行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　　　※
※　固定試驗成本下最佳退化試驗之設計　※
※　　　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：個別型計畫

計畫編號：NSC 89－2118－M－032－004

執行期間：88 年 8 月 1 日至 89 年 7 月 31 日

計畫主持人：吳碩傑

共同主持人：張春桃

執行單位：私立淡江大學統計學系

中 華 民 國　89 年 9 月 12 日

行政院國家科學委員會專題研究計畫成果報告

固定試驗成本下最佳退化試驗之設計

Optimal Degradation Design under Determined Cost of Experiment

計畫編號：NSC 89-2118-M-032-004

執行期限：88年8月1日至89年7月31日

主持人：吳碩傑　淡江大學統計學系

共同主持人：張春桃　淡江大學統計學系

計畫參與人員：楊清淇、陳芳易　淡江大學統計學系

1 中文摘要

隨著科技的進步，許多壽命試驗在結束時常常只收集到很少產品壽命資料，甚至可能沒有一個測試元件發生故障。在這種情況下，我們便無法使用傳統的可靠度分析方法來評估產品的可靠度。最近被討論的一種方法則是在壽命試驗的同時去測量導致產品故障的退化特徵，而這些產品機能的退化量常包含許多可靠度的訊息。一般說來，我們必須建立產品退化量和時間的關係，然後利用此關係來評估產品的可靠度。具有隨機係數的非線性迴歸模式常被用來描述產品的退化路徑。如果我們能得到此模式中參數的估計值，那麼我們便可估計出產品的壽命分配。而壽命分配百分位數估計值變異數的大小主要受到試驗中樣本數的大小、量測時間的選定及總試驗時間長短等三個變數的影響。因此這三項變數的決定將關係到壽命分配估計的精確性。然而在實際的試驗中礙於經費，往往有試驗成本的限制，此限制導致這三個變數的決定受到約制。爲使試驗符合實際狀況，本研究將探討在既定的試驗成本之下，該如何擬定一個決定退化試驗樣本大小、量測時間和總試驗時間的方法，以使所得之壽命分配百分位數估計值的變異數最小。最後，我們將利用眞實的資料來驗證本研究的可行性。

關鍵詞：退化試驗；非線性整數規劃；非線性混合效果模式；可靠度

Abstract

Degradation test is a useful technique to provide information about the lifetime of highly reliable products. We obtain the degradation measurements over time in such a test. In general, the degradation data are modeled by a nonlinear regression model with random coefficients. If we can obtain the estimates of parameters under the model, then the failure time distribution can be estimated. However, in order to obtain a precise estimate of the percentile of failure time distribution, one needs to design an optimal degradation test. Therefore, this study proposes an approach to determine the number of units to test, inspection frequency, and termination time of a degradation test under a determined cost of experiment such that the variance of estimator of percentile of failure time distribution is minimum. The method will be applied to an actual example.

Keywords: Degradation experiment; Nonlinear integer programming; Nonlinear mixed-effects model; Reliability.

2 Introduction

With today's high technology, manufacturers face increasingly intense global competition. To remain profitable, they are challenged to design, develop test and produce high reliability products. Thus, some life tests result in few or no failures in a short life testing time. In such cases, it is difficult to analyze lifetime data with traditional reliability studies. This circumstance applies to a variety of materials and products such as metals, insulations, semiconductors and electrical devices.

To study this type of reliability data, a recent approach is to obtain degradation measurements of product performance over time. Product performance is usually measured in terms of physical properties. We call these kinds of physical properties, "degradation mechanisms". From an engineering point of view, the common degradation mechanisms include fatigue, cracks, corrosion and oxidation. Examples are loss of tread on rubber tires and degradation of the active ingredient of a drug as a result of chemical reactions with oxygen and water, and by microbial, etc. To conduct a degradation test, one has to prespecify a threshold level of degradation, obtain measurements of degradation at different times, and define that failure occurs when the amount of degradation for a test unit exceeds this level. Thus, these degradation measurements may provide some useful information to assess reliability. (*e.g.*, see Nelson 1990, Chapter 11; Meeker and Escobar 1993).

In the literature, Lu and Meeker (1993) considered a nonlinear mixed-effect model and used a two-stage method to obtain estimates of the percentile of failure time distribution. Lu, Park and Yang (1997) proposed a model with random regression coefficients and standard-deviation function for analyzing linear degradation data from semiconductors. Wu and Shao (1999) established the asymptotic properties of the (weighted) least squares estimators under the nonlinear mixed-effect model. They used these properties to obtain point estimates and approximate confidence intervals for percentiles of the failure time distribution. Wu and Tsai (2000) used the optimal fuzzy clustering method to modify the two-stage method. Their procedure can robust estimation of the underlying parameters and failure time distribution. These studies focus on estimating the parameters in the degradation model and the percentile of failure time distribution.

