixture.

**Maintenance Sample Generation Process Analysis of Complex System**

As the major index to the character of maintainability, maintenance time is a function of design but affected by various personnel and logistic factors. Mixture distribution indicates that several distributions are present in the data. These distributions usually stem from different repair actions, different sets of maintenance personnel, maintenance instructions or different environments[10].

According to MIL-HDBK-505, item levels from the simplest division-‘part’, to the most complex division-‘system’, include subassembly, assembly, unit, group, set and subsystem[11]. For a complex system, it may have different types of failure modes due to its complex structure and composition. Thus, its corresponding maintenance tasks may have different time distribution functions caused by different repair actions.

Suppose a system has \( g \) types of maintenance tasks and their relative failure rates are \( \lambda_1, \ldots, \lambda_g (\lambda_1 + \cdots + \lambda_g = 1) \). \( \{T_1, \ldots, T_n\} \) denotes a random maintenance time sample set of size \( n \). Let \( C_j \) denote a \( g \)-dimensional component-label vector, where the \( ith \) element of \( C_j \), \( C_{ij} = (C_j)_i \), is defined to be one or zero, according to whether the sample \( T_j \) is from maintenance task \( i \) or not. Thus, \( C_j \) is distributed according to a multinomial distribution consisting of one draw on \( g \) categories with probabilities \( \lambda_1, \ldots, \lambda_g \); that is,

\[
P(C_j = c_j) = \lambda_1^{c_1} \lambda_2^{c_2} \cdots \lambda_g^{c_g}
\]

Eq. (2) can also be written as:

\[
C_j \sim Mult_g(1, \lambda_1, \ldots, \lambda_g)
\]
Suppose that the conditional density of $T_j$ given $C_{ij} = 1$ is $f_{i}(t_j)$ which is the time probability density function of maintenance task $i$. Then, the unconditional density of $T_j$ is given by $p(t_j)$ which has the $g$-component mixture form, that is:

$$p(t_j) = \lambda_i f_{i}(t_j) + \cdots + \lambda_g f_{g}(t_j)$$  \hfill (4)

The mean and variance of $T_j$ are[12]:

$$E[T_j] = \mu = \sum_{i=1}^{g} \lambda_i \mu_i$$  \hfill (5)

$$E[(T_j - \mu)^2] = d^2 = \sum_{i=1}^{g} \lambda_i ((\mu_i - \mu)^2 + d_i^2)$$  \hfill (6)

where $\mu_i$ and $d_i$ are the mean and variance of time distribution function of maintenance task $i$.

**MTTR Verification of Complex System**

**Assuming Log-Normal Distribution Instead of Mixture Distribution**

Assume that the design goal maintainability index $H_0$ for the specified item is $MTTR = \mu_0$ and required maintainability index $H_1$ is $MTTR = \mu_1$. The producer’s risk is $\alpha$ and consumer’s risk is $\beta$.

According to MIL-STD-471A, if no specific assumption concerns the distribution of maintenance time, the sample size is given by:

$$n = \left( \frac{Z_{\alpha} + Z_{\beta}}{\mu_1 - \mu_0} \right)^2 \hat{d}^2$$  \hfill (7)

where $\hat{d}^2$ is a prior estimate of the variance of maintenance time.

If the maintenance time can be adequately described by a log-normal distribution, the sample size is given by:

$$n = \left( \frac{Z_{\alpha} \mu_0 + Z_{\beta} \mu_1}{(\mu_1 - \mu_0)^2} \right)^2 (e^{\hat{\sigma}^2} - 1)$$  \hfill (8)

where $\hat{\sigma}^2$ is a prior estimate of the variance of the logarithms of maintenance time.

The decision criteria are:

Accept $H_0$, if $\bar{X} \leq \mu_0 + Z_{\alpha} \frac{\hat{d}}{\sqrt{n}}$; Reject $H_0$, if $\bar{X} > \mu_0 + Z_{\beta} \frac{\hat{d}}{\sqrt{n}}$  \hfill (9)

where $\bar{X}$ is the sample mean, and $\hat{d}$ is the sample standard deviation.

