Advances in Reliability, Safety and Security, Part 3 – Kołowrocki, Dąbrowska (eds.)
 ISBN 978-83-68136-15-9 (printed), ISBN 978-83-68136-02-9 (electronic)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Reliability Demonstration Test Design Method With Missing Data

Dian Zhang[a], Ping Jiang[a], Yunyan Xing[a], Jianjun Qi[b]

[a]*National University of Defense Technology,Changsha, China*
[b]*Beijing special engineering design institution,Beijing, China*

Abstract

Reliability demonstration test (RDT) is a test to verify whether the reliability of the product meets the reliability requirements during the product development, and it is usually required to be carried out before product delivery. During the product development process, data collection mechanism is sometimes not strictly imposed, which often results in test data missing. And the data missing usually has great impact on product reliability assessment and RDT design as well. In this paper, we study the RDT design method of products in cases of two types of missing data, namely left censored data and interval censored data, which are common in engineering practice. Firstly, the AMSAA model is used to model the product reliability growth process; and then the MCEM algorithm is used to estimate the unknown parameters of the model to obtain the product life distribution; Finally, we derive the timed censoring test plan for the product based on the producer's risk and consumer's risk of the test. Compared with the traditional RDT plan, the derived RDT plan has the advantages of shorter test time and lower test risks.

Keywords: reliability demonstration test RDT, missing data, type-I censored test, MCEM

1. Introduction

A reliability demonstration test (RDT) is a test to verify that the reliability of a product meets the reliability requirements of the product development, and the results of the test can provide a basis for identifying the reliability of the product. Reliability plays an essential role in ensuring the performance of high reliability products, such as satellites, cars and drones. Reliability requirements for these complex products are becoming more stringent, making it necessary to perform RDTS during development. In engineering practice, test data is often missing due to negligence of test personnel, equipment failure, etc. when the test data of a product is missing, errors will occur when using the incomplete data to evaluate the reliability of the product, and the RDT plan developed is also inaccurate. Therefore, it is necessary to study the RDT design of products in the cases of missing data.

Traditional RDT design test plans are based on failure data in development testing. In practices, they are from the standards, such as China's GJB-899A and the United States' MIL-STD-781, which provide detailed RDT plans for exponentially distributed products. However, for many high reliable products, due to the advancement of production technology, and testing according to the test plans in the standard usually requires longer test time. Meanwhile, during the product development, lots of test information, such as reliability test data during the development stage, can provide references for the RDT design of the product. Therefore, many scholars have studied how to use product test information to facilitate designing the RDT of product. Martz (Martz and Waller, 1979) proposed the use of Bayesian prior distribution to estimate the life distribution of the product, and then set the reliability demonstration test based on the risk of the user, the use of a Bayesian prior distribution would be more consistent with the actual lifetime of the product than the classical lifetime distribution estimate, which was also confirmed in the subsequent study. The prior distribution is determined by the percentile method in this paper, which is easy to implement but does not fully exploit the product test data, resulting in a less plausible choice of prior distribution. Based on the multi-level Bayesian method and using reliability test information from the product development process, Zhang (Zhang and Tian, 2005) proposed a method for formulating an exponential product reliability acceptance test plan, which has the advantage of shorter test time compared with the traditional test plan. Yang (2009) introduced a zero-failure sampling identification method that is widely used in the engineering field, called Bogey Testing, it is to take certain samples from a batch of products for testing. if all samples do not fail, the batch of products has passed the identification, but this type of trial typically requires larger sample sizes. Wilson (Wilson and Malcolm, 2019) proposed to incorporate the historical data of the product into the prior distribution of the product, and studied the RDT scheme of binary distribution products and Weibull distribution products. Feng (Feng and Liu, 2012) determines the RDT plan based on test losses under the constraints of the average risk criterion under the condition of fixed-number censoring tests for exponential products. Xu (Xu et al., 2017) proposes to use information from the product development stage to design RDT solutions based on the power law process. This method can significantly reduce test costs. Kleyner (Kleyner et al., 2015) proposed applying the Bayes method to design RDT when product design changes, using data collected from trials of earlier versions of the product before the design change, avoiding repeated trials of the redesigned product, which can reduce the required sample size. however, its method only considers the use of success-failure test data, and the design change coefficient is determined arbitrarily, and the two types of risks of the test are not considered. Jeon (Jeon and Ahn, 2018) proposed a sequential sampling Bayes method, which shows the advantage of small samples in RDT, but is only suitable for identification of successful and failed products. Yuan (Yuan and Li, 2014) proposed the Bayes method for reliability identification in the multi-stage development process of products, combining the test data from the previous stage with the test data from the current stage to formulate a RDT plan for the product. Li (Li et al., 2017) and Zhang (Zhang et al., 2023) designed the RDT method of the product based on the reliability growth process of successful and failed products and index products respectively, and based on the test information in the development stage. Jiang (Jiang et al., 2022) proposed a method to design a system RDT plan based on product subsystem test data. Compared with the traditional RDT plan, it has the advantage of shorter test time.