There are some important questions about how to design a degradation test to provide a more precise estimate of the percentile of failure time distribution. Important questions include how to determine the number of units that should be tested, the number of inspections on each unit, and a termination time for the degradation experiment. Boulanger and Escobar (1994) presented a method for determining the selection of stress levels, sample size at each level and the times at which to measure the devices. However, their method is discussed under a pre-determined termination time. Yu and Tseng (1998) explained why it is not appropriate to fix the termination time in advance and proposed an intuitively appealing procedure to determine an appropriate termination time for an accelerated degradation test.

One practical problem arising from designing a degradation test is the budget of experiment. The size of budget always affects the decisions of number units to test, number of inspections and termination time and hence, affects the precision of estimating the failure time distribution. In this study, we are going to integrate these factors and the limited cost of experiment to construct a mathematical model. Then we will use the method of nonlinear integer programming to find the optimal solution of the total number of units to test, the times at which to measure the units, and the termination time of a degradation experiment. Thus, we can set up an optimal degradation test and then obtain an estimate of the percentile of failure time distribution with minimum variance. We will apply the proposed method to an numerical example.

The rest of the paper is organized as follows: Section 3 describes a nonlinear mixed-effect model for degradation data and gives the estimation procedure. Section 4 proposes a procedure to determine the number of units to test, the number of inspections on each unit and the termination time for a degradation test. Section 5 applies the proposed procedure to a numerical example. Some conclusions are in Sectoin 6.

3 Degradation Model and Parameter Estimation

In a degradation test, product performance is obtained as it degrades over time and different product units may have different performance. Thus, a statistical model for a degradation test consists of (1) a relationship between degradation measurement and time, and (2) a distribution that describes an individual product unit's characteristics. The general approach is to model the degradation of the individual units using the same functional form and the differences between individual units using random effects. The model proposed by Lu and Meeker (1993) is:

$$y_{ij} = \eta(t_{ij}; \boldsymbol{\alpha}, \boldsymbol{\beta}_i) + \varepsilon_{ij}, \qquad (1)$$

$i = 1, \ldots, n$ and $j = 1, \ldots, m_i \leq s$, where η is the actual level of degradation of the unit under study and is a given function nonlinear in $(\boldsymbol{\alpha}, \boldsymbol{\beta}_i)$; t_{ij} is the time of the jth measurement for the ith unit; y_{ij} represents the level of degradation actually observed at time t_{ij}; $\boldsymbol{\alpha}$ denotes the vector of fixed-effect parameters; $\boldsymbol{\beta}_i$ represents the vector of the ith unit random effects; ε_{ij}'s are *i.i.d.* measurement errors with mean 0 and variance σ_ε^2; s is the prespecified largest number of measurements for all units. We assume that $\{\varepsilon_{ij}\}$ and $\{\boldsymbol{\beta}_i\}$ are independent.

Suppose that η in (1) is a continuous and differentiable function for any fixed $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. We assume that the degradation is not reversible. Hence, without lost of generality, we assume that $\eta(t; \boldsymbol{\alpha}, \boldsymbol{\beta})$ is a strictly increasing function of t. The product performance degrades over time and when actual degradation η reaches a prespecified critical value η_c, failure occurs. Therefore, the failure time of a product, denoted by $T = T(\eta_c)$, is equal to the solution of $\eta(t; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \eta_c$. Furthermore, we assume that there exist constants η_1 and η_2 satisfying $\eta_2 < \eta_c < \eta_1$ such that $P[\ \lim_{t \to 0} \eta(t; \boldsymbol{\alpha}, \boldsymbol{\beta}) < \eta_2] = 1$ and $P[\ \lim_{t \to \infty} \eta(t; \boldsymbol{\alpha}, \boldsymbol{\beta}) > \eta_1] = 1$; that is, no failure occurs before the test starts and the lifetime of a unit is finite. Under these assumptions, $\eta(t;\boldsymbol{\alpha},\boldsymbol{\beta}) = \eta_c$ has a unique and finite solution for any given $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

There is an important relation between the failure time distribution and the degradation distribution, which is useful for estimating percentiles of the failure time distribution.