In the case of assuming log-normal distribution instead of mixture distribution, the producer’s risk is:

under $H_0$: $P[\bar{X} > \mu_0 + Z_{\alpha} \frac{\hat{d}}{\sqrt{n}}]$  \hfill (10)

Eq. (10) can be written as:
The consumer’s risk is:

$$P[\bar{X} > \mu_0 + Z_{1-\alpha} \frac{d}{\sqrt{n}}] = P\left[\frac{\bar{X} - \mu_0}{d/\sqrt{n}} > \frac{\hat{d}}{d/\sqrt{n}} - \mu_0 \right] = P\left[\frac{\bar{X} - \mu_0}{d/\sqrt{n}} > z_{1-\alpha} \frac{d}{\sqrt{n}} - \mu_0 \right] = 1 - \Phi(z_{1-\alpha}) = \alpha$$

under $H_1$: $P[\bar{X} \leq \mu_0 + Z_{1-\alpha} \frac{d}{\sqrt{n}}]$  \hspace{1cm} (11)

Eq. (11) can be written as:

$$P[\bar{X} \leq \mu_0 + Z_{1-\alpha} \frac{d}{\sqrt{n}}] = P\left[\frac{\bar{X} - \mu_0}{d/\sqrt{n}} \leq \frac{\mu_0 + Z_{1-\alpha} \frac{d}{d/\sqrt{n}} - \mu_1}{d/\sqrt{n}} \right]$$

$$= P\left[\frac{\bar{X} - \mu_0}{d/\sqrt{n}} \leq Z_{1-\alpha} + \frac{\sqrt{n}}{d} (\mu_0 - \mu_1)\right]$$

$$= \Phi[Z_{1-\alpha} + \frac{\sqrt{n}}{d} (\mu_0 - \mu_1)]$$

According to Eqs. (7) and (8), the sample sizes may be different between two distribution assumptions. Thus, the result indicates that, the producer’s risk will remain unchanged and consumer’s risk may have a change if log-normal distribution is assumed when in fact the maintenance time has a mixture distribution. Thus, during MTTR verification, it is important to check whether the maintenance time of specified item conforms to a log-normal distribution.

**Combining Maintenance time Samples from Similar Maintenance Tasks**

According to MIL-STD-471A, the test methods for MTTR verification all require a minimum or fixed sample size based on different distribution assumptions[1]. However, development budgets and schedules may not allow enough time to obtain the required number of maintenance tasks, especially for the equipment with high reliability.

For a complex system, there may be much similarities among some of its maintenance tasks, such as the maintenance tasks belonging to equipment which has almost same structure. These maintenance tasks may have time probability density functions with parameters varying a little. Thus, it might be a feasible method to combine their maintenance time samples for verification in order to meet the required sample size.

Suppose there are two maintenance tasks A and B, which show much similarities regarding their repair actions. They all conform to log-normal distribution with parameters varying a constant; that is:

$$\ln T_A \sim N(\theta_A, \sigma_A^2) \quad \ln T_B \sim N(\theta_A + g, \sigma_A^2 + h)$$

and

$$\mu_A = e^{\theta_A + \frac{1}{2} \sigma_A^2}, d_A^2 = e^{2\theta_A + \sigma_A^2}(e^{\sigma_A^2} - 1), \mu_B = e^{(\theta_A + g) + \frac{1}{2}(\sigma_A^2 + h)} e^{2(\theta_A + g)(\sigma_A^2 + h)}(e^{(\sigma_A^2 + h)} - 1)$$

where $g$ and $h$ are constants.

Thus, if the maintenance time sample sets of task A and B are combined, it will conform to a mixture distribution; that is,

$$f(t_{AB}) = \lambda_A f(t_A) + \lambda_B f(t_B)$$

where $f(t_A)$ and $f(t_B)$ are the maintenance time probability density functions of task A and B and
\[
\hat{\lambda}_A = \frac{w_A}{w_A + w_B} \quad \hat{\lambda}_B = \frac{w_B}{w_A + w_B}
\]

(15)

where \( w_A \) and \( w_B \) are the failure rates of corresponding failure modes of task A and B.