Regarding the situation of missing data, scholars have conducted more detailed research on product reliability assessment when missing data occurs. Based on the Bayes principle, Wang (Wang and Wang, 2001) studied the product life estimation method when data is missing in the timed censoring test. Tian (Tian and Liu, 2015) studied the approximate Bayes estimation of the product life of Weibull products in the case of missing data in the test. In addition, many scholars have studied missing data processing methods from the perspective of filling in missing data. Dempster (Dempster et al., 1977) proposed the expectation maximization algorithm, which was widely used in processing missing data. Many improved methods are also proposed, such as Stochastic Expectation Maximization algorithm (Celeux and Diebolt, 1985), Monte Carlo Expectation Maximization algorithm (Wei and Martin, 1990), etc.

The missing data will affect the reliability evaluation of the product and the RDT design of the product. In this paper, we study how to design product RDT based on missing data in cases of two types of missing data: left censored data and interval censored data. First, the type of missing data is clarified; Secondly, since the missing data mainly affects the parameter estimation of the reliability growth model in the development, this article uses the MCEM algorithm to estimate the parameters of the reliability growth model; Then the Bayes method was used again to determine the life distribution of the product, and a time censored test plan for the product was designed based on two types of risks; Finally, the feasibility of the RDT design method in the case of missing data was verified based on actual cases.

2. Problem formulation

2.1. Type of data

In product development phase testing, the situation of missing data often occurred when evaluating product reliability. According to the different periods when missing data occurs, missing data can be divided into three types: left censored data, interval censored data and right censored data, and different types of missing data have different handling methods. In this paper, we will introduce the two main types of missing data that exist in engineering practice, namely left censored data and interval censored data. To facilitate subsequent calculations, we use Y to represent the complete fault data, X represents the actual observed fault data set and Z represents the missing data set.

The missing data types diagram is shown in Figure 1, where the symbol "X" on the coordinate axis indicates missing data.

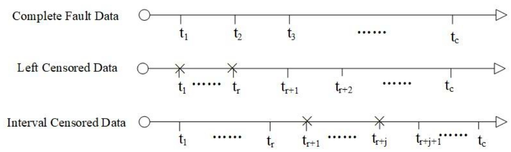

Fig. 1 The missing data types diagram.

A: The left censored data

The left censored data refers to the type of missing data in which early test data is missing during the products development stage. We use t_i to represent the moment when the *i*-th fault occurs, and assume that a total of *c* faults during the development stage, the complete fault data can be expressed as

$$Y = \{t_1, t_{2,\ldots,} t_c\}\ . \tag{1}$$

If the number of missing data is *r*, $(0 < r < c)$, the observed data can be expressed as

$$X = \{t_{r+1}, t_{r+2,\ldots,} t_c\}\ . \tag{2}$$

The left censored data can be expressed as

$$Z = \{t_1, t_{2,\ldots,} t_r\}\ . \tag{3}$$

B: The interval censored data

The interval censored data refers to the type of missing data in which test data at one or more stages in the middle of the product development phase is missing. If the number of missing data is $j, (0 < j < c)$, the observed data can be expressed as

$$X = \{t_1, \ldots, t_r, t_{r+j+1}, \ldots, t_c\}\ . \tag{4}$$

The Interval Censored Data can be expressed as

$$Z = (t_{r+1}, t_{r+2,\ldots,} t_{r+j})\ . \tag{5}$$

2.2. Objective

The aim of this paper is to derive an RDT design method for products whose lifetimes follow an exponential distribution with a failure rate λ . And reliability function as

$$\begin{cases} f(t) = \lambda e^{-\lambda t} \\ R(t) = e^{-\lambda t} \end{cases} \tag{6}$$

2.3. Test plan

The reliability demonstration test of a product is used to verify whether its reliability meets the development requirements. Since type-I censored (time censored) test is mostly applied in practices, this paper concentrates on deriving the test plans (T, C) for type-I censored test.

