Optimal Experimental Design

for Accelerated Life Testing and Design Evaluation

by

Tao Yang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved September 2013 by the
Graduate Supervisory Committee:

Rong Pan, Chair
Douglas Montgomery
Connie Borror
Steve Rigdon

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

Nowadays product reliability becomes the top concern of the manufacturers and customers always prefer the products with good performances under long period. In order to estimate the lifetime of the product, accelerated life testing (ALT) is introduced because most of the products can last years even decades. Much research has been done in the ALT area and optimal design for ALT is a major topic. This dissertation consists of three main studies. First, a methodology of finding optimal design for ALT with right censoring and interval censoring have been developed and it employs the proportional hazard (PH) model and generalized linear model (GLM) to simplify the computational process. A sensitivity study is also given to show the effects brought by parameters to the designs. Second, an extended version of I-optimal design for ALT is discussed and then a dual-objective design criterion is defined and showed with several examples. Also in order to evaluate different candidate designs, several graphical tools are developed. Finally, when there are more than one models available, different model checking designs are discussed.

DEDICATION

I would dedicate this dissertation to:

- to both my parents Hongbin Zhao and Yonggen Yang who always support me and encourage me,
- and to all the people who helped me.

ACKNOWLEDGEMENTS

I would like to thank my advisor: Dr. Rong Pan. I still remembered the time back to 2009 fall semester I was looking for a advisor who can guide me for Ph.D program and Dr. Pan gave me an chance to work with him even I didn't have any previous experience in reliability area. During the past 4 years, Dr. Pan provided me countless valuable suggestions and help on my research and always encouraged me when I met problems in research and life. I will be forever grateful for your time and efforts.

To my committee members, Dr. Douglas Montgomery, you are one of my favorite professors because of your teaching style and the magic that you can always explain complicated things in easy terms; Dr. Borror gave me some suggestions on my research, especially her interesting questions for my proposal brought me some insight for potential extension of my current research topic; Dr. Rigdon also provided me a lot of help and he flied thousands of miles to Tempe to attend my oral test of my proposal. All my committee have provided me the opportunity to do outstanding research and potential future directions.

Additionally, I would like to thank National Science Foundation (NSF) and all my committee chair and members who financially supported my research and life at Arizona State University.

To my hang out buddies Joe Xu, and Bozhi Li, thanks for the dinners, movie nights, soccer games, and road trips; To my friend and senior Xiaotian Zhuang, for your encouragement and bringing me a valuable internship opportunity in Chicago; To Yang Pang, thanks for numerous dinner treats and interesting discussions of political sciences; To all my friends at industrial engineering department, Peng Yue, Mengqi Hu, Ning Wang, Zhouyang Zhou, Lei Liu, Yinbo He, Ting Wu, Can Cui, Min Zhang, Zixin Wang, Congze Su, Jie Chu, Zhen Zhao, Shuluo Ning, Na Zou, Chao Li and more. Also to my

former roommates Moxuan Zhu, Dona Montes, Candice Eisenfeld and neighbours Jun Zhang and Youyuan Jiang! ASU will not be the same without you guys!

And to anyone else I forgot to mention here! I will forever appreciate your support!

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Overview

Reliability is defined as the ability of a system or component to perform its required functions under stated conditions for a specified period of time. An easier definition is: reliability is the quality over time (Condra (1993)). Nowadays, with the rapid development of advanced technologies, new products with interesting features have been delivered to the market frequently and manufacturers have to respond to the market quickly which means more competitions to them. On the other hand, consumers become more concerned about the reliability of the products because they want to own products with long lifetime. Therefore, the manufacturers could not take the risk of losing customers due to the poor quality and reliability of their products and have to invest money and manpower to continuously improve the reliability of products. The popular quality and reliability improvement methods include Advanced Product Quality Planning (APQP), New Product Introduction (NPI), Failure Modes and Effects Analysis (FMEA), Statistical Process Control (SPC), and Root Cause Analysis (RCA), Failure Reporting, Analysis and Correction Action System (FRACAS), and Reliability, Availability, Maintainability, and Safety (RAMS). In all these methods, in order to find the root causes of the failures and estimate the lifetime of a product, testing plays an important role. Since a newly developed product can last years even decades, the naive method of testing sample units at the normal use condition is not practical due to the cost and time limitation. Therefore, Accelerated Life Testing (ALT) is widely used to shorten the testing time of new products. Basically, ALTs increase certain test stresses, like temperature, humidity, voltage, and cycles to accelerate the failure of test units. However, even for ALTs the testing time could be too long for observing all failures so censoring has to be

introduced into tests.

There are several different censoring methods and the most popular ones are right censoring and interval censoring. Right censoring means that the testing time is prefixed so when the time reaches the prefixed testing time, the test is terminated. The advantage of right censoring is that for those failure units which fail before the censoring time, exact failure times can be collected. However, in many situations exact failure times are difficult to obtain due to measurement availability, cost, and other constraints. Therefore, interval censoring is introduced which can be easily implemented. The general idea of interval censoring is given certain number of test units, the total testing time is divided into several time intervals and at the end of each interval, the number of failures will be counted. However, despite the ease of the ALT application, interval censoring also brings 'side effects' because the exact failure times of test units are unknown and the only information obtained is in which intervals the units fail. This feature will cause the reduction of the accuracy of the estimation of product's lifetime.

In Montgomery (2012), the statistical design refers to the process of planning the experiment so that appropriate data will be collected and analyzed by statistical methods, resulting in valid and objective conclusions. Typical experimental designs include standard factorial design and optimal design. A factorial experiment is an experiment whose test plan consists of two or more factors, each with discrete possible values or levels, and whose experimental units take on all possible combinations of these levels across all such factors. The fractional factorial design, as indicated by its name, includes partial runs from the full factorial designs. Some other designs related to the factorial designs are composite designs and Plackett-Burman designs. Optimal designs are a class of experimental designs that are optimal with respect to some statistical criteria. The most popular criteria are alphabetic optimalities including A-optimality, D-optimality, T-optimality, G-optimality, and I-optimality. There are some other alphabetic optimality

criteria that can be derived or updated from the basic optimality criteria.

Developing optimal design for ALTs is an interesting topic. Maximum likelihood estimation (MLE) is usually used to estimate the ALT model parameters with different censoring situations and this approach involves the calculation of first and second partial derivatives of a log-likelihood function. With the censoring situation, it will be more difficult to solve but the Generalized Linear Model (GLM) could be applied here to simplify the problem. The details of the GLM method will be given in later chapters.

Two other important topics are how to evaluate the different designs and how to distinguish different models by a design. The first problem could be solved by applying some evaluation criteria and also some graphic tools. The second problem could be handled by some specific design like T-optimal design which will be discussed later.

Overall, this dissertation will explain the methods of finding an optimal design based on different optimal criteria with different censoring scenarios. The cases of right censoring and interval censoring will be discussed. After obtaining the optimal designs, how to compare and evaluate these designs with other alternative designs would be explained. Finally, details of applying optimal criteria such as D_s-optimality to distinguish competing ALT models will be discussed.

1.2 Motivation

Accelerated life testing has become an important part of modern reliability and how to obtain good test plans/designs and evaluate them under ALT is a practical topic. In this dissertation, these problems will be addressed.

First, this research is motivated by the work from Monroe et al. (2011). Most experimental designs in the books have been done with the normal use condition settings and DOE books seldom mentioned the ALT. Therefore, it would be an interesting topic to

consider the ALT and DOE together. By reviewing the literature, a lot of work has been done in experimental design for ALTs. The earliest one can be traced back to Chernoff (1962). Most of the work related with ALT and DOE (optimal design) was done in late 80s and early 90s. Work from Monroe et al. (2011) is different compared with the previous research because they implemented the Generalized Linear Model (GLM) method to find the optimal design under ALT. As we know that the big difference between the normal testing and ALT is the design region of ALT may not cover the use condition which requires experimenters to extrapolate the data from the design region back to the use region. As mentioned before, MLE will be used to estimate the parameters of the ALT model but it may cause a lot of computation loads. GLM is a better way to simplify the computational process. Therefore it would be an interesting topic to develop and derive some further work based on Monroe et al. (2011)'s research.

Second, right censoring may be the most popular method due to its simplicity. However, in reality it is difficult to observe the exact failure times for the test units and instead, interval censoring is actually more often to be applied in the real ALT. Much research work has been done with right censoring cases but interval censoring cases are seldom discussed before. Therefore, finding optimal design for ALTs with interval censoring is appealing. In addition, another interesting topic is the use condition region of ALTs. For the simplicity, the use condition is usually considered as a single, fixed point in a lot of previous studies. However, in the real world the use condition is seldom limited as a single point/condition but a region. So finding optimal designs for ALTs with a pre-defined use condition region would be a valuable technique which is more applicable and close to real cases.

The third motivation comes from a question: when there is more than one design available, how to choose the designs? Once some candidate designs have been developed based on different optimal criteria, some of them are easy to be implemented, some have

good statistically features, so it is necessary to evaluate them and choose the best ones. There are several ways to evaluate designs and the traditional methods are usually single value criteria like optimal efficiency values. In addition, graphical evaluation methods are also popular due to their visualization. They are more straightforward methods and usually more comprehensive compared with single value criteria. Variance Dispersion Graphs (VDG) and Fraction of Design Space (FDS) plots are popular graphical tools for design evaluation. However, under ALT not many graphical tools have been introduced and it is necessary to derive and develop some graphical methods based on existing approaches. Sometimes there are several competing models for failure mechanism under ALT and how to distinguish different models by a design is also an appealing topic.

1.3 Research Objectives

- Find optimal designs by applying the Generalized Linear Model (GLM) approach under ALT with right censoring;
- Find optimal designs by applying the Generalized Linear Model (GLM) approach under ALT with interval censoring;
- Use or develop different single value criteria and graphical tools to evaluate different designs;
- Find designs to distinguish different failure models under ALT scenario.

1.4 Dissertation Organization

The remainder of this dissertation is structured as follows. A literature review is summarized in Chapter 2. The review includes previous and current research in accelerated life testing (ALT), optimal designs, generalized linear model(GLM), design evaluation, and model separation. Chapter 3 introduces the theoretical work associated

with the GLM-based approach toward planning ALT optimal designs with right censoring and interval censoring. Chapter 4 develops some single value criteria and graphical tools to evaluate different ALT designs. Model checking and discrimination are discussed in Chapter 5. Finally, Chapter 6 is a conclusion for the previous chapters and also gives an insight for the future potential research.

Chapter 2

REVIEW OF RELATED LITERATURE

In this chapter, a literature review is provided with the major topics in this dissertation. The review has five parts: accelerated life testing, experimental designs, generalized linear model, design evaluation, and model discrimination.

2.1 Accelerated Life Testing

Accelerated life testing (ALT) is the process of testing a product by subjecting it to conditions (stress, strain, temperatures, etc.) in excess of its normal service parameters in an effort to uncover faults and potential modes of failure in a short amount of time (Wikipedia). The reason for applying ALT is the good quality and reliability of new products which have long lifetimes under normal use conditions. With the consideration of limited time and cost, in order to observe some failures during the test, the levels of stress factors have to be increased or decreased out of their normal use value region. Two models are needed for designing and analyzing ALTs, one is the acceleration model and the other one is the lifetime distribution.

Acceleration models are fit to the data to describe the effect that the variables have on the failure-causing process for the testing over the range of accelerating stress variables. The typical accelerating stress variables include temperature, voltage, humidity, pressure, or any other stress related to the failure mechanism. The purpose is to test at high stress levels to speed up the failure processes and to establish the functional relationship between failure parameters and stresses. The acceleration models can be classified into two categories, physical acceleration models and empirical acceleration models. Physical accelerations models are developed under well-understood failure mechanisms based on physical/chemical theory. On the other hand, empirical models are

an alternative if it is impossible to develop a model based on physical/chemical theory. In previous research, physical acceleration models are more common and some typical models will be introduced next.

The first acceleration model mentioned here is the Arrhenius model. It describes the effects that temperature has on the rate of a simple chemical reaction. The Arrhenius model has been widely used and has the form

$$R(temp) = \gamma_0 \exp(\frac{-E_a}{k_B \times tempK}) \tag{2.1}$$

in Meeker & Escobar (1998) where R is the reaction rate and $tempK$ is temperature in the absolute Kelvin scale, k_B is Boltzmann's constant, and E_a is the activation energy. The parameters E_a and γ_0 are product or material characteristics. One thing that should be noticed is the Arrhenius relationship does not apply to all temperature acceleration situations.

The Eyring model is another temperature related acceleration model. It is introduced by Glasstone et al. (1941) and it represents the amount of energy needed to move an electron to the state where the processes of chemical reaction or diffusion or migration can take place (Tobias & Trindade (2011)). The reaction rate of the Eyring model can be written as

$$R(temp) = \gamma_0 \times A(.) \times exp(\frac{-E_a}{k_B \times tempK}) \tag{2.2}$$

where again E_a and γ_0 are product or material characteristics. The term $A(.)$ is a function which could represent the effects of temperature or other stress factors. It depends on how the modeler defines the problem. Sometimes $A(.)$ could be a inverse power law function and is used to describe the voltage effects.

The last acceleration model discussed here is the inverse power law (IPL) model.

The inverse power relationship can be described as

$$AF(volt) = AF(volt, volt_U, \beta_1) = \frac{T(volt_U)}{T(volt)} = (\frac{volt}{volt_U})^{-\beta_1} \tag{2.3}$$

where $T(volt)$ and $T(volt_U)$ are the failure times at test condition and use condition respectively. β_1 is a parameter and usually has $\beta_1 < 0$. $AF(volt)$ is the accelerated factor. The inverse power law relationship describes the effects of voltage to the failure and can be part of the Eyring model we mentioned above.

Another important part of ALT modeling is the lifetime distribution model. The lifetime distribution is used to describe the failure rate at specific time point. The popular ones are the exponential distribution (which is a simple version of Weibull distribution), Weibull distribution, and lognormal distribution. Details of these three popular distributions are given below and they are originally described in Meeker & Escobar (1998).

The exponential distribution is denoted as $T \sim EXP(\theta, \gamma)$ and it has cumulative density function (cdf), probability density function (pdf), and hazard function (hf) written as

$$\begin{aligned} F(t;\theta,\gamma) &= 1 - exp(-\frac{t-\gamma}{\theta}), t > r \\ f(t;\theta,\gamma) &= \frac{1}{\theta} exp(-\frac{t-\gamma}{\theta}), t > r \\ h(t;\theta,\gamma) &= \frac{1}{\theta}, t > r \end{aligned}$$

where $\theta > 0$ is a scale parameter and γ is a location parameter and $t > \gamma$.

The Weibull distribution has cdf

$$F(t;\eta,\beta) = 1 - exp[-(\frac{t}{\eta})^{\beta}]$$

where $t > 0$. However, for the Weibull distribution it is usually reparameterized by the smallest extreme value (SEV) distribution. If T has a Weibull distribution then

$Y = log(T) \sim SEV(\mu, \sigma)$, where $\sigma = 1/\beta$ is the scale parameter and $\mu = \log(\eta)$ is the location parameter. Now the Weibull distribution can be indicated as $T \sim WEIB(\mu, \sigma)$ and the cdf, the pdf, and the hf can be written as

$$F(t;\mu,\sigma) = \Phi_{sev}[\frac{log(t)-\mu}{\sigma}],$$
$$f(t;\mu,\sigma) = \frac{1}{\sigma t}\phi_{sev}[\frac{log(t)-\mu}{\sigma}] = \frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1}exp[(\frac{t}{\eta})^{\beta}],$$
$$h(t;\mu,\sigma) = \frac{1}{\sigma exp(\mu)}[\frac{t}{exp(\mu)}]^{1/\sigma-1} = \frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1}.$$

When variable T belongs to a lognormal distribution, it is indicated as $T \sim LOGNOR(\mu, \sigma)$, and $Y = log(T) \sim NOR(\mu, \sigma)$. It has cdf and pdf written as

$$F(t;\mu,\sigma) = \Phi_{nor}[\frac{log(t)-\mu}{\sigma}],$$
$$f(t;\mu,\sigma) = \frac{1}{\sigma t}\phi_{nor}[\frac{log(t)-\mu}{\sigma}]$$

where ϕ_{nor} and Φ_{nor} are pdf and cdf for the standardized normal distribution.

With these fundamental models and distributions, much research has been done in the ALT area. Nelson (2005a) gave a comprehensive review of the research related to ALT and it classified those papers into several categories: the basis of ALT, the history of test plans, general AT models including ALT and ADT, failure time and stress model, development of test plans, optimal criteria, censoring, constraints, and other topics. The second part of the review from Nelson (2005b) listed references of all the important papers mentioned in the previous one.

2.2 Experimental designs

Statistical design history

The original experimental design is from the field of agriculture which is used to find out the influential factors to the yield. Fisher is the pioneer who first introduced the

ideas of factorial design and analysis of variance and the details of his research work can be found in (Fisher, 1958, 1966).

For any experimental design, the factors can be classified as controllable factors and uncontrollable factors. Usually, the controllable factors which experimenters are interested in can be changed in certain ranges and have major influences on the response. On the other hand, uncontrollable factors can be further classified into two subgroups. One is called unknown and uncontrollable factor and another is called known and uncontrollable factor. The former type of factor is handled by randomization so that their effects on the response can be canceled out by the random order of the runs. The latter type can be handled by a technique called blocking. The idea of blocking is to run each test unit with uncontrollable variable in different blocks which can eliminate the uninterested effects in the experiment. Therefore, in order to deal with the nuisance factor, randomized complete block design (RCDB) is introduced and early research include Kempthorne (1952) and Wilk (1955).

As mentioned before, the factorial design is first introduced by Fisher and Yates also made important contributions to modify the idea. The factorial design is usually denoted as n^k where n is the level of each factor and k is the number of factors. The advantage of factorial designs has several aspects. The biggest one is it has less runs compared with one-factor-at-a-time designs. In addition, a factorial design is needed to detect the interaction effects in the experiments. Most of the time, an early stage model with main effects and their interactions will be built first and in order to give some protection against effects brought by quadratic terms, some additional center points could be added to the designs. By adding some more axial runs to the design, the classical central composite design is formed. A standard full factorial design is useful but sometimes the experiment has put some constraints including time and limited resource. Therefore, running a full factorial design may be expensive and not available sometimes

and fractional factorial designs are developed in order to use less runs to estimate more terms in the model. Confounding and resolution are two important terms for fractional factorial designs and the Plackett & Burman (1946) designs and Box & Behnken (1960) designs are two famous examples.

In early 1980s, product quality improvement became an interesting topics to the western world and the research and practical work from (Taguchi, 1987, 1991) had great impact on manufacturing. Basically, a Taguchi design is a crossed array design incorporating controllable variables and uncontrollable variables. However, a Taguchi design also has some defects including not addressing the interactions between the controllable variables and not having a clear statistical model format. Box et al. (1988) and Hunter (1989) discussed some details about the features and defects about the Taguchi design.

After late 1980s and 1990s, the concepts of experimental design were not limited in the agriculture and manufacturing industry, but widely spread to other fields including business, healthcare, psychology, and others. For example, Kuhfeld et al. (1994) applied the D-optimal design and discussed orthogonal, nonorthogonal design algorithms for marketing research. Viney et al. (2005) investigated the design properties of discrete choice experiments and developed three experimental design approaches in health care. Kirk (2012) explained how to apply experimental design to analyze problems in behavioral science. In the reliability field, Chapter 13 of Wu & Hamada (2009) explains how to use designed experiments for improving reliability.

Optimal design criteria

The standard factorial design and classical designs like central composite design and Box-Behnken design are widely used. However, in some circumstances it is hard or impossible to apply them. First, the design region of experiments may not be a regular

space. For example, the two factor experimental design region is not a square and the three factor design region is not a cube. Second, the response surface model may not be a hierarchical model which means some terms are not included in the model. For example, a full second order model for a two factor experiment should have 6 terms but experimenters exclude one of the squared term due to certain pre-knowledge and only discuss a model with 5 terms. Third, due to the limitations of time and money, experiments can only have small number of runs which the classical designs may not be applicable. In order to tackle the design problems with constraints mentioned above, Kiefer (1959, 1961) developed the concepts of optimal design. Optimal designs are certain types of designs with the respect of different optimal criteria. Several popular optimal criteria will be introduced below.

The most widely used optimality is D-optimality. A D-optimal design focuses on good model parameter estimations and trying to maximize the determinant of moment matrix as shown below (Myers et al. (2009)),

$$\max |M(\xi)| = \max |\frac{X(\xi)'X(\xi)}{N}| \tag{2.4}$$

where $X(\xi)$ is the design matrix of an experimental design ξ and N is the sample size.

The purpose of A-optimality is to make a design that minimizes the variances of model coefficients by minimizing the trace of inverse information matrix which can be written as

$$\min Trace(X(\xi)'X(\xi))^{-1} \tag{2.5}$$

where $X(\xi)$ is the design matrix of an experimental design ξ.

The G-optimality criterion is used to minimize the maximum value of prediction variance over the design region and can be written as

$$\min \max \frac{Nx'(X(\xi)'X(\xi))^{-1}x}{\sigma^2} \tag{2.6}$$

where N is sample size and x is a point from design region and $X(\xi)$ is the design matrix of an experimental design ξ.

The I-optimality criterion is used to minimize the average prediction variance over the whole design region and can be written as

$$\min \frac{\int_{\Omega} x'(X(\xi)'X(\xi))^{-1}xdx}{S_{\Omega}} \tag{2.7}$$

where x is a point from design region and $X(\xi)$ is the design matrix of an experimental design ξ and Ω is the design region and S is the area of design region.

The four optimality criteria mentioned above are called alphabetic optimality criteria because they are all named after a letter. There are also some other alphabetic optimality criteria that are variations of the basic criteria we discussed above. For example, V-optimality is a special case of I-optimality which only consider a subset of points in the design region and D_s-optimality only checks the subset of model coefficients. Several other derived alphabetic optimality criteria will be discussed in detail in later chapters. Since under many situations, the optimal designs are not standard designs and the design structures are unknown, experimenters have to find those designs with the help of computers.