According to Eqs. (5) and (6), the mean \( \mu_{AB} \) and variance \( d_{AB}^2 \) are:

\[
\mu_{AB} = \hat{\lambda}_A \mu_A + \hat{\lambda}_B \mu_B \quad (16)
\]

\[
d_{AB}^2 = \hat{\lambda}_A ((\mu_A - \mu_{AB})^2 + d_A^2) + \hat{\lambda}_B ((\mu_B - \mu_{AB})^2 + d_B^2)
\]

(17)

Assume that the design goal maintainability index for the combined data set is \( H_0 : \mu_{AB} = \mu_0 \) and required maintainability index is \( H_1 : \mu_{AB} = \mu_1 \). The producer’s risk is \( \alpha \) and consumer’s risk is \( \beta \).

If \( H_0 \) is true, according to Eqs. (13) and (16),

\[
\hat{\lambda}_A \mu_{A(H_0)} + \hat{\lambda}_B \mu_{B(H_0)} = \mu_0 \Rightarrow \hat{\lambda}_A e^{\theta_{H_0} \frac{1}{2} \sigma_A^2} + \hat{\lambda}_B e^{\theta_{H_0} \frac{1}{2} \sigma_B^2} = \mu_0 
\]

\[
\Rightarrow \theta_{A(H_0)} = \ln \left( \frac{\mu_0}{\hat{\lambda}_A e^{\frac{1}{2} \sigma_A^2} + \hat{\lambda}_B e^{\frac{1}{2} \sigma_B^2}} \right)
\]

(18)

and

\[
\mu_{A(H_0)} = e^{\theta_{H_0} \frac{1}{2} \sigma_A^2}, d_{A(H_0)}^2 = \mu_{A(H_0)}^2 (e^{\sigma_A^2} - 1) \quad ; \quad \mu_{B(H_0)} = e^{\theta_{H_0} \frac{1}{2} \sigma_B^2}, d_{B(H_0)}^2 = \mu_{B(H_0)}^2 (e^{\sigma_B^2} - 1)
\]

(19)

and

\[
d_{AB(H_0)}^2 = \hat{\lambda}_A ((\mu_{A(H_0)} - \mu_0)^2 + d_{A(H_0)}^2) + \hat{\lambda}_B ((\mu_{B(H_0)} - \mu_0)^2 + d_{B(H_0)}^2)
\]

(20)

Thus, the distribution of \( \bar{X} \) is:

\[
\bar{X} \sim N(\mu_0, \frac{d_{AB(H_0)}^2}{n_{AB}})
\]

(21)

If \( H_1 \) is true, then

\[
\theta_{A(H_1)} = \ln \left( \frac{\mu_1}{\hat{\lambda}_A e^{\frac{1}{2} \sigma_A^2} + \hat{\lambda}_B e^{\frac{1}{2} \sigma_B^2}} \right)
\]

(22)

and

\[
\mu_{A(H_1)} = e^{\theta_{H_1} \frac{1}{2} \sigma_A^2}, d_{A(H_1)}^2 = \mu_{A(H_1)}^2 (e^{\sigma_A^2} - 1) \quad ; \quad \mu_{B(H_1)} = e^{\theta_{H_1} \frac{1}{2} \sigma_B^2}, d_{B(H_1)}^2 = \mu_{B(H_1)}^2 (e^{\sigma_B^2} - 1)
\]

(23)

and

\[
d_{AB(H_1)}^2 = \hat{\lambda}_A ((\mu_{A(H_1)} - \mu_1)^2 + d_{A(H_1)}^2) + \hat{\lambda}_B ((\mu_{B(H_1)} - \mu_1)^2 + d_{B(H_1)}^2)
\]

(24)

Thus, the distribution of \( \bar{X} \) is:
\[\bar{X} \sim N\left(\mu, \frac{d^2_{AB(H_0)}}{n_{AB}}\right)\]  

(25)

Then, under the asymptotic normality of \(\bar{X}_n\), \(n_{AB}\) can be defined as:

\[n_{AB} = \left(\frac{Z_{1-\alpha}d_{AB(H_0)}^2 + Z_{1-\beta}d_{AB(H_1)}^2}{\mu_1 - \mu_0}\right)^2\]  

(26)

The sample size ratio of task A to B is same as the failure rate ratio; that is,

\[n_A = \lambda_A n_{AB} ; n_B = \lambda_B n_{AB}\]  

(27)

The decision criteria are:

Accept \(H_0\), if \(\bar{X} \leq \mu_0 + Z_{1-\alpha} \frac{\hat{d}}{\sqrt{n_{AB}}}\); Reject \(H_0\), if \(\bar{X} > \mu_0 + Z_{1-\alpha} \frac{\hat{d}}{\sqrt{n_{AB}}}\)  

(28)

If the test result of the combined data set is considered as the one of maintenance task A, the samples of task B may bring changes to the real test risks.