2.4. Hypothesis test

In an RDT, the decision to accept or reject certain batch of products is made by executing the following hypothesis test, restricted to predetermined producer's risk and consumer's risk.

$$H_0 : \lambda < \lambda_0,\ \ H_1 : \lambda > \lambda_1 \tag{7}$$

For exponentially distributed products, λ_0 is the failure rate specified in the development requirements. And when the discrimination ratio of the test is $d = \lambda_1 / \lambda_0$, we can know the lower limit of product failure rate λ_1 .

3. Life distribution determination

For products with missing test data in the development stage, our idea of obtaining the life distribution of the product is as follows. Firstly, we use the AMSAA model to model the product development stage to obtain product reliability estimates; Secondly, based on the principle of maximum entropy, the Bayesian method is used to obtain the prior distribution of product failure rate; Finally, we use the conjugate prior distribution to infer the posterior distribution of the failure rate.

3.1. AMSAA model

Since the defects exposed in the test will be improved during the product development process, the product development process is also a reliability growth process. We use the AMSAA model to model the reliability growth process in the development phase because the exponential distribution is a continuous lifetime distribution. The fault intensity function of the AMAA model is

$$\lambda(t)=\frac{\gamma}{\varepsilon}(\frac{t}{\varepsilon})^{\gamma-1} \tag{8}$$

Where ε is the scale parameter and γ is the shape parameter, and $\varepsilon>0,\gamma>0$.

We assume that the tests in the product development stage are all type-I censored tests. The time when the first fault occurs is t_i and the censoring time of the test is t_c.If a total of n faults occur within the censoring time, the likelihood function of the test data is

$$\begin{aligned} L &= f(t_1,t_2,\ldots,t_n,t_c) \\ &= f_1(t_1)\times\prod_{i=2}^{n} f(t_i|t_1,\ldots,t_{i-1})\times P\{N(t_n,t_c)=0\} \\ &= (\frac{\gamma}{\varepsilon})^n\times\prod_{i=1}^{n}(\frac{t_i}{\varepsilon})^{\gamma-1}\times\exp[-(\frac{t_n}{\varepsilon})^{\beta}]\times\exp[(\frac{t_n}{\varepsilon})^{\gamma}-(\frac{t_c}{\varepsilon})^{\gamma}] \\ &= (\frac{\gamma}{\varepsilon})^n\exp[-(\frac{t_c}{\varepsilon})^{\beta}]\prod_{i=1}^{n}(\frac{t_i}{\varepsilon})^{\gamma-1} \end{aligned} \tag{9}$$

The joint probability density of $t_1,t_2,\ldots,t_c$ is

$$\begin{aligned} f(t_1,t_2,\ldots,t_c) &= [\prod_{i=1}^{n}\lambda(t_i)]\exp[-m(t)] \\ &= [\prod_{i=1}^{n}\lambda(t_i)]\exp[-\int_0^{t_c}\lambda(t)dt] \\ &= \frac{\gamma^n}{\varepsilon^{n\gamma}}\exp[-(\frac{t_c}{\varepsilon})^{\gamma}]\prod_{i=1}^{n}t_i^{\gamma-1} \end{aligned} \tag{10}$$

3.2. Parameter estimation

Since the lack of experimental data will affect the parameter estimation of the AMSAA model, we use the Monte Carlo Expectation Maximization (MCEM) algorithm, which is currently the most widely used to deal with missing data, to solve the parameter estimation of the model.

The MCEM algorithm is an effective algorithm that uses the Monte Carlo method to improve the original EM algorithm to solve parameter estimation in the case of missing data. The basic principle of the EM algorithm is to use an iterative method to find the maximum likelihood estimate of the unknown parameters in the model. The algorithm consists of two steps: E step (Expectation step) is to find the conditional expectation of the log-likelihood function of the complete data under the current parameter estimation; The M step (Maximization Step) is to updates parameter estimates by maximizing expectations. These two steps are repeated in an iterative process until the algorithm converges and maximum likelihood estimates of the unknown parameters are obtained. The MCEM algorithm uses the Monte Carlo method to randomly sample from the conditional distribution function in the E step of the EM algorithm, and then uses the sample mean calculated from the sampled samples to replace the conditional expectation value of the function, which can solve the difficult problem of conditional expectation calculation in EM algorithm.