A widely used algorithm for optimal design is called point exchange. Basically, the experimenters start to select a set of points from a candidate set as an initial design. Next, select an point out of the current design and replace with a point from the pool which can improve the optimality criterion. Then repeat the last step until there is no improvement found. The advantage of the point exchange method is it is easy to understand and apply but on the other hand, point exchange cannot always guarantee to find an optimal design and may end up at a local optimality case. In addition, the experimenter may not know the pool of candidate points before they start to build the designs and this could be another issue for point exchange method. The point exchange

algorithm was first introduced by Mitchell (1974).

Another approach, similar to the point exchange, is the coordinate exchange. It also starts with an initial design. Then in each iteration, the algorithm tries to find a single coordinate in each point in the design and replace it with a new coordinate which can improve the optimality of the criterion. After many iterations, the design will converge when any changes made to it will not improve the optimality. Compared with the point exchange algorithm, the coordinate exchange algorithm does not need a candidate pool and usually have better convergence speed. This algorithm was first described in Meyer & Nachtsheim (1995).

The definitions of alphabetical optimal criteria and two typical search algorithms have been discussed. Now how to compare the available designs and find out the best ones is an interesting topic. Design efficiency is a scaled value and changes from 0 to 1. D-efficiency and G-efficiency are the two common ones which are defined as

$$D_{eff} = \{\frac{M(\xi)}{M(\xi^*)}\}^{1/p} \tag{2.8}$$

$$G_{eff} = \frac{\bar{d}(\xi^*)}{\bar{d}(\xi)} = \frac{p}{\bar{d}(\xi)} \tag{2.9}$$

where M is the information matrix and p is the number of parameters in the model. $\bar{d}()$ is a maximal function over the design region. These two efficiency formulas are adopted from Atkinson et al. (2007). Design efficiencies are easy to calculate and used to compare the current designs with the optimal designs. However, since the design efficiency is a single value criterion, it sometimes may not well present the properties of designs over the whole design region. Therefore, some other graphical tools are developed which are straightforward to compare different designs. Two common tools are variance dispersion graph (VDG) and fraction of design space (FDS) which will be discussed in detail in later chapters.

2.3 Generalized Linear Model

The most important statistical model for data analysis is the linear regression model and it is usually described as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

where y is the response variable, x_i represents the regressor variable, and β_i represents the unknown parameter. This model is called a linear model because the expectation of response variable y is a linear function of unknown parameters of β_i. In addition, the error terms ε belong to a normal distribution and they are independent to each other.

A nonlinear model is an extension of the linear model which usually explains some complicated mathematical relationship between the response and the regressors. The nonlinear model is usually written as

$$y = f(x, \beta) + \varepsilon$$

where the response y is not a linear function of the parameters β. The nonlinear model also assumes that the responses belong to a normal distribution and they are independent to each other.

The linear and nonlinear models have been briefly discussed above and both of them have an important assumption that the response variables belong to the normal distribution. However, a lot of times the response variables may not be continuous variable. For example, the binary variables could be the response which can only take values of 0 or 1. This is a common situation when the regression model is trying to classify things into two groups. In addition, response variable may only take values between 0 and 1 which is not uncommon that the regression model is trying to find out the probability of an event happening. Therefore, the generalized linear model (GLM) was

developed to fit the model with the response variable belongs to a more general distribution. Nelder & Wedderburn (1972) first explained theory of GLM and more details were discussed in McCullagh & Nelder (1989). Myers et al. (2002) gave a clear description about the GLM and the details are inherited here. As mentioned above, the responses in GLM usually belong to a general distribution which is called the exponential family and some commonly-used distributions include the normal, binomial, Poisson, and exponential distributions all belong to this family. Now in order to build up a relationship between the response variable of a linear model and the true response variable which belongs to an exponential family, a link function is used to fulfill the task. This relationship can be written as

$$g(\mu_i) = g[E(y_i)] = \eta_i = x_i'\beta \tag{2.10}$$

where $g()$ is the link function and x is a vector of regressors and β is a vector of unknown parameters. For link functions, there are a lot of options available and the criterion of choosing a link function for GLM is based on the response data, modeler's preference, and other requirements. The typical link functions include identity link, logistic link, reciprocal link. They are given by

$$\eta_i = \mu_i$$
$$\eta_i = ln(\frac{P}{1-P})$$
$$\eta_i = \frac{1}{\mu_i}.$$

Some other link functions will be discussed in later chapters. Now the input to a link function is usually called linear predictor and it is defined as

$$\eta_i = x'\beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

In conclusion, GLMs model the relationship of linear regressors with the response variables belong to a general distribution. A typical GLM has three components,

- a distribution from the exponential family,
- a linear predictor,
- a link function.

Since the introduction by Nelder & Wedderburn (1972), GLM has been applied to different research topics and survival analysis is one of them. Aitkin & Clayton (1980) fitted the model with censored data by using exponential, Weibull, and extreme distributions from GLM. Later, Holford (1980) and Bennett (1983) discussed the log-linear models and log-logistic models for survival data respectively. Hurley (1985) used GLM to do the survival analysis with two different types of failures. Later, the application of GLM has extended to the area of accelerated life testing (ALT). Barbosa & Louzada-Neto (1994) considered a Weibull failure distribution and a log link function to build the relationship between stress variables and response variable and finally estimated the mean lifetime of the units under use stress. Wang & Kececioglu (2000) presented an algorithm to obtain the MLE of model parameters from testing data with Weibull log-linear model. Lee & Pan (2010) analyzed the step-stress ALT with exponential distribution and applied GLM approach to estimate the parameters.

2.4 Design evaluation

Because of the different optimality criteria, experimenters may need to compare the different available designs and choose the best one out of them. There are a lot of ways to compare designs based on computational algorithms, efficiency, and other standards. Box & Draper (1959) discussed two problems which are how to fit a polynomial model close to the true model by least square method over the design region and how to easily detect the gap between the polynomial model built by the experimenters and true model. Two examples were given and the optimal design derived were two level

fractional design with center points. Cook & Nachtsheim (1980) compared existing algorithms for computer generating D-optimal design and modified the Fedorov algorithm which reduced the computational time. Bohning (1986) introduced an vertex-exchange method to find exact D-optimal design and compared the convergence speed with other existing iteration methods. Fedorov (1989) discussed the case where the number of runs had restrictions and proposed an heuristic algorithm based on the point exchange method. Nguyen & Miller (1992) reviewed and compared the exchange algorithms and suggested a modified Fedorov exchange algorithm.

As described in previous sections, the design efficiencies are great criteria to compare different candidate designs. D-efficiency, G-efficiency, and other derived efficiencies have been widely used. However, there are some disadvantages of design efficiencies since they are single value criteria and sometimes may not fully represent the situations over the whole design region and they are hard to explain. For example, for G-efficiency its definition is to minimize the maximal prediction variance in the design region. However, suppose there are two designs, one has smaller maximal prediction variance but the rest area in design region have similar prediction variance values to the maximum. The other design has a slightly larger maximal prediction variance in design region compared with the first design but the rest area all have relative smaller prediction variance values. Therefore, the average prediction variance of the first design will be larger than the second design and most of the times the second design will be preferable even the first one is preferred by G-efficiency. This situation actually led to the development of I-optimality criterion which is used to represent the average prediction variance over the whole design region. In the purpose of better illustration, some graphical tools have been developed to compare different designs. Two examples are the variance dispersion graph (VDG) and fraction of design space (FDS) graph.

The original idea of variance dispersion graph (VDG) was presented in

Giovannitti-Jensen & Myers (1989) and VDG is a performance comparison tool of competing designs on a fixed design space. The first step of generating a VDG is finding all the available design points with same distance r from the center of design region and calculating the maximum, minimum, and average prediction variances. The second step is to increase the radius of r and record the maximum, minimum, and average prediction variances. The last step is to plot the data points which should have three different curves. One exception is when the design is rotatable and the maximum, minimum, and average prediction will be the same and three curves will converge into one. The VDG allows experimenters to compare designs and check their overall variance properties over the whole design region. Since VDG is a quite popular graphical tool, much research has been done to apply the tool and modify it for different purposes. Piepel et al. (1993) applied VDG to irregular design regions. Borkowski (1995) applied the VDG to the mixture designs. Trinca & Gilmour (1998, 1999) compared blocked response surface designs by using VDG and derived a difference variance dispersion graph (DVDG) to help in the choice of response surface designs. Goldfarb, Borror, et al. (2004) discussed the design factors in mixture experiment and some other independent variables which will change without affections of other independent variables or mixture variables. In order to evaluate this mixture-process designs, three dimensional VDG were built. Liang, Anderson-Cook, Robinson, & Myers (2006) considered the prediction properties of split-plot designs and used three dimensional VDG to demonstrate comprehensive study with the traditional single value criteria like D-, V-criteria.

Fraction of design space (FDS) plots were first developed by Zahran et al. (2003) and it gave the percentage of design space where the scaled prediction variance is less than or equal to a pre-defined value. The desirable curve of FDS plots are flat which means most of the design space have small prediction variances. Goldfarb, Anderson-Cook, et al. (2004) developed the FDS plots for mixture experiments and used them to help identify

the important factors which affect the scaled prediction variance (SPV). Ozol-Godfrey et al. (2005) used the FDS plots to compare design robustness by SPV values. FDS plots can show the SPV distributions of potential designs over the same design space which will give experimenters better understanding of features of these designs. J. Li et al. (2009) compared three designs including central composite design (CCD), small composite design (SCD), and minimum run resolution (MinRes) V design with different axial length choices. The fraction of design (FDS) plots were used to represent the prediction variances of these three designs and help the experimenters to make selection. Liang, Anderson-Cook, & Robinson (2006) applied FDS plots to the split-plot designs and analyzed the whole design over the whole design region. In addition, the curves of FDS plots are also sliced into different subregion groups in order to satisfy certain design constraints. The authors demonstrated their method by using FDS plots with variance ratio to compare two CCD split-plot examples. Ozol-Godfrey et al. (2008) raised the issue that the designs under GLM are given by the experimenters which may be misspecified. Therefore, applying FDS plots to several designs with different sets of parameters can help experimenters evaluate the design robustness. Jang & Anderson-Cook (2011) built a ridge regression model for mixture experiment in order to handle the multicollinearity.

2.5 Model discrimination

For optimal design of accelerated life testing (ALT), the response variable and stress factors usually have a nonlinear relationship which is quite different from the optimal design of linear relationship in terms that: the model parameters need to be pre-specified based on previous experience. Therefore, the optimal designs for ALT are model dependent and the quality of the model settings will decide the quality of generated designs. However, many potential models are available for the experimenters to choose from and how the designs will be affected by model selections and how to distinguish

different models by certain designs become interesting topics to explore. For the first question, it can be viewed as the robust design problem and the experimenters prefer the designs which would be less affected by the unknown parameters values. The second question is a model selection problem and the experimenters want to know if some of the terms in the model are necessary to explain the data. For example, a two stress factor model could have only two main effect terms, or it could also have the interaction term, or even more that it can include all the second order terms. Therefore, experiments are generated to distinguish different models.

Atkinson & Fedorov (1975a,b) described the experimental designs for discriminating between rival regression models, especially for T-optimal designs. Hill (1978) reviewed the several methods for model discrimination designs including Box and Hill's procedure, Fedorov's procedure, and Atkinson's procedure. Jones et al. (2007) proposed several criteria including SA, MPD, EPD for gauging the capability of a design for model discrimination. Agboto et al. (2010) discussed the existing methods like T-optimality and several new criteria to construct optimal two-level model discriminating designs for screening experiments. However, the previous two papers only discussed the linear model situation. Dette & Titoff (2009) derived several new properties of optimal designs with respect to the T-optimality and also demonstrated that in nested linear models the number of support points of T-optimal designs is usually too small to estimate all parameters in the full model. Biedermann et al. (2011) developed optimal design theory for additive partially nonlinear regression models and generalised their results to parameter robust optimality criteria, called Bayesian and standardised maximin D-optimality. DeLeon & Atkinson (1991) used numerical methods to find non-sequential optimal designs, which can be used both for the construction of designs and for checking the optimality of proposed designs. Chaloner (1984) discussed optimal Bayesian experimental designs for estimation and prediction in linear models. An optimal Bayesian

design for the nonlinear problem with a single explanatory variable is considered in Chaloner (1993) and a literature review on Bayesian experimental design is given by Chaloner & Verdinelli (1995). DuMouchel & Jones (1994) modified D-optimality with a Bayesian paradigm and handled the dependent problem on assumed models. Another interesting work from Waterhouse et al. (2006) considered a problem when two rival GLMs for a binomial response and compared designs based on four different optimal criteria.

Chapter 3

Optimal Accelerated Life Test Planning With Interval Censoring

Accelerated life testing (ALT) is widely used in industry to obtain the lifetime estimate of a product which is expected to last years or even decades. It is important to find an effective experimental design of ALT with the consideration of certain optimality criteria. In this paper, we discuss a new approach to designing ALT test plans when readout data (i.e., interval censoring) are collected. In this chapter, we utilize the proportional hazard (PH) model for a failure time distribution, and formulate a generalized linear model (GLM) for censored data. The optimal design is obtained such that the prediction variance of the expected product lifetime at the products' use condition is minimized.

3.1 Introduction

Background and Motivation

New products will be tested in order to check their functionality, safety, and reliability before they are released to the market. However, most products nowadays can last years, even decades, so the regular life tests are not the options for them. Accelerated life testing (ALT) is introduced in order to shorten the product's lifetime. In ALT certain stress variables, such as temperature, humidity, and voltage, are set to the higher than the normal stress level so that experimenters can expect more failures from test units within a limited testing period. Readout data are very common in ALTs, as due to measurement availability, cost, and other constraints, experimenters are often not able to monitor test units continuously; instead, they inspect the test units periodically according to an inspection plan. Therefore, the product failure times are interval censored; i.e., we know a

test unit failed in a certain time interval, but we do not know the exact failure time. The ALT with interval censoring is easy to implement, but interval censoring also brings the side effect of losing the exact failure time information into experimental planning and data analysis. This situation may reduce the accuracy of the product lifetime prediction, and it also complicates the way of finding the expected information matrix for constructing optimal test plans.

The conventional way to develop an optimal ALT plan is to formulate the likelihood function, and then derive the expected information matrix for test planning. When interval censoring is applied, the likelihood function of an interval censored failure time is the probability of observing a failure between the lower and upper bounds of the time interval. This causes the total likelihood function to be overly complicated and its expected information matrix is hard to obtain. In this chapter, we propose to use the proportional hazard (PH) model for failure time distribution with stress variables (experimental factors). In general, the PH model is semi-parametric, thus it is more flexible compared to the traditional failure time regression model. More specifically, when the baseline hazard function is defined, this model will become many different distributions, such as Weibull distribution and gamma distribution, which are commonly used in reliability data analysis. Given the proportional hazard rate property of the PH model, we can formulate the total likelihood function of censored data by a generalized linear model (GLM) formulation. The GLM was developed by Nelder & Wedderburn (1972) and its details can be found in McCullagh & Nelder (1989) and Myers et al. (2002). Typically there are three components in a GLM formulation:

- a distribution from the exponential family,
- a linear predictor, and
- a link function.

With the help of GLM, the information matrix can be easily derived, and then it can assist experimenters to find the optimal test plan and to assess the properties of the optimal plan quickly.

To plan an optimal ALT, some statistical criteria need to be achieved. For example, we may want to maximize the determinant of expected information matrix (as the overall variance of model parameter estimates will be minimized), then it is a D-optimal design. Other optimal criteria include A-optimality, G-optimality, V-optimality, I-optimality, etc. (Myers et al. (2009)). In this chapter, we are interested in the asymptotic variance of the expected product lifetime at the use condition. A test plan that minimizes the product lifetime prediction variance at its use condition is called the U_c-optimal design in Monroe et al. (2010). As the U_c-optimality minimizes prediction variance, this optimal design gives us more confidence on predicting a product's lifetime at its use condition.

In conclusion, compared with the conventional ALT optimal plans which are generated under the accelerated failure time (AFT) model, our PH model assumes that stress variables accelerate the failure rate. However, it is well known that, through some re-parameterizations, these two models are equivalent for the exponential or Weibull failure time distribution. The benefit of applying the PH model is that we can formulate the problem by a generalized linear model (GLM), so the statistical inference of regression coefficients can be easily obtained. By using this formulation, we are ready to plan an ALT with more than one stress variable and with interaction of stress variables, which can not be conveniently done by the conventional method.

Previous work

The literature on optimal ALT designs is vast. In this section we mention some important developements that had been done in this field, particularly for a test with censoring, but it is not intended to be a comprehensive review of ALT. For that, readers

may refer to Nelson (2005a,b). Nelson & Kielpinski (1976) discussed the theory of optimal ALT plans for estimating a simple linear relationship between a stress factor and product lifetime that has a normal or lognormal distribution, when censoring is considered. Nelson & Meeker (1978) and Meeker & Nelson (1975) applied the maximum likelihood theory for designing optimal ALT plans, while assuming the product lifetime has a Weibull or smallest extreme value distribution. Tang et al. (1999) presented the method for finding optimal ALT plans for censored two-parameter exponential distribution. However, these papers only discussed ALTs with a single stress factor. Escobar & Meeker (1995) introduced the ALT planning with censoring for two stress factors. Park & Yum (1996) and Sitter & Torsney (1995) presented the optimal test plans with two stress factors. In addition, Xu & Fei (2007) and C. Li & Fard (2007) consider step-stress accelerated life testing plans for two stress variables. To deal with interval censoring, Finkelstein & Wolfe (1985) and Finkelstein (1986) developed a PH regression model for analyzing interval censored data. However, we found very few papers directly addressed the optimal test plans with the consideration of readout data. Islam & Ahmad (1994) and Yum & Choi (1989) developed the optimal designs under periodic inspection and type-I censoring. Seo & Yum (1991) extended the optimal ALT plans under intermittent inspection scheme. Tse et al. (2008) gave the optimal ALT designs under interval censoring with random removals. In Ng et al. (2004), the ways of finding optimal test plans for different optimal criteria by deriving the expected Fisher information matrix were discussed in detail. Other extensions of ALT plans can be found in Pascual (2007), Pascual (2008) and Liu & Qiu (2011). They derive the the plans for ALT with independent competing risks.

All previous work in optimal ALT designs follows the direct approach of deriving the expected information matrix from the likelihood function of failure times. Recently, Monroe et al. (2011) and Pan & Yang (2011) proposed to use the generalized linear model

(GLM) approach to approximate the information matrix for finding optimal ALT plans. Their methods had been applied on the ALT with type-I censoring. We will modify and improve this approach in this chapter, so as to plan ALTs with interval censoring. In Woods et al. (2006) the reasons of why exact (optimal) design is more efficient than the traditional factorial design were discussed. Aitkin & Clayton (1980) showed how to fit the regression model with censored survival data by the use of exponential, Weibull, and extreme value distribution in GLM. In addition, Barbosa et al. (1996) analyzed ALT results using a piecewise exponential distribution with the GLM approach. To see the difference between the optimal experimental designs for GLM and those for typical linear models, one may refer to Chipman & Welch (1996), in which the authors used the asymptotic covariance matrix to develop an analogous D-optimality criterion and made the comparisons between GLM D-optimal designs and linear regression D-optimal designs. Dror & Steinberg (2006) proposed a heuristic method, based on clustering a set of local optimal designs, for constructing robust designs for multivariate generalized linear models, and they also discussed sequential experimentation methods. Moreover, M. Yang et al. (2011) develop a new approach to identifying optimal designs for multi-factor logistics and probit models under different optimal criteria. In this chapter, we will apply the proportional hazard model on failure times so as to transform the problem of ALT planning with interval censoring into the problem of experimental design for GLM.

3.2 Methodology

Optimality criterion

First, we introduce the optimality criterion for finding the optimal experimental design in this chapter. The benefits of applying optimal experimental design principles on the ALT planning had been demonstrated in Monroe & Pan (2008). In this chapter, we consider the U_c-optimality, which is to seek the smallest asymptotic variance of the failure

time prediction at the product's use condition among all test plans.

Define a moment matrix as (Myers et al. (2002))

$$\mathbf{M} = \frac{\mathbf{X}(\xi)' \cdot \mathbf{X}(\xi)}{n}, \tag{3.1}$$

where $\mathbf{X}$ is the design matrix of an experimental design ξ and n is sample size. The design matrix $\mathbf{X}$ has the form of

$$X(\xi) = \begin{bmatrix} 1 & x_{1,1} & \dots & x_{1,p} \\ 1 & x_{2,1} & \dots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \dots & x_{n,p} \end{bmatrix},$$

where there are p covariates and each row in the design matrix stands for a testing condition (a point in the experimental design region).

For a linear model, the inverse of M contains the variances and covariances of regression coefficients. The asymptotic variance of the response prediction at a point in the design region is determined by $\mathbf{x}'(\mathbf{X}(\xi)'\mathbf{X}(\xi))^{-1}\mathbf{x}$, where $\mathbf{x}$ is the vector of the covariates at this specific point. In ALT, we are interested in the lifetime prediction under the product's actual use condition, so it is the point $\mathbf{x}_{use}$. However, notice that this point typically is located outside of the experimental design region and extrapolation is needed in order to compute the prediction variance.

Moreover, regression models for lifetime data are not linear models. In addition, when the failure time censoring exists, it makes the situations more complicated. Thus, the above formula needs to be modified for lifetime prediction variance estimation. If we can reformulate the lifetime regression problem as a GLM problem, then the U_c-optimality becomes as solving the following optimization,

$$\arg\min_{\xi} \ \mathbf{x}_{use}{}' \cdot (\mathbf{X}(\xi)'\mathbf{W}\mathbf{X}(\xi))^{-1} \cdot \mathbf{x}_{use}, \tag{3.2}$$

where **W** is a weight matrix associated with the variance of response variable in the GLM formulation.