Assume that the real design goal maintainability index is \(H_0: \mu_A = \mu_0\) and required maintainability index is \(H_1: \mu_A = \mu_1\). The producer’s risk is \(\alpha\) and consumer’s risk is \(\beta\).

If \(H_0\) is true, then

\[\theta_{A(H_0)} = \ln \mu_0 - \frac{1}{2} \sigma_A^2\]  

(29)

and

\[\mu_{B(H_0)} = e^{(\theta_{A(H_0)} + \frac{1}{2} \sigma_A^2)}\sigma_A^2_{AB(H_0)} = \mu_0^2(e^{\sigma_A^2} - 1); \quad \mu_{AB(H_0)} = \lambda_A \mu_0 + \lambda_B \mu_{B(H_0)}; \quad d_{AB(H_0)}^2 = \mu_{AB(H_0)}^2(e^{\sigma_A^2} - 1)\]  

(30)

and

\[d_{AB(H_0)}^2 = \lambda_A ((\mu_0 - \mu_{AB(H_0)})^2 + d_{AB(H_0)}^2) + \lambda_B ((\mu_{B(H_0)} - \mu_{AB(H_0)})^2 + d_{B(H_0)}^2)\]  

(31)

the distribution of \(\bar{X}\) is:

\[\bar{X} \sim N(\mu_{AB(H_0)}, \frac{d_{AB(H_0)}^2}{n_{AB}})\]  

(32)

Thus, the real producer’s risk is:

\[\alpha = P[\bar{X} > \mu_0 + Z_{1-\alpha} \frac{\hat{d}}{\sqrt{n_{AB}}}]\]  

(33)

If \(H_1\) is true, then

\[\theta_{A(H_1)} = \ln \mu_1 - \frac{1}{2} \sigma_A^2\]  

(34)

and

\[\mu_{B(H_1)} = e^{(\theta_{A(H_1)} + \frac{1}{2} \sigma_A^2)}\sigma_A^2_{AB(H_1)} = \mu_1^2(e^{\sigma_A^2} - 1); \quad \mu_{AB(H_1)} = \lambda_A \mu_1 + \lambda_B \mu_{B(H_1)}; \quad d_{AB(H_1)}^2 = \mu_{AB(H_1)}^2(e^{\sigma_A^2} - 1)\]  

(35)

and
\[ d^2_{AB(H_1)} = \lambda_A((\mu_1 - \mu_{AB(H_1)})^2 + d^2_{A(H_1)}) + \lambda_B((\mu_{B(H_1)} - \mu_{AB(H_1)})^2 + d^2_{B(H_1)}) \]  

(36)

the distribution of \( \bar{X} \) is:

\[ \bar{X} \sim N(\mu_{AB(H_1)}, \frac{d^2_{AB(H_1)}}{n_{AB}}) \]  

(37)

Thus, the real consumer’s risk is:

\[ \beta = P(\bar{X} \leq \mu_0 + Z_{1-\alpha} \frac{d}{\sqrt{n_{AB}}}) \]  

(38)

Case Study

In this section, two sets of maintenance time samples from *Defense Communication System/Satellite Control Facility Interface System (DSIS) Maintainability Demonstration Report* are combined for MTTR verification. The equipment has much redundancy which allows on-line maintenance by “reconfiguration” the system by patching[13]. Table 1 shows the probability density functions of the two data sets[2].