A: The left censored data

For the Left Censored Data, we can know the log-likelihood function of the complete data Y is

$$\ln L(Y;\varepsilon,\gamma)=n\ln\varepsilon-n\ln\gamma+(\gamma-1)\sum_{i=1}^{r}\ln t_i+(\gamma-1)\sum_{j=r+1}^{n}\ln t_j-(\frac{t_c}{\varepsilon})^{\gamma} \tag{11}$$

Then we use the MCEM algorithm to calculate the conditional expectation of the complete data. E Step: Calculate the conditional expectation of complete data Y under observed data X:

$$E\{\ln L(Y;\varepsilon,\gamma)\big|X\} = n\ln\gamma - n\gamma\ln\varepsilon + (\gamma-1)\sum_{i=1}^{r}\ln t_i + (\gamma-1)E\{\sum_{j=r+1}^{n}\ln t_j\big|X\} - (\frac{t_c}{\varepsilon})^{\gamma} \quad (12)$$

M Step: Update parameter estimates using maximizing conditional expectation. Let the estimated value of the k-th step parameter be (ε^k,γ^k), $(\varepsilon^{k+1},\gamma^{k+1})$ is obtained by maximizing (12). Find the partial derivatives of ε and γ respectively for (12), and set the partial derivative equal to 0 to get the updated value of the parameter:

$$\begin{cases} \varepsilon^{k+1} = \dfrac{t_c}{n^{\frac{1}{\gamma^{k+1}}}} \\ \gamma^{k+1} = \dfrac{n}{n\ln t_c - \sum_{i=r+1}^{n}\ln t_i - E\{(\sum_{j=1}^{r}\ln t_j)\big|X,\varepsilon^k,\gamma^k\}} \end{cases} \quad (13)$$

The missing sample data sampled using the Monte Carlo method is $(t_1',t_2',\ldots,t_r')^j$, and s is the number of sampling samples, and the conditional expectation of replacing it with the sampling sample mean is

$$E\{(\sum_{j=1}^{r}\ln t_j)\big|X,\varepsilon^k,\gamma^k\} = \frac{1}{s}\sum_{j=1}^{s}\sum_{i=1}^{r}\ln t_i' \quad (14)$$

B: The interval censored data

The log-likelihood function of the complete data Y is

$$\ln L(Y;\varepsilon,\gamma) = n\ln\gamma - n\gamma\ln\varepsilon + (\gamma-1)\sum_{i=r+1}^{r+j}\ln t_i + (\gamma-1)(\sum_{i=1}^{r}\ln t_i + \sum_{k=i+j+1}^{n}\ln t_k) - (\frac{t_c}{\varepsilon})^{\gamma} \quad (15)$$

E Step: Calculate the conditional expectation of complete data Y under observed data X:

$$E\{L(Y;\varepsilon,\gamma)\big|X\} = n\ln\gamma - n\gamma\ln\varepsilon + (\gamma-1)\sum_{i=r+1}^{r+j}\ln t_i + (\gamma-1)E\{(\sum_{i=1}^{r}\ln t_i + \sum_{k=i+j+1}^{n}\ln t_k)\big|X\} - (\frac{t_c}{\varepsilon})^{\gamma} \quad (16)$$

M Step: Update parameter estimates using maximizing conditional expectation.

$$\begin{cases} \varepsilon^{k+1} = \dfrac{t_c}{n^{\frac{1}{\gamma^{k+1}}}} \\ \gamma^{k+1} = \dfrac{n}{n\ln t_c - \sum_{i=r+1}^{n}\ln t_i - E\{(\sum_{i=r+1}^{r+j}\ln t_i)\big|X,\varepsilon^k,\gamma^k\}} \end{cases} \quad (17)$$

The conditional expectation of replacing it with the sampling sample mean is

$$E\{(\sum_{i=r+1}^{r+j}\ln t_i)\big|X,\varepsilon^k,\gamma^k\} = \frac{1}{s}\sum_{i=1}^{j}\sum_{k=r+1}^{r+j}\ln t_i' \quad (18)$$

C: Parameter estimation algorithm

The steps for parameter estimation using the MCEM algorithm are as follows:

Table 1. MCEM algorithm for parameter estimation.

Step	MCEM algorithm for parameter estimation
1	Set the initial value (ε^0,γ^0) of the unknown parameter and set the algorithm convergence conditions;
2	sampling to generate s groups of missing data $(t_1',t_2',\ldots,t_r')^i, i=1,2,\ldots,s$;
3	Calculate conditional expectations;
4	Update parameter values $(\varepsilon^{k+1},\gamma^{k+1})$;
5	Determine whether the algorithm converges, and if it converges, output parameter estimates $(\varepsilon',\gamma')=(\varepsilon^{k+1},\gamma^{k+1})$; otherwise, return the updated values to the initial values, repeat the algorithm until convergence, and output parameter estimates.