Model derivation

In this section, we will present the GLM approach to find optimal designs of ALT with interval censoring. The GLM formulation of interval censored survival data was initially discussed in Collett (2003) and the details are given below. Suppose that all test units enter a test at time 0 and during the course of the test, they are inspected at times $t_1, t_2, \ldots$ until t_k, where $t_1 < t_2 < \cdots < t_k$. Furthermore, let $t_0 = 0$ and $t_{k+1} = \infty$. If a test unit has failed in the j^{th} interval, then its failure time has $t_{j-1} \leq t < t_j$, for $j = 1, 2, \ldots, k$. Let p_{ij} be the probability of failure of the i^{th} test unit in the j^{th} interval and let π_{ij} be the conditional probability of the i^{th} test unit failed in the j^{th} interval given that it survived at t_{j-1}, i.e.,

$$p_{ij} = P(t_{j-1} \leq T_i < t_j),$$

and

$$\begin{aligned} \pi_{ij} &= P(T_i < t_j \mid T_i \geq t_{j-1}) \\ &= \frac{P(t_{j-1} \leq T_i < t_j)}{P(T_i \geq t_{j-1})} \\ &= \frac{p_{ij}}{R_i(t_{j-1})}, \end{aligned}$$

with $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$. $R(.)$ is a reliability function. For a test unit that have survived to the last inspection time t_k, we have

$$p_{i,k+1} = P(T_i \geq t_k) = R_i(t_k) = 1 - \sum_{l=1}^{k} p_{il},$$

and

$$\pi_{i,k+1} = P(T_i \geq t_k \mid T_i \geq t_k) = 1.$$

It then follows that

$$1-\pi_{ij}=1-\frac{P(t_{j-1}\le T_i<t_j)}{P(T_i\ge t_{j-1})}=\frac{P(T_i\ge t_j)}{P(T_i\ge t_{j-1})},$$

so

$$(1-\pi_{i1})(1-\pi_{i2})\cdots(1-\pi_{i,j-1})=P(T_i\ge t_{j-1}),$$

and

$$p_{ij}=(1-\pi_{i1})(1-\pi_{i2})\cdots(1-\pi_{i,j-1})\pi_{ij} \tag{3.3}$$

for $j=2,3,\ldots,k+1$, with $p_{i1}=\pi_{i1}$.

We let r_{ij} be an indicator variable for whether or not the i^{th} test unit is failed in the t_j interval and let s_{ij} be another indicator variable for whether or not the i^{th} test unit survives by the time t_j, i.e.,

$$r_{ij}=\begin{cases}1 & \text{when } t_{j-1}\le T_i<t_j\\ 0 & \text{otherwise}\end{cases}$$

and

$$s_{ij}=\begin{cases}1 & \text{when } T_i\ge t_j\\ 0 & \text{otherwise}\end{cases},$$

then

$$s_{ij}=r_{i,j+1}+r_{i,j+2}+\cdots+r_{i,k+1}$$

for $j=1,2,\ldots,k$.

The sample likelihood of r_{ij}'s is

$$\prod_{i=1}^{n}\prod_{j=1}^{k+1}p_{ij}^{r_{ij}}.$$

Substituting p_{ij} by (3.3), the likelihood function becomes

$$\prod_{i=1}^{n}\prod_{j=1}^{k+1}\left[(1-\pi_{i1})(1-\pi_{i2})\cdots(1-\pi_{i,j-1})\pi_{ij}\right]^{r_{ij}}.$$

This function can be written as

$$\prod_{i=1}^{n} \pi_{i1}^{r_{i1}} \left[(1-\pi_{i1})\pi_{i2}\right]^{r_{i2}} \cdots \left[(1-\pi_{i1})\cdots(1-\pi_{ik})\pi_{i,k+1}\right]^{r_{i,k+1}},$$

which reduces to

$$\prod_{i=1}^{n} \left[\pi_{i,k+1}^{r_{i,k+1}} \prod_{j=1}^{k} \pi_{ij}^{r_{ij}} (1-\pi_{ij})^{s_{ij}} \right]. \tag{3.4}$$

As $\pi_{i,k+1} = 1$, (3.4) becomes

$$\prod_{i=1}^{n} \left[\prod_{j=1}^{k} \pi_{ij}^{r_{ij}} (1-\pi_{ij})^{s_{ij}} \right]. \tag{3.5}$$

The above likelihood function has the same likelihood function form as from random variables Y_{ij} such that $Y_{ij} \sim \text{Binomial}(r_{ij}+s_{ij}, \pi_{ij})$. Since we have n units and k intervals in a test, there are nk corresponding binomial variables. Based on the definitions of r_{ij} and s_{ij}, their sum will be 1 if the i^{th} test unit survives at the time t_{j-1} and 0 if it fails before t_{j-1}, i.e.,

$$r_{ij}+s_{ij} = \begin{cases} 1 & \text{when } T_i \geq t_{j-1} \\ 0 & \text{otherwise} \end{cases}.$$

Therefore, we have

$$\begin{aligned} E[r_{ij}+s_{ij}] &= 0 \times P(T_i < t_{j-1}) + 1 \times P(T_i \geq t_{j-1}) \\ &= P(T_i \geq t_{j-1}) = 1 - \sum_{l=1}^{j-1} p_{il}. \end{aligned} \tag{3.6}$$

Note that

$$1-\pi_{ij} = P(T_i \geq t_j \mid T_i \geq t_{j-1}) = \frac{R_i(t_j)}{R_i(t_{j-1})}. \tag{3.7}$$

When the proportional hazard model is applied, (3.7) can be written as

$$1-\pi_{ij} = \left[\frac{R_0(t_j)}{R_0(t_{j-1})}\right]^{\exp(\eta_i)}, \tag{3.8}$$

where $R_0()$ is baseline reliability function and η_i is a linear predictor, i.e., $\eta_i = \mathbf{x}_i\beta$. Taking two steps of logarithm, we find that

$$\begin{aligned} \log\{-\log(1-\pi_{ij})\} &= \eta_i + \log[\log\{R_0(t_{j-1})/R_0(t_j)\}] \\ &= \eta_i + \gamma_i \end{aligned} \tag{3.9}$$

where $\gamma_i = \log[\log\{R_0(t_{j-1})/R_0(t_j)\}]$ and it is irrelevant to stress factors. Therefore, if we apply the GLM formulation on the binomial likelihood function, the link function is a complementary log-log link function.

The iterative weighted least square (IWLS) procedure is used to find the maximum likelihood estimation of a GLM (Myers et al. (2002)). The weight matrix $\mathbf{W}$ is a diagonal matrix and the diagonal elements can be derived by finding the approximate variance of the linear predictor, which is

$$\mathbf{W} = Diag\{\Delta V \Delta\}, \tag{3.10}$$

where $\Delta = diag\{\frac{d\theta_i}{d\eta_i}\}$ and V is the variance of Y_{ij}. The θ_i is the natural location parameter of the distributions from exponential family. The Δ term is introduced into the weight matrix because the log-log link function is not a canonical link function for the binomial distribution. As r_{ij} and s_{ij} are random variables, we use their expectations to replace these

variables in V. We derive the weight matrix from the (3.10) as

$$\begin{aligned}
\mathbf{W}_{adj} &= Diag\{\frac{d(\log\frac{\pi_{ij}}{1-\pi_{ij}})}{d(\log(-\log(1-\pi_{ij})))}\cdot V \\
&\quad \cdot\frac{d(\log\frac{\pi_{ij}}{1-\pi_{ij}})}{d(\log(-\log(1-\pi_{ij})))}\} \\
&= Diag\{(-\frac{\log(1-\pi_{ij})}{\pi_{ij}})\cdot V\cdot(-\frac{\log(1-\pi_{ij})}{\pi_{ij}})\} \\
&= Diag\{\frac{\log(1-\pi_{ij})}{\pi_{ij}}\cdot E[r_{ij}+s_{ij}]\cdot\pi_{ij}\cdot(1-\pi_{ij}) \\
&\quad \cdot\frac{\log(1-\pi_{ij})}{\pi_{ij}}\} \\
&= Diag\{\frac{[\log(1-\pi_{ij})]^2(1-\pi_{ij})}{\pi_{ij}}E[r_{ij}+s_{ij}]\} \qquad (3.11)
\end{aligned}$$

The weight matrix here becomes $\mathbf{W}_{adj}$ because we use the expected sample size of the binomial distribution, $E[r_{ij}+s_{ij}]$ from (3.6), to replace the actual sample size.

Now, based on the U_c-optimality, the prediction variance can be written as

$$PV_{use} = \mathbf{x}'_{use}\cdot[\mathbf{X}'\mathbf{W}_{adj}\mathbf{X}]^{-1}\cdot\mathbf{x}_{use} \qquad (3.12)$$

where $\mathbf{X}$ is the design matrix and $\mathbf{x}_{use}$ is the vector of the use condition.

We apply the nonlinear optimization procedure in SAS software to find the U_c-optimal design. Readers need to pay attention to the format of the design matrix $\mathbf{X}$. If we have n test units and k time intervals, then the dimension of $\mathbf{X}$ will be $nk\times(p+1)$ where the p is the number of stress factors used in ALT. For a specific test unit, its testing conditions are the same in different time intervals. In other words, the ALT only have n different test units, but, because of the k intervals, these n test units are virtually treated as nk units. Now, we have developed a general method that can handle different lifetime distributions, such as exponential, Weibull and so forth, with interval censoring.

Weibull and exponential distributions

Assuming a Weibull distribution for product's lifetime distribution, in this section we will show the formulation of the adjusted weight matrix for finding the optimal test plan, given an inspection plan of the test. The failure function of Weibull distribution is

$$F(t) = 1 - e^{-t^{\alpha}\lambda},$$

where α and λ are the shape and scale parameters, respectively. According to the PH model, the stress factors will affect λ only by a log-linear function, i.e.,

$$\log \lambda = \eta = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p,$$

where η is the linear predictor and x_i's are the stress factors. Thus, for a test unit we have

$$F(t) = 1 - e^{-t^{\alpha} e^{\eta_i}}.$$

The reliability function is $R(t) = 1 - F(t)$ and replace it into (3.7),

$$1 - \pi_{ij} = \frac{e^{-t_j^{\alpha} e^{\eta_i}}}{e^{-t_{j-1}^{\alpha} e^{\eta_i}}}. \tag{3.13}$$

Assume the shape parameter α is known. Let $y_j = t_j^{\alpha}$, then we have

$$\pi_{ij} = 1 - \frac{e^{-t_j^{\alpha} e^{\eta_i}}}{e^{-t_{j-1}^{\alpha} e^{\eta_i}}} = 1 - \{\frac{e^{-y_j}}{e^{-y_{j-1}}}\}^{e^{\eta_i}}. \tag{3.14}$$

With (3.11), (3.12) and (3.14), we can do the optimization and find the U_c-optimal test plan.

Now consider a special case of $\alpha = 1$, so the failure time distribution becomes exponential distribution,

$$F(t) = 1 - e^{-te^{\eta_i}}.$$

Suppose all the time intervals have the same length, Δt, then (3.13) can be written as

$$1-\pi_{ij}=\frac{e^{-t_j e^{\eta_i}}}{e^{-t_{j-1}e^{\eta_i}}} \quad = \quad \frac{e^{-j\Delta t e^{\eta_i}}}{e^{-(j-1)\Delta t e^{\eta_i}}}=e^{-e^{\eta_i}\Delta t}, \tag{3.15}$$

and the W_{adj} matrix can be derived as

$$\begin{aligned} W_{adj} &= Diag[\frac{(1-\pi_{ij})[log(1-\pi_{ij})]^2}{\pi_{ij}}\cdot E[r_{ij}+s_{ij}]] \\ &= Diag[\frac{e^{-e^{\eta_i}\Delta t}\cdot[-e^{\eta_i}\Delta t]^2}{1-e^{-e^{\eta_i}\Delta t}}\cdot e^{-e^{\eta_i}(j-1)\Delta t}] \\ &= Diag[\frac{\Delta t^2\cdot e^{2\eta_i-\Delta t\cdot(e^{\eta_i}+e^{\eta_i(j-1)})}}{1-e^{-e^{\eta_i}\Delta t}}]. \end{aligned} \tag{3.16}$$

3.3 Examples, comparison, and sensitivity analysis

We first use a simple example from Tse et al. (2008), and compare our U_c-optimal test plan from the GLM approach with their equal spaced (ES) and equal probability (EP) test plans. The ES test plan means the testing time is equally divided into several intervals and the EP test plan makes the failure probability equal in each test interval. We then demonstrate the capabilities of our method by applying it on a multiple-stress factor ALT and performing the sensitivity study of the optimal plan to model parameter assumptions.

Comparison with a traditional method

Tse et al. (2008) used the second derivative of the total likelihood function to obtain the design information matrix and to develop the equal spaced (ES) and equal probability (EP) deigns for U_c-optimal test plans. Only one stress factor was considered in Tse et al. (2008) and the lifetime distribution was given as a Weibull distribution with the shape parameter being 2.

In order to make comparisons between the optimal design generated from the GLM approach and optimal designs under ES and EP from Tse et al. (2008), we use the

Table 3.1. U_c-optimal design with type-I censoring $P_u = 0.001, P_h = 0.9$

number of intervals	ES			EP			GLM		
k	s_1^*	α_1^*	PV	s_1^*	α_1^*	PV	s_1^*	α_1^*	PV
2	0.453	0.85	0.41	0.452	0.85	0.41	0.464	0.86	0.41
3	0.457	0.85	0.41	0.456	0.85	0.41	0.459	0.86	0.41
5	0.460	0.85	0.40	0.459	0.85	0.41	0.462	0.86	0.40
10	0.461	0.85	0.40	0.460	0.85	0.40	0.463	0.86	0.40

Table 3.2. U_c-optimal design with type-I censoring $P_u = 0.001, P_h = 0.5$

number of intervals	ES			EP			GLM		
k	s_1^*	α_1^*	PV	s_1^*	α_1^*	PV	s_1^*	α_1^*	PV
2	0.367	0.90	0.67	0.367	0.90	0.67	0.392	0.86	0.67
3	0.367	0.90	0.67	0.367	0.90	0.67	0.393	0.86	0.67
5	0.367	0.90	0.66	0.367	0.90	0.66	0.393	0.86	0.66
10	0.368	0.90	0.66	0.368	0.90	0.66	0.393	0.86	0.66

same assumptions from their paper and apply the ES inspection scheme to get the optimal design. The comparison is given in Tables 3.1 and 3.2. In these two tables, P_u and P_h stand for the percentile of failure at the use condition and the high stress condition and k is the number of intervals in the test. s_1^* stands for the coded stress level, α_1^* stands for the percentage of the test units allocated at the α_1^* level, and PV stands for the prediction variance at the use condition. An important thing should be noticed here is that in this comparison, the coded stress variable s_1 of the use condition and the high stress level condition are set as 0 and 1 in order to match the terms in Tse et al. (2008).

One can see that the results from our GLM method and Tse's ES, and EP are very similar. The small differences in test plans come from the computation roundup errors. However, when there is more than one stress factor, the traditional method has to reformulate the information matrix, which becomes very complicated, while the GLM approach can easily address the problem by adding new stress variables in the linear predictor. Moreover, the interaction effect of stress factors can be investigated by the GLM approach without further difficulty.

An example with two stress factors

Suppose there is an electronic part whose lifetime belongs to an exponential distribution and its lieftime is affected by the temperature and humidity. The use condition of this electronic part is set as $30°C$ and 25%. Under the ALT test, the temperature ranges from 60 $°C$ to $110°C$ and the relative humidity level ranges from 60% to 90%. The natural stress of temperature and humidity can be presented as $S_1 = 11605/T$, with temperature in degree Kelvin, and $S_2 = \log(h)$, with relative humidity as a percentage. Following the notation used in Monroe et al. (2010), we let the design space of this experimentation to be a unit square and the use condition to be located at the first quadrant. The transformation is given by $x_{ji} = \frac{S_{ji} - S_{jH}}{S_{jL} - S_{jH}}$ where S_{jH} and S_{jL} are the high and low level of stress factor j; thus, the highest stress level is transformed to (0, 0) and the lowest stress level is transformed to (1, 1). The linear predictor is given below, while we note that an interaction effect of temperature and humidity is considered in this model.

$$\eta = -4.086x_1 - 1.476x_2 + 0.01x_1x_2. \tag{3.17}$$

Let the total testing time be 30 hours and set k equal-length inspection intervals during the test, where $k = 2$, 5, 10, and 30. We program the optimal design for GLM in SAS, based on the method from Atkinson et al. (2007). Tables 3.3 to 3.6 show the optimal test plans for these experiments. In these tables, four ALT testing conditions are selected and their temperature and humidity settings are given. The allocation column lists the number of test units to be allocated at each testing condition, with the assumption of totally 100 test units. The next two columns are the failure probability at the end of testing period and the lifetime prediction variance at the use condition. Figures 3.1 to 3.4 provide the contour plots of prediction variance under these optimal test plans. The unit square region in the lower-left corner of each graph is the experimental design region. The

Table 3.3. U_c-optimal design with 2 intervals

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	108.72	0.022	60.00	1.000	12	0.99	11.86
2	94.55	0.280	85.65	0.097	31	0.99	
3	64.22	0.905	60.00	1.000	44	0.16	
4	60.00	1.000	90.00	0.000	13	0.40	

Table 3.4. U_c-optimal design with 5 intervals

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	60.00	1.000	11	0.99	8.32
2	103.00	0.124	90.00	0.000	21	1.00	
3	67.00	0.843	60.00	1.000	49	0.20	
4	60.00	1.000	90.00	0.000	19	0.40	

Table 3.5. U_c-optimal design with 10 intervals

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	90.00	0.000	15	1.00	7.37
2	110.00	0.000	60.00	1.000	10	0.99	
3	66.64	0.851	60.00	1.000	54	0.19	
4	60.74	0.984	90.00	0.000	21	0.42	

selected testing conditions are marked by circles and the circle diameter is corresponding to the sample size allocation at such condition. The use condition is marked by a rectangle, which is located outside of the experimental design region, and the locations with equal prediction variance are outlined by contour lines.

Table 3.6. U_c-optimal design with 30 intervals

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	90.00	0.000	12	1.00	6.94
2	110.00	0.000	60.00	1.000	12	0.99	
3	67.82	0.825	60.00	1.000	52	0.21	
4	61.74	0.961	90.00	0.000	24	0.45	

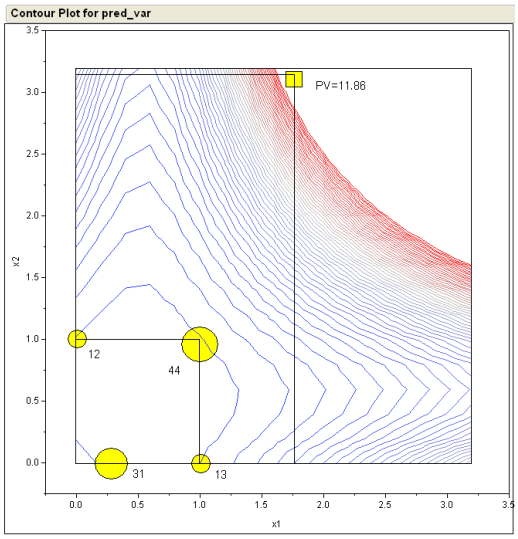

Figure 3.1. Contour plot of U_c-optimal design with 2 intervals

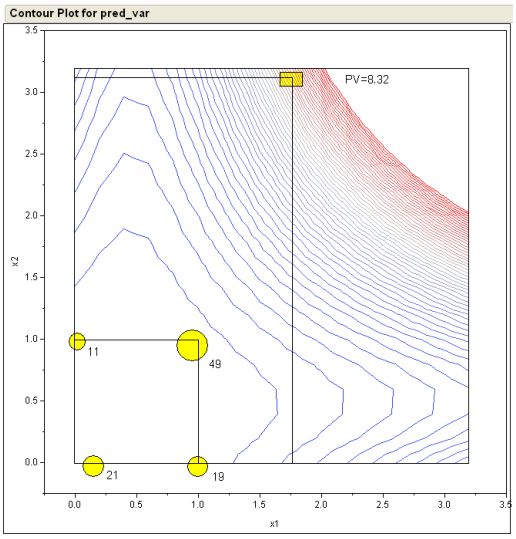

Figure 3.2. Contour plot of U_c-optimal design with 5 intervals

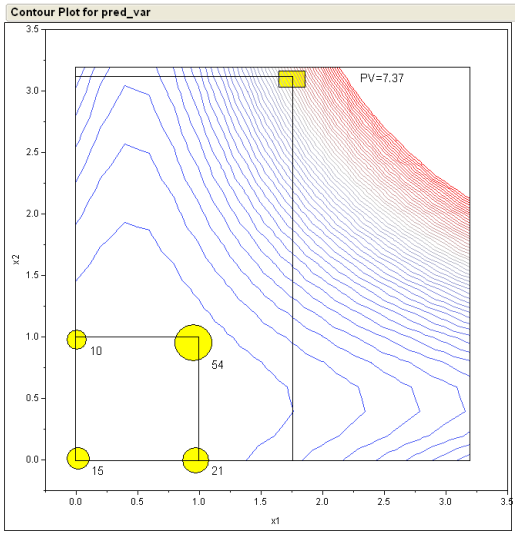

Figure 3.3. Contour plot of U_c-optimal design with 10 intervals

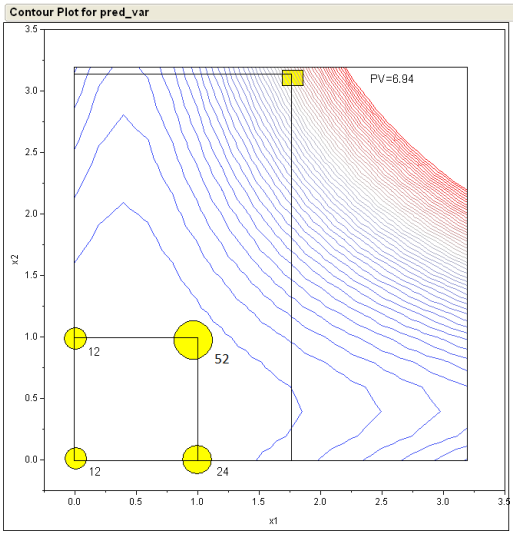

Figure 3.4. Contour plot of U_c-optimal design with 30 intervals

From these tables and figures readers can see that with increasing number of inspection intervals, the lifetime prediction variance at the use condition is decreasing. This is expected, as more intervals means that more precise information of failure times can be obtained. Notice that the number of test units being allocated at the location of the lowest combination in coded variables x_1 and x_2 (or, the highest stress level) shows a decreasing trend when the number of inspection intervals increases, i.e., 31 at (0.280, 0.097), 21 at (0.124, 0.000), 15 at (0.000, 0.000), and 12 at (0.000, 0.000) in Tables 3.3 to 3.6, respectively. This happens because with fewer number of intervals failure time information is less available, and correspondingly, sample size needs to be increased. This phenomenon is more prominent at a higher stress level. Also, fewer intervals (or longer interval time) causes the problem to be more nonlinear, thus the highest testing stress condition may not be located at the low-left corner of the design region, as it often appears on the optimal experimental design for linear models. In fact, the actual number of testing conditions found by our algorithm are more than four; however, in order to have an easier implementation of any optimal test plan, we round up the coded test location coordinates to 0.001, which aggregates these locations to four distinct points. Moreover, in all of these plans one can see that most test units are allocated at the upper-right corner of the design region, which corresponds to the lowest testing stress level. This is because 1) it is the closest point in the design region to the use condition point, so more failure information at this point will help reduce the prediction variance at the use condition; 2) as fewer failures are expected at low stress level, we need to increase the sample size.