Lemma 1. *Let $F_\eta(x|t)$ be the degradation distribution of η given time t and $F_T(t|x)$ be the failure time distribution of $T = T(x)$ given degradation value x. Under the previously described assumptions, $F_T(t|x) = 1 - F_\eta(x|t)$.*

The percentiles of the failure time distribution are very important characteristics and elemental in reliability analysis. Let t_p be the $100p$th percentile of the failure time distribution. One can obtain percentiles of the failure time distribution using Lemma 1, *i.e.*, t_p can be obtained by solving $1 - F_\eta(\eta_c|t) = p$ in t. Suppose that the solution is $t_p = g(\eta_c, \boldsymbol{\theta})$, where g is a known function and $\boldsymbol{\theta}$ is a vector of unknown parameters. If $\boldsymbol{\theta}$ is estimated by $\hat{\boldsymbol{\theta}}_n$ based on y_{ij}'s, then t_p is estimated by $\hat{t}_p = g(\eta_c, \hat{\boldsymbol{\theta}}_n)$. Thus, it is essential to estimate $\boldsymbol{\theta}$.

Suppose that the distribution of $\boldsymbol{\beta}$ is $\pi(\boldsymbol{\beta}|\boldsymbol{\phi})$, where π is a known function and $\boldsymbol{\phi}$ is an unknown $q \times 1$ parameter vector. Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\phi})'$ and $h(t_{ij};\boldsymbol{\theta}) = E_{\boldsymbol{\beta}}[\eta(t_{ij};\boldsymbol{\alpha},\boldsymbol{\beta})]$. Define $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{im_i})'$, $\boldsymbol{h}_i(\boldsymbol{\theta}) = (h(t_{i1};\boldsymbol{\theta}), \ldots, h(t_{im_i};\boldsymbol{\theta}))'$ and $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{im_i})'$. Then, model (1) can be written as the following heteroscedastic nonlinear model:

$$\boldsymbol{y}_i = \boldsymbol{h}_i(\boldsymbol{\theta}) + \boldsymbol{e}_i, \quad i = 1, \ldots, n.$$

Denote the parameter space by Θ and the unknown true parameter by $\boldsymbol{\theta}_0$. The least squares estimator of $\boldsymbol{\theta}_0$ based on data $\{\boldsymbol{y}_i\}_{i=1}^n$ is a vector $\hat{\boldsymbol{\theta}}_n \in \Theta$ minimizing

$$Q_n(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{h}_i(\boldsymbol{\theta}))'(\boldsymbol{y}_i - \boldsymbol{h}_i(\boldsymbol{\theta})).$$

Under some regularity conditions, Wu and Shao (1999) showed the following theorem.

Theorem 1.

(i) $\hat{\boldsymbol{\theta}}_n \longrightarrow \boldsymbol{\theta}_0$ *almost surely as* $n \to \infty$.

(ii) $\boldsymbol{D}_n^{-1/2}(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \longrightarrow N(\boldsymbol{0}, \boldsymbol{I}_{p+q})$ *in distribution, where* $\boldsymbol{D}_n(\boldsymbol{\theta}) = \boldsymbol{A}_n^{-1}(\boldsymbol{\theta})(\sum_{i=1}^n \boldsymbol{H}_i(\boldsymbol{\theta})\Sigma_i(\boldsymbol{\theta})\boldsymbol{H}_i'(\boldsymbol{\theta}))\boldsymbol{A}_n^{-1}(\boldsymbol{\theta})$, $\boldsymbol{H}_i(\boldsymbol{\theta}) = \partial \boldsymbol{h}_i'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$, $\boldsymbol{A}_n(\boldsymbol{\theta}) = \sum_{i=1}^n \boldsymbol{H}_i(\boldsymbol{\theta})\boldsymbol{H}_i'(\boldsymbol{\theta})$, *and* $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$ *is the covariance matrix of* $\boldsymbol{e}_i$. *The* (j,k)*th entry of* $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$ *is*

$$\sigma_{ijk}(\boldsymbol{\theta}) = \begin{cases} \sigma_\eta(t_{ij};\boldsymbol{\theta}) + \sigma_\varepsilon^2 & \text{if } j = k \\ \sigma_\eta(t_{ij}, t_{ik};\boldsymbol{\theta}) & \text{if } j \neq k, \end{cases}$$

where $\sigma_\eta(t_{ij};\boldsymbol{\theta}) = Var_{\boldsymbol{\beta}_i}[\eta(t_{ij};\boldsymbol{\alpha},\boldsymbol{\beta}_i)]$ *and* $\sigma_\eta(t_{ij}, t_{ik};\boldsymbol{\theta}) = Cov_{\boldsymbol{\beta}_i}[\eta(t_{ij};\boldsymbol{\alpha},\boldsymbol{\beta}_i), \eta(t_{ik};\boldsymbol{\alpha},\boldsymbol{\beta}_i)]$.