After obtaining the estimated value (ε',γ') of the unknown parameter of the AMSAA model, we can obtain the fault intensity function of the model as

$$\lambda'(t)=\frac{\gamma'}{\varepsilon'}(\frac{t}{\varepsilon'})^{\gamma'-1} \tag{19}$$

The cumulative number of product failures is

$$P\{N(T)=n\}=\frac{(\varepsilon^{-\gamma}t^{\gamma})^{n}}{n!}e^{-(\varepsilon^{-\gamma}t^{\gamma})} \tag{20}$$

If a timely correction strategy is adopted during the product development stage, a total of m tests are conducted, and m faults occur, then for a product that has m faults, the probability of the m-th and m-1 faults of the product can be calculated the times are:

$$\begin{cases} t_m=\left(\dfrac{m}{\varepsilon^{-\gamma}}\right)^{\frac{1}{\gamma}} \\ t_{m-1}=\left(\dfrac{m-1}{\varepsilon^{-\gamma}}\right)^{\frac{1}{\gamma}} \end{cases} \tag{21}$$

The estimated failure rate of the product after m times of testing is

$$\hat{\lambda}_m=\frac{1}{t_m-t_{m-1}}=\frac{\varepsilon^{-1}}{m^{1/\gamma}-(m-1)^{1/\gamma}} \tag{22}$$

3.3. Prior distribution

For products whose lifetimes follow an exponential distribution, we assume that the prior distribution of products failure rates follows a gamma distribution $Ga(\alpha,\beta)$, and

$$\begin{cases} \pi(\lambda)=\dfrac{\mu^{\theta}\lambda^{\theta-1}}{\Gamma(\theta)}e^{-\mu\lambda} \\ \Gamma(\theta)=\int_0^{\infty}\lambda^{\theta-1}e^{-\lambda}d\lambda \end{cases} \tag{23}$$

where μ is the shape parameters, θ is the scale parameters. We use the moment estimation method to find the prior distribution of the products.

$$\begin{cases} E(\lambda)=\int_0^{\infty}\lambda\dfrac{\mu^{\theta}\lambda^{\theta-1}}{\Gamma(\theta)}e^{-\mu\lambda}d\lambda=\dfrac{\theta}{\mu} \\ E(\lambda^2)=\int_0^{\infty}\lambda^2\dfrac{\mu^{\theta}\lambda^{\theta-1}}{\Gamma(\theta)}e^{-\mu\lambda}d\lambda=\dfrac{\theta(\theta+1)}{\mu^2} \end{cases} \tag{24}$$

We take the estimate of the failure rate obtained from the converted data as the mean value of the prior distribution, then

$$E(\lambda_{\mathrm{m}})=\hat{\lambda}_m \tag{25}$$

According to the principle of maximum entropy, the prior distribution of the failure rate can be expressed as

$$\pi(\lambda_m)=\frac{e^{u\lambda_m}}{\int_0^{\infty}e^{u\lambda_m}d\lambda_m} \tag{26}$$

where u is the coefficient to be determined and is a constant, according to (25) and (26), we can get

$$E(\lambda_m)=\int_{-\infty}^{+\infty}\lambda_m\pi(\lambda_m)d\lambda_m=\frac{\int_0^{\infty}\lambda_m e^{u\lambda_m}d\lambda_m}{\int_0^{\infty}e^{u\lambda_m}d\lambda_m}=\hat{\lambda}_m \tag{27}$$

From this, the value of the constant u can be solved, which in turn yields, from which the second order moment of the failure rate prior distribution can be obtained as

$$E({\lambda_m}^2) = \frac{\int_0^\infty {\lambda_m}^2 e^{u\lambda_m} d\lambda_m}{\int_0^\infty e^{u\lambda_m} d\lambda_m} \tag{28}$$

Since we choose the gamma distribution as the a priori distribution of the failure rate, then according to the above equation we can get

$$\begin{cases} E(\lambda_m) = \dfrac{\theta}{\mu} \\ E({\lambda_m}^2) = \dfrac{\theta(\theta+1)}{\mu^2} \end{cases} \tag{29}$$

and

$$\begin{cases} \hat{\theta} = \dfrac{[E(\lambda_m)]^2}{E(\lambda_m^2) - [E(\lambda_m)]^2} \\ \hat{\mu} = \dfrac{E(\lambda_m)}{E(\lambda_m^2) - [E(\lambda_m)]^2} \end{cases} \tag{30}$$

According to the above formula, the prior distribution of product failure rate λ_m after m tests can be obtained as $Ga(\hat{\theta}, \hat{\mu})$.