Now we compare the interval censoring case with a right censoring case with the same model assumption, as shown in Table 3.7. One can see that the optimal test plan for interval censoring becomes more and more close to the optimal plan for right censoring, as the number of intervals increases. Actually when we run an extreme case with 120 intervals during 30 hours testing period, the optimal test plan is exactly the same as that of

Table 3.7. U_c-optimal design with type-I censoring

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	90.00	0.000	12	1.00	6.91
2	110.00	0.000	60.00	1.000	12	0.99	
3	67.86	0.824	60.00	1.000	52	0.21	
4	62.05	0.954	90.00	0.000	24	0.46	

a right censoring case. It validates the test plan generated for interval censoring.

Sensitivity study

The linear predictor in (3.17) in our model plays a very important role in determining the optimal test plan. The coefficients specified in the linear predictor are from previous experimental results and engineering experience. However, the true model may not be the same as the model pre-specified; therefore, a study of the sensitivity of the optimal test plan to a misspecified model is necessary. Suppose the original coefficient values in (3.17) are used for finding an optimal test plan. However, the real values of these coefficients are not the same as those in (3.17). Assuming that the real value of each stress's coefficient may vary $\pm 20\%$ from the used value, we have 9 combinations of these coefficient values, or 9 true linear predictor models. We will compare the performance of the assumed optimal test plan to the true optimal plan and discuss the impact of the misspecified model. In this study, we let the number of inspection intervals $k = 10$ and the shape parameter $\alpha = 1.2$. The results are presented in Table 3.8, where the column of ideal PV gives the lifetime prediction variance if the right model is specified and the column of actual PV gives the prediction variance when the predictor model is misspecified.

In Table 3.8, the ideal PV is always smaller than the actual PV. This is because the

Table 3.8. PV of designs with different combination for stress coefficient

No.	comb	coefficient of x_1	coefficient of x_2	ideal PV	actual PV	% Change
0	(Orig,Orig)	-4.086	-1.476	5.08	5.08	0.0%
1	(+20%,Orig)	-4.900	-1.476	7.73	8.58	11.0%
2	($-$20%,Orig)	-3.269	-1.476	3.43	3.52	2.6%
3	(+20%,+20%)	-4.900	-1.771	8.76	10.25	17.0%
4	($-$20%,+20%)	-3.269	-1.771	3.78	3.88	2.6%
5	(+20%,$-$20%)	-4.900	-1.181	6.85	7.37	7.6%
6	($-$20%,$-$20%)	-3.269	-1.181	3.14	3.29	4.8%
7	(Orig,+20%)	-4.086	-1.771	5.78	5.86	1.4%
8	(Orig,$-$20%)	-4.086	-1.181	4.45	4.53	1.8%

the "ideal PV" is calculated from the true optimal test plan derived from the true predictor function, while the "actual PV" is calculated from the assumed optimal test plan from the misspecified predictor function. One can see that their differences are relatively small ($< 8\%$) in 6 out of 8 cases, which indicates that, under these cases, even if the model is misspecified, the resulted test plan still give a robust performance. However, for the cases that the coefficient of x_1 changes in the positive direction when the coefficient of x_2 is kept the same or moving to the same direction, the PV differences are larger than 10%. This can be explained from the predictor function, in (3.17), as one can see that x_1 has a much larger influence on η. The result implies that at the test planning stage, we should pay more attention to the coefficient of temperature and specify it as accurately as possible.

The above example shows how the wrongly estimated coefficients in the linear predictor will affect ALT results. Therefore, a reliability engineer may prefer a test plan that is robust to a group of possible models, instead of a single model. It is common that at the test planning stage, the engineer assumes that the coefficients used in the predictor function may take values from an interval with certain distribution, or it can be discretized to a set of values using probability weights. For example, let the coefficient of x_1 be chosen from -4.486, -4.286, -4.086, -3.886, and -3.686 with the probability of 0.1, 0.2, 0.4, 0.2, and 0.1, respectively, and the coefficient of x_2 be chosen from -1.876, -1.676, -1.476, -1.276, and -1.076 with the probability of 0.1, 0.2, 0.4, 0.2, and 0.1

Table 3.9. Weighted average U_c-optimal design with 5 intervals

Test	Temperature		Humidity		Alloc	FP	PV
Cond	Natural	Std	Natural	Std			
i	$^\circ$ C	x_1	%	x_2	n_i		
1	110.00	0.000	60.00	1.000	10	0.99	8.59
2	103.04	0.123	90.00	0.000	21	0.99	
3	66.73	0.849	60.00	1.000	50	0.19	
4	60.00	1.000	90.00	0.000	19	0.40	

respectively. Then, we can seek a test plan that can minimize the average prediction variance over the range of the coefficient values. We call it a weighted average U_c-optimal test plan. The optimal criterion of the weighted average U_c-optimal design can be written as

$$\min \sum_{c1} \sum_{c2} p \cdot \mathbf{x}_{use}{}' \cdot (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \cdot \mathbf{x}_{use} \tag{3.18}$$

where c_1 and c_2 are the index of the available values for the coefficients and p is the prior probability for each coefficient combination.

Continuing our example, we compute the weighted average prediction variance based on the 25 coefficient combinations. Also, assume that the shape parameter is 1 and there are 5 inspection intervals in the test. The weighted average U_c-optimal test plan is given in Table 3.9 and the contour plot is in Figure 3.5. Suppose the coefficient values $(-4.086, -1.476)$ are the true values, then given the weighted average U_c-optimal plan from Table 3.9, the real prediction variance at the use condition is 8.356, which is close to the predicted value, 8.59.

In addition, the comparisons among the weighted average U_c-optimal plans with different probability distributions are presented in Table 3.10. With the same discretizing schemes for x_1 and x_2 as above, we compare plans for three different weighting schemes, equal weighting (0.2, 0.2, 0.2, 0.2, 0.2), symmetric weighting (0.1, 0.2, 0.4, 0.2, 0.1), and concentrated weighting, or fixed coefficient value, (0, 0, 1, 0, 0), respectively. The last case is assumed using the true coefficient values. We also vary the number of intervals

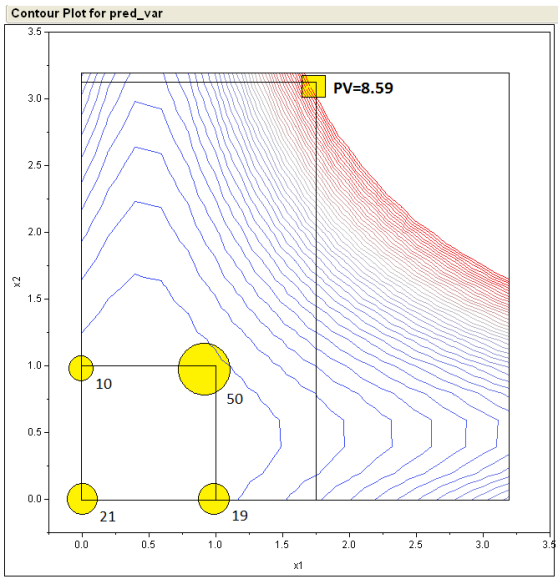

Figure 3.5. Weighted average U_c-optimal design with 5 intervals

from 2 to 10. From this table we can see that Case B, which has larger weight on the true coefficient values, provides a better test plan than Case A, where all possible coefficient combinations have the equal weight. In other words, when the prior probabilities of these coefficients are more accurate, the test plan will be closer to the true optimal plan. At the same time, when the number of inspection intervals increases the prediction variance will decrease monotonically.

3.4 Conclusion

In this chapter, we develop a GLM approach to constructing optimal ALT test plans when failure times of test units are expected to be interval censored. The optimal criterion is selected to minimize the prediction variance of the product's expected lifetime at its use condition. A PH regression model is assumed for the failure time distribution, as it allows the GLM formulation to be derived from the total likelihood function of readout data. This assumption indeed encompasses a wide range of failure time distributions, with Weibull and exponential distributions as two special cases. Compared to the conventional

Table 3.10. Comparisons among weighted average U_c-optimal test plans

Interval	Case A: equal weighting				Case B: symmetric weighting				Case C: concentrated weighting			
2	Location		Alloc	PV	Location		Alloc	PV	Location		Alloc	PV
i	x_1	x_2	n_i		x_1	x_2	n_i		x_1	x_2	n_i	
1	0.048	1.000	13	12.48	0.042	1.000	12	12.18	0.022	1.000	12	11.86
2	0.284	0.096	29		0.285	0.087	29		0.280	0.097	31	
3	0.897	1.000	45		0.908	1.000	46		0.905	1.000	44	
4	1.000	0.000	13		1.000	0.000	13		1.000	0.000	13	
Interval	Case A				Case B				Case C			
5	Location		Alloc	PV	Location		Alloc	PV	Location		Alloc	PV
i	x_1	x_2	n_i		x_1	x_2	n_i		x_1	x_2	n_i	
1	0.000	1.000	10	8.68	0.000	1.000	10	8.59	0.000	1.000	11	8.32
2	0.125	0.000	20		0.123	0.000	21		0.124	0.000	21	
3	0.848	1.000	51		0.849	1.000	50		0.843	1.000	49	
4	1.000	0.000	19		1.000	0.000	19		1.000	0.000	19	
Interval	Case A				Case B				Case C			
10	Location		Alloc	PV	Location		Alloc	PV	Location		Alloc	PV
i	x_1	x_2	n_i		x_1	x_2	n_i		x_1	x_2	n_i	
1	0.000	0.000	14	7.64	0.000	0.000	14	7.51	0.000	0.000	15	7.37
2	0.000	1.000	10		0.000	1.000	10		0.000	1.000	10	
3	0.841	1.000	55		0.846	1.000	55		0.851	1.000	54	
4	0.980	0.000	21		0.984	0.000	21		0.984	0.000	21	

approach to ALT test planning, our GLM approach has the advantages on both the easiness of computation and the ability of handling more complex models, such as a model with more than one stress factor and with the interaction effect of two stress factors, or the sensitivity study of a test plan to its model parameter specification, etc. These have been demonstrated in the examples above.

Chapter 4

Design and Evaluation of Accelerated Life Testing Plans with Dual Objectives

4.1 Introduction

Background and Motivation

Accelerated life testing (ALT) is widely accepted in the manufacturing industry for predicting a product's lifetime that could last years or even decades. The general idea of ALT is to elevate the stress level of an environmental stress factor (e.g., temperature, humidity or pressure) to a higher-than-normal level so that a good number of product failures can be observed within the testing period and the failure time distribution can be effectively estimated. However, even with an ALT, cost, equipment capability and other practical experimental constraints are often critical to designing a feasible and statistically efficient test plan. For instance, failure time censoring is almost unavoidable in ALT due to the constraint of limited testing period or the method used for inspecting test units. Two common censoring types are right censoring and interval censoring. If a test unit survives after the testing period or it is removed from the test before failure, then its failure time is right censored. If a test unit is periodically inspected, its exact failure time cannot be observed and it is interval censored. Experimenters must consider the effect of censoring when selecting an ALT test plan. Moreover, the physical/engineering model, such as Arrhenius model, Eying model, or Peck model, etc., for modeling the relationship between mean lifetime and environmental stress factors is typically a nonlinear function. For a multi-factor ALT experiment, it is not easy to find a test plan that is both practically feasible and statistically optimal.

The techniques of alphabetic optimal experimental design have been applied on ALT planning (e.g. Park & Yum (1996), McGree & Eccleston (2009)). The conventional

approach to obtaining an optimal test plan is by constructing a sample likelihood function of failure times based on the failure time distribution assumption and then deriving the expected Fisher information matrix. However, when censoring exists, the likelihood function may become quite complicated and its information matrix cannot be easily computed. To alleviate the computation and to show the intrinsic features of optimal test plans, Monroe et al. (2011) and Pan & Yang (2011) proposed a proportional hazard (PH) model approach, which treats the optimal ALT planning problem as an experimental design problem for a generalized linear model (GLM). We will take this approach to deriving optimal ALT plans in the rest of the chapter.

An optimally designed ALT may have the purpose of optimizing either model parameter estimation or making precise prediction at the use condition stress level. An estimation-oriented test plan tries to optimize an optimality criterion that is related to the quality of model parameter estimation. For example, the D-optimal criterion, which considers the general variance of parameter estimators, is one of the most popular optimality criterion for model estimation. A prediction-oriented ALT test plan focuses on the quality of model-based prediction, which is typically the prediction of product's mean lifetime or a low percentile of its lifetime under its use condition. The U_c-optimality, proposed in Monroe et al. (2010), assumes that product's use condition is a single combination of stress factors and levels. In practice, a product's use environment may vary according to its application and/or time, so its use stress condition is no longer a fixed point. Instead, it is a use region. For example, the operational temperature of a CPU may vary from $35°C$ to $50°C$ depending on computing demands. Therefore, we would like to find an optimal test plan that can minimize the average prediction variance over the entire use condition region, and we define it as *I-optimality*. In addition, a good ALT plan should take into consideration that any optimal test plan is model and parameter dependent; i.e., the optimal plan can only be obtained after the product's life-stress model

and model parameters are fully specified. The experimenter may want to assess the estimation property of a test plan, such as its D-efficiency, as well as its prediction property at the same time so as to choose a balanced plan.

In this chapter, we introduce a dual-objective optimization process based on the I- and D-optimality to provide experimenters the most efficient ALT plan. We develop several graphical tools – the fraction of use space (FUS) plot, efficiency plot, and Pareto frontier plot – for evaluating and comparing ALT test plans. In the rest of this section the previous work of optimal ALT planning is reviewed. The I-optimality is then introduced in Section 2 and we study the effect of interval censoring on I-optimal test plans. The dual objective optimization is formulated in Section 3, followed by the discussion of test plan evaluation using graphical tools in Section 4. Finally, this chapter is concluded in Section 5.

Previous work

The literature on ALT and ALT planning is vast. One may refer to Nelson (2005a,b) for a summary of ALT literature up to 2005. Notably, Meeker & Nelson (1975) and Nelson & Meeker (1978) applied the maximum likelihood theory on designing optimal ALT plans, while assuming the product lifetime has a Weibull or extreme value distribution. Nelson & Kielpinski (1976) discussed the theory of optimal ALT plans for establishing a simple linear relationship between a single stress factor and product lifetime that has a normal or lognormal distribution, and they also considered censoring. In Escobar & Meeker (1995), Park & Yum (1996), and Sitter & Torsney (1995), optimal test plans with two stress factors were discussed and right censoring was adopted. Islam & Ahmad (1994) developed the optimal designs under periodic inspection and type-I censoring. More recently, Tang et al. (2002) considered two alternative methods for planning optimal ALTs with three stress levels. Tse et al. (2008) provided the optimal

ALT designs under the interval censoring with random removals.

The conventional approach to developing optimal test plans for ALT is by applying the accelerated failure time (AFT) model to obtain the likelihood function of failure observations. However, due to censoring and the nonlinear relationship between failure time and stress factors, obtaining the expected information matrix from the log-likelihood function of AFT model becomes a tedious task for many commonly used failure time distributions. In Monroe et al. (2011) and Pan & Yang (2011), a new approach based on the proportional hazard (PH) model was proposed. This approach converts the multi-stress ALT planning problem to an experimental design problem for GLM. Their method has been applied to designing optimal ALT plans with right censoring and interval censoring. For the theories and applications of the PH model, one may also refer to, e.g., Finkelstein & Wolfe (1985), Finkelstein (1986), and Elsayed & Jiao (2002).

In the design of experiments literature, Myers et al. (2009) described the I-optimal design as an experimental design that minimizes the average model prediction variance over the entire design space. For an ALT experiment, the region of use stress levels is usually located outside of the experimental design space, so the model-based extrapolation has to be conducted in order to infer the failure time distribution under use condition. This implies that the model parameter estimation property of a test plan is very important too, since any bias in parameter estimation can be exaggerated by extrapolation. In fact, in some ALT applications experimenters are more interested in establishing the product's lifetime-stress model than predicting its failure time at use stress level (e.g., Monroe & Pan (2008)). Therefore, it is necessary to investigate a multi-objective optimization process to find a test plan with desirable estimation and prediction properties. Konak et al. (2006) pointed out that a reasonable solution to a multi-objective problem was to investigate a set of solutions, with each of which satisfying the objectives at an acceptable level without being dominated by any other solutions, i.e., a Pareto

optimal solution set. In this chapter, we formulate a dual-objective optimization problem with the consideration of both I- and D-efficiency of an ALT plan.

Graphical methods are very attractive to experimenters for assessing the performance of an experimental design and for comparing different designs. Giovannitti-Jensen & Myers (1989) proposed the variance dispersion graph (VDG) as a performance comparison tool of competing designs on a fixed design space. Zahran et al. (2003) developed the plot of fraction of design space (FDS), which shows the distribution of scaled prediction variance (SPV) in a design region. As aforementioned, the use stress region in ALT is often located outside of the experimental design region. Therefore, we propose a plot of the fraction of use space (FUS), in analogy to FDS. In addition, the efficiency plot and the Pareto frontier plot will be used to compare different test plans.

4.2 I-optimal Test Plan

Optimal criterion

In this chapter we consider the I-optimality, which, among all test plans, seeks the one that minimizes the integral of variance of failure time prediction over the product's use condition region. It is an extension of the U_c-optimality defined in Monroe et al. (2010). The design matrix of an experimental design (a test plan) ξ has the form of

$$X(\xi) = \begin{bmatrix} 1 & x_{1,1} & \dots & x_{1,p} \\ 1 & x_{2,1} & \dots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \dots & x_{n,p} \end{bmatrix},$$

where each column corresponds to a term in the linear predictor (including the intercept term as the first column) and each row corresponds to the testing condition of a single test unit. Thus, this matrix specifies how the experiment will be conducted.

A generalized linear model consists of three components (see McCullagh & Nelder (1989)): (1) the distribution of response variable that is from the exponential family; (2) a linear predictor of independent variables such as $\eta = \mathbf{x}\beta$; and (3) a link function that connects the linear predictor to the response distribution parameter, e.g., $\eta = g(\mu)$, where μ is the mean of response variable. Given a design matrix $\mathbf{X}$, the asymptotic variance-covariance matrix of the estimators of the linear predictor coefficients is found to be proportional to $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$, where $\mathbf{W}$ is a diagonal matrix with the diagonal elements being weights, and the specification of the weight matrix depends on the link function and the censoring type. At the use condition, $\mathbf{x}_{use}$, the asymptotic variance of response prediction is then proportional to $\mathbf{x}'_{use}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{x}_{use}$. A test plan ξ^* that minimizes the asymptotic prediction variance at the use condition is called the U_c-optimal design.

To consider a region of use conditions, we define the I-optimal test plan to be the plan, among all test plans, that minimizes the average prediction variance over the entire use condition region. That is, we are looking for a test plan such that

$$\xi^* := \min_{\xi} \frac{\int_{\Omega} \mathbf{x}_{use}{}' \cdot (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \cdot \mathbf{x}_{use} d\mathbf{x}_{use}}{S_{\Omega}}, \tag{4.1}$$

where Ω is the use condition region and S_{Ω} is the area of the use condition region.

GLM formulation

In this section, we briefly describe how to convert an optimal ALT planning problem to the problem of experimental design for GLM. More details, along with an example, are given in Appendix A.

By the proportional hazard assumption used in the PH model, the failure rate function and the reliability function can be written as $\lambda(t) = \lambda_0(t)e^{\mathbf{x}'\beta}$ and $R(t) = e^{-\Lambda_0(t)\exp(\mathbf{x}'\beta)} = R_0(t)^{\exp(\mathbf{x}'\beta)}$, respectively. Here, $\lambda_0(t)$ is the baseline hazard

function, $\Lambda_0(t) = \int_0^t \lambda_0(\tau)d\tau$ is the baseline cumulative hazard function, and $R_0(t) = e^{-\Lambda_0(t)}$ is the baseline reliability function.

For the case of interval censoring, suppose that all test units are inspected at times $t_1, t_2, \ldots, t_k$, where $t_1 < t_2 < \cdots < t_k$. Denote r_{ij} as a binary variable indicating whether or not the i^{th} test unit fails in the inspection interval from t_{j-1} to t_j, and m_{ij} as another binary variable indicating whether or not the i^{th} test unit survives at the time t_{j-1}. It is shown in Collett (2003) that the log-likelihood function of failure observations during the course of testing can be expressed as

$$L = \sum_{i=1}^{n} \sum_{j=1}^{k} \left[r_{ij} \log \pi_{ij} + (m_{ij} - r_{ij}) \log(1 - \pi_{ij}) \right], \tag{4.2}$$

where π_{ij} is the conditional probability of the i^{th} test unit failing in the time interval (t_{j-1}, t_j). A log-log link function between π_{ij} and the linear predictor η_i, $\eta_i = \mathbf{x}_i'\beta$, can be established as

$$\log[-\log(1 - \pi_{ij})] = \eta_i + \log \left[\log \frac{R_0(t_{j-1})}{R_0(t_j)} \right]. \tag{4.3}$$

The second term on the right hand side of (4.3) is an offset term, as it is irrelevant to stress factors. This formulation can be viewed as the sample likelihood function of the binomial random variable r_{ij} with a log-log link function. The weight matrix is a diagonal matrix with its diagonal elements being $w_i = m_{ij}(1 - \pi_{ij})[\log(1 - \pi_{ij})]^2 / \pi_{ij}$ with $i = 1, 2, ..., n$ and $j = 1, 2, ..., k$. Using this formulation, T. Yang & Pan (2013) derived the U_c-optimal ALT plans.

For the case of right censoring, exact failure times, t_i's, are observed from the ALT test, unless the test unit is censored. Without loss of generality, we assume a common censoring time t_c, where $t_c > t_i$ for all i's. Let c_i be a binary variable indicating whether or not the failure time of the i^{th} test unit is censored. It is shown in Aitkin & Clayton (1980) that the sample log-likelihood function can be expressed as

$$L = \sum_{i=1}^{n} [c_i \log \mu_i - \mu_i] + \sum_{i=1}^{n} \left[c_i \frac{\log \lambda_0(t_i)}{\log \Lambda_0(t_i)} \right], \tag{4.4}$$

where μ_i is the value of the cumulative hazard function of the i^{th} test unit at its failure time and

$$\log \mu_i = \eta_i + \log \Lambda_0(t_i). \tag{4.5}$$

Notice that the second term in the right hand side of the log-likelihood function above is a constant, unrelated to stress factors. Therefore, to maximize the likelihood function, we may ignore this term. Then, this likelihood function can be viewed as a Poisson distribution with a log link function. The weight matrix is given by a diagonal matrix with its diagonal elements as $w_i = \mu_i$ with $i = 1, 2, .., n$. For the derivation of the optimal test plans with exponential and Weibull failure time distributions, one may refer to Monroe et al. (2011) and Pan & Yang (2011).

We code the optimization process of finding optimal ALT plans in SAS. The computer program is available from the author Tao Yang upon request.

Effects of censoring on I-optimal test plans

In Monroe et al. (2010) an ALT experiment with two stress factors was described. Engineers are interested in studying the lifetime of an electronic part, which is assumed to have an exponential distribution and its mean lifetime is affected by temperature and humidity. During the ALT experiment temperature can be varied from $60°C$ to $110°C$ and relative humidity can be varied from 60% and 90%. The nominal use condition is set at $30°C$ and 25% relative humidity, but it is known that the actual use condition has the temperature range as from $20°C$ to $40°C$ and the relative humidity from 20% to 30%. All experimental runs are scheduled to last 30 hours.

Following the notation in Monroe et al. (2010), we standardize the ALT experimental region of natural testing stresses (i.e., the inverse of temperature in degrees Kelvin and the natural logarithm of relative humidity) as a unit square with each coded

stress variable having values from 0 to 1. The highest test stress level is coded as 0 and the lowest test stress level is coded as 1, so the use stress level, which is outside of this experimental region, falls in the first quadrant. The nominal use condition becomes (1.758, 3.159) and the use region is a 0.557×1 rectangle. The derivation of the GLM formulation of this example is given in Appendix A. We assume that test units are periodically inspected to determine whether or not they have failed and these inspection intervals are all equal in time. There are a total of 100 test units and an optimal test plan is to allocate them to four distinct testing conditions (as there are four regression coefficients to be estimated; see the linear predictor function in Appendix). Table 4.1 lists the I-optimal designs when there are 2, 5, 10, or 30 inspections. The last two columns in Table 4.1 represent the expected proportion of failed units among all test units allocated at such testing condition, and the average prediction variance of the test plan.

Table 4.1. I-optimal designs

Interval No. k	Test Cond. i	Temperature Natural $^\circ C$	Temperature Std x_1	Humidity Natural %	Humidity Std x_2	Alloc n_i	FP	PV
2	1	64.13	0.907	60.00	1.000	44	0.16	12.92
	2	60.00	1.000	90.00	0.000	13	0.40	
	3	108.66	0.023	60.00	1.000	12	0.99	
	4	94.50	0.281	85.75	0.094	31	0.99	
5	1	66.91	0.845	60.00	1.000	49	0.20	9.03
	2	60.00	1.000	90.00	0.000	19	0.40	
	3	110.00	0.000	60.00	1.000	11	0.99	
	4	102.88	0.126	89.12	0.000	21	0.99	
10	1	66.64	0.851	60.00	1.000	53	0.19	8.04
	2	60.91	0.980	90.00	0.000	21	0.41	
	3	110.00	0.000	60.00	1.000	10	0.99	
	4	110.00	0.000	90.00	0.000	16	0.99	
30	1	66.59	0.826	60.00	1.000	52	0.21	7.50
	2	60.00	0.962	90.00	0.000	24	0.45	
	3	110.00	0.000	60.00	1.000	12	0.99	
	4	110.00	0.000	90.00	0.000	12	0.99	

From Table 4.1 one can see that when the inspection becomes more frequent, the I-optimal design gives a smaller average prediction variance. This is anticipated, because more inspection intervals will provide more precise failure time information from the test.

One can also see that when the inspection becomes more frequent, more test units will be allocated to the lowest stress level (testing condition 1) and fewer units to the highest stress level (test condition 4). The lowest stress level is closer to the use condition region than other testing conditions and it is expected to have fewer failure observations as the stress is less severe than others. Therefore, the optimal test plan allocates more test units to the lower test condition. When more failure information can be obtained by increasing the number of intervals, the optimal plan tends to have even more units allocated to this condition. In Figure 4.1 we plot the contour lines of prediction variance over the experimental design region and the use region for the cases of $k = 2$ and $k = 30$. One can see that the contour lines become less dense over the use region when the number of intervals increases, which indicates smaller prediction variance, and the distribution of test units in the experimental region shifts to the lowest testing condition (the contour lines are at the same level for both figures).

4.3 Dual-Objective Test Plan

It is known that optimal ALT plans are model dependent, so these optimal plans are local optimal plans; that is, the test plan depends on the specific model used in planning. Typically some physical models, such as Arrhenius model, are used to describe the relationship between product lifetime and environmental stress factors. The model parameter value used in the lifetime-stress model must be specified by engineers before planning the optimal experimentation, and they are called the planning values. Although these planning values are assumed known, but, in fact, they are also of interest to experimenters. Model estimation is an important task to ALT experimenters, because the model-based extrapolation is unavoidable for inferring failure distribution at the product's use stress level. Therefore, besides of obtaining a test plan that can minimize the prediction variance at the use region, it is also desired that the plan has a good model

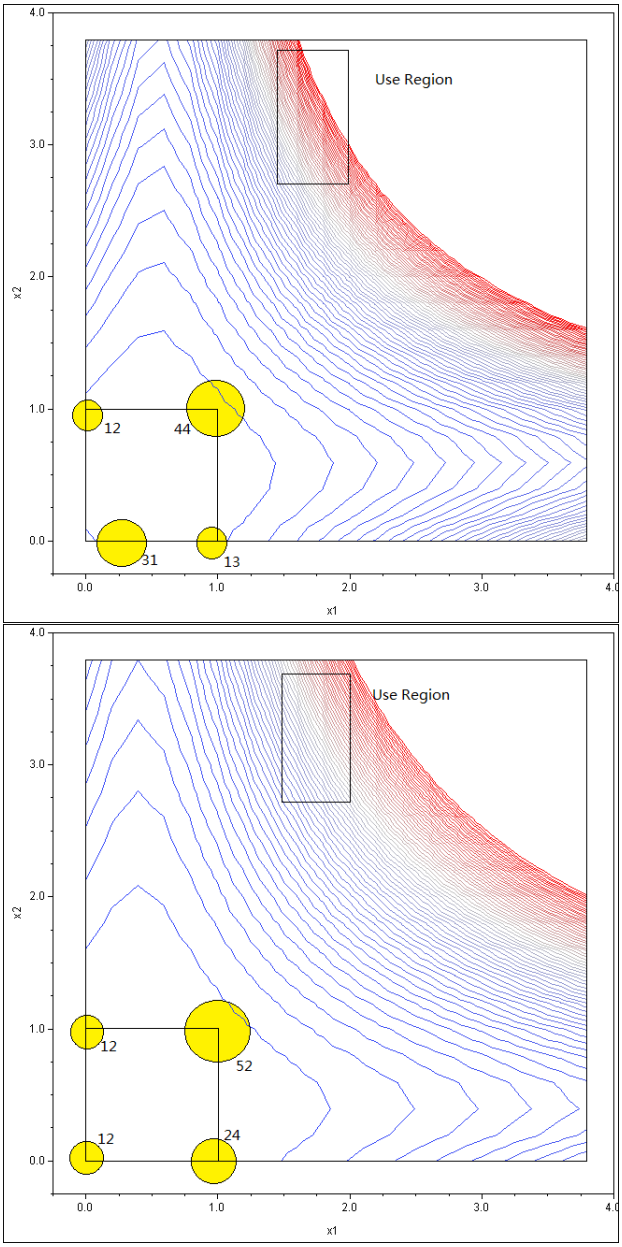

Figure 4.1. Contour plots of I-Optimal designs with $k = 2$ (top) and $k = 30$ (bottom)

parameter estimation property. Meeker & Escobar (1998) (Table 20.5, page 553) proposed a split design for the ALT with two stress variables, in which two optimal design points (at the highest stress level and a lower stress level) are first found to minimize the prediction variance at the use stress level, then the design point at the lower stress level is split to two design points according to the D-optimality criterion so that the test plan may cover a

larger experimental region and model parameters can be better estimated.

In this chapter we propose a dual-objective formulation for ALT planning, so we can study the trade-off between the I- and D-optimalities of a test plan and provide experimenters alternative test plans to suit their specific needs. The goal of dual optimization is to find the test plans that are located on the Pareto efficiency frontier; i.e., there is no any other plan that is more efficient with both I- and D-optimalities than the plans on the frontier. Let Det_{opt} be the determinant of a D-optimal design and Var_{opt} the average prediction variance of an I-optimal design. Suppose that the determinant value of the information matrix of a test plan, ξ, is Det_{ξ} and its average prediction variance value is Var_{ξ}. We want to find a test plan such that it can maximize the combination of D-efficiency and I-efficiency as

$$\max_{\xi}\{\lambda \times \frac{Var_{opt}}{Var_{\xi}} + (1-\lambda) \times \frac{Det_{\xi}}{Det_{opt}}\}, \tag{4.6}$$

where λ is the weight assigned to the I-optimality and $0 \leq \lambda \leq 1$.

Using the previous example and letting λ be 0.1, 0.3, 0.5, 0.7, or 0.9, we find the test plans listed in Table 4.2 that maximize (4.6). From this table, one can see that when the λ value increases, I-optimality becomes more important, so the average prediction variance will decrease and the test plan has more test units allocated at the lowest stress level (testing condition 1).

4.4 Test Plan Evaluation

In this section we describe several graphical tools that are used to evaluate the estimation and prediction properties of an ALT plan and to help experimenters compare different test plans. These tools are illustrated by examples.

Table 4.2. Dual-objective designs with 5 equal inspection intervals

Weight	Test	Temperature		Humidity		Alloc	FP	PV	Det
	Cond	Natural	Std	Natural	Std				
λ	i	$^\circ$ C	x_1	%	x_2	n_i			
0.1	1	75.97	0.650	60.00	1.000	35	0.38		
	2	62.53	0.943	90.00	0.000	21	0.47		
	3	110.00	0.000	60.00	1.000	19	0.99	10.63	7524
	4	96.09	0.251	90.00	0.000	25	0.99		
0.3	1	74.78	0.675	60.00	1.000	37	0.35		
	2	62.27	0.949	90.00	0.000	21	0.46		
	3	110.00	0.000	60.00	1.000	18	0.99	10.21	7196
	4	96.25	0.248	90.00	0.000	24	0.99		
0.5	1	72.85	0.716	60.00	1.000	41	0.31		
	2	61.83	0.959	90.00	0.000	20	0.45		
	3	110.00	0.000	60.00	1.000	16	0.99	9.71	6331
	4	96.68	0.240	90.00	0.000	23	0.99		
0.7	1	71.26	0.750	60.00	1.000	42	0.28		
	2	61.35	0.970	90.00	0.000	19	0.43		
	3	110.00	0.000	60.00	1.000	16	0.99	9.52	5928
	4	97.64	0.222	90.00	0.000	23	0.99		
0.9	1	67.91	0.823	60.00	1.000	48	0.28		
	2	60.00	1.000	90.00	0.000	20	0.40		
	3	110.00	0.000	60.00	1.000	12	0.99	8.95	3856
	4	100.41	0.171	90.00	0.000	20	0.99		

Fraction of use space plot

Although the I-efficiency can provide the average prediction variance over the use stress region, it is more informative to experimenters if the distribution of prediction variance over the whole use region can be displayed. As a hypothetical case, a test plan that has large prediction variances over the most part of its use stress region may have very small variances over a small portion of the use region, which keeps the average prediction variance low. Obviously, this plan is worse than a plan that keeps the prediction variances relatively low over most of the use stress region. The fraction of design space (FDS) plot, developed in Zahran et al. (2003), displays the fraction of an experimental design region that is under certain prediction variance values and it is good at assessing the overall prediction performance on the entire design region. A good design should have

a flat FDS curve. We extend this idea to the use stress region. We plot the prediction variance versus the fraction of use stress region that has less than or equal to this prediction variance value, and call it the fraction of use space (FUS) plot.

Using the previous example, the four I-optimal designs, corresponding to four different inspection intervals, in Table 4.1 are evaluated by their FUS plots, as shown in Figure 4.2. In this figure, the horizontal axis represents the fraction of use region that has the value of prediction variance that is less than or equal to the value on the vertical axis. It is clear from the figure that when there are more intervals the prediction variances over the whole use stress region are reduced and the FUS curve becomes flatter. Figure 4.3 depicts the FUS curves for the both I-optimal and D-optimal test plans with 5 inspection intervals. We notice that the curve of I-optimal test plan is uniformly lower than the curve of D-optimal test plan, which indicates that this test plan is well behaved by keeping the prediction variance low over the entire use stress region.

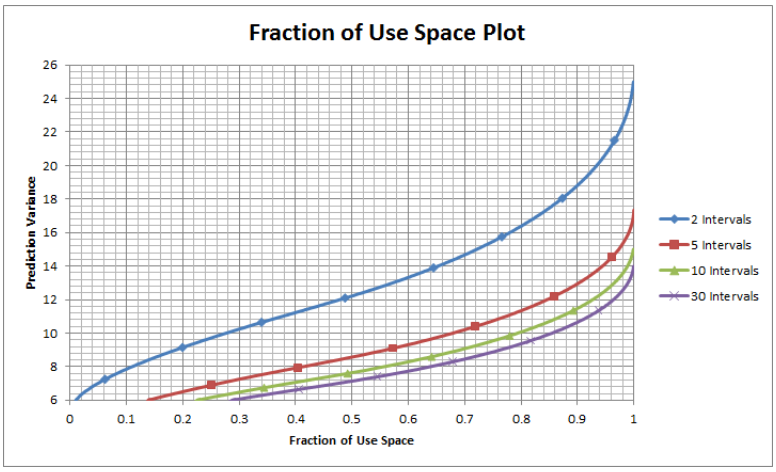

Figure 4.2. FUS Plots of I-optimal test plans with different intervals

In addition, the dual-objective designs in Table 4.2 are evaluated by their FUS plots, as shown in Figure 4.4. As expected, when the weight value in (4.6) increases the prediction variance of the FUS curve decreases.

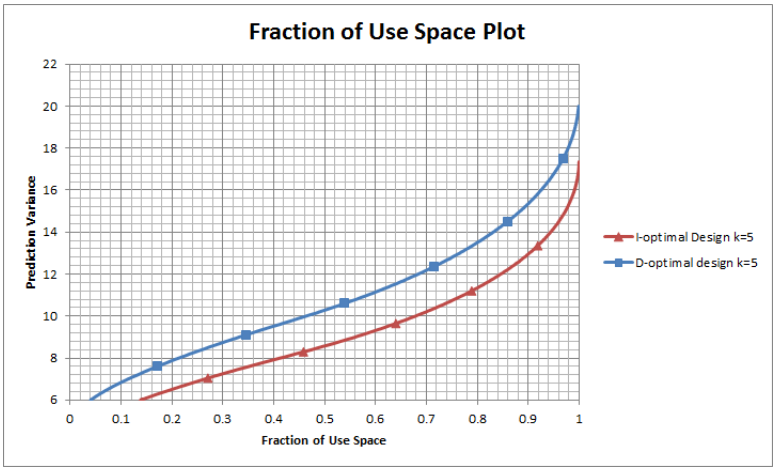

Figure 4.3. FUS Plots of I-optimal test plan and D-optimal test plan

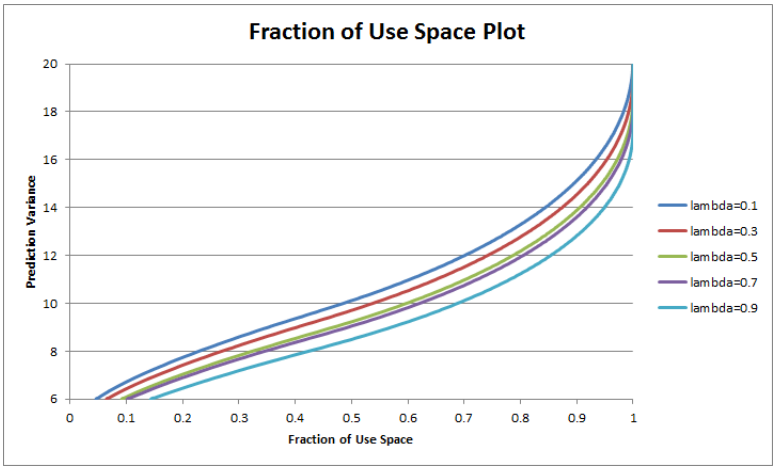

Figure 4.4. FUS Plots of dual objective design

Efficiency plot

To evaluate the D-efficiency and I-efficiency simultaneously, we plot them against the λ value in the dual-objective optimization formulation. At the two extremes when $\lambda = 0$ and $\lambda = 1$, the test plan is basically a D-optimal test plan and I-optimal test plan, respectively. However, for any other λ values the efficiency plot will illustrate the trade-off between the two optimality criteria for the optimal test plan. Table 4.3 gives both efficiency values for the previous example with various weight values and with five

inspection intervals. As one can see from Figure 4.5, the D-efficiency decreases more rapidly than the increase of I-efficiency when the weight changes from 0 to 1. Therefore, an I-optimal test plan may not be desirable to experimenters as it possesses a poor ability of estimating model parameters. Based on the experimenter's goal, a test plan that carefully balances its prediction and estimation abilities can be found from this plot.

Table 4.3. I-efficiency and D-efficiency of dual-objective designs with different weights

λ	Det	D-eff	PV	I-eff
0	7696.96	1.00	10.80	0.83
0.1	7523.92	0.98	10.63	0.84
0.2	7555.80	0.98	10.43	0.86
0.3	7195.92	0.93	10.21	0.88
0.4	6818.83	0.89	10.06	0.89
0.5	6331.19	0.82	9.71	0.92
0.6	6263.95	0.81	9.51	0.94
0.7	5928.14	0.77	9.52	0.94
0.8	5164.94	0.67	9.23	0.97
0.9	3855.65	0.50	8.95	0.99
1.0	3031.50	0.39	8.93	1.00

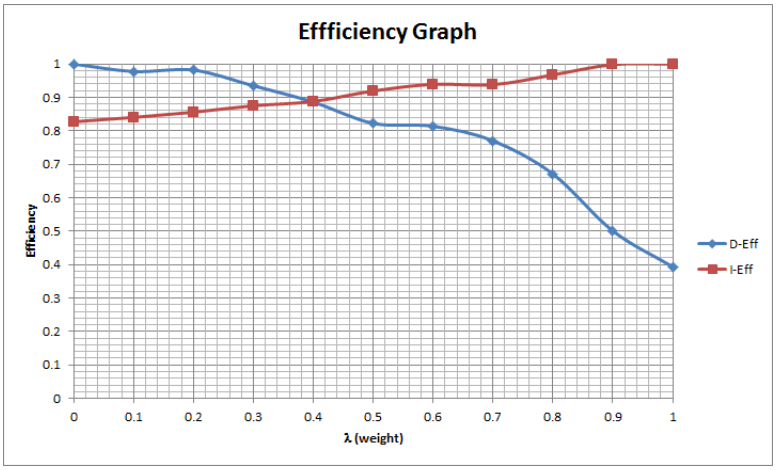

Figure 4.5. Efficiency plots of multi-objective designs

Pareto frontier

A test plan may not achieve two optimalities at the same time, but we can find a plan that is not dominated by any other plans with both optimal criteria. Such solutions

form the the Pareto frontier and any other solution that is not on the frontier can be improved by moving it toward the frontier. The Pareto frontier can be found by solving the combined optimization problem of (4.6) with different weights and selecting non-dominated solutions. In this section, we use the Pareto frontier to compare our dual-objective design with Meeker & Escobar (1998)'s approach of optimal split design (Table 20.5, page 553). The ALT experimentation is performed on a type of insulation with two stress factors – voltage and temperature. The main objective of the experiment is to predict a low percentile of failure time of the insulation at its use stress condition, which is $120^{\circ}C$ and 80 volts. The highest stress level can be set as $260^{\circ}C$ and 200 volts. It is assumed that a Weibull distribution with the shape parameter of 1.485 is good for the failure time distribution and the location parameter is a function of the two stress factors.

When the lifetime-stress function is fully specified, it is convenient to treat the linearly combined multiple stresses as one stress factor. By Meeker and Escobar's approach, one will first find the lowest test stress level of the combined stress factor and the test unit allocation on this level (as the highest stress level is always used), and then these test units can be split to two design points that are located on the edges of design region and have the same expected product lifetime as the low testing stress condition. The optimal split is determined by the secondary objective, i.e., to maximize the estimation property of the test plan. Thus, a 3-design point test plan is obtained with one design point located at the highest stress levels of both stress factors and the other two design points at two boundary lines of the design region. Figure 4.6 plots the optimal split design for the example above.

Using the dual-objective optimization we find the Pareto frontier of this problem. As shown in Figure 4.7, the optimal split solution is not on the frontier, so it is possible to improve this test plan by either increasing its D- or I-efficiency without sacrificing the other type of efficiency. In particular, we find that the estimation ability of the optimal

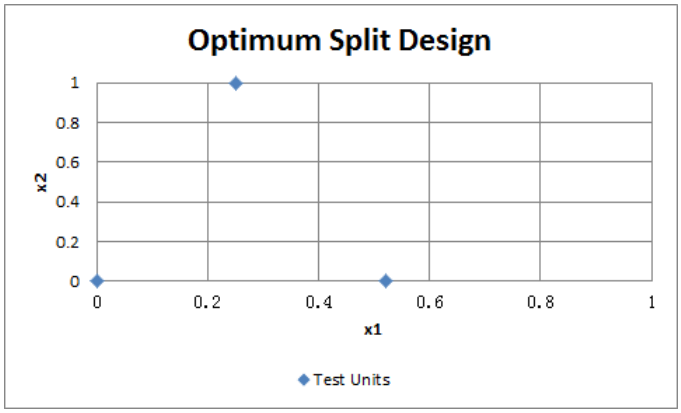

Figure 4.6. Optimal split design for the example in Meeker and Escobar (1998) (Table 20.5, page 553)

split test plan is poor. We can easily improve it without compromising the prediction ability of the test plan. Two such plans, along with the optimal split plan are presented in Table 4.4. The test unit allocations and the contours of prediction variance of these plans are shown in Figure 4.8. Note that the use condition in this example is a single point, so the dual-objective designs is based on D-optimality and U_c-optimality which is a special case of I-optimal design. Our test plans have four testing conditions, instead of three as in the optimal split plan, thus they cover a larger portion of the experimental design region and have a better parameter estimation property.

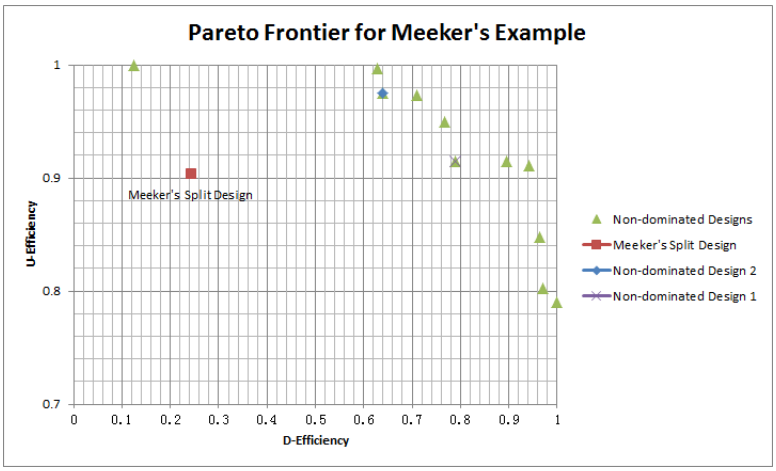

Figure 4.7. Pareto frontier

Table 4.4. Optimal test plans for the example in Meeker and Escobar (1998)

Test Plan	Test Cond i	Voltage Natural Volts	Voltage Std x_1	Temperature Natural °C	Temperature Std x_2	Alloc n_i	Det	PV
Opt. Split	1	124.00	0.522	260	0.000	36	1422	0.3100
	2	159.05	0.250	120.00	1.000	25		
	3	200.00	0.000	260.00	0.000	39		
Non-dominated 1	1	200.00	0.000	260	0.000	22	4864	0.2946
	2	200.00	0.000	120.00	1.000	4		
	3	173.36	0.156	120.00	1.000	27		
	4	132.30	0.451	260.00	0.000	47		
Non-dominated 2	1	200.00	0.000	260	0.000	20	3662	0.2853
	2	200.00	0.000	120.00	1.000	1		
	3	168.82	0.185	120.00	1.000	31		
	4	130.25	0.468	260.00	0.000	48		

4.5 Conclusion

In this study we investigate the PH model-based approach to optimal ALT planning. We consider both the censoring of failure times and the varying use stresses. To provide experimenters a set of alternative test plans with desirable model prediction and model estimation properties, we propose a dual-objective optimization process for finding the best plan. The FUS plot, efficiency plot, and Pareto frontier plot are developed to evaluate the prediction and estimation abilities of a test plan and to help experimenters compare the competing plans. The examples used in this chapter demonstrate the advantages of our approach over the conventional approach to optimal ALT planning.

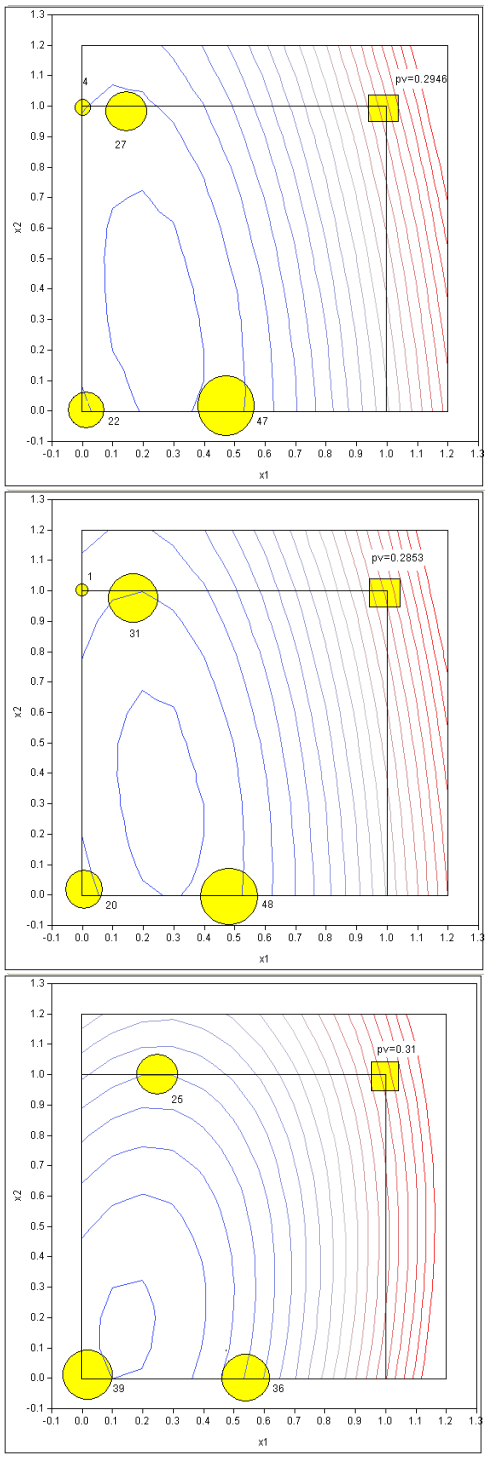

Figure 4.8. Two dual-objective test plans with $\lambda = 0.62$ (top) and $\lambda = 0.80$ (middle) and the optimal split test plan (bottom)

Chapter 5

Optimal design of ALT plans for acceleration model checking

5.1 Introduction

Background and motivation

Accelerated Life Testing (ALT) is a popular method to shorten the products' lifetime. In ALT, certain stress factors, such as temperature, humidity, voltage, are set to be at high stress levels so more failures from test units can be observed in limited testing period. For ALT, censoring is a important feature for failure data collection and two most popular ones are right censoring and interval censoring. For right censoring case, ALT is terminated at a pre-determined time and exact failure time of test units can be observed for those who fail before the pre-determined time and the failure time of the rest of test units are censored. For the interval censoring case, ALT is also terminated at a pre-determined time but the testing period is divided into several time intervals and only the number of failures from each interval can be counted. For these two censoring mechanisms, the advantage of right censoring is that exact failure time are available which can bring more information from the experimental design, but the shortcoming is also obvious because it is hard to implement in reality due to the difficulty of exact failure time observation; in contrast, interval censoring only needs to count the failure number from each interval and is easier to implement but the side effect is less information are available to find the optimal designs.

The conventional way of developing an optimal ALT plan is to formulate the likelihood function and then derive the expected information matrix for test planning. When different censoring methods are applied, the likelihood function becomes complicated and is hard to obtain. In order to find a better way to solve this problem,

T. Yang & Pan (2013), Pan & Yang (2013) suggested using the proportional hazard (PH) model for a failure time distribution with stress variables. The PH model is semiparametric and so is more flexible than the traditional failure time regression model. Given the proportional hazard function, the total likelihood function of censored data can be formulated by using a generalized linear model (GLM) formulation. The GLM was originally developed by Nelder & Wedderburn (1972) and work from McCullagh & Nelder (1989), and Myers et al. (2002) have given more details about it. The GLM formulation usually has three parts: a distribution from the exponential family; a linear predictor; a link function.

In order to plan an optimal design with the GLM formulation we developed, different statistical efficiency criteria may be applied. For example, D-optimality, A-optimality, G-optimality, I-optimality, etc and the definition of these optimal criteria could be found in Myers et al. (2009).

The optimal designs we got are local optimal because of the nonlinear relationship between the response variable and stress variables. Therefore, the accuracy of the optimal designs depend on the estimation of model parameters and model structures in the linear predictor of a GLM. Most of the time, only limited information is available for the experimenter about the models and a best guess based on historical data should be made. For example, for an ALT with two stress variables the full second order model includes intercept, two main effects, one interaction, and two squared terms. The terms in the reduced model are a subset of the full model which may only have intercept, two main effects, and interaction. Therefore, we want to develop a design which can distinguish the two competing models. D_s-optimality is a good choice to do model checking and it focuses on estimating a subset of the parameters as precisely as possible. However, Atkinson et al. (2007) mentioned that the potential disadvantage of D_s-optimality is that the effort is concentrated on checking whether the reduced model is true, rather than on

estimating the parameters of that model. In order to handle this problem, a dual objective design of D-optimality and D_s-optimality could be considered here to trade off efficiency of estimation of the parameters in the reduced model against efficiency for the model checking parameters in the full model but not in the reduced model. In addition, a lot of times only limited information is available for the model so a Bayesian design could be useful.

In this chapter, the previous research work about model checking and discrimination are briefly reviewed first. Then several examples of designs with a full model and reduced models by the GLM approach are given and comparisons are made to show the importance of how different models would affect the designs and their features. Next, details of D_s-optimal design and dual objective design are discussed. And last, a Bayesian design for model checking and estimation is discussed.

Previous work

The details of how to develop optimal designs under ALT with different censoring situations have been discussed in T. Yang & Pan (2013), Pan & Yang (2013). Examples of D-optimal designs, U_c-optimal design, and I-optimal designs, etc. are shown and some tools for design evaluations are developed. Atkinson & Fedorov (1975a,b) described the experimental designs for discriminating between rival regression models, especially for T-optimal designs. Hill (1978) reviewed the several methods for model discrimination designs including Box and Hill's procedure, Fedorov's procedure, and Atkinson's procedure. Jones et al. (2007) proposed several criteria including SA, MPD, EPD for gauging the capability of a design for model discrimination. Agboto et al. (2010) discussed the existing methods like T-optimality and several new criteria to construct optimal two-level model discriminating designs for screening experiments. However, the

previous two papers only discussed the linear model situation. Dette & Titoff (2009) derived several new properties of optimal designs with respect to the T-optimality and also demonstrated that in nested linear models the number of support points of T-optimal designs is usually too small to estimate all parameters in the full model. Biedermann et al. (2011) developed an optimal design theory for additive partially nonlinear regression models and generalized their results to parameter robust optimality criteria, called Bayesian and standardised maximin D-optimality. DeLeon & Atkinson (1991) used numerical methods to find non-sequential optimal designs, which can be used for both the construction of designs and for checking the optimality of proposed designs. Chaloner (1984) discussed optimal Bayesian experimental designs for estimation and prediction in linear models. An optimal Bayesian design for nonlinear problem with a single explanatory variable is considered in Chaloner (1993) and a literature review on Bayesian experimental design is done by Chaloner & Verdinelli (1995). DuMouchel & Jones (1994) modified D-optimality with the Bayesian paradigm and handled the dependent problem on assumed models. Another interesting work from Waterhouse et al. (2006) considered a problem with two rival GLMs for a binomial response and compared the designs based on four different optimal criteria.

5.2 Optimal designs with different models

In T. Yang & Pan (2013) and Pan & Yang (2013), the development of the GLM approach for finding optimal designs with right censoring and interval censoring has been discussed. However, in previous research they supposed the model of linear predictor is the 'best guess' they can get and did not compare the designs based on different models. In this section, examples of U_c-optimal design under ALT with right censoring are given and different models are chosen to develop the experiments.

The background setting of the example is from T. Yang & Pan (2013) and for the

convenience, we also describe it here. Suppose there is an electronic part whose lifetime belongs to an exponential distribution and its lifetime is affected by temperature and humidity. The use condition of this electronic part is set as $30°C$ and 25%. Under the ALT test, the temperature ranges from $60\ °C$ to $110°C$ and the relative humidity level ranges from 60% to 90%. The natural stress of temperature and humidity can be presented as $S_1 = 11605/T$, with temperature in degree Kelvin, and $S_2 = \log(h)$, with relative humidity as a percentage. Let the design space of this experimentation to be a unit square and the use condition to be located at the first quadrant. The transformation is given by $x_{ji} = \frac{S_{ji} - S_{jH}}{S_{jL} - S_{jH}}$ where S_{jH} and S_{jL} are the high and low level of stress j; thus, the highest stress level is transformed to (0, 0) and the lowest stress level is transformed to (1, 1). There are three linear predictors available for GLM and shown below,

$$\eta = \begin{cases} -4.086x_1 - 1.476x_2 \\ -4.086x_1 - 1.476x_2 + 0.01x_1x_2 \\ -4.086x_1 - 1.476x_2 + 0.01x_1x_2 + 0.03x_1^2 + 0.03x_2^2 \end{cases}$$

The first two models are reduced models and the last one is a full model and the generalized model form could be written as

$$\eta = \mathbf{X_r}\beta_\mathbf{r} + \mathbf{X_s}\beta_\mathbf{s} \tag{5.1}$$

where $\mathbf{X_r}$ and $\mathbf{X_s}$ are the design matrix of the required and the secondary terms respectively. $\beta_\mathbf{r}$ and $\beta_\mathbf{s}$ stand for the coefficient parameters for required and secondary terms respectively.

The I-optimality is defined as

$$\xi^* := \arg\min_{\xi} \frac{\int_\Omega \mathbf{x}_{use}{}' \cdot (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \cdot \mathbf{x}_{use} d\mathbf{x}_{use}}{S_\Omega}, \tag{5.2}$$

where Ω is the use condition region and S_Ω is the area of the use condition region and $\mathbf{X}$ is the design matrix.

Table 5.1. I-optimal design with only main effects under right censoring

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	90.00	0.000	28	0.99	0.48
2	81.64	0.533	60.00	1.000	67	0.54	
3	62.89	0.935	90.00	0.000	5	0.48	

Table 5.2. I-optimal design with main and interaction effects under right censoring

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	90.00	0.000	12	0.99	7.04
2	110.00	0.000	60.00	1.000	12	0.99	
3	67.82	0.825	60.00	1.000	52	0.21	
4	62.00	0.955	90.00	0.000	24	0.45	

Table 5.3. I-optimal design with main, interaction, and quadratic effects under right censoring

Test Cond i	Temperature Natural ° C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	FP	PV
1	110.00	0.000	90.00	0.000	15	0.99	46.76
2	110.00	0.000	71.95	0.522	9	0.99	
3	93.34	0.303	74.14	0.449	48	0.98	
4	88.87	0.389	60.00	1.000	11	0.75	
5	88.00	0.406	90.00	0.000	1	0.98	
6	72.85	0.716	60.00	1.000	16	0.31	

Now suppose the temperature of use condition varies between 1.458 to 2.058 and the humidity of use condition varies varies between 2.859 to 3.459 in coded variables. The failures are right censored and the censoring time is 30 hours. Now based on the three different models of η, the results of I-optimal designs are given in Tables 5.1 to 5.3. From Tables 5.1 to 5.3, we can see that the three I-optimal designs have 3, 4, and 6 design locations respectively and this is due to the models which have 3, 4, and 6 parameters (including intercept). However, if the true model is the first model which has only main effects, what kind of effects will be brought by the designs from the wrong models under true model settings? Here FUS (Fraction of Use Space) plots in Figure 5.1 will be given in order to show the difference. A FUS plot is a graphical tool which plots the prediction

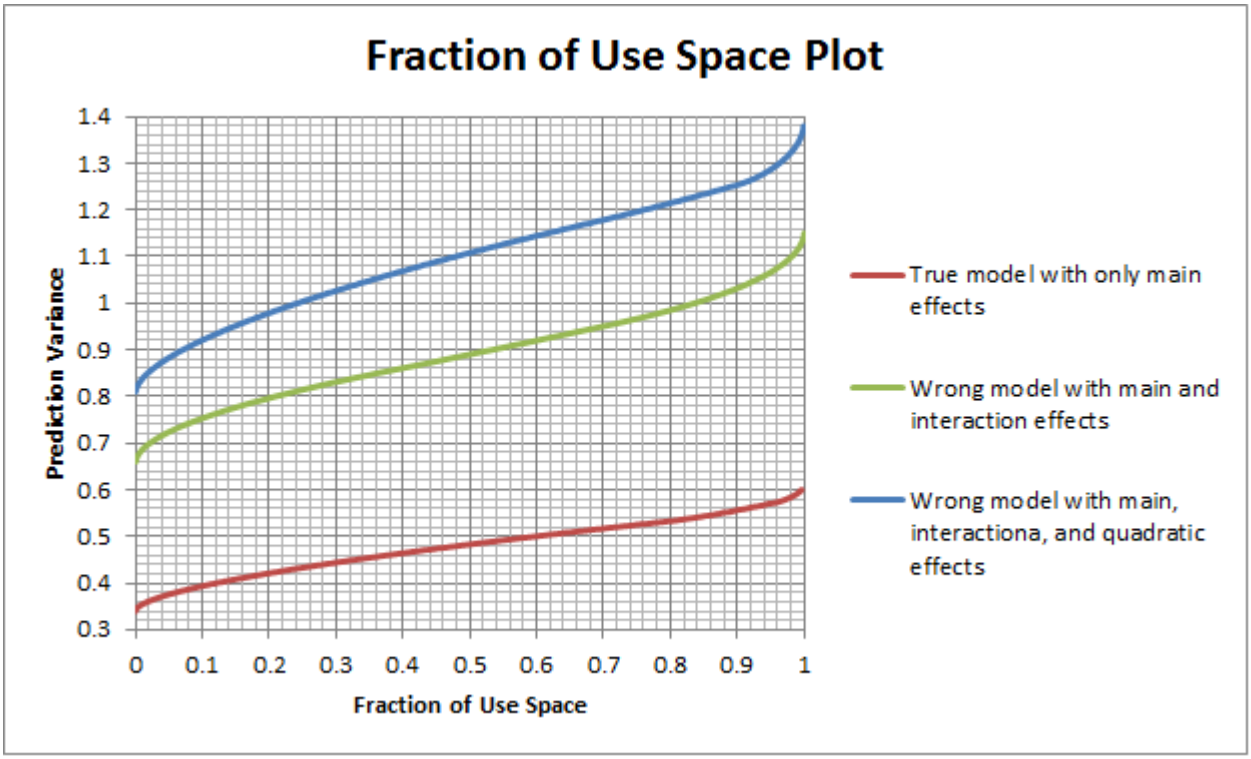

Figure 5.1. FUS plots with different models

variance versus the fraction of use stress region that has less than or equal to this prediction variance value. More details about FUS plot can be found in Pan & Yang (2013). From Figure 5.1 we know that wrong models will affect the prediction accuracy in the use region. Also if the wrong model has more or fewer terms compared with the true model, then the prediction power would be less accurate and this can be proved by the curve of the wrong model with main, interaction, and quadratic effects is higher than the curve of wrong model with main, and interaction effects.

In addition, extrapolation is also an important feature of ALT. The extrapolation will exaggerate the difference between two models as the use condition moving away from use region. Now suppose we use the last two linear predictors of η in this section and give their mean time to failure (hours) based on three different use conditions. The results are shown in Table 5.4 and it is obvious that extrapolation will be a serious

Table 5.4. D_s-optimal design: checking x_2, x_2^2

Model No.	Mean time to failure		
	use condition (1, 1)	use condition (1.3, 2)	use condition (1.758, 3.159)
2	257.75	3781.21	89164.78
3	242.74	3187.84	131970.90

problem when use condition is far away from design region and wrong model is used. Therefore, it is important to choose the right model before developing the optimal design and this is a motivation to find some model checking and discrimination designs to distinguish rival models.

5.3 D_s-optimal design

When the experimenters have interest in estimating a subset of s parameters as precisely as possible, the D_s-optimal design is a good choice. Suppose we have a linear predictor that can be written as

$$\eta = \mathbf{X}\beta = \mathbf{X_1}\beta_1 + \mathbf{X_2}\beta_2 \tag{5.3}$$

And the information matrix of the design can be partitioned as

$$M(\xi) = \begin{bmatrix} M_{11}(\xi) & M_{12}(\xi) \\ M'_{12}(\xi) & M_{22}(\xi) \end{bmatrix},$$

where $M_{11}(\xi)$ and $M_{22}(\xi)$ are the information matrices corresponding to the subsets 1 and 2. The D_s-optimality criterion under GLM formulation to check parameters subset β_2 could be defined as

$$\xi^* := \max_{\xi} |\mathbf{X'_2 W X_2} - \mathbf{X'_2 W X_1}(\mathbf{X'_1 W X_1})^{-1}\mathbf{X'_1 W X_2}| = \frac{|\mathbf{M}(\xi)|}{|\mathbf{M_{11}}(\xi)|} \tag{5.4}$$

Now suppose we have 3 different scenarios which the pairs of full model and reduced model can be written as

$$Scenario1 \begin{cases} \text{reduced model: } -4.086x_1 - 1.476x_2 + 0.01x_1x_2 \\ \text{full model: } -4.086x_1 - 1.476x_2 + 0.01x_1x_2 + 0.03x_1^2 + 0.03x_2^2 \end{cases}$$

$$Scenario2 \begin{cases} \text{reduced model: } -1.476x_2 + 0.01x_1x_2 + 0.03x_2^2 \\ \text{full model: } -4.086x_1 - 1.476x_2 + 0.01x_1x_2 + 0.03x_1^2 + 0.03x_2^2 \end{cases}$$

$$Scenario3 \begin{cases} \text{reduced model: } -4.086x_1 + 0.01x_1x_2 + 0.03x_1^2 \\ \text{full model: } -4.086x_1 - 1.476x_2 + 0.01x_1x_2 + 0.03x_1^2 + 0.03x_2^2 \end{cases}$$

Based on the three scenarios, we want to check the parameters in the reduced models and the three D_s-optimal designs are shown in Tables 5.5 to 5.7 and the designs are also illustrated in Figures 5.2 to 5.4.

From the Table 5.5 and Figure 5.2, we know that in order to check the quadratic effects x_1 and x_2 in the model, the D_s-optimal design allocates more test units at the center point of the all design boundaries and also the center of the design region which has the largest number of test units as 40. From Table 5.6 and Figure 5.3, almost all the test units are allocated on the x_1 axis in order to check the effects only brought by the x_1. Table 5.7 and Figure 5.4 is similar to the previous case but the only difference is that almost all the test units are allocated on x_2 axis. In general, the D_s-optimality is a good criteria to check the interesting terms in the model. However, the D_s-optimality focuses on checking whether the reduced model is true, rather than on estimating the parameters of the model. Therefore in next section, a dual objective design criterion will be given to tackle this problem.

Table 5.5. D_s-optimal design: checking x_1^2, x_2^2

Test Cond i	Temperature Natural $^\circ$ C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i
1	110.00	0.000	90.00	0.000	4
2	110.00	0.000	73.50	0.470	10
3	110.00	0.000	60.00	1.000	2
4	91.40	0.340	60.00	1.000	11
5	88.82	0.390	74.11	0.450	40
6	86.78	0.430	90.00	0.000	13
7	60.00	1.000	90.00	0.000	3
8	60.00	1.000	76.58	0.370	14
9	60.00	1.000	60.00	1.000	3

Table 5.6. D_s-optimal design: checking x_1, x_1^2

Test Cond i	Temperature Natural $^\circ$ C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i
1	110.00	0.000	90.00	0.000	24
2	110.00	0.000	74.72	0.430	1
3	104.88	0.090	60.00	1.000	1
4	84.27	0.480	90.00	0.000	41
5	74.54	0.680	60.00	1.000	1
6	60.00	1.000	90.00	0.000	23

Table 5.7. D_s-optimal design: checking x_2, x_2^2

Test Cond i	Temperature Natural $^\circ$ C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i
1	110.00	0.000	90.00	0.000	26
2	110.00	0.000	77.21	0.350	1
3	110.00	0.000	72.60	0.500	46
4	110.00	0.000	60.00	1.000	25
5	79.82	0.570	77.21	0.350	1
6	66.68	0.850	90.00	0.000	1

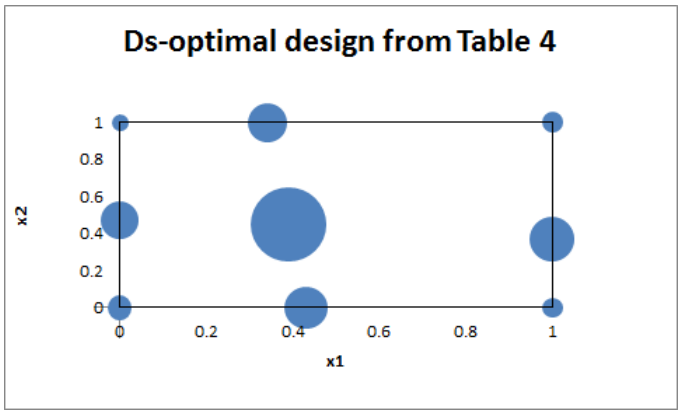

Figure 5.2. D_s-optimal design: checking x_1^2 and x_2^2

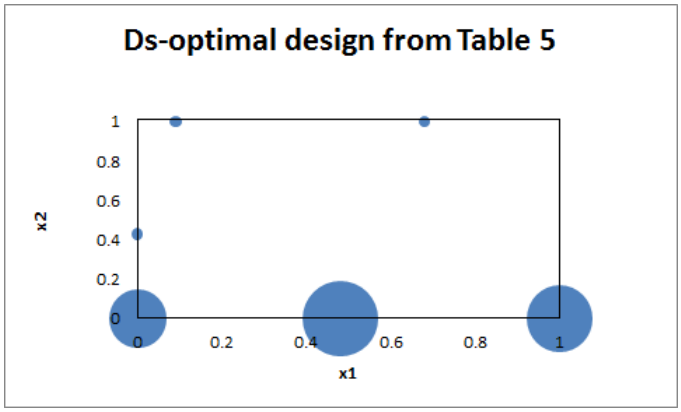

Figure 5.3. D_s-optimal design: checking x_1 and x_1^2

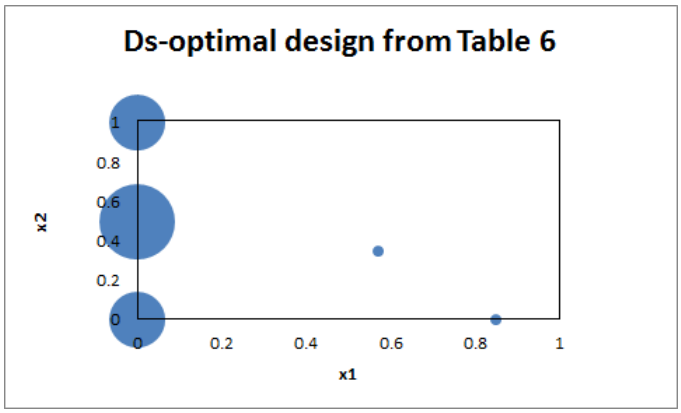

Figure 5.4. D_s-optimal design: checking x_2 and x_2^2

5.4 Dual objective design

In previous section, D_s-optimal designs have been discussed and are good choices for checking the interesting subset of model parameters. However, in equation (5.1) if we already know that the required terms $\mathbf{X_r}$ are important in the model and want to check the effects of secondary terms $\mathbf{X_s}$, D_s-optimality may no longer be a good option. Therefore, a dual objective design of D- and D_s-optimalities are introduced here. The D-optimality is used to estimate the required terms in the model and D_s-optimality is used to check the potential terms in the model. The dual objective design was described by Atkinson et al. (2007) and the criterion is defined as

$$\begin{aligned} \xi^*: &= \arg\max_{\xi}[\frac{\kappa}{r}\log|M_{11}(\xi)| + \frac{1-\kappa}{s}\{\log|M(\xi)| - \log|M_{11}(\xi)|\}] \\ &= \arg\max_{\xi}[\frac{(p\kappa - r)}{rs}\log|M_{11}(\xi)| + \frac{1-\kappa}{s}\log|M(\xi)|] \end{aligned} \tag{5.5}$$

where r is the number of required terms in the full model and s is the number of secondary terms and $p = r + s$. κ is between 0 and 1.

In addition, in the first row of equation (5.5), the first part is for the D-optimal design with the required terms and the second part is for the D_s-optimal design of the secondary terms. Now we use the first scenario from the last section to develop three dual objective designs with $\kappa = 0.2, 0.5, 0.8$. The three designs are given in Tables 5.8 to 5.11 and they are also plotted in Figures 5.5 to 5.7.

From equation (5.5), we know that with the increase of κ, the focus of dual objective design shifts from D_s-optimal design to D-optimal design. There are three special situations for dual objective designs which are D_s-optimal design when $\kappa = 0$; D-optimal design when $\kappa = 1$ for β_r; and D-optimal design for β when $\kappa = \frac{r}{p}$. From Figures 5.5 to 5.7, when $\kappa = 0.2$, more test units are allocated at the center of each boundary and the center of the design region. This is due to the strong emphasis on

Table 5.8. Dual objective design with $\kappa = 0.2$

Test Cond i	Temperature Natural $^\circ$C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	D_s-eff	D-eff
1	110.00	0.000	90.00	0.000	5	1.00	0.49
2	110.00	0.000	72.90	0.490	9		
3	110.00	0.000	60.00	1.000	3		
4	90.37	0.360	60.00	1.000	12		
5	88.82	0.390	74.11	0.450	35		
6	85.77	0.450	90.00	0.000	13		
7	60.00	1.000	90.00	0.000	6		
8	60.00	1.000	75.02	0.420	13		
9	60.00	1.000	60.00	1.000	4		

Table 5.9. Dual objective design with $\kappa = 0.5$

Test Cond i	Temperature Natural $^\circ$C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	D_s-eff	D-eff
1	110.00	0.000	90.00	0.000	8	1.00	0.63
2	110.00	0.000	72.90	0.490	11		
3	110.00	0.000	60.00	1.000	8		
4	88.82	0.390	60.00	1.000	9		
5	88.31	0.400	73.50	0.470	28		
6	85.27	0.460	90.00	0.000	12		
7	60.00	1.000	90.00	0.000	10		
8	60.00	1.000	74.41	0.440	6		
9	60.00	1.000	60.00	1.000	8		

Table 5.10. Dual objective design with $\kappa = 0.8$

Test Cond i	Temperature Natural $^\circ$C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	D_s-eff	D-eff
1	110.00	0.000	90.00	0.000	8	0.66	0.82
2	110.00	0.000	74.11	0.450	11		
3	110.00	0.000	60.00	1.000	8		
4	86.78	0.430	72.90	0.490	9		
5	82.77	0.510	60.00	1.000	28		
6	82.28	0.520	90.00	0.000	12		
7	66.68	0.850	60.00	1.000	10		
8	60.00	1.000	90.00	0.000	6		
9	60.00	1.000	72.01	0.520	8		

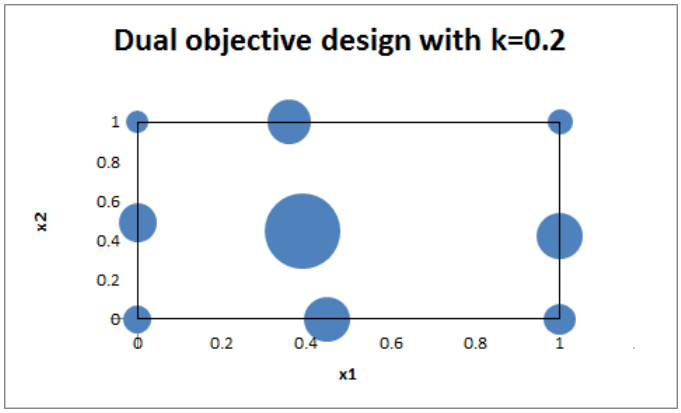

Figure 5.5. Dual objective design with $\kappa = 0.2$

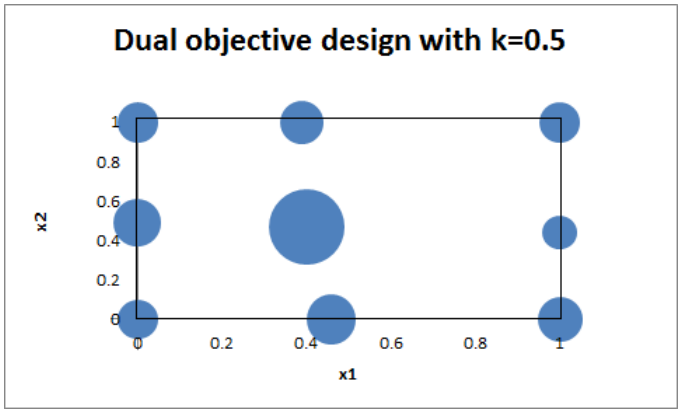

Figure 5.6. Dual objective design with $\kappa = 0.5$

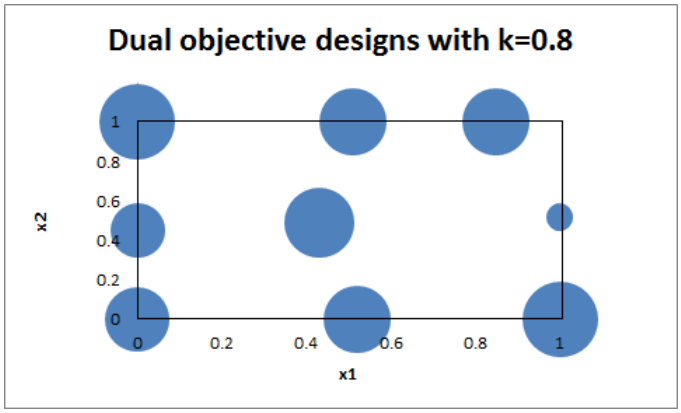

Figure 5.7. Dual objective design with $\kappa = 0.8$

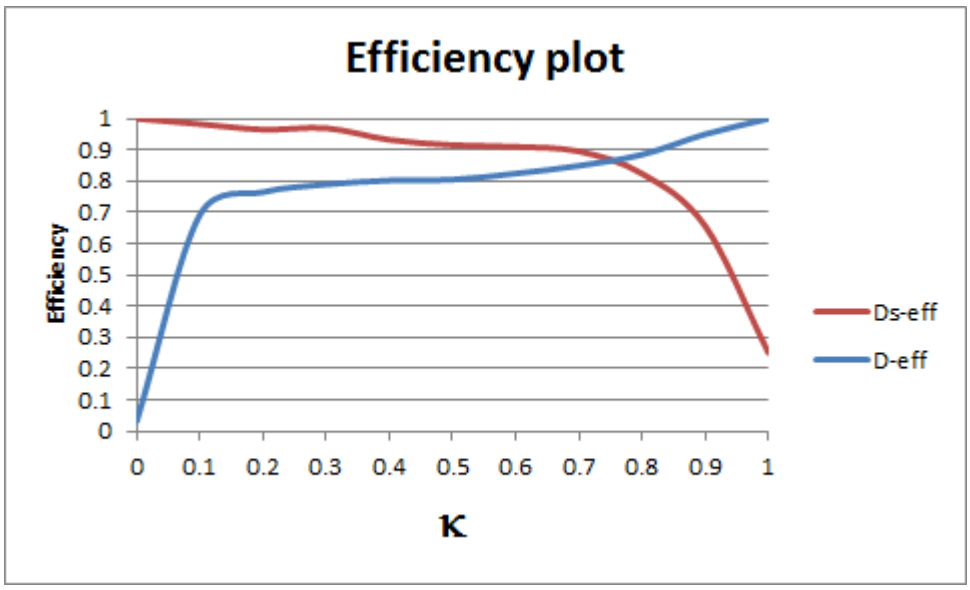

Figure 5.8. Efficiency plot

D_s-optimality to check the quadratic terms x_1^2 and x_2^2. When $\kappa = 0.5$, all 9 different locations almost have received similar number of test units except the center point of the design region and this design is a more balanced one. For $\kappa = 0.8$, more test units are allocated at four corners and it is caused by the D-optimality trying to estimate β_r. In order to better illustrate the effects of κ, a efficiency plot of D_s-efficiency for β_s and D-efficiency for β_r is given in Figure 5.8. From this figure, we know that when κ is greater than 0.3, the D-efficiency will be high and when κ is less than 0.8, the D_s-efficiency will be high. Therefore, the κ value could be chosen between 0.3 and 0.8 to get a good estimation of required terms and check the potential terms at the same time.

Meeker & Escobar (1998) discussed the statistical optimum design and introduced the compromise design which uses more stress levels and reduces the statistical efficiency. The benefit of the compromise designs is stated in Meeker & Escobar (1998) and they tend to be more robust to misspecification of unknown inputs and they allow one to estimate model parameters even if there are no failures at one level of the accelerating variables. Since the dual objective design introduced in this chapter is also a robust design

Table 5.11. Dual objective design with $\kappa = 0.8$

Design	Test Cond i	Voltage Natural % C	Voltage Std x_1	Temperature Natural $^\circ$	Temperature Std x_2	Alloc n_i	D_s-eff	D-eff	PV
compromise Split	1	124.00	0.520	260.00	0.000	27	0.36	0.65	3.23
	2	145.00	0.350	260.00	0.000	12			
	3	170.00	0.180	260.00	0.000	16			
	4	158.00	0.260	120.00	1.000	27			
	5	185.00	0.090	120.00	1.000	8			
	6	200.00	0.000	158.00	0.660	10			
Dual Objective $\kappa = 0.2$	1	200.00	0.000	260.00	0.000	9	0.99	0.71	1.18
	2	200.00	0.000	120.00	1.000	15			
	3	153.33	0.290	146.44	0.760	52			
	4	131.21	0.460	260.00	0.000	24			
Dual Objective $\kappa = 0.8$	1	200.00	0.000	260.00	0.000	16	0.95	0.89	1.61
	2	200.00	0.000	120.00	1.000	22			
	3	168.04	0.190	120.00	1.000	13			
	4	162.00	0.230	133.90	0.870	22			
	5	134.87	0.430	260.00	0.000	27			

to the unknown parameters, we will use the example of a two stress factor ALT on page 555 from Meeker & Escobar (1998) and applied the same settings to generate several dual objective designs and compare them with the compromise design. The dual objective designs consider if the interaction term should be included or not in the model. The results can be seen from Table 5.11. From the result of the comparison, dual objective designs have better D-efficiency and D_s-efficiency and smaller prediction variances than the compromise design. Therefore, dual objective design is a good option to find a robust experimental design when uncertainty exists in the model.

5.5 Parsimonious model checking design

In this section, a Bayesian method of model checking will be discussed and is originally introduced by DuMouchel & Jones (1994). Atkinson et al. (2007) gave a more detailed explanation for this method. In a lot of situations, the experimenters are pretty sure about the main and interaction effects in the model but not quite sure about if the quadratic terms should be included in the model or not. Therefore, the parameters β of the

model could be divided into two subgroups. One is β_r for the required terms and the other is β_s for secondary terms. Now the joint prior distribution of the all parameters is

$$\beta \sim N\left(\begin{bmatrix} \beta_r \\ 0_s \end{bmatrix}, \begin{bmatrix} \gamma^2 I_r & 0_{r\times s} \\ 0_{s\times r} & \tau^2 I_s \end{bmatrix}\right) = N\left(\begin{bmatrix} \beta_r \\ 0_s \end{bmatrix}, D(\beta)\right) \tag{5.6}$$

In order to design experiments, we require the prior information matrix for β, that is we require the inverse of the dispersion matrix $D(\beta)$. As $\gamma \to \infty$

$$\{D(\beta)\}^{-1} \to \frac{1}{\tau^2}K, \tag{5.7}$$

where

$$K = \begin{pmatrix} 0_{r\times r} & 0_{r\times s} \\ 0_{s\times r} & I_s \end{pmatrix} \tag{5.8}$$

So the posterior information matrix for β, given design ξ, is

$$\widetilde{M}(\xi) = N_0 K + NM(\xi) \tag{5.9}$$

where $N_0 = \frac{\sigma^2}{\tau^2}$ and this equation can be normalized as

$$M_\alpha(\xi) = (1-\alpha)K + \alpha M(\xi) \tag{5.10}$$

where

$$\alpha = \frac{N}{N+N_0} = \frac{N\tau^2}{\sigma^2 + N\tau^2}. \tag{5.11}$$

Now for our GLM case, we only need to modify the information matrix $M(\xi)$ and then maximize the determinant of $M_\alpha(\xi)$ will give us a design for model checking. In Tables 5.12 to 5.14 and Figures 5.9 to 5.11, three Bayesian D-optimal designs with $\alpha = 0.1, 0.5, 0.9$ are given. The DB values in the tables mean that the determinant values of the designs.

From the results we notice that when α is approaching 0, the design is more like a D-optimal design for the required terms in the model and when α is approaching to 1,

Table 5.12. Bayesian D-optimal design with $\alpha = 0.1$

Test Cond i	Temperature Natural $^\circ$ C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	DB
1	110.00	0.000	90.00	0.000	20	1.99
2	110.00	0.000	60.00	1.000	19	
3	76.44	0.640	60.00	1.000	35	
4	67.14	0.840	90.00	0.000	26	

Table 5.13. Bayesian D-optimal design with $\alpha = 0.5$

Test Cond i	Temperature Natural $^\circ$ C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	DB
1	110.00	0.000	90.00	0.000	12	569.49
2	110.00	0.000	74.41	0.440	7	
3	110.00	0.000	60.00	1.000	17	
4	87.80	0.410	72.90	0.490	14	
5	80.31	0.560	90.00	0.000	11	
6	80.31	0.560	60.00	1.000	3	
7	73.60	0.700	60.00	1.000	21	
8	60.00	1.000	90.00	0.000	15	

Table 5.14. Bayesian D-optimal design with $\alpha = 0.9$

Test Cond i	Temperature Natural $^\circ$ C	Std x_1	Humidity Natural %	Std x_2	Alloc n_i	DB
1	110.00	0.000	90.00	0.000	11	4100.8
2	110.00	0.000	73.20	0.480	10	
3	110.00	0.000	60.00	1.000	13	
4	87.29	0.420	72.90	0.490	20	
5	84.77	0.470	60.00	1.000	8	
6	83.77	0.490	90.00	0.000	13	
7	64.44	0.900	60.00	1.000	10	
8	60.00	1.000	90.00	0.000	13	
9	60.00	1.000	72.31	0.510	2	

then the design becomes a D-optimal design for all the terms. This explains that when $\alpha = 0.1$, there are only four locations for test units and it is similar to the design with only main and interaction effects. On the other hand, once the α is over 0.5, then the designs reduced test units from the corners of design region and added them to the boundaries and center which can be used to estimate the potential quadratic terms. Therefore, the Bayesian D-optimal criteria is a good way to find a model checking design with some prior information.

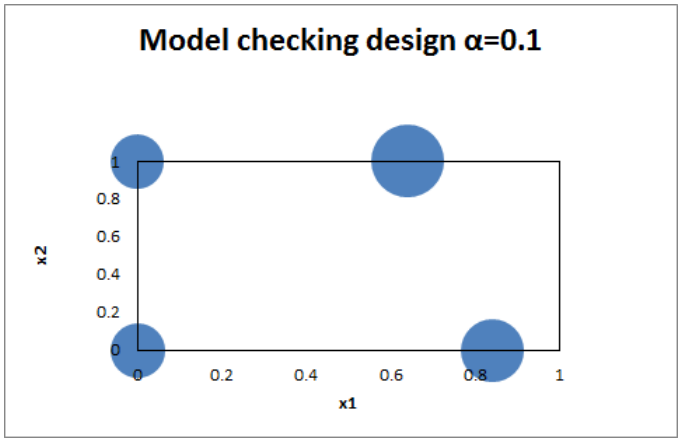

Figure 5.9. Dual objective design with $\kappa = 0.2$

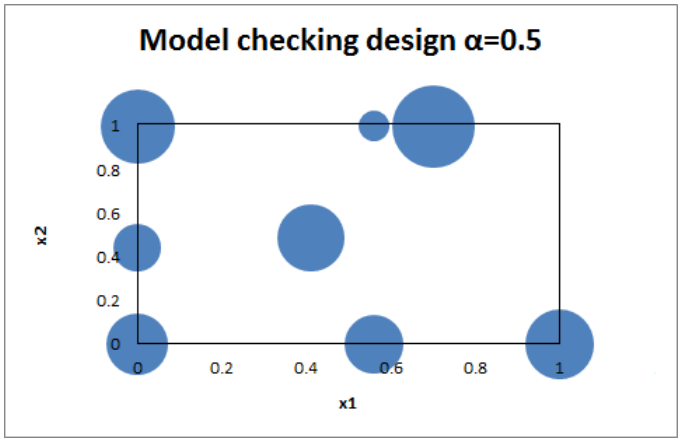

Figure 5.10. Dual objective design with $\kappa = 0.5$

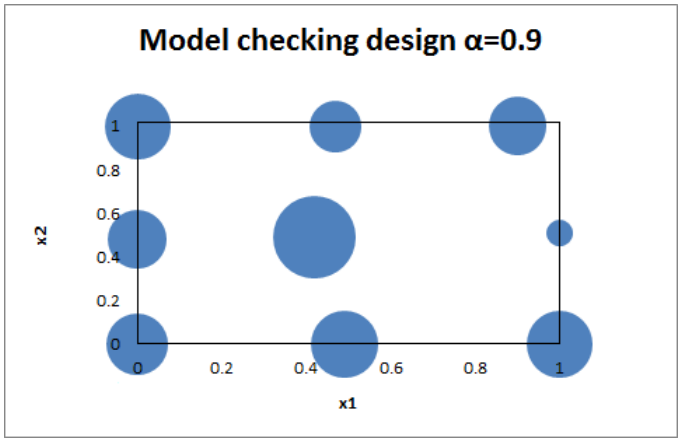

Figure 5.11. Dual objective design with $\kappa = 0.8$

Chapter 6

CONCLUSIONS AND FUTURE RESEARCH

In this dissertation, the D-optimal design, U_c-optimal design, and I-optimal designs have been developed by the GLM approach with right censoring and interval censoring. In addition, model sensitivity analysis, graphical tools for model evaluation, and dual objective design concept have been introduced. At last, model checking and discrimination are discussed.

Future work could be Bayesian design, more general distributions, and type-II censoring.

6.1 Conclusion

In Chapter 3, the GLM approach of finding the optimal design under ALT with right censoring and interval censoring is introduced. Some D-optimal designs, U_c-optimal designs, and I-optimal designs are generated in order to show the advantage of the GLM approach. Then the U_c-optimality is proved to be a robust criterion with the wrongly estimated model parameters. Chapter 4 starts with the discussion of the I-optimal designs which is an extension of the U_c-optimality. However, the U_c-optimality and I-optimality only consider the prediction variance in the use condition or use region. Therefore, a dual objective design is explained based on I- or U_c-optimality and D-optimality and the experimenters have to make trade-off between these optimal criteria. Now with several optimality criteria available, choose designs by applying graphical methods like the FUS, efficiency plot, and Pareto frontier would be straightforward. Model checking and discrimination are important when the experimenters are not sure about the model structure and this is addressed by the D_s-optimal design, dual objective optimal design with D- and D_s-optimality, and parsimonious design introduced in Chapter 5.

6.2 Future work

The research in this dissertation has highlighted several types of the optimal designs and the graphical methods for design evaluations associated with accelerated life testing. Some extension of this work include:

- deriving the U_c- or I-optimal criteria for other distributions such as lognormal;
- deriving the U_c- or I-optimal criteria for other censoring plans such as Type II censoring;
- developing an optimal criterion which generate robust designs to multiple censoring plans;
- incorporating Bayesian methods which assign the prior distributions to the model parameters and finding a robust design;
- constructing robust designs from the clusters of different designs corresponding to the different combinations of models.

REFERENCES

Agboto, V., Li, W., & Nachtsheim, C. (2010). Screening designs for model discrimination. *Journal of Statistical Planning and Inference*, *140*(3), 766 - 780.

Aitkin, M., & Clayton, D. (1980). The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using glim. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *29*(2), pp. 156-163.

Atkinson, A., Donev, A., & Tobias, R. (2007). *Optimum experimental designs, with SAS*. Oxford University Press, USA.

Atkinson, A., & Fedorov, V. (1975a). The design of experiments for discriminating between two rival models. *Biometrika*, *62*(1), pp. 57-70.

Atkinson, A., & Fedorov, V. (1975b). Optimal design: Experiments for discriminating between several models. *Biometrika*, *62*(2), pp. 289-303.

Barbosa, E., Colosimo, E., & Louzada-Neto, F. (1996, dec). Accelerated life tests analyzed by a piecewise exponential distribution via generalized linear models. *Reliability, IEEE Transactions on*, *45*(4), 619 -623.

Barbosa, E., & Louzada-Neto, F. (1994). Analysis of accelerated life tests with Weibull failure distribution via generalized linear models. *Communications in Statistics - Simulation and Computation*, *23*(2), 455-465.

Bennett, S. (1983). Log-logistic regression models for survival data. *Journal of the Royal Statistical Society. Series C*, *32*(2), 165-171.

Biedermann, S., Dette, H., & Woods, D. C. (2011). Optimal design for additive partially nonlinear models. *Biometrika*, *98*(2), 449 - 458.

Bohning, D. (1986). A vertex-exchange-method in D-optimal design theory. *Metrika*.

Borkowski, J. (1995). *New variance dispersion graphs for mixture experiments.*

Box, G., & Behnken, D. (1960). Some new three level designs for the study of quantitative variables. *Technometrics*, *2*, 455-476.

Box, G., Bisgaard, S., & Fung, C. (1988). An explanation and critique of Taguchi's contributions to quality engineering. *Quality and Reliability Engineering International*, *4*, 123-131.

Box, G., & Draper, N. (1959). A basis for the selection of a response surface design. *Journal of the American Statistical Association*, *54*(287), 622-654.

Chaloner, K. (1984). Optimal Bayesian experimental design for linear models. *The Annals of Statistics*, *12*(1), pp. 283-300.

Chaloner, K. (1993). A note on optimal Bayesian design for nonlinear problems. *Journal of Statistical Planning and Inference*, *37*(2), 229 - 235.

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, *10*(3), pp. 273-304.

Chernoff, H. (1962). Optimal accelerated life designs for estimation. *Technometrics*, *4*(3), pp. 381-408.

Chipman, H., & Welch, H. (1996). D-optimal design for generalized linear models. *unpublished*.

Collett, D. (2003). *Modelling survival data in medical research* (2nd éd.). Chapman and Hall/CRC.

Condra, L. W. (1993). *Reliability improvement with design of experiments*. Marcel Dekker.

Cook, R. D., & Nachtsheim, C. J. (1980). A comparison of algorithms for constructing exact D-optimal designs. *Technometrics*, *22*(3), 315-324.

DeLeon, P. d., & Atkinson, A. C. (1991). Optimum experimental design for discriminating between two rival models in the presence of prior information. *Biometrika*, *78*(3), pp. 601-608.

Dette, H., & Titoff, S. (2009). Optimal discrimination designs. *The Annals of Statistics*, *37*(4), pp. 2056-2082.

Dror, H. A., & Steinberg, D. M. (2006). Robust experimental design for multivariate generalized linear models. *Technometrics*, *48*(4), 520-529.

DuMouchel, W., & Jones, B. (1994). A simple Bayesian modification of D-optimal designs to reduce dependence on an assumed model. *Technometrics*, *36*(1), pp. 37-47.

Elsayed, E. A., & Jiao, L. (2002). Optimal design of proportional hazards based accelerated life testing plans. *International Journal of Materials and Product Technology*, *17*(5-6), 411-424.

Escobar, L. A., & Meeker, W. Q. (1995). Planning accelerated life tests with two or more experimental factors. *Technometrics*, *37*(4), pp. 411-427.

Fedorov, V. (1989). Optimal design with bounded density: Optimization algorithms of the exchange type. *Journal of Statistical Planning and Inference*.

Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, *42*(4), pp. 845-854.

Finkelstein, D. M., & Wolfe, R. A. (1985). A semiparametric model for regression analysis of interval-censored failure time data. *Biometrics*, *41*(4), pp. 933-945.

Fisher, R. (1958). *Statistical methods for research workers* (13th éd.). Edinburgh : Oliver and Boyd.

Fisher, R. (1966). *The design of experiments* (8th éd.). New York : Hafner Publishing Company.

Giovannitti-Jensen, A., & Myers, R. H. (1989). Graphical assessment of the prediction capability of response surface designs. *Technometrics*.

Glasstone, S., Laidler, K. J., & Eyring, H. (1941). *The theory of rate processes*. New York City : McGraw-Hill Book Co.

Goldfarb, H. B., Anderson-Cook, C. M., Borror, C. M., & Montgomery, D. C. (2004, 04). Fraction of design space plots for assessing mixture and mixture-process designs. *Journal of Quality Technology*, *36*(2), 169-179.

Goldfarb, H. B., Borror, C. M., Montgomery, D. C., & Anderson-Cook, C. M. (2004). Three-dimensional variance dispersion graphs for mixture-process experiments. *Journal of Quality Technology*, *36*(1), 109–124.

Hill, P. D. H. (1978). A review of experimental design procedures for regression model discrimination. *Technometrics*, *20*(1), pp. 15-21.

Holford, T. R. (1980). The analysis of rates and of survivorship using log-linear models. *Biometrics*, *36*(2), 299-305.

Hunter, J. (1989). Let's all beware the Latin square. *Quality Engineering*, *1*, 453-465.

Hurley, M. A. (1985). An application of generalized linear models to survival analysis with two types of failure. *Journal of the Royal Statistical Society. Series C*, *34*(3), 273-281.

Islam, A., & Ahmad, N. (1994). Optimal design of accelerated life tests for the Weibull distribution under periodic inspection and type i censoring. *Microelectronics Reliability*, *34*(9), 1459 - 1468.

Jang, D.-H., & Anderson-Cook, C. M. (2011). Fraction of design space plots for evaluating ridge estimators in mixture experiments. *Quality and Reliability Engineering International*, *27*(1), 27–34.

Jones, B. A., Li, W., Nachtsheim, C. J., & Ye, K. Q. (2007). Model discrimination and another perspective on model-robust designs. *Journal of Statistical Planning and Inference*, *137*(5), 1576 - 1583.

Kempthorne, O. (1952). *The design and analysis of experiments*. New York : Wiley.

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society B*, *21*, 272-304.

Kiefer, J. (1961). Optimum designs in regression problems. *Annals of Mathematical Statistics*, *32*, 298-325.

Kirk, R. E. (2012). *Experimental design: Procedures for the behavioral sciences* (4th éd.). SAGE Publications, Inc.

Konak, A., Coit, D. W., & Smith, A. E. (2006). Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety*, *91*(9), 992 - 1007.

Kuhfeld, W. F., Tobias, R. D., & Garratt, M. (1994). Efficient experimental design with marketing research applications. *Journal of marketing research*, *31*(4), 545-557.

Lee, J., & Pan, R. (2010). Analyzing step-stress accelerated life testing data using generalized linear models. *IIE Transactions*, *42*(8), 589-598.

Li, C., & Fard, N. (2007, march). Optimum bivariate step-stress accelerated life test for censored data. *Reliability, IEEE Transactions on*, *56*(1), 77 -84.

Li, J., Liang, L., Borror, C., Anderson-Cook, C., & Montgomery, D. (2009). Graphical summaries to compare prediction variance performance for variations of the central composite design for 6 to 10 factors. *Quality Technology & Quantitative Management*, *6*(4), 433-449.

Liang, L., Anderson-Cook, C., & Robinson, T. (2006). Fraction of design space plots for split-plot designs. *Quality and Reliability Engineering International*, *22*(3), 275-289.

Liang, L., Anderson-Cook, C., Robinson, T., & Myers, R. (2006). Three-dimensional variance dispersion graphs for split-plot designs. *Journal of Computational and Graphical Statistics*, *15*(4), 757-778.

Liu, X., & Qiu, W. S. (2011, dec.). Modeling and planning of step-stress accelerated life tests with independent competing risks. *Reliability, IEEE Transactions on*, *60*(4), 712 -720.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd éd.). Chapman and Hall/CRC.

McGree, J. M., & Eccleston, J. A. (2009). Investigating design for survival models. *Metrika*, *72*(3), 295-311.

Meeker, W. Q., & Escobar, L. A. (1998). *Statistical methods for reliability data* (1st éd.). Wiley-Interscience.

Meeker, W. Q., & Nelson, W. (1975, dec.). Optimum accelerated life-tests for the Weibull and extreme value distributions. *Reliability, IEEE Transactions on*, *R-24*(5), 321 -332.

Meyer, R. K., & Nachtsheim, C. J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental design. *Technometrics*, *37*(1), 60-69.

Mitchell, T. J. (1974). An algorithm for the construction of D-optimal experimental designs. *Technometrics*, *16*(2), 203-210.

Monroe, E. M., & Pan, R. (2008). Experimental design considerations for accelerated life tests with nonlinear constraints and censoring. *Journal of Quality Technology*, *40*(4), 355-367.

Monroe, E. M., Pan, R., Anderson-Cook, C., Montgomery, D. C., & Borror, C. M. (2010). Sensitivity analysis of optimal designs for accelerated life testing. *Journal of Quality Technology*, *42*(2), 121-135.

Monroe, E. M., Pan, R., Anderson-Cook, C., Montgomery, D. C., & Borror, C. M. (2011). A generalized linear model approach to designing accelerated life test experiments. *Quality and Reliability Engineering International*, *27*(4), 595-607.

Montgomery, D. C. (2012). *Design and analysis of experiments* (8th éd.). Wiley.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2009). *Response surface methodology: Process and product optimization using designed experiments* (3rd éd.). Wiley.

Myers, R. H., Montgomery, D. C., & Vining, G. G. (2002). *Generalized linear models : with applications in engineering and the sciences* (2nd éd.). New York : J. Wiley, c2002.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), pp. 370-384.

Nelson, W. (2005a). A bibliography of accelerated test plans. *Reliability, IEEE Transactions on*, *54*(2), 194-197.

Nelson, W. (2005b). A bibliography of accelerated test plans part ii - references. *Reliability, IEEE Transactions on*, *54*(3), 370-373.

Nelson, W., & Kielpinski, T. J. (1976). Theory for optimum censored accelerated life tests for normal and lognormal life distributions. *Technometrics*, *18*(1), pp. 105-114.

Nelson, W., & Meeker, W. Q. (1978). Theory for optimum accelerated censored life tests for Weibull and extreme value distributions. *Technometrics*, *20*(2), pp. 171-177.

Ng, H. K. T., Chan, P. S., & Balakrishnan, N. (2004). Optimal progressive censoring plans for the Weibull distribution. *Technometrics*, *46*(4), 470-481.

Nguyen, N., & Miller, A. J. (1992). A review of some exchange algorithms for constructing discrete D-optimal designs. *Computational Statistics and Data Analysis*, *14*(4), 489 - 498.

Ozol-Godfrey, A., Anderson-Cook, C., & Montgomery, D. C. (2005, 07). Fraction of design space plots for examining model robustness. *Journal of Quality Technology*, *37*(3), 223-235.

Ozol-Godfrey, A., Anderson-Cook, C., & Robinson, T. J. (2008). Fraction of design space plots for generalized linear models. *Journal of Statistical Planning and Inference*, *138*(1), 203 - 219.

Pan, R., & Yang, T. (2011). A GLM approach to optimal ALT test plans for Weibull distribution with type-I censoring. *The Proceedings of SAE 2011 World Congress, SAE*.

Pan, R., & Yang, T. (2013). Design and evaluation of accelerated life testing plans with dual objectives. *accepted at Journal of Quality Technology*.

Park, J.-W., & Yum, B.-J. (1996). Optimal design of accelerated life tests with two stresses. *Naval Research Logistics (NRL)*, *43*(6), 863–884.

Pascual, F. (2007). Accelerated life test planning with independent Weibull competing risks with known shape parameter. *Reliability, IEEE Transactions on*, *56*(1), 85 -93.

Pascual, F. (2008). Accelerated life test planning with independent weibull competing risks. *Reliability, IEEE Transactions on*, *57*(3), 435 -444.

Piepel, G. F., Anderson, C. M., & Redgate, P. E. (1993). Variance dispersion graphs for designs on. polyhedral regions - revisited. *Proceedings of the Selection on Physical and Engineering Sciences, Ameican Statistical Association*, 102-107.

Plackett, T. L., & Burman, J. (1946). The design of optimum multifactorical experiments. *Biometrika*, *33*, 305-325.

Seo, S.-K., & Yum, B.-J. (1991). Accelerated life test plans under intermittent inspection and type-i censoring: The case of Weibull failure distribution. *Naval Research Logistics (NRL)*, *38*(1), 1–22.

Sitter, R. R., & Torsney, B. (1995). Optimal designs for binary response experiments with two design variables. *Statistica Sinica*, *5*, 495-419.

Taguchi, G. (1987). *System of experimental design: engineering methods to optimize quality and minimize cost*. White Plains, NY : UNIPUB.

Taguchi, G. (1991). *Introduction to quality engineering. Asian Productivity Organization*. White Plains, NY : UNIPUB.

Tang, L. C., Goh, T. N., Sun, Y. S., & Ong, H. L. (1999). Planning accelerated life tests for censored two-parameter exponential distributions. *Naval Research Logistics (NRL)*, *46*(2), 169–186.

Tang, L. C., Tan, A. P., & Ong, S. H. (2002). Planning accelerated life tests with three constant stress levels. *Computers & Industrial Engineering*, *42*(24), 439 - 446.

Tobias, P. A., & Trindade, D. (2011). *Applied reliability* (3rd éd.). Chapman and Hall/CRC.

Trinca, L. A., & Gilmour, S. G. (1998, 10). Variance dispersion graphs for comparing blocked response surface designs. *Journal of Quality Technology*, *30*(4), 314-327.

Trinca, L. A., & Gilmour, S. G. (1999). Difference variance dispersion graphs for comparing response surface designs with applications in food technology. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *48*(4), 441–455.

Tse, S.-K., Ding, C., & Yang, C. (2008). Optimal accelerated life tests under interval censoring with random removals: the case of Weibull failure distribution. *Statistics*, *42*(5), 435 - 451.

Viney, R., Savage, E., & Louviere, J. (2005). Empirical investigation of experimental design properties of discrete choice experiments in health care. *Health Econ*, *14*(4), 349-362.

Wang, W., & Kececioglu, D. (2000). Fitting the Weibull log-linear model to accelerated life-test data. *Reliability, IEEE Transactions on*, *49*(2), 217-223.

Waterhouse, T., Woods, D., Eccleston, J., & Lewis, S. (2006). Design selection criteria for discrimination between nested models for binomial data. *S3RI Methodology Working Papers*.

Wilk, M. (1955). *Linear models and randomized experiments*. Thèse de doctorat non publiée, Iowa State College.

Woods, D. C., Lewis, S. M., Eccleston, J. A., & Russell, K. G. (2006). Designs for generalized linear models with several variables and model uncertainty. *Technometrics*, *48*(2), 284-292.

Wu, C. F. J., & Hamada, M. S. (2009). *Experimentsl planning, analysis, and optimization* (2nd éd.). Wiley.

Xu, H.-Y., & Fei, H.-L. (2007, sept.). Planning step-stress accelerated life tests with two experimental variables. *Reliability, IEEE Transactions on*, *56*(3), 569 -579.

Yang, M., Zhang, B., & Huang, S. (2011). Optimal designs for generalized linear models with multiple design variables. *Statistica Sinica*, *21*, 1415-1430.

Yang, T., & Pan, R. (2013). A novel approach to optimal ALT planning with interval censoring. *Reliability, IEEE Transaction on,*, *to appear on June Issue*.

Yum, B.-J., & Choi, S.-C. (1989). Optimal design of accelerated life tests under periodic inspection. *Naval Research Logistics (NRL)*, *36*(6), 779–795.

Zahran, A., Anderson-Cook, C., & Myers, R. H. (2003, 10). Fraction of design space to assess prediction capability of response surface designs. *Journal of Quality Technology*, *35*(4), 377-386.

APPENDIX A

A GLM Example

A GLM Example

The PH model is assumed for the hazard function of failure time. Thus, the hazard function and reliability function are respectively,

$$\lambda(t) = \lambda_0(t)e^{\mathbf{x}'\beta}, \tag{A.1}$$

and

$$R(t) = e^{-\Lambda(t)} = R_0(t)^{\exp(\mathbf{x}'\beta)}, \tag{A.2}$$

where $\lambda_0(t)$ is the baseline hazard function, $\Lambda(t)$ is the cumulative hazard