We now return to the original statistical inference in degradation analysis, *i.e.*, using the degradation data to make inference for percentiles of the failure time distribution. From Lemma 1, t_p is a solution of $1 - F_\eta(\eta_c|t) = p$ and is a function of η_c and $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\phi})'$, say $t_p = g(\eta_c, \boldsymbol{\theta})$. Substituting the least squares estimate $\hat{\boldsymbol{\theta}}_n$ for $\boldsymbol{\theta}$, we obtain an estimate $\hat{t}_p = g(\eta_c, \hat{\boldsymbol{\theta}}_n)$. Since t_p is a function of $\boldsymbol{\theta}$, the asymptotic normality of $\hat{t}_p$ can be derived by the Taylor expansion and the asymptotic normality of $\hat{\boldsymbol{\theta}}_n$. Hence, the asymptotic variance of $\hat{t}_p$ is

$$Var(\hat{t}_p) = \left[\frac{\partial g(\eta_c, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]' \boldsymbol{D}_n(\boldsymbol{\theta}) \left[\frac{\partial g(\eta_c, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]. \quad (2)$$

The estimate of the asymptotic variance of $\hat{t}_p$ can be obtained by substituting $\hat{\boldsymbol{\theta}}_n$ into equation (2).

4 Optimal Degradation Design

To obtain a precise estimate of the percentile of the failure time distribution, frequently asked questions include "How many units do I need to test?", "How long do I need to run the degradation test?" or "How many times do I need to inspect the units in an experiment?" Simply put, more test units, more test time, and more number of measurements will generate more information, which improves the precision of estimates. However, the limited cost of experiment does not allow us to do so. Thus, the problem of obtaining a precise estimate of the percentile of failure time distribution under a limited cost of experiment is an important issue to the reliability analyst.

There are a lot of decision variables that affect the cost of experiment and the precision of the estimation of the percentile of failure time distribution. The most important three decision variables are: (1) the number of test units, (2) the number of inspections on each unit, and (3) the termination time of the degradation experiment. Let n denote the number of units on test. For each unit, let t be the interval between inspections (or inspection frequency) and let k be the number of inspections. Then the termination time is $(k-1)t$. The cost of experiment consists of the following three parts.

1. Sample cost. This is the cost of test units. Let C_s be the cost of a test unit. Then the total sample cost is $n * C_s$.

2. Inspection cost. The inspection cost includes the cost of using inspection equipment and material. It also depends on the number of test units and the number of inspections. Let C_m denote the cost of one inspection on one test unit. Since the total number of inspections is $n * k$, the total inspection cost is $n * k * C_m$.

3. Operation cost. The cost consists of the salary of operators, utility, and depreciation of test equipment, etc. It is proportional to the degradation-testing time. Let C_e be the operation cost in the time interval between two inspections. Then the total operation cost is $(k-1) * t * C_e$.

Therefore, the total cost of experiment is:

$$C_T = n * C_s + n * k * C_m + (k-1) * t * C_e.$$

The objective is to obtain a precise estimate of the percentile of failure time distribution; that is, to minimize the asymptotic variance of $\hat{t}_p$. Note that the asymptotic variance of $\hat{t}_p$ in (2) is a function of t, k and n. For simplicity, we can write $G(t,k,n) = Var(\hat{t}_p)$. Then, the optimal degradation design problem consists of finding t, k and n that minimize the asymptotic variance of $\hat{t}_p$. However, the selection of t, k and n is restricted to the budget of experiment, say, C_r. Hence, the optimal degradation design problem can be expressed as follows.

$$\begin{aligned} \min \quad & G(t,k,n) \\ \text{s.t.} \quad & n * C_s + n * k * C_m \\ & \quad + (k-1) * t * C_e \le C_r, \\ & k, n \in N, \quad \text{and} \quad k \ge 2, \end{aligned}$$

where N is the set of positive integers. In order to find the optimal solution for the problem of nonlinear integer programming, we set up the following algorithm.

1. Give the values of parameters $\boldsymbol{\theta}$ and σ_ε^2, and set the values of C_s, C_m, C_e and C_r.

2. Let $n_b = \left[\dfrac{C_r}{C_s + 2C_m}\right]$, where $[x]$ is the greatest integer that is less than or equal to x.

3. Set $n = 2$.

4. Let $k_b^n = \left[\dfrac{C_r - nC_s}{nC_m}\right]$.

5. For all $k \in N$, $2 \le k \le k_b^n$, let $t_b^k = \dfrac{C_r - nC_s - nkC_m}{(k-1)C_e}$ and calculate $G(t_b^k, k, n)$.