3.4. Posterior distribution

In the Bayesian method, the conjugate prior distribution is the most common prior distribution, and we also use the conjugate prior distribution as the prior distribution of the failure rate. Since the conjugate prior distribution and posterior distribution have corresponding forms, it is easy to calculate the posterior distribution, so we use the conjugate prior to infer the posterior distribution of the product. Assuming that the time when the product fails in the next test is t', the likelihood function of the failure rate is

$$L(t') = \lambda e^{-\lambda t'} \tag{31}$$

The posterior distribution of the product failure rate is

$$\pi(\lambda) = Ga(\theta', \mu'), \tag{32}$$

where $\theta' = \theta + 1, \mu' = \mu + t'$.

4. RDT plans design

For N samples in the test, we assume that event K means that all samples pass the RDT. In the type-I censored test, the event K equals that each of the N samples has the estimated failure time t_{ri} greater than the test censored time. When the test samples are from the same batch of products, we consider them to have the same life distribution, then the estimated failure time of each product is also the same. namely $t_{r1} = t_{r2} = ... = t_{rN} = t_r$.

$$P(K) = C_N^r P^r(t_{ri} < T) P^{N-r}(t_{ri} \ge T) = C_N^r (1 - e^{-\lambda T})^r (e^{-\lambda T})^{N-r} \tag{33}$$

In the RDT, the producer's risk is expressed as the probability that the product's failure rate meets the development requirements but the products are rejected by the test, which is expressed as follows

$$\begin{aligned} \alpha &= 1 - P(r \le C | \lambda \le \lambda_0) \\ &= 1 - \frac{\sum_{r=0}^{C} P(K, \lambda \le \lambda_0)}{P(\lambda \le \lambda_0)} \\ &= 1 - \frac{\sum_{r=0}^{C} \int_0^{\lambda_0} [C_N^r (1 - e^{-\lambda T})^r (e^{-\lambda T})^{N-r}] \pi(\lambda) d\lambda}{\int_0^{\lambda_0} \pi(\lambda) d\lambda} \end{aligned} \tag{34}$$

The consumer's risk of RDT is expressed as the risk of the product failure rate is higher than the development requirements but the products are accepted by the test, which is expressed as follows

$$\beta = P(r \le C | \lambda \ge \lambda_1) = \frac{\sum_{r=0}^{C} P(K, \lambda \ge \lambda_1)}{P(\lambda \ge \lambda_1)} = \frac{\sum_{r=0}^{C} \int_{\lambda_1}^{\infty} [C_N^r (1-e^{-\lambda T})^r (e^{-\lambda T})^{N-r}] \pi(\lambda) d\lambda}{\int_{\lambda_1}^{\infty} \pi(\lambda) d\lambda} \tag{35}$$

For the type-I censored test, when given the maximum acceptable producer's risk and consumer's risk, we can get

$$\begin{cases} 1 - \dfrac{\sum_{r=0}^{C} \int_{0}^{\lambda_0} [C_N^r (1-e^{-\lambda T})^r (e^{-\lambda T})^{N-r}] \pi(\lambda) d\lambda}{\int_{0}^{\lambda_0} \pi(\lambda) d\lambda} \le \alpha_0 \\ \dfrac{\sum_{r=0}^{C} \int_{\lambda_1}^{\infty} [C_N^r (1-e^{-\lambda T})^r (e^{-\lambda T})^{N-r}] \pi(\lambda) d\lambda}{\int_{\lambda_1}^{\infty} \pi(\lambda) d\lambda} \le \beta_0 \end{cases} \tag{36}$$

When the maximum number of failures C allowed by the test and the number of test samples N is given, we can obtain the censored time T of the test based on the above formula, and thus obtain the test plan (T, C).

5. Case study

We use the test data from the development stage of a certain type of aircraft generator for case study, it is assumed that the first three test data are missing, and the specific test data are shown in the Table 2.

Table 2. Case data of a certain type of aircraft generator.

Failure times	1	2	3	4	5	6	7	8	9	10	11	12	13
Failure time	55*	166*	205*	341	488	567	731	1308	2050	2453	3115	4017	4596

This generator was tested 13 times, with an "*" indicating missing data. The data type of this experiment is left censored data, and the first three data are missing.

According to the content of Section 3, we can use the MCEM algorithm to calculate the point estimate of the unknown parameters of the AMSAA model is $\varepsilon = 0.5227$, $\gamma = 0.1420$. We can obtain the estimated failure rate of the generator after 13 tests is $\lambda = 0.0012$, the prior distribution can be obtained as $Ga(\theta, \mu) = (1, 833.333)$, the posterior distribution of the product failure rate is $\pi(\lambda) = (2, 1412.333)$.

According to the content of Section 4, we can get that when the number of test samples $N = 1$ is and the allowable number of failures $C = 0$, if the development requirement for product reliability is that the failure rate is not higher than $\lambda^* = 0.002$, then when the test identification ratio is $d = 2$, the changes in the producer's risk and the consumer's risk over time are shown in Figure 1. If the censoring time when the risks of the producer and the consumer are equal is selected as the censored time of the RDT, the censored time of the timed censoring test of the product is $T = 297$, and the test plan is $(T, C) = (297, 0)$. This means that if the product does not fail within the censored time, the reliability of the product is considered to meet the development requirements and has passed the RDT; if not, it means that the reliability of the product has not met the development requirements and needs to be continuously improved.

In order to test the advancement of the RDT plan obtained in this paper, we compared it with the timed censored test plan in China's current RDT standard GJB899A (Table 3). The type-I censored test plan we choose is numbered 20 in GJB899A, and the sample size of the test is $N = 1$. According to the results in the table, it can be seen that when the test development requirements and the test identification ratio are the same, the test plan obtained in this paper has the advantages of lower test risk and shorter test time.

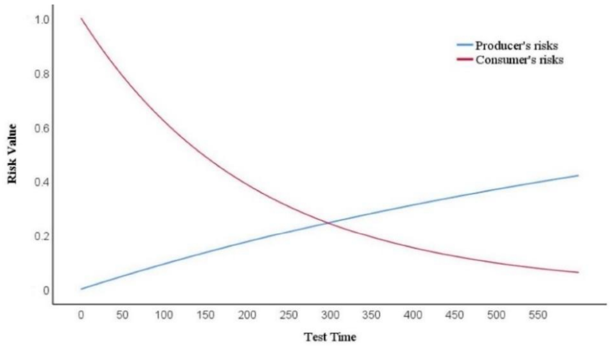

Fig. 2. Two types of risk change over test time.

Table 3. Comparison of RDT plan.

	Test plan in this paper	Test plan in GJB899A
Research requirements	$\lambda_0 = 0.002$	$\lambda_0 = 0.002$
Identification ratio	$d = 2$	$d = 2$
Judgment Criteria	Number of failures $r < 1$	Number of failures $r < 1$
Producer's risks and Consumer's risks	$\alpha = 0.24$ $\beta = 0.24$	$\alpha = 0.3$ $\beta = 0.3$
Test censored time	$T = 297$	$T = 1433$
Test plan	$(T, C) = (297, 0)$	$(T, C) = (1433, 0)$

According to section 4, the RDT plan of the product under different test risk selections when the identification ratio is $d = 2$ and the sample size is $N = 1$ can be obtained (Table 4).

Table 4. RDT plans at different test risks.

Producer's risks	Consumer's risks	Test censored time	RDT plan
0.1	0.594	$T = 109$	$(T, C) = (109, 0)$
0.2	0.32	$T = 234$	$(T, C) = (234, 0)$
0.3	0.165	$T = 382$	$(T, C) = (382, 0)$
0.24	0.24	$T = 297$	$(T, C) = (297, 0)$
0.215	0.3	$T = 212$	$(T, C) = (212, 0)$
0.274	0.2	$T = 295$	$(T, C) = (295, 0)$
0.367	0.1	$T = 491$	$(T, C) = (491, 0)$

For the same product, the larger the number of samples participating in the test, the shorter the time required for the test, so we can shorten the test time by increasing the test sample size. According to section 4, we can obtain the test plan when the test sample size of the product RDT is 2 and 3 (Table 5).

Table 5. RDT plans for different sample sizes.