<table>
<thead>
<tr>
<th></th>
<th>( \theta )</th>
<th>( \sigma^2 )</th>
<th>( \lambda )</th>
<th>( \ln T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. On-line maintenance</td>
<td>2.45</td>
<td>0.68</td>
<td>0.5</td>
<td>( N(2.45,0.68) )</td>
</tr>
<tr>
<td>B. Off-line maintenance</td>
<td>2.72</td>
<td>0.63</td>
<td>0.5</td>
<td>( N(2.72,0.63) )</td>
</tr>
</tbody>
</table>

Assume that the design goal maintainability index \( H_0 \) for the combined data set is \( \mu_{AB} = 15 \text{min} \) and required maintainability index \( H_1 \) is \( \mu_{AB} = 20 \text{min} \) and the producer’s and consumer’s risks are \( \alpha = \beta = 0.05 \).

According to Eq.(26), the required sample size is:

\[ n_{AB} = \left( \frac{1.65 \times 14.56 + 1.65 \times 19.42}{20 - 15} \right)^2 = 125.7 \approx 126 \]

and the required sample size for maintenance task A and B are: \( n_A = n_B = 0.5 \times 126 = 63 \)

In fact, the real specified item is the on-line maintenance task. Thus, the design goal maintainability index \( H_0 \) is \( \mu_A = 15 \text{min} \) and required maintainability index \( H_1 \) is \( \mu_A = 20 \text{min} \) and the producer’s and consumer’s risks are \( \alpha = \beta = 0.05 \).

If \( H_0 \) is true, according to Eqs. (29) ~ (32), the distribution of sample mean is:

\[ \bar{X} \sim N(17.08,2.18) \]

Thus, according to Eq. (33), the real producer’s risk is:

\[ \alpha = P(\bar{X} > 15 + 1.65 \times \frac{212.09}{\sqrt{126}}) = 0.48 \]

If \( H_1 \) is true, according to Eqs. (34) ~ (37), the distribution of sample mean is:
Thus, according to Eq. (38), the real consumer’s risk is:

$$\beta = P(\bar{X} \leq 15 + 1.65 \times \frac{212.09}{\sqrt{126}}) = 0.002$$

Figure 1. The relationship between $g$ and risk values ($h = -0.05$).

Figure 2. The relationship between $h$ and risk values ($g = 0.27$).

The above results indicate that if the two maintenance time sample sets are combined for MTTR verification of on-line maintenance task, the real producer’s risk increases the number of about 0.43 while consumer’s risk decreases the number of about 0.048. It means that it is difficult to accept the equipment if it doesn’t meet the requirement but easy to reject it for not meeting the requirement when, in fact, it has. Thus, during MTTR verification, if the equipment is rejected for the hypothesis test, further investigation need to be made to confirm whether it is acceptable. Fig. 1 and Fig. 2 show the risk values with different values of $g$ and $h$.

If the same hypothesis test is performed on the on-line maintenance task independently, the sample size is:
n_A = \left( \frac{Z_0 \mu_0 + Z_1 \mu_1}{(\mu_1 - \mu_0)} \right)^2 \left( e^{\sigma^2} - 1 \right) = \frac{(1.65 \times 15 + 1.65 \times 20)^2}{(20 - 15)^2} \left( e^{0.68} - 1 \right) = 129.9 \approx 130

Thus, nearly half the number of sample size can be saved if on-line and off-line maintenance sample sets are combined for maintainability verification.

Conclusions

Maintenance time of complex system sometimes conforms to mixture distribution due to its various maintenance tasks. During maintainability demonstration, as it is difficult to perform goodness of fit analysis with the limited sample size, there will be error in test risks caused in assuming log-normal distribution when in fact it has mixture distribution. On the other hand, for the maintenance tasks of complex system, there may be much similarities between some of them. Thus, the samples of similar maintenance tasks can be used as a supplement to the specified task’s samples to implement maintainability verification.

Regarding the issues mentioned above, firstly, the maintenance sample generation process of complex system was analyzed and the mixture distribution model of maintenance time was proposed. Then, this paper analyzed the changes in two types of test risks when assuming log-normal distribution instead of mixture distribution. Finally, a method to perform maintainability verification using maintenance time data from similar maintenance task was proposed. A case study using the maintenance time data of satellite communication interface system was presented as an example of the proposed method.

References

[3] E. Camozu, A Study of the application of the lognormal and gamma distribution to corrective maintenance repair time data.