6. Let $F(n) = G(t_b^{k^n}, k^n, n) = \min_{2 \le k \le k_b^n} G(t_b^k, k, n)$.

7. Set $n = n + 1$. If $n \le n_b$ go to Step 4, else go to Step 8.

8. Let $F(n^*) = \min_{2 \le n \le n_b} F(n) = G(t^*, k^*, n^*) = \min_{2 \le n \le n_b} G(t_b^{k^n}, k^n, n)$.

9. (t^*, k^*, n^*) is the optimal solution.

Remark 1. In Step 1, the values of $\boldsymbol{\theta}$ and σ_ε^2 are usually unknown. Thus, we can use prior information or data from a pilot test to get their estimates by using the estimation method in Section 3.

5 Numerical Example

The example consists of designing a degradation test for a metal film resistor. No failures are expected to occur during an experiment. But there is some concern about changes over time in resistance, an important characteristic of the metal film resistor. The resistance is an increasing function of time. If resistance increases too far from its original value, the failure of metal film resistor could occur.

Wu and Shao (1999) reported and analyzed the results of a previous degradation test. A total of 200 metal film resistors were tested, and the resistance of each resistor was measured at five different times during the experiment. There were no failures, where a failure was defined as the ratio of resistance at time t to initial resistance that is beyond 1.02. Wu and Shao (1999) used the following model to describe the degradation over time t_j of a resistor i:

$$y_{ij} = \beta_i t_j^\alpha + \varepsilon_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, m,$$

where y_{ij} = (observed resistance at time t_j/ observed resistance at time t_1) $-$ 1, t_j is the measurement time, α is a fixed effect parameter, and ε_{ij} is random error with mean 0 and variance σ_ε^2. The random effects β_i's are independently distributed as an exponential distribution with mean λ.

Following Section 3, t_p has an explicit form:

$$t_p = \left(\frac{-\eta_c}{\lambda \log(p)}\right)^{\frac{1}{\alpha}}.$$

The estimates of λ, α and σ_ε^2 obtained by Wu and Shao (1999) are $\hat{\lambda} = 0.0021$, $\hat{\alpha} = 0.4626$ and $\hat{\sigma}_\varepsilon = 0.00001$, respectively. We use their results in the design of our new experiment. Suppose that the values of cost parameters are as follows: $C_s = 40$ (dollar) $C_m = 0.07$ (dollar), $C_e = 50.0$ (dollar/thousand hours), and $C_r = 4000.0$ (dollar). By using the algorithm in Section 4, we can obtain the optimal degradation design as follows.

$$t^* = 0.077, \quad ,k^* = 6, \quad n^* = 49.$$

That is, the optimal number of test units is 49, the optimal number of inspections is 6, and the optimal termination time is $0.077 * 1000 * (6 - 1) = 385$ hours. In Table 1, we also list some optimal solutions under different values of C_r.

Table 1: The optimal solutions of (t^*, k^*, n^*) under some values of C_r

C_r	(t^*, k^*, n^*)
2000	(0.077, 6, 49)
3500	(0.149, 5, 86)
5000	(0.139, 3, 124)
8000	(0.384, 3, 198)

6 Conclusion

Determining appropriate number of test units, number of inspections, and termination time under limited cost of experiment is an important decision problem for experimenters when conducting a degradation test. We use the method of nonlinear integer programming and propose an algorithm to set up an optimal degradation design. Under this optimal degradation design, we obtain an estimate of the percentile of failure time distribution with minimum variance. This approach is very intuitive, and can be easily explained to engineers.

References

Boulanger, M. and Escobar, L. A. (1994). Experimental design for a class of accelerated degradation tests, *Technometrics*, **36**, 260-272.

Lu, C. J. and Meeker, W. Q. (1993). Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, **35**, 161-174.

Lu, J.-C., Park, J. and Yang, Q. (1997). Statistical inference of a time-to-failure distribution derived from linear degradatoin data, *Technometrics*, **39**, 391-400.

Meeker, W. Q. and Escobar, L. A. (1993). A review of recent research and current issues in accelerated testing, *International Statistical Review*, **61**, 147-168.

Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis.* Wiley, New York.

Wu, S.-J. and Shao, J. (1999). Reliability analysis using the least squares method in nonlinear mixed-effect degradation models, *Statistica Sinica*, **9**, 855-877.

Wu, S.-J. and Tsai, T.-R. (2000). Estimation of time-to-failure distribution derived from a degradation model using fuzzy clustering, *Quality and Reliability Engineering International*, accepted.

Yu, H.-F. and Tseng, S.-T. (1998). On-line procedure for terminating an accelerated degradation test, *Statistica Sinica*, **8**, 207-220.