The Pennsylvania State University

The Graduate School

Department of Industrial and Manufacturing Engineering

BAYESIAN TWO-STAGE ACCELERATED LIFE TESTS UNDER

DISTRIBUTION MISSPECIFICATION

A Thesis in

Industrial Engineering

by

Xiujie Zhao

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science

August 2015

The thesis of Xiujie Zhao was reviewed and approved* by the following:

Enrique del Castillo
Distinguished Professor of Industrial Engineering and Professor of Statistics
Thesis Advisor

M. Jeya Chandra
Professor of Industrial Engineering

Harriet B. Nembhard
Professor of Industrial Engineering
Interim Department Head of Industrial and Manufacturing Engineering

*Signatures are on file in the Graduate School.

Abstract

Accelerated life testing (ALT) is a common way to predict the lifetime of an item at use condition by subjecting items to extreme stress conditions that accelerate the occurrence of failures. Weibull or lognormal lifetime assumptions are most widely used in ALT. However, under many circumstances, it is not easy to determine which distribution to assume. This thesis proposes a two-stage Bayesian based design method to construct an ALT plan that is robust with respect to distribution misspecification and parameter uncertainty. In the method, the first stage experiment gives posteriors of parameters and distribution information, which are then used in the second stage experiment to get the final updated posteriors. The final posteriors are used to optimize the final ALT plan, which provides the stress levels and the allocation of the items under test to each stress level. The objective is to minimize the weighed expected prediction variance of a given quantile of the lifetime at use conditions. Our numerical experiments demonstrate the efficiency and increased robustness of the resulting Bayesian two stage ALT plans compared to single stage ALT plans.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to take this opportunity to express my sincere gratitude to my thesis advisor, Dr. Enrique del Castillo. I am very grateful to his the continuous support and help for my curriculum and research. His guidance helps me in all the time of research and writing of this thesis.

I am also thankful for my thesis reader, Dr. M. Jeya Chandra for his encouragement, insightful comments to make the thesis better.

My sincere thanks also go to Dr. Michael Akritas, Dr. Jose A. Ventura, Dr. Sanjay Joshi and many faculty members in Industrial Engineering and Statistics, for offering valuable help in my curriculum.

I appreciate the help from my fellow mates : Xue Wang, Wendian Wan, Yu'an Wang and Yue Xie, who studied, discussed and worked with me with great patience.

Finally, I would like to thank my family: my parents Yongyuan Zhao and Shumin Zhang, my old brother Xiulin and sister-in-law Juan Sun for loving me and supporting me throughout all my life. I also want to thank my girlfriend, Ran Wang, who gives unconditional support to me in recent years.

Chapter 1

Introduction

1.1 Accelerated Lifetime Testing

In reliability studies, lifetimes of many products are extremely long, making it very difficult to conduct life tests under normal conditions. Accelerated Life Testing (ALT) is the process of testing a product by subjecting it to higher levels of stresses (stress, strain, temperatures, voltage, vibration rate, pressure etc.) in excess of its use conditions to produce failures faster in an effort to reduce cost and duration of testing. Then the results are analyzed to extrapolate to the use condition, hence obtain an estimation of lifetime at use condition. Nowadays ALT is widely used in reliability research and during the design and testing of new products.

Most accelerated life tests are based on some known acceleration relationships, including physical acceleration relationships and empirical acceleration relationships [14]. The failure mechanisms may be either well-understood or not, thus the models are being updated by physical/chemical scientists and reliability researchers continuously. One common such relationship popular to model acceleration is the Arrhenius relationship, which we will use in our study and will be introduced in Chapter 2.

Data obtained from ALT usually involve right censoring (some of the units haven't failed at the end of a test). Two types of right censoring are commonly considered in practice:

- Type I censoring: or "time censoring", is to remove unfailed units from test

at a specific time, denoted as t_c in this study. Type I censoring is due to constraints on the duration of lifetime tests, thus tests have to be terminated before all the units have failed.

- Type II censoring: or "failure censoring", means that a life test has to be terminated after a specific number, often denoted as r, of test units fail. Type II censoring is due to constraints on the number of units available in life tests, such censoring mechanism is usually applied on expensive units to control the cost of testing.

The reason we are considering censoring data is that censoring data contribute different likelihood and different kind of censoring provides different Fisher information on the parameters.

Regression models are frequently used to explain the relationship between accelerating regressors and lifetimes, log-location-scale regression models (applicable to log-location-scale distributions, including Weibull, lognormal, loglogistic etc.) are especially widely used. A simple regression model is an example,

$$\Pr(T \le t) = F(t; \mu, \sigma) = F(t; \beta_0, \beta_1, \sigma) = \Phi\left[\frac{\log(t) - \mu}{\sigma}\right] \qquad (1.1)$$

where $\mu = \beta_0 + \beta_1 z$ and Φ is the CDF of assumed log-lifetime distribution. The quantile function of the model is

$$\log\left[t_p(z)\right] = \mu + \Phi^{-1}(p)\sigma = \beta_0 + \beta_1 z + \Phi^{-1}(p)\sigma \qquad (1.2)$$

where Φ^{-1} is the inverse CDF of the assumed log-lifetime distribution and p is some specific percentile determined by researchers or test designers. The accelerating regressor z is usually coded or transformed from real accelerating factor based on acceleration relationships.

Before conducting a life test, test designers need to plan the test. Usually to plan an ALT one need to specify,

- The levels of accelerating regressor(s), $z_i, i = 1, 2, ..., l$, where l is number of levels. The regressors are determined by accelerating variables (in original measure)

- The proportion of test units to allocate at each accelerating level, $\pi_i, i = 1, 2, ..., l$, corresponding to z_i respectively, where $\sum_{i=1}^{l} \pi_i = 1$

Some common ALT planning criteria are as follows [14],

- Estimate a particular quantile t_p, or in our study $\log t_p$, in the lower tail of the failure-time distribution under use conditions. Hence the most important criterion is to minimize the asymptotic variance or standard error of $\log(\hat{t}_p)$, the ML estimator of $\log t_p$, denoted as Ase $\left[\log(\hat{t}_p)\right]$.

- The Fisher information matrix I_θ may also be used for planning an ALT, in this case an important purpose of the test is to estimate the parameters (β's, σ, etc.) more precisely. For example, if we maximize the determinant of I_θ, a "D-Optimal" design is obtained, instead, if tr (I_θ) is maximized, we will get an "A-Optimal" design for the planning.

- In recent years, the model robustness with respect to the acceleration models, failure time distribution and regression models have become popular in the literature [18][30][31]. Many different criteria are used in this field.

The constraints on planning ALT may include

- The length (duration) of test.

- The total number of units to be tested.

- Constraints on accelerating levels. We cannot conduct tests under impossible or excessively costly accelerating conditions, thus there may be lower and(or) upper bounds on the accelerating levels of the factors.

1.2   Weibull and lognormal Failure Times

Weibull and lognormal failure time models describe a broad range of ALT situations. In the earlier stage of ALT research, Kielpinski and Nelson [11], Nelson and Kielpinski [22] investigated the ALT plans to estimate the lifetime of Class-B insulation for electric motors when either a normal lifetime or lognormal lifetime is assumed. Nelson and Meeker [23] studied the lifetime of insulating fluid at use

voltage of 15kV by assuming the lifetime is Weibull or extreme value distributed. In the following explorations, Meeker [12] compared the models based on Weibull and lognormal lifetimes, Pascual [25] introduced ALT planning under distribution misspecification between the two distributions.

Weibull distribution was introduced by Waloddi Weibull [29] in 1951. The distribution function (CDF)

$$F(x; \lambda, k) = \begin{cases} 1 - e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (1.3)$$

The Weibull distribution has two parameters, scale parameter λ and shape parameter k. The shape parameter k makes Weibull distribution different and more flexible than exponential distribution:

- A value of $k < 1$ indicates that the failure rate decreases over time. This happens if there is "infant mortality".

- A value of $k = 1$ indicates that the failure rate is constant over time. This case coincides with a exponential distribution.

- A value of $k > 1$ indicates that the failure rate increases over time. This happens if there is an "aging" process, such as corrosion or fatigue.

Hence the hazard function of Weibull distribution is monotone. Weibull distribution describes many failure models. Figure 1.1 illustrates the CDF under a shared parameter λ but different parameter k's. Weibull lifetimes have such properties that we can use on log-lifetime. If a random variable T follows a Weibull distribution, $T \sim \text{Weibull}(\lambda, k)$, then the natural logarithm of T follows a smallest extreme value(SEV) distribution,

$$\log(T) \sim \text{SEV}(\mu, \sigma)$$

where $\mu = \log 1/\lambda$, $\sigma = 1/k$. The distribution function (CDF) of SEV distribution is

$$\Phi_{SEV}(x) = \begin{cases} 1 - \exp\left[-\exp(x)\right] & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (1.4)$$
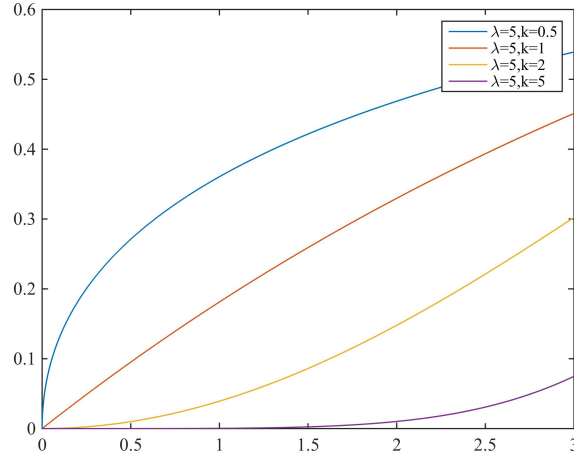
Figure 1.1. Plots of distribution function of Weibull distribution with different k