Sample size: N = 2			Sample size: N = 3		
Producer's risks	Consumer's risks	RDT plan (T, C)	Producer's risks	Consumer's risks	RDT plan (T, C)
0.1	0.59	(54.0)	0.1	0.59	(36,0)
0.2	0.33	(117,0)	0.2	0.33	(78,0)
0.3	0.165	(191.0)	0.3	0.165	(127,0)
0.24	0.24	(149,0)	0.24	0.24	(99,0)
0.21	0.3	(127,0)	0.21	0.3	(85,0)
0.27	0.2	(170,0)	0.27	0.2	(114,0)

6. Conclusion

In this paper, we study the design method of product RDT with missing data. First, the type of missing data is clarified; Secondly, since the missing data mainly affects the parameter estimation of the reliability growth model in the development stage, this article uses the MCEM algorithm to estimate the parameters of the reliability growth model; Then the Bayes method was used again to determine the life distribution of the product, and a timed censoring test plan for the product was designed based on two types of risks; Finally, the feasibility of the RDT design method in the case of missing data was verified based on actual cases, and comparing with the traditional RDT plan, it is shown that the RDT plan obtained in this paper has the advantages of shorter test time and lower test risk.

There are many areas can improve in the research in this paper. We study exponential distribution products with a relatively simple life distribution, but the RDT design methods with missing data from other distributions such as Weibull distribution are also worth studying. What's more, comparing with the traditional test plan, our test plans are shorter test time and lower test risk, but we can also observe that our test plan is not significant enough in reducing experimental risks. There are some studies have shown that integrating prior information such as expert information in RDT design can reduce the test risks, so in the future, we will study how to integrate more prior information into RDT design in order to design a RDT plan with lower test risk.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant No.72271239.

References

Andre, K., Elmore, D., Boukai, B. 2015. A Bayesian Approach to Determine Test Sample Size Requirements for Reliability Demonstration Retesting after Product Design Change. Quality Engineering 27, 289–295.

Celeux, G., Diebolt, J. 1985. The SEM algorithm: a probabilistic teacher algorithm derived from the em algorithm for the mixture problem. Computational Statistics Quarterly 02, 73-82.

Dempster, A.P., Laird, N. M., Rubin, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 39, 1-38.

Feng, W.Z. Liu, Q. 2012. Bayes reliability test design for exponential products based on test loss. Journal of Systems Engineering 27, 568-574.

Jiang, P., Zhao, Q., Xiao, H., Wang, B , Xing, Y. Y. 2022.A reliability demonstration test plan derivation method based on subsystem test data. Computers and Industrial Engineering 170, 108325.

Jongseon, J., Suneung, A. 2018. Bayesian Methods for Reliability Demonstration Test for Finite Population Using Lot and Sequential Sampling. Sustainability 10, 1-11.

Li, D.W., Ruan, M. Z., You, K. K. 2017. Research on reliability qualification test program of weapon system based on reliability growth. Journal of Military Engineering 38, 1815-1821.

Martz, Jr.H-F., Waller, R-A. 1979. A Bayesian Zero-Failure (BAZE) Reliability Demonstration Testing Procedure. Journal of Quality Technology 03, 128-138.

Tian, T., Liu, J. H. 2015. Approximate Bayes estimation of characteristic life at normal stress level for Weibull distribution with missing data from fixed truncated-tailed step-plus-life test. Applied Mathematics vol(28), 777-781.

Wang, N.S., Wang, L.L. 2001. Bayes estimation of exponentially distributed parameters in the absence of fixed truncated data. Applied Probability Statistics 03, 229-236.

Wei, G., Martin, A. 1990. Tanner. A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. Journal of the American Statistical Association 85, 699-704.

Wilson, K.-J., Farrow, M. 2021. Assurance for sample size determination in reliability demonstration testing. Statistics 63, 523-535.

Xu, J.Y., Yu, D., Xie, M., Hu, Q. 2017. An Approach for Reliability Demonstration Test Based on Power-Law Growth Model. Quality and Reliability Engineering International 33, 1719-1730.

Yang, G.B. 2009. Reliability Demonstration Through Degradation Bogey Testing. IEEE Transactions on Reliability 58, 604-610.

Yuan, K., Li, X. G. 2014. Bayesian Approach to Determine the Test Plan of Reliability Qualification Test. Engineering Asset Management - Systems, Professional Practices and Certification.

Zhang, D., Xing, Y. Y., Jiang, P. 2023. Reliability qualification test program based on reliability growth. Systems Engineering and Electronics 45, 3699-3705.

Zhang, Z.H., Tian, Y.M. 2005. GUO Shang-Feng. Research on exponential product reliability acceptance test program. System Engineering and Electronic Technology 04, 753-756