Lognormal distribution is another distribution describing many failure models. The CDF of lognormal distribution is

$$F(x; \mu, \sigma) = \Phi\left(\frac{\log x - \mu}{\sigma}\right) \tag{1.5}$$

where Φ is cumulative distribution function of standard normal distribution. Similar to Weibull distribution, if a random variable T follows a lognormal distribution, $T \sim \text{LogN}(\mu, \sigma)$, then the natural logarithm of T follows a normal distribution,

$$\log(T) \sim N(\mu, \sigma)$$

A significant difference of lognormal distribution from Weibull is that the hazard function of lognormal distribution may be not monotone, as Figure 1.2 shows. From the figure, the hazard rate may be increasing at first, then decreasing or just decreasing. Weibull distribution and lognormal distribution can be very similar to each other when the parameters are set particularly, some examples are shown on Figure 1.3.
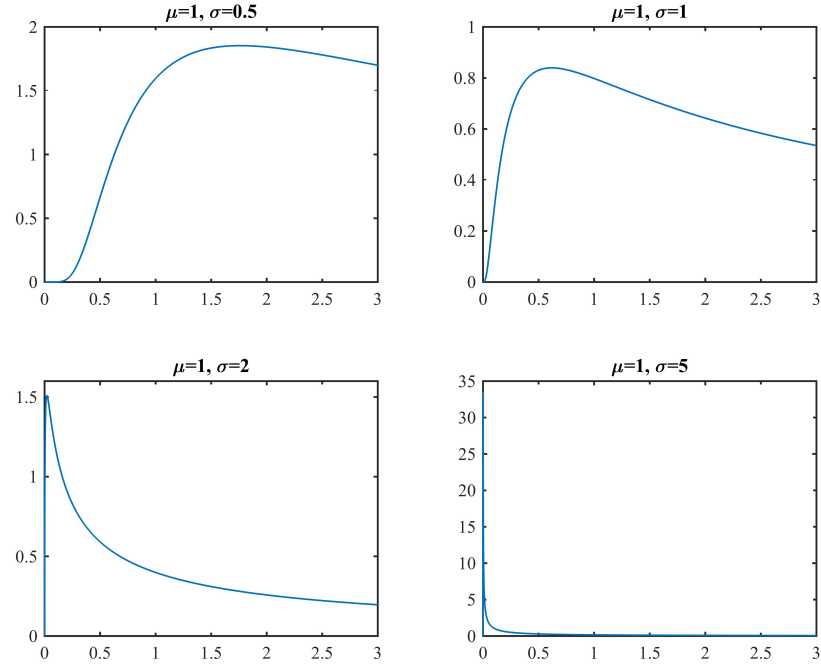
Figure 1.2. Hazard function of lognormal distribution under different σ
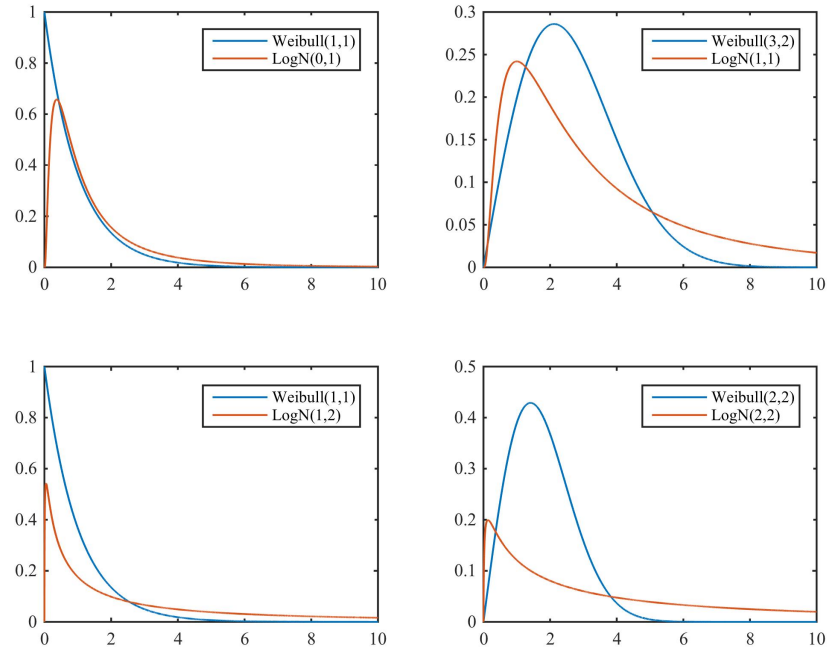
Figure 1.3. Plots of Weibull and lognormal distribution with different parameters

1.3 Previous Research on ALT Planning

Research on accelerated life testing has a history of over fifty years. Towards the end of 1970's, reliability researchers began to investigate how to evaluate and design ALT plans.

1.3.1 Traditional ALT Planning Methods

In the 20th century, most ALT planning investigations were based on the assumption that lifetime follows one particular distribution. Weibull and lognormal distributions have been most widely studied in the last three decades. Nelson and Kiepimski [22], as well as Nelson and Meeker [23] proposed the theories for optimum ALT planning for censored data. As the model described previously, these theories aimed at minimizing $\text{Ase}\left[\log(\hat{t}_p(z_0))\right]$, where z_0 is the regressor corresponding to use condition of the product. For β_0, β_1 and σ, the asymptotic inverse of the Fisher information matrix is

$$\Sigma = F^{-1} = \begin{bmatrix} \text{Var}(\hat{\beta}_0) & \text{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \text{Cov}(\hat{\beta}_0, \hat{\sigma}) \\ \text{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \text{Var}(\hat{\beta}_1) & \text{Cov}(\hat{\beta}_1, \hat{\sigma}) \\ \text{Cov}(\hat{\beta}_0, \hat{\sigma}) & \text{Cov}(\hat{\beta}_1, \hat{\sigma}) & \text{Var}(\hat{\sigma}) \end{bmatrix} \tag{1.6}$$

Denote

$$z_i = \frac{x_i - x_0}{x_H - x_0} \tag{1.7}$$

where x_i is the acceleration factor at level i, note that x_i may also be coded from an acceleration relationship, such as Arrhenius equation, hence the transformation from x_i to z_i is only to constrain the regressor between 0 and 1. Then denote

$$\zeta_i = \frac{\log(t_c) - \beta_0 - \beta_1 z_i}{\sigma} \tag{1.8}$$

where t_c is the specified right censoring time for Type I censoring. For Type II censoring, $\zeta_i = \Phi^{-1}(p_{c,i})$ [25], where $p_{c,i}$ is the specified proportion of failures at z_i. The study showed that the Fisher information matrix of an observation at z_i, F_{z_i}

is

$$F_{z_i} = \begin{bmatrix} A(\zeta_i) & z_i A(\zeta_i) & B(\zeta_i) \\ z_i A(\zeta_i) & z_i^2 A(\zeta_i) & z_i B(\zeta_i) \\ B(\zeta_i) & z_i B(\zeta_i) & C(\zeta_i) \end{bmatrix} \tag{1.9}$$

where $A(), B(), C()$ are functions of ζ_i [6]. If π_i is the proportion of allocation to z_i, then the Fisher Information matrix for the design D is

$$I_{\beta_0, \beta_1, \sigma}(D) = n \sum_{i=1}^{l} \pi_i F_{z_i} \tag{1.10}$$

The inverse of the Fisher Information matrix is the variance-covariance matrix of $(\beta_0, \beta_1, \sigma)$, denoted as Σ. By Σ, Ase $\left[\log(\hat{t}_p(z_0)) \right]$ can be easily derived then we can optimize on z_i's and π_i's. To solve the optimization problem, we need to evaluate Ase $\left[\log(\hat{t}_p(z_0)) \right]$ for a large number of times.

Unfortunately, the functions $A(), B(), C()$ are very difficult to compute at that time since it involves considerable numerical integration, making the optimization problem very time-costing and hard to solve. In the early 90's, Escobar and Meeker [9] introduced an algorithm (AS292) to compute the Fisher information matrix for the extreme value, normal and logistic distribution. The algorithm is an extension of AS218 [8], AS218 only considered the case of smallest extreme value distribution (the case of Weibull lifetime). Based on the algorithms, Escobar and Meeker [10] summarized the planning of accelerated life tests with one or two factors by illustrating a Weibull lifetime example. The algorithm introduced a scaled Fisher information matrix for an observation. For $\boldsymbol{\theta} = (\beta_0, \sigma, \beta_1)'$,

$$I_{\boldsymbol{\theta}}(D) = \frac{n}{\sigma^2} \sum_{i=1}^{l} \pi_i \mathcal{F}_i \tag{1.11}$$

where $\mathcal{F}_i$ is the scaled Fisher information matrix,

$$\mathcal{F}_i = \begin{bmatrix} f_{11}(\zeta_i) & f_{12}(\zeta_i) & f_{11}(\zeta_i) z_i \\ f_{12}(\zeta_i) & f_{22}(\zeta_i) & f_{12}(\zeta_i) z_i \\ f_{11}(\zeta_i) z_i & f_{12}(\zeta_i) z_i & f_{11}(\zeta_i) z_i^2 \end{bmatrix} \tag{1.12}$$

where f_{11}, f_{12}, f_{22} can be derived numerically from the algorithm. As computer

science and numerical methods advance, the computation of the Fisher information matrix for ALT parameters becomes faster, making the optimization on ALT planning possible. In Chapter 3, numerical experiment will show the optimization using MATLAB.

Besides traditional ALT planning, other accelerated tests have also been being investigated. Nelson [21] investigated ALT step-stress models and data analysis, Bai, Kim and Lee [2] illustrated a simple example of ALT step-stress model with censoring.

Meeker and Escobar [15] proposed accelerated degradation tests and described the modeling and evaluation, and made a comparison of the results of accelerated degradation analysis and traditional accelerated life test failure-time analysis. Tang, Yang and Xie [28] studied the planning of step-stress accelerated degradation test.

1.3.2   ALT Planning Based on Bayesian Methods

All the research on ALT planning described in the last section is non-Bayesian, parameters are maximum likelihood estimated. However, in practice, before the tests, some lifetime acceleration properties are not determined, making the selection of models and corresponding parameters of ALT very difficult. Hence assigning prior distributions to the parameters becomes a potential way to plan better accelerated life tests.

Chaloner and Larntz [4] put a point mass prior distribution on parameters β's and σ, which is very similar to Bayesian designs, allowing for prior uncertainty in the parameters. In this investigation, prior probabilities of Weibull or lognormal model, λ_1 and λ_2 are specified. In fact, they use the priors to average the optimum plans, assuming that the prior information will not be used in inference. Erkanli and Soyer [7] presented optimum Bayesian ALT designs for exponential distribution with no censoring, in this study, the authors adopted the curve fitting method of Monte Carlo experiments, which was introduced by Müller and Parmigiani [20] and Müller [19]. Al-Hussaini and Alla H [1] presented Bayesian Estimation of the parameters of ALT models. Zhang and Meeker [33][32] followed the general Bayesian design framework by Polson [26], and showed how to evaluate the effect

of the prior information on planning. Shi and Meeker [27] investigated Bayesian methods for accelerated destructive degradation test planning using similar ideas.

1.3.3 Modeling: Asymptotic Bias and Asymptotic Mean Square Error

Even though there are large amount of non-Bayesian or Bayesian methods based ALT investigations, very few of them studied the model robustness with respect to acceleration regression models or assumed distributions. Nelson [24] remarked that robust ALT planning researches are lacking, which should be addressed.

Pascual and Montepiedra [25] showed how to compute the asymptotic bias (ABias) and the asymptotic mean square error (AMSE) of observed bias for the lognormal and Weibull ALT models when the wrong distribution is specified to plan the ALT. The asymptotic mean square error of $\log(\hat{t}_p(z_0))$ is defined as,

$$\text{AMSE}\left[\log(\hat{t}_p(z_0))\Big|M_i\right] = \text{AVar}\left[\log(\hat{t}_p(z_0))\Big|M_i\right] + \left\{\text{ABias}\left[\log(\hat{t}_p(z_0))\Big|M_i\right]\right\}^2$$

(1.13)

where M_i is the model that is assumed. If the model is true, then $\text{ABias}\left[\log(\hat{t}_p(z_0))\Big|M_i\right] = 0$. If unlimited number of test units are given, it means the sample size N is sufficiently large, $\text{AVar}\left[\log(\hat{t}_p(z_0))\Big|M_i\right]$ can be made arbitrarily small. However, if a model selection error is made, $\text{ABias}\left[\log(\hat{t}_p(z_0))\Big|M_i\right]$ can not be near 0 even if the sample size is large. Thus optimization can be done on either $\text{AMSE}\left[\log(\hat{t}_p(z_0))\Big|M_i\right]$ or $\text{ABias}\left[\log(\hat{t}_p(z_0))\Big|M_i\right]$. The result recommends to minimize $\text{ABias}\left[\log(\hat{t}_p(z_0))\Big|M_i\right]$ only if the sample size is relatively large.

However, the research is based on Meeker and Hanh's 4:2:1 planning [16], which is basically planning for the possible quadratic acceleration model, in case the linear acceleration model fails. The planning makes the optimization problem an simple one-dimensional search on one level of acceleration. Hence the method may be not general for many other problems.

Monroe and Pan et al. [18] and Yang and Pan [30] used a generalized linear model framework to conduct a sensitivity analysis of optimal designs for ALT, then proposed optimal accelerated life test planning with interval censoring.

1.3.4 Bayesian Model Averaging

Yu and Chang [31] applied bayesian model averaging (BMA) for quantile estimation in ALT. Denote Δ be the quantity of interest, after collecting some data $\boldsymbol{y}$, BMA gives the posterior distribution of Δ as,

$$f(\Delta|\boldsymbol{y}) = \sum_{k=1}^{K} f(\Delta|\boldsymbol{y}, M_k)\text{Pr}(M_k|\boldsymbol{y}) \tag{1.14}$$

where $M_1, M_2, ..., M_K$ are the candidate models. To compute $\text{Pr}(M_k|\boldsymbol{y})$, Beyes theorem gives that,

$$\text{Pr}(M_k|\boldsymbol{y}) \propto \text{Pr}(M_k)f(\boldsymbol{y}|M_k) \tag{1.15}$$

where $\text{Pr}(M_k)$ is the prior probability that M_k is the true model. Let $\boldsymbol{\theta}_k$ denote the vector of the parameters in model M_k with prior $f(\boldsymbol{\theta}_k|M_k)$, then

$$f(\boldsymbol{y}|M_k) = \int f(\boldsymbol{\theta}_k|M_k)f(\boldsymbol{y}|\boldsymbol{\theta}_k, M_k)\text{d}\boldsymbol{\theta}_k \tag{1.16}$$

where $f(\boldsymbol{y}|\boldsymbol{\theta}_k, M_k)$ is the likelihood function. By using BMA, a Bayesian weighted function is obtained, which can be optimized later to increase robustness to models.

More pitfalls of of accelerated tests can be found in [13][17], which were summarized by William Q. Meeker and many other reliability researchers and engineers.

1.4 Motivation for the Two-Stage Bayesian ALT Planning Methods

Previous ALT planning research, either non-Bayesian or Bayesian, is based on information that is given or supposed. However, it is usually the case that we don't know exactly the acceleration lifetime properties behind the product. When a new product is to be tested, it is possible that we are not sure about the lifetime distribution and the acceleration parameters. This motivates our approach where we build a two-stage design in which the second stage uses the information obtained in the first stage.

Bayesian Two-Stage Design Model

2.1 Preliminaries

Preliminaries and notations used in the rest of this thesis are introduced in this section.

2.1.1 Basic Model Notations

As in Zhang and Meeker's study [33], consider a linear acceleration regression model with only one regressor z,

$$\log t_p(z) = \gamma_0 + \gamma_1 z + \sigma \Phi^{-1}(p) \tag{2.1}$$

where Φ is the distribution function (CDF) of SEV distribution or standard normal distribution. Then denote Model 1 as M_1, that is,

$$\log t_p(z) = \gamma_{01} + \gamma_{11} z + \sigma_1 \Phi_{SEV}^{-1}(p) \tag{2.2}$$

where Φ_{SEV} is the CDF of smallest extreme value distribution and denote Model 2 as M_2, that is,

$$\log t_p(z) = \gamma_{02} + \gamma_{12} z + \sigma_2 \Phi_N^{-1}(p) \tag{2.3}$$

where Φ_N is the CDF of standard normal distribution.

We are considering both models in the same problem, hence the two models

share the same desired percentile p. For each of the models M_1 and M_2, respectively, the parameters are denoted as,

$$\boldsymbol{\theta}_{01} = (\gamma_{01}, \sigma_1, \gamma_{11})'$$
$$\boldsymbol{\theta}_{02} = (\gamma_{02}, \sigma_2, \gamma_{12})' \tag{2.4}$$

2.1.2 Arrhenius Relationship

The Arrhenius relationship is a widely used model describing the effect temperature has on a chemical reaction rate. The relationship can be written as,

$$\mathcal{R}(\text{temp}) \propto \exp\left(-E_a \times \frac{11605}{\text{temp(K)}}\right) \tag{2.5}$$

where $\mathcal{R}$ is the reaction rate and $\text{temp(K)} = \text{temp}(^\circ\text{C}) + 273.15$ is the temperature in absolute Kelvin scale, and E_a is a parameter that is determined by material characteristics. Hence the Arrhenius acceleration factor is

$$\mathcal{AF}(\text{temp}, \text{temp}_0) = \Psi(z) = \frac{\mathcal{R}(\text{temp})}{\mathcal{R}(\text{temp}_0)} = \exp\left[-E_a\left(\frac{11605}{\text{temp}_0(\text{K})} - \frac{11605}{\text{temp(K)}}\right)\right] \tag{2.6}$$

where temp_0 is the temperature under use condition. When $\text{temp} > \text{temp}_0$, $\mathcal{AF} > 1$, thus higher temperature helps increasing reaction rate. Recall the model in Equation 2.1,

$$t_p(z) = \underbrace{\exp(\gamma_0 + \sigma\Phi^{-1}(p))}_{t_p(0)} \times \exp(\gamma_1 z) = \frac{t_p(0)}{\Psi(z)} \tag{2.7}$$

[6] where $\Psi(z) = -\exp(\gamma_1 z)$. To make the model an AL model, we can make

$$\beta_1 = E_a$$

$$z = \frac{11605}{\text{temp(K)}} - \frac{11605}{\text{temp}_0(\text{K})} \tag{2.8}$$

hence $\Psi(0) = 1$, ensuring the model is an AL model.

2.1.3   Standardization of Variables

The proposed acceleration regression model is only valid in a range of z, (z_U, z_H), where z_U is the use level and z_H is the highest level we can set in the test. Testing beyond z_H is either impossible practically or breaking the accelerating model, hence the extrapolation is dangerous, the planning levels should be set in the range (z_U, z_H). For simplicity, the accelerating variable is standardized, denote ξ as the variable after standardization,

$$\xi = \frac{z}{z_H - z_U} = \left(\frac{11605}{\text{temp}(\text{K})} - \frac{11605}{\text{temp}_U(\text{K})} \right) \bigg/ \left(\frac{11605}{\text{temp}_H(\text{K})} - \frac{11605}{\text{temp}_U(\text{K})} \right) \quad (2.9)$$

where temp_U and temp_H denote the use-level and highest level of temperatures respectively, so that $\xi \in [0, 1]$. In addition, the parameters need to be transformed corresponding to ξ, denote

$$\log t_p(\xi) = \beta_0 + \beta_1 \xi + \sigma \Phi^{-1}(p) \quad (2.10)$$

In this transformation, let

$$\beta_0 = \gamma_0$$

$$\beta_1 = (z_H - z_U)\gamma_1 \quad (2.11)$$

Hence the model is equivalent to that in equation 2.1, and $\mu_U = \beta_0$ and $\mu_H = \beta_0 + \beta_1$ are the location parameters of the log-lifetime distribution at use condition level and highest level respectively. At the stage we can denote M_1 and M_2 in the standardized way,

$$M_1 : \log t_p(\xi) = \beta_{01} + \beta_{11} \xi + \sigma_1 \Phi_{SEV}^{-1}(p) \quad (2.12)$$

with parameters $\boldsymbol{\theta}_1 = (\beta_{01}, \sigma_1, \beta_{11})'$. Similarly,

$$M_2 : \log t_p(\xi) = \beta_{02} + \beta_{12} \xi + \sigma_2 \Phi_N^{-1}(p) \quad (2.13)$$

with parameters $\boldsymbol{\theta}_2 = (\beta_{02}, \sigma_2, \beta_{12})'$

2.2 Final Objective Function

To design the two-stage ALT planning, first we need to determine the final objective function. Denote the function as $C(D)$, where D is the ALT plan to be optimized. Thus

$$
\begin{aligned}
C(D) \quad = \quad & E_{\boldsymbol{\theta}_1|D} \left\{ \mathrm{AVar}_{\boldsymbol{\theta}_1|\boldsymbol{t},D} \left[\log t_p(0)|M_1 \right] \right\} \mathrm{Pr}(M_1|\boldsymbol{t}) \\[2mm]
& +E_{\boldsymbol{\theta}_2|D} \left\{ \mathrm{AVar}_{\boldsymbol{\theta}_2|\boldsymbol{t},D} \left[\log t_p(0)|M_2 \right] \right\} \mathrm{Pr}(M_2|\boldsymbol{t})
\end{aligned}
\tag{2.14}
$$

where

- $\mathrm{AVar}_{\boldsymbol{\theta}_1|\boldsymbol{t},D} \left[\log t_p(0) \Big| M_1 \right]$ is the asymptotic posterior variance of $\log t_p(0)$ after observing data $\boldsymbol{t}$ in the first stage by assuming that M_1 is the true model, i.e., the Weibull case.

- $\mathrm{AVar}_{\boldsymbol{\theta}_2|\boldsymbol{t},D} \left[\log t_p(0) \Big| M_2 \right]$ is the asymptotic posterior variance of $\log t_p(0)$ after observing data $\boldsymbol{t}$ in the first stage by assuming that M_2 is the true model, i.e., the lognormal case.

- $\mathrm{Pr}(M_1|\boldsymbol{t})$ is the posterior of probability that M_1 is the true model after observing data $\boldsymbol{t}$, i.e., the Weibull case.

- $\mathrm{Pr}(M_2|\boldsymbol{t})$ is the posterior of probability that M_2 is the true model after observing data $\boldsymbol{t}$, i.e., the lognormal case.

The objective function is the weighed expectation of posterior asymptotic variances by posterior weights. The reason why there are expectations on the asymptotic variance is that $\boldsymbol{\theta}_1|\boldsymbol{t}$ and $\boldsymbol{\theta}_2|\boldsymbol{t}$ follows posterior distributions, instead of being assumed. By minimizing the objective function, a final ALT plan will be obtained after observing data in the first stage.

2.3 First Stage Design

The first stage design of the ALT is completely based on priors, no data are collected before the design. Thus determining the priors is the key preparation for the first stage design. This is analogous to choosing "planning parameters" in frequentist ALT plans.

2.3.1 Priors

Priors of parameters $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and the information which distribution is more likely to be true need to be specified. To describe the information regarding the distribution, a Bernoulli random variable is good to explain the model selection event. Denote

$$U = \begin{cases} 0 & \text{if } M_1 \text{ is true} \\ 1 & \text{if } M_2 \text{ is true} \end{cases} \tag{2.15}$$

The distribution has one parameter, denotes as a, making

$$\Pr(U = 0) = a$$

Priors of ALT model parameters have no general disciplines. Zhang and Meeker [33] proposed that available prior information can be quantified in terms of a joint prior distribution of $\boldsymbol{\theta}$, denoted as $w(\boldsymbol{\theta})$. Actually the priors are usually given by experienced engineers, and the parameters are commonly believed to be independent. Let S denote the variance-covariance matrix of the prior distribution of $\boldsymbol{\theta}$. Hence S^{-1} is the prior precision matrix of $\boldsymbol{\theta}$. In the examples in Chapter 3, some further details will be given.

For the Bernoulli parameter a in formula 2.15, several candidate prior distributions may be considered.

- Degenerate distribution. To use degenerate distribution as prior for a, we only need to specify a constant c making $\Pr(a = c) = 1$, thus the prior is simply a constant with no variance. For example, as showed in the upper-left plot in Figure 2.1, $c = 0.5$.

- Discrete distribution. Instead of specifying only one constant c, we can specify several possible values for a and assign probabilities to them, making the prior distribution a discrete distribution. The possible values of the distribution should be inside $[0, 1]$.

- Continuous Distribution. Compared with degenerate and discrete distribution, continuous distribution provides more flexible choices for the prior. Distribution defined on an interval may be considered as the prior,

– Uniform distribution. Uniform distribution on [0,1] corresponds to a diffuse or approximately noninformative prior distribution. As in the upper right plot in Figure 2.1, the prior does not express strong preference to the value of a.

– Triangular distribution. Triangular distribution has a single mode, showing a preference of a value of a. As in the lower right plot in Figure 2.1, the distribution must be defined on $[0, 1]$

– Beta distribution. Beta distribution is a family of continuous probability distributions defined on interval $[0, 1]$ parametrized by two positive shape parameters, making it a model for the random behavior of percentages or proportions, as in the lower right plot of Figure 2.1 and Figure 2.2. From the figures, we can see by modifying the two parameters, a very flexible range of shapes are obtained from the distribution.
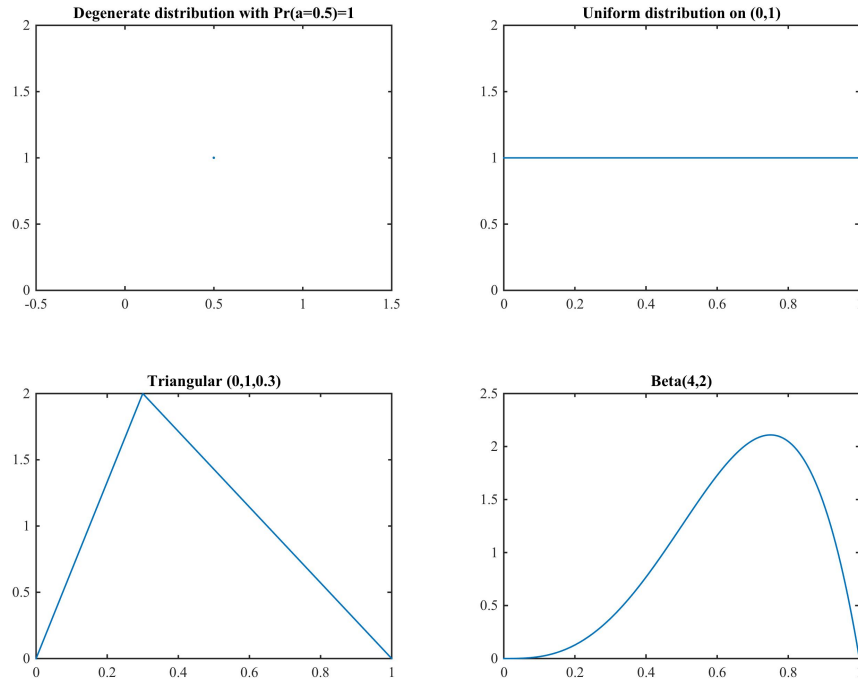
Figure 2.1. Examples of PDF plots for candidate prior distribution for a
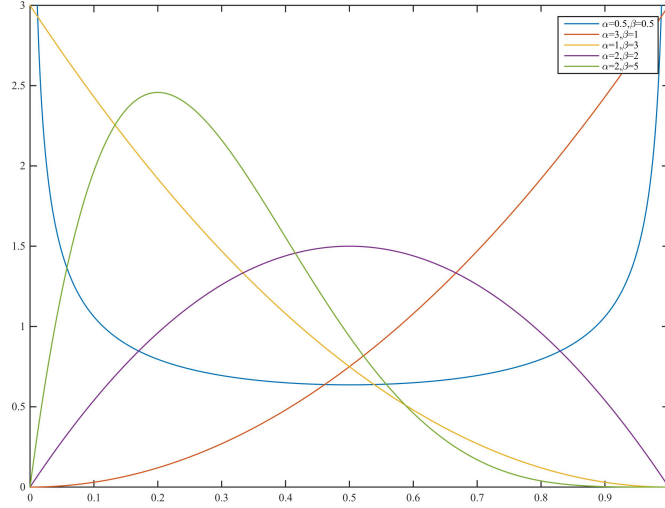
Figure 2.2. PDF Plots of Beta distribution with different α and β

2.3.2   Optimization at the First Stage

The first stage ALT planning is figured out by optimization based on the priors. Denote the objective function for the first stage as C_0, and denote

$$C_{01}(D) = \mathrm{AVar}_D(\log t_p(0)|M_1)$$

$$C_{02}(D) = \mathrm{AVar}_D(\log t_p(0)|M_2) \tag{2.16}$$

If the prior for a is degenerate or discrete,

$$C_0 = \Pr(a = c) \sum_{\text{all } c} \left\{ c \int C_{01}w(\boldsymbol{\theta}_1)\mathrm{d}\boldsymbol{\theta}_1 + (1 - c) \int C_{02}w(\boldsymbol{\theta}_2)\mathrm{d}\boldsymbol{\theta}_2 \right\} \tag{2.17}$$

where c is the possible values a can be. If the prior for a is continuous,

$$C_0 = \int_0^1 p(a) \left[a \int C_{01}w(\boldsymbol{\theta}_1)\mathrm{d}\boldsymbol{\theta}_1 + (1 - a) \int C_{02}w(\boldsymbol{\theta}_2)\mathrm{d}\boldsymbol{\theta}_2 \right] \mathrm{d}a \tag{2.18}$$

Denote

$$\boldsymbol{c}_1 = (1, \Phi_{SEV}^{-1}(p), 0)'$$

$$c_2 = (1, \Phi_N^{-1}(p), 0)'$$

Hence by model in equation 2.12 and 2.13,

$$C_{01}(D) = c_1' \mathrm{AVar}_D(\boldsymbol{\theta}_1) c_1$$

$$C_{02}(D) = c_2' \mathrm{AVar}_D(\boldsymbol{\theta}_2) c_2 \tag{2.19}$$

To calculate the objective function for the first stage, we need to get the asymptotic variance of $\boldsymbol{\theta}$ first, Clyde, Müller and Parmigiani [5], and Berger [3] showed that when the sample size is relatively large (in ALT's the sample size is usually greater than 30), a multivariate normal distribution gives reasonable approximation for variances. Then denote $\boldsymbol{\theta} = (\beta_0, \sigma, \beta_0)'$, for both models, the following steps are applicable.

$$\mathrm{AVar}_D(\boldsymbol{\theta}) \approx [S^{-1} + I_{\boldsymbol{\theta}}(D)]^{-1} \tag{2.20}$$

where as in Chapter 1, S^{-1} is the prior precision matrix of $\boldsymbol{\theta}$, $I_{\boldsymbol{\theta}}(D)$ is the Fisher information for plan D. For $I_{\boldsymbol{\theta}}(D)$, the equation 1.11 still holds. Similar to equation 1.12 the scaled Fisher Information after standardization for each observation is,

$$\mathcal{F}_i = \begin{bmatrix} f_{11}(\zeta_i) & f_{12}(\zeta_i) & f_{11}(\zeta_i)\xi_i \\ f_{12}(\zeta_i) & f_{22}(\zeta_i) & f_{12}(\zeta_i)\xi_i \\ f_{11}(\zeta_i)\xi_i & f_{12}(\zeta_i)\xi_i & f_{11}(\zeta_i)\xi_i^2 \end{bmatrix} \tag{2.21}$$

We assume Type I right censoring, thus

$$\zeta_i = \frac{\log(t_c) - \beta_0 - \beta_1 \xi_i}{\sigma} \tag{2.22}$$

The elements in $\mathcal{F}_i$ is calculated as,

$$f_{11}(\zeta_i) = \Psi_0(\zeta_i) + \eta(\zeta_0)$$

$$f_{12}(\zeta_i) = \Psi_1(\zeta_i) + \zeta_0\eta(\zeta_0)$$

$$f_{22}(\zeta_i) = \Psi_2(\zeta_i) + \zeta_0^2\eta(\zeta_0) \tag{2.23}$$

where $\zeta_0 = (-\beta_0 - \beta_1\xi_i)/\sigma$,

$$\Psi_i(\zeta_i) = \int_{\zeta_0}^{\zeta_i} [1 + xH(x)]^i \, H(x)^{2-i} g(x)\mathrm{d}x \quad i = 0, 1, 2$$

$$H(x) = \frac{g'(x)}{g(x)} + \frac{g(x)}{1 - G(x)}$$

$$\eta(x) = \frac{g(x)}{[1 - G(x)]\, G(x)}$$

where $g(x)$ and $G(x)$ is the PDF and CDF of assumed log-lifetime distribution, in the study, it is either SEV or standard normal distribution. Now the objective function for first stage can be derived and optimized.

A brief introduction of constraints in ALT planning has been given in section 1.1. Because the objective of first stage test is mainly updating the prior information, both time and cost should be constrained. A small proportion of total test units should be used in the first stage. In addition, it is not appropriate for the first stage test to take a long time because this may cause less time in the second stage, even leading to some biases in the results. However, the setting of constrains is a tradeoff. When more resource is used in the first stage, more precise posterior will be obtained, helping us updating the prior more efficiently, on the other hand, the remaining resource in the second stage reduces. The setting of censoring time, which is the total test time of the first stage, denoted as t_{c1}, should ensure that at higher levels of acceleration, a relatively small proportion of test units are expected to be right censored. For example, at the highest level of acceleration, if there is only 10% of test units failed when the test is over, the planning is not considered to be a good one.

2.3.3   Results

After planning the ALT in stage 1, an AL test is expected to be conducted, experiment data should be collected, denoted as $\boldsymbol{t}$. The size of $\boldsymbol{t}$ is n_1, number of test units in the first stage. For each element in $\boldsymbol{t}$, there are two properties:

- Corresponding ξ. Each data element must be labeled with ξ, the standardized variable of the accelerating factor.

- Censored or not. If a test unit failed in test time, t_{c1}, the observation is exact, otherwise it is right censored. Denoted as δ_i, for each of the observation,

$$\delta_i = \begin{cases} 1 & \text{if the observation is not censored} \\ 0 & \text{if the observation is right censored} \end{cases} \tag{2.24}$$

2.4   Second Stage Experiment

Once the data are obtained in the first stage, we recommend planning and conducting the second stage experiment as soon as possible, ensuring the condition of test units remain the same and racing for more testing time.

2.4.1   Updating the Priors

We need to update two parts of the priors to get posteriors: acceleration model parameters and distribution specification parameter a.

For the acceleration model parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, very similar processes to update the priors are conducted. In this section $\boldsymbol{\theta}_1$, corresponding to model M_1, is used as an example to illustrate the steps. As in section 2.3.1, the joint distribution for $\boldsymbol{\theta}_1$ is denoted as $w(\boldsymbol{\theta}_1)$, the posterior can be expressed as,

$$w(\boldsymbol{\theta}_1|\boldsymbol{t}) = \frac{L(\boldsymbol{t}|\boldsymbol{\theta}_1)w(\boldsymbol{\theta}_1)}{\int L(\boldsymbol{t}|\boldsymbol{\theta}_1)w(\boldsymbol{\theta}_1)\mathrm{d}\boldsymbol{\theta}_1} \tag{2.25}$$

where $L(\boldsymbol{t}|\boldsymbol{\theta}_1)$ is the likelihood of $\boldsymbol{t}$ evaluated at $\boldsymbol{\theta}_1$, the likelihood is calculated as follows,

$$L(\boldsymbol{t}|\boldsymbol{\theta}_1) = \prod_{i=1}^{n} \left[\frac{1}{\sigma_1 \xi_i} \phi_{SEV} \left(\frac{\log(t_i) - \beta_{01} - \beta_{11}\xi_i}{\sigma_1} \right) \right]^{\delta_i} \left[1 - \Phi_{SEV} \left(\frac{\log(t_i) - \beta_{01} - \beta_{11}\xi_i}{\sigma_1} \right) \right]^{1-\delta_i} \tag{2.26}$$

In addition to obtaining the posterior distribution of $\boldsymbol{\theta}_1$, we also need to specify the posterior variance of $\boldsymbol{\theta}_1$ to compute the variance of $\log t_p(0)$. Similar to equation 2.20,

$$\mathrm{Var}_{\boldsymbol{\theta}_1|\boldsymbol{t},D}(\boldsymbol{\theta}_1) = [S_1^{-1} + \hat{I}_{\boldsymbol{\theta}_1}(D)]^{-1} \tag{2.27}$$

where $\hat{I}_{\boldsymbol{\theta}_1}(D)$ is the Fisher information matrix evaluated at $\hat{\boldsymbol{\theta}}_1$, maximum likelihood estimator of $\boldsymbol{\theta}_1$ estimated by the data in the first stage. By the relationship,

$$\text{Var}_{\boldsymbol{\theta}_1|\boldsymbol{t},D}\left[\log t_p(0)\Big|M_1\right] = \boldsymbol{c}_1'\text{Var}_{\boldsymbol{\theta}_1|\boldsymbol{t},D}(\boldsymbol{\theta}_1)\boldsymbol{c}_1 \tag{2.28}$$

where $\boldsymbol{c}_1 = (1, \Phi_{SEV}^{-1}(p), 0)'$, the posterior of variance of $\log t_p(0)$ assuming that M_1 is true is obtained. By replacing $\boldsymbol{\theta}_1$ with $\boldsymbol{\theta}_2$, and change the SEV distribution to standard normal distribution, the posteriors for M_1 can be derived by conducting the same previous steps.

Besides updating the information regarding acceleration model parameters, the distribution specification parameter a is also updated. Similar to the equation 2.25,

$$p(a|\boldsymbol{t}) = \frac{L(\boldsymbol{t}|a)p(a)}{\int L(\boldsymbol{t}|a)p(a)\mathrm{d}a} \tag{2.29}$$

where $L(\boldsymbol{t}|a)$ is the likelihood of $\boldsymbol{t}$ evaluated at a, which can be calculated as,

$$L(\boldsymbol{t}|a) = \prod_{i=1}^{n}\left\{a\left[\frac{1}{\sigma_1\xi_i}\phi_{SEV}\left(\frac{\log(t_i)-\beta_{01}-\beta_{11}\xi_i}{\sigma_1}\right)\right]^{\delta_i}\left[1-\Phi_{SEV}\left(\frac{\log(t_i)-\beta_{01}-\beta_{11}\xi_i}{\sigma_1}\right)\right]^{1-\delta_i} + \right.$$
$$\left.(1-a)\left[\frac{1}{\sigma_2\xi_i}\phi_N\left(\frac{\log(t_i)-\beta_{02}-\beta_{12}\xi_i}{\sigma_2}\right)\right]^{\delta_i}\left[1-\Phi_N\left(\frac{\log(t_i)-\beta_{02}-\beta_{12}\xi_i}{\sigma_2}\right)\right]^{1-\delta_i}\right\} \tag{2.30}$$

where we can use the the MLE of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ to replace the model parameters to get the likelihood of the data, that is,

$$L(\boldsymbol{t}|a) = \prod_{i=1}^{n}\left\{a\left[\frac{1}{\hat{\sigma}_1\xi_i}\phi_{SEV}\left(\frac{\log(t_i)-\hat{\beta}_{01}-\hat{\beta}_{11}\xi_i}{\hat{\sigma}_1}\right)\right]^{\delta_i}\left[1-\Phi_{SEV}\left(\frac{\log(t_i)-\hat{\beta}_{01}-\hat{\beta}_{11}\xi_i}{\hat{\sigma}_1}\right)\right]^{1-\delta_i} + \right.$$
$$\left.(1-a)\left[\frac{1}{\hat{\sigma}_2\xi_i}\phi_N\left(\frac{\log(t_i)-\hat{\beta}_{02}-\hat{\beta}_{12}\xi_i}{\hat{\sigma}_2}\right)\right]^{\delta_i}\left[1-\Phi_N\left(\frac{\log(t_i)-\hat{\beta}_{02}-\hat{\beta}_{12}\xi_i}{\hat{\sigma}_2}\right)\right]^{1-\delta_i}\right\} \tag{2.31}$$

Thus we can get the posterior distribution of a.

2.4.2 Explanations of the Objective Function and Modification

Recall the objective function 2.14, what weighed the expected asymptotic variance are not probability densities, but constant probabilities. In fact it is more appropriate to take double expectation respect to both $\boldsymbol{\theta}$'s and a. However, when conducting the numerical experiment, the integration for the expectation is very computationally intensive, thus what is included the objective function is $\Pr(M_1|\boldsymbol{t})$ and $\Pr(M_2|\boldsymbol{t})$. By taking expectation on the posterior a,

$$\Pr(M_1|\boldsymbol{t}) = \int a p(a|\boldsymbol{t}) \mathrm{d}a$$

$$\Pr(M_2|\boldsymbol{t}) = 1 - \Pr(M_1|\boldsymbol{t}) \tag{2.32}$$

At this stage we can compute all the elements in function 2.14, so that optimization for the second stage is ready to be conducted.

2.5 Assumptions and Settings for Optimization

Usually if a linear model is assumed as an accelerated lifetime model, it is reasonable to only set two levels of acceleration to estimate the desired lifetime, on the other hand, if the model is suspected to be quadratic or higher-order polynomial, more levels need to be added to the test. In this study we are assuming a linear model, thus we only need to optimize on four variables, ξ_L, π_L and ξ_H, π_H corresponding to low level and high level of the test respectively.

Chapter 3

Numerical Example

3.1 Software Utilized

The numerical experiments were conducted mainly in MATLAB 2014b, Statistics and Machine Learning Toolbox was employed to assist the computation. Some numerical ALT model maximum likelihood estimation was conducted in R version 3.1.3.

3.2 Adhesive Bond Test Background

The example originates from Meeker and Hahn [16]. Engineerings were investigating the reliability of adhesive bond, aiming at estimate the 0.1 quantile of the failure-time distribution at the use temperature 50°C. The 0.1 quantile of the lifetime was expected to be more than 10 years. Hence an accelerating lifetime test was needed to demonstrate this, and there were 300 test units for testing. The failure process of the adhesive bond was believed to be simple chemical degradation process, which can be modeled by Arrhenius relationship. The acceleration model was assumed is to be linear and after standardization it satisfies,

$$\mu(\xi_i) = 9.35715 - 4.64533\xi_i \tag{3.1}$$

In the model, engineers believe the parameters are near $\beta_0 = 9.35715, \beta_1 = -4.64533$, in the example[16] the lifetime distribution was assumed to be Weibull

and the shape parameter was thought to be near $k = 1.667$, which made $\sigma = 1/k = 0.6$. To conduct our two stage experiment, we need another model based on lognormal lifetime distribution, thus denote,

$$\boldsymbol{\theta}_1 = (9.35715, 0.6, -4.64533)'$$

$$\boldsymbol{\theta}_2 = (9.35715, 0.76953, -4.64533)' \tag{3.2}$$

and

$$\boldsymbol{c}_1 = (1, \Phi_{SEV}^{-1}(p), 0)'$$

$$\boldsymbol{c}_2 = (1, \Phi_N^{-1}(p), 0)'$$

We set $\sigma = 0.75953$ for the lognormal model because the standard deviation σ of log lifetime is σ_1 when it is lognormal model and $\pi\sigma_2/\sqrt{6}$ when it is a Weibull one [25]. Thus in summary,

$$M_1 : \log t_p(\xi) = 9.35715 - 4.64533\xi + 0.6\Phi_{SEV}^{-1}(p)$$

$$M_2 : \log t_p(\xi) = 9.35715 - 4.64533\xi + 0.76953\Phi_N^{-1}(p) \tag{3.3}$$

3.3 Case 1: Prior of a is Degenerately Distributed

The simplest case is that the prior of a is degenerately distributed, and we don't imply any information in $\boldsymbol{\beta}$'s. Hence for $\boldsymbol{\beta}$'s, it is a likelihood method. We assume the right censoring time is 183 days for each of the two stages. And there are 100 test units for stage 1 and 200 for stage 2, denoted $N_1 = 100, N_2 = 200$

3.3.1 First Stage Planning

Because prior of a is assumed to be degenerately distributed. The objective function in the first stage is simplified as,

$$C_0 = \mathrm{AVar}_D(\log t_p(0)|M_1)p_1 + \mathrm{AVar}_D(\log t_p(0)|M_1)(1 - p_1) \tag{3.4}$$

where p_1 is the value of a which has probability 1. Firstly for each of model M_1 and M_2, we have the numerically optimal plans respectively as the following tables show.

Condition	Level(Standardized)	Level (TEMPC)	Proportion	Number
i	ξ_i	°C	π_i	n_i
Use	0	50	0	0
Low	0.6821	94.67	0.706	71
High	1	120	0.294	29

$\mathrm{AVar}_D(\boldsymbol{\theta}_1|M_1)=0.4339$

Table 3.1. Optimal design for M_1 in stage 1 when $N_1 = 100$

Condition	Level(Standardized)	Level (TEMPC)	Proportion	Number
i	ξ_i	°C	π_i	n_i
Use	0	50	0	0
Low	0.6496	92.26	0.660	66
High	1	120	0.340	34

$\mathrm{AVar}_D(\boldsymbol{\theta}_1|M_2)=0.4144$

Table 3.2. Optimal design for M_2 in stage 1 when $N_1 = 100$

We can find that the optimal plans for M_1 and M_2 are slightly different. The log-normal model (M_2) suggests lower temperature on the low level and less allocation at it, while the Weibull model (M_1) suggests the opposite.

By varying the value of p_1 from 0 to 1, different optimal plans are expected to be obtained, Figure 3.1 and 3.2 show the plot of optimal plan to minimize C_0 in formula 3.4 vs. p_1. Because the ξ for high level is always 1 in the planning, only ξ's and π's of low level are demonstrated in the figures.

From the two figures, we can find that when the prior prefers Weibull distribution (when p_1 increases), the optimal planning suggests higher temperature and higher allocation for the low level. However, the patterns of increase for ξ_L and π_L are different. In Figure 3.1, ξ_L vs. p_1 is an increasing convex function, when p_1
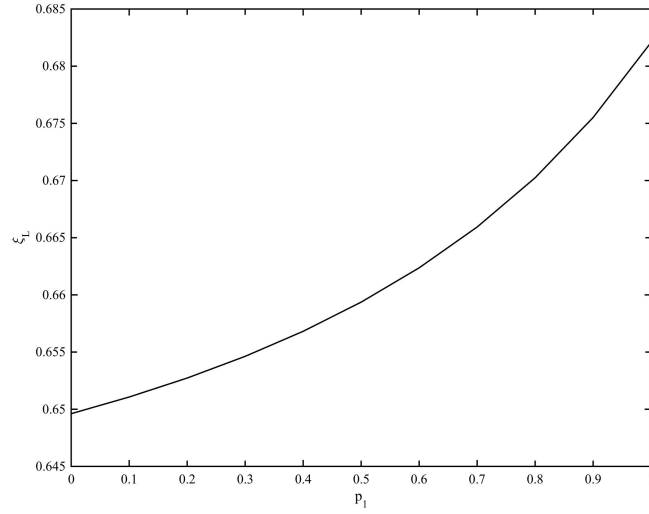
Figure 3.1. Plot of optimal ξ_L vs. p_1
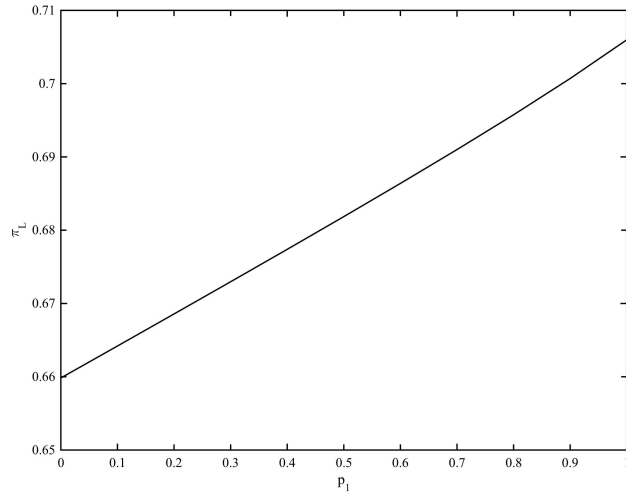
Figure 3.2. Plot of optimal π_L vs. p_1

increases linearly, the increase in the optimal ξ_L will become faster. In the other Figure 3.2, the function is also convex, however, it is very near to a linear function.

Figure 3.3 shows how the weighed asymptotic variance of $\log t_p(0)$ varies when p_1 is between 0 and 1. The function is an increasing concave function.
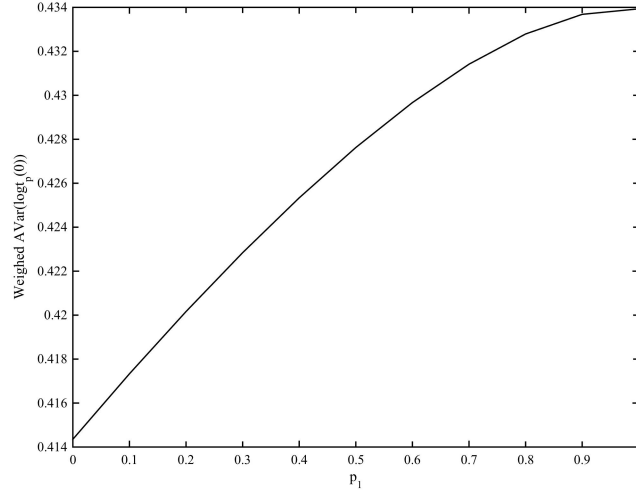
Figure 3.3. Plot of optimal weighed $\mathrm{AVar}_D(\log t_p(0))$ vs. p_1

In the first stage, we set $a = 0.5$ as an example, that is to say, the priors can be written as follows,

$$\Pr(M_1) = 0.5$$

$$\Pr(M_2) = 0.5 \tag{3.5}$$

Under the assumption, the optimal design for the first stage is as follows,

Condition	Level(Standardized)	Level (TEMPC)	Proportion	Number
i	ξ_i	°C	π_i	n_i
Use	0	50	0	0
Low	0.6594	92.99	0.6819	68
High	1	120	0.3181	32
$C_0(D)$=0.4276				

Table 3.3. Optimal design at stage 1 (Case 1)

3.3.2 Simulation of Data from the First Stage

After the planning of the first stage, a test is expected to be conducted. To illustrate how the methods work in practice, we simulate some data as the test

data. The steps of simulation can be found in Chapter 19 in Meeker and Escobar's book [14], and the settings of the simulation were as follows,

Variable	Setting
β_0	9.35715
σ	0.76953
β_1	-4.64533
Distribution	lognormal
Number of simulations	100

Table 3.4. Setting of simulation of data for stage 1 (Case 1)

The simulation settings indicate that the true model is M_2. However, at this stage, whether the model is correct is unknown. And for $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, we need to use maximum likelihood estimation to move to the next step. The maximum likelihood estimates from the experimental data are as follows,

	Parameter	MLE	Standard Error	95% CI
$\boldsymbol{\theta}_1$	β_{01}	9.03340	0.796	(7.47,10.59)
	σ_1	0.503767	0.078	(0.372,0.683)
	β_{11}	-4.07940	0.832	(-5.71,-2.45)
$\boldsymbol{\theta}_2$	β_{02}	8.65502	0.548	(-7.58,9.73)
	σ_1	0.637065	0.083	(0.484,0.813)
	β_{22}	-3.91318	0.599	(-5.09,-2.74)

Table 3.5. MLEs from the simulated data (Case 1)

3.3.3 Update the Prior for Distribution Specification

Because we assume a degenerate prior for $\boldsymbol{\theta}$'s, hence the posterior is also degenerate and is the same as the priors. For the prior in equation 3.5, we can update as follows,

$$\Pr(M_1|\boldsymbol{t}) = \frac{\Pr(M_1)L(\boldsymbol{t}|M_1)}{\Pr(M_1)L(\boldsymbol{t}|M_1) + \Pr(M_2)P(\boldsymbol{t}|M_2)} \tag{3.6}$$

where $L(\boldsymbol{t}|M_1)$ and $L(\boldsymbol{t}|M_2)$ is the likelihood of $\boldsymbol{t}$ evaluated at M_1 and M_2 respectively. From the data we have simulated, we can compute

$$L(\boldsymbol{t}|M_1) = \exp(-54.152)$$

$$L(t|M_2) = \exp(-48.620) \tag{3.7}$$

The likelihood for M_2 is apparently larger than M_1, which results in,

$$\Pr(M_1|\boldsymbol{t}) = 0$$

$$\Pr(M_2|\boldsymbol{t}) = 1 \tag{3.8}$$

3.3.4 Second Stage Planning and Comments

Because $\Pr(M_1|\boldsymbol{t}) = 0$, we are minimizing only over the lognormal model based on MLEs, which will result the plan shown in Table 3.6 in the second stage, noting that$N_2 = 200$. The plan is reasonable because in the simulation, we simulate based on the fact that M_2 is true (lognormal is the true lifetime distribution).

Condition	Level(Standardized)	Level (TEMPC)	Proportion	Number
i	ξ_i	°C	π_i	n_i
Use	0	50	0	0
Low	0.650	92.29	0.6599	66
High	1	120	0.3401	34
$C(D){=}0.1336$				

Table 3.6. Final optimal design (Case 1)

3.3.5 Comparison and Comments

We compared the plan with the traditional statistically optimum plan in Meeker and Escobar's book [14] (labeled as "Traditional" in the following content) and Zhang and Meeker [33]'s Bayesian method which only uses the priors, which is labeled as "One Stage Bayesian". By computing the mean square errors of $\log t_p(0)$ from the simulated data, the comparison is as the table shows.

Method	MSE[$\log t_p(0)$]
Traditional	0.1522
One Stage Bayesian	0.1522
Two Stage Bayesian	0.1321

Table 3.7. Comparison with other methods (Case 2)

In this case, the first two methods provide the same plan. Because our method helps eliminating the wrong model, its result is better than those of the traditional planning and one stage bayesian method. This is the simplest case of the two stage design, although it would be atypical to specify degenerate distributions over priors. In Bayesian studies, it is more common to specify some non-degenerate information on the priors, making the posterior also non-degenerate. However, the example shows that the two-stage design helps eliminating the wrong model, at least, the wrong distribution of the model. A more complex and practical example will be introduced in the next section.

3.4 Case 2: Priors are Non-degenerately Distributed

In this case, the priors of the acceleration model parameters, $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and the distribution specification parameter a are non-degenerately distributed.

3.4.1 Setting of Priors

For the priors of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, uniform distributions are set as priors, and in addition, we assume the three parameters in each of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are independent with each other. And for a, a uniform distribution on $[0, 1]$ is assumed to be the prior. Hence the priors are as showed in the following table

	Variable	Distribution
	β_{01}	Uniform$(8, 11)$
$\boldsymbol{\theta}_1$	σ_1	Uniform$(0.55, 0.65)$
	β_{11}	Uniform$(-5, -3)$
	β_{02}	Uniform$(8, 11)$
$\boldsymbol{\theta}_2$	σ_2	Uniform$(0.72, 0.82)$
	β_{12}	Uniform$(-5, -3)$
	a	Uniform$[0, 1]$

Table 3.8. Priors for stage 1 (Case 2)

The variance-covariance matrices for $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are

$$S_1 = S_2 = \begin{bmatrix} 3/4 & 0 & 0 \\ 0 & 1/1200 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \tag{3.9}$$

Thus the precision matrices are

$$S_1^{-1} = S_2^{-1} = \begin{bmatrix} 4/3 & 0 & 0 \\ 0 & 1200 & 0 \\ 0 & 0 & 3 \end{bmatrix} \tag{3.10}$$

3.4.2 First Stage Planning

In this case the first stage objective function is as function 2.18. In the function, we can employ numerical integration to get $\int C_{01}w(\boldsymbol{\theta}_1)d\boldsymbol{\theta}_1$ and $\int C_{02}w(\boldsymbol{\theta}_2)d\boldsymbol{\theta}_2$, and by conducting the optimization, we get the plan as follows,

Condition	Level(Standardized)	Level (TEMPC)	Proportion	Number
i	ξ_i	°C	π_i	n_i
Use	0	50	0	0
Low	0.66	93	0.68	68
High	1	120	0.32	32
$C_0(D)$=0.1183				

Table 3.9. Optimal design at stage 1(Case 2)

We can find that the plan is very near to the plan when $p_1 = 0.5$ in Case 1. The reason is that the prior of a is uniform distribution, which also indicates no preference between the models.

3.4.3 Simulation of Data from the first stage

By using the same simulation settings as in Table 3.4, however we need to conduct another simulation, even though the settings are completely the same. From the simulated data, the MLEs are as follows,

	Parameter	MLE	Standard Error	95% CI
	β_{01}	10.5917	1.163	(8.31,12.87)
$\boldsymbol{\theta}_1$	σ_1	0.510603	0.079	(0.376,0.693)
	β_{11}	-5.75797	1.193	(-8.10,-3.42)
	β_{02}	9.61111	0.689	(8.26,10.96)
$\boldsymbol{\theta}_2$	σ_1	0.625	0.087	(0.476,0.821)
	β_{22}	-5.04995	0.727	(-6.48,-3.62)

Table 3.10. MLEs from the simulated data (Case 2)

3.4.4 Second Stage Planning

The priors of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are updated as in function 2.25, and updating for a results in,

$$\Pr(M_1|\boldsymbol{t}) = E\left[p(a|\boldsymbol{t})\right] = 0.3622$$

$$\Pr(M_2|\boldsymbol{t}) = 1 - \Pr(M_1|\boldsymbol{t}) = 0.6378 \tag{3.11}$$

Thus we have all the elements in the objective function, the final plan is as the following table shows,

Condition	Level(Standardized)	Level (TEMPC)	Proportion	Number
i	ξ_i	°C	π_i	n_i
Use	0	50	0	0
Low	0.7212	97.61	0.6547	65
High	1	120	0.3453	35
$C(D)$=0.0713				

Table 3.11. Final optimal design (Case 2)

3.4.5 Comparison and Comments

For case 2, a similar comparison is conducted as in section 3.3.5. The result is as follows,

Method	MSE$[\log t_p(0)]$
Traditional	0.2055
One Stage Bayesian	0.1942
Two Stage Bayesian	0.1707

Table 3.12. Comparison with other methods (Case 2)

From the result we can find that the bayesian methods are better than the traditional planning because priors increase the amount of information, helping the planning and making the mean square error smaller. Compared with the one stage bayesian, our two stage method also shows an obvious smaller mean square error by reducing the posterior probability of the wrong model.

In the case, the two stage design also tells us the lognormal model is more likely to be true, as in equation 3.11. Thus it is also efficient to specify the distribution by Bayesian methods in this case. Compared to case 1, the objective weighed expected asymptotic variance is much lower, because we have priors for all the parameters,

including acceleration model parameters and distribution specification parameter a, making the than that in case 1, resulting a much smaller objective.

Chapter **4**

Conclusions

4.1 Summary of Modeling and Experiment Results

We have presented a Bayesian two-stage ALT planning methods under linear models with one experimental variable. From the experiment results, the two-stage design is very efficient to specify the lifetime distribution, avoiding distribution misspecification. In addition, we find that by adding some prior information on the parameters and combine it with data, the posterior will result in much smaller asymptotic variance of desired quantile of lifetime at use condition.

The two-stage design costs more time to conduct tests under the same experimental constraints. The second stage can be repeated to make use of all of the priors and posteriors before the stage, thus there are continuous modifications on the planning.

4.2 Discussion and Future Research

The method described in this thesis can be extended to ALT planning with more experimental variables, and also with more complex models, such as quadratic and higher order polynomial. However, even under the simple assumption, the numerical integration and optimization takes considerable computation, especially when the priors are non-degenerate, and this is the reason that we made modification

in section 2.4.2. Without the modification, the modeling is more precise and better results are expected, but better numerical optimization methods to solve the problem will be required.

In addition, in both of the two steps, there is a large sample approximation made, other methods, such as simulation methods, are also of interest to be investigated as they allow to estimate the small sample behavior of the two stage plans.

Bibliography

[1] E. K. Al-Hussaini and A. H. Abdel-Hamid, "Bayesian estimation of the parameters, reliability and hazard rate functions of mixtures under accelerated life tests," *Communications in Statistics-Simulation and Computation*, vol. 33, no. 4, pp. 963–982, 2004.

[2] D. Bai, M. Kim, and S. Lee, "Optimum simple step-stress accelerated life tests with censoring," *Reliability, IEEE Transactions on*, vol. 38, no. 5, pp. 528–532, 1989.

[3] J. O. Berger, *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.

[4] K. Chaloner and K. Larntz, "Bayesian design for accelerated life testing," *Journal of Statistical Planning and Inference*, vol. 33, no. 2, pp. 245–259, 1992.

[5] M. Clyde, P. Müller, and G. Parmigiani, *Optimal design for heart defibrillators*. Springer, 1995.

[6] E. Del Castillo, *Lecture notes: IE 532 Reliability Engineering*. Dept. of Industrial and Manufacturing Engineering, Penn State University, 2014.

[7] A. Erkanli and R. Soyer, "Simulation-based designs for accelerated life tests," *Journal of Statistical Planning and Inference*, vol. 90, no. 2, pp. 335–348, 2000.

[8] L. A. Escobar and W. Q. Meeker, "Algorithm as 218: elements of the fisher information matrix for the smallest extreme value distribution and censored data," *Applied Statistics*, pp. 80–86, 1986.

[9] ——, "Algorithm as 292: Fisher information matrix for the extreme value, normal and logistic distributions and censored data," *Applied Statistics*, pp. 533–540, 1994.

[10] ——, "Planning accelerated life tests with two or more experimental factors," *Technometrics*, vol. 37, no. 4, pp. 411–427, 1995.

[11] T. J. Kielpinski and W. Nelson, "Optimum censored accelerated life tests for normal and lognormal life distributions," *Reliability, IEEE Transactions on*, vol. 24, no. 5, pp. 310–320, 1975.

[12] W. Q. Meeker, "A comparison of accelerated life test plans for weibull and lognormal distributions and type i censoring," *Technometrics*, vol. 26, no. 2, pp. 157–171, 1984.

[13] W. Q. Meeker, L. Escobar *et al.*, "Pitfalls of accelerated testing," *Reliability, IEEE Transactions on*, vol. 47, no. 2, pp. 114–118, 1998.

[14] W. Q. Meeker and L. A. Escobar, *Statistical methods for reliability data*. John Wiley & Sons, 1998.

[15] W. Q. Meeker, L. A. Escobar, and C. J. Lu, "Accelerated degradation tests: modeling and analysis," *Technometrics*, vol. 40, no. 2, pp. 89–99, 1998.

[16] W. Q. Meeker and G. J. Hahn, *How to plan an accelerated life test: Some practical guidelines*. ASQC Statistics Division, 1985.

[17] W. Q. Meeker, G. Sarakakis, A. Gerokostopoulos, N. Doganaksoy, D. B. Hall, G. J. Hahn, D. L. Kugler, J. Lawless, R. Pan, D. C. Trindade *et al.*, "More pitfalls of accelerated tests," *Journal of Quality Technology*, vol. 45, no. 3, p. 213, 2013.

[18] E. M. Monroe, R. Pan, C. M. Anderson-Cook, D. C. Montgomery, and C. M. Borror, "Sensitivity analysis of optimal designs for accelerated life testing," *Journal of quality technology*, vol. 42, no. 2, p. 121, 2010.

[19] P. Müller, "Simulation based optimal design," *Handbook of Statistics*, vol. 25, pp. 509–518, 2005.

[20] P. Müller and G. Parmigiani, "Optimal design via curve fitting of monte carlo experiments," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1322–1330, 1995.

[21] W. Nelson, "Accelerated life testing-step-stress models and data analyses," *Reliability, IEEE Transactions on*, vol. 29, no. 2, pp. 103–108, 1980.

[22] W. Nelson and T. J. Kielpinski, "Theory for optimum censored accelerated life tests for normal and lognormal life distributions," *Technometrics*, vol. 18, no. 1, pp. 105–114, 1976.

[23] W. Nelson and W. Q. Meeker, "Theory for optimum accelerated censored life tests for weibull and extreme value distributions," *Technometrics*, vol. 20, no. 2, pp. 171–177, 1978.

[24] W. B. Nelson, "A bibliography of accelerated test plans," *IEEE Transactions on Reliability*, vol. 54, no. 2, pp. 194–197, 2005.

[25] F. G. Pascual and G. Montepiedra, "Lognormal and weibull accelerated life test plans under distribution misspecification," *Reliability, IEEE Transactions on*, vol. 54, no. 1, pp. 43–52, 2005.

[26] N. Polson, "A bayesian perspective on the design of accelerated life tests," *Advances in Reliability*, pp. 321–330, 1993.

[27] Y. Shi and W. Q. Meeker, "Bayesian methods for accelerated destructive degradation test planning," *Reliability, IEEE Transactions on*, vol. 61, no. 1, pp. 245–253, 2012.

[28] L.-C. Tang, G. Yang, and M. Xie, "Planning of step-stress accelerated degradation test," in *Reliability and Maintainability, 2004 Annual Symposium-RAMS*. IEEE, 2004, pp. 287–292.

[29] W. Weibull, "A statistical distribution function of wide applicability," *Journal of Applied Mechanics*, vol. 103, pp. 293–299, 1952.

[30] T. Yang and R. Pan, "A novel approach to optimal accelerated life test planning with interval censoring," *Reliability, IEEE Transactions on*, vol. 62, no. 2, pp. 527–536, 2013.

[31] I. Yu, C.-L. Chang *et al.*, "Applying bayesian model averaging for quantile estimation in accelerated life tests," *Reliability, IEEE Transactions on*, vol. 61, no. 1, pp. 74–83, 2012.

[32] Y. Zhang and W. Q. Meeker, "Bayesian life test planning for the weibull distribution with given shape parameter," *Metrika*, vol. 61, no. 3, pp. 237–249, 2005.

[33] ——, "Bayesian methods for planning accelerated life tests," *Technometrics*, vol. 48, no. 1, pp. 49–60, 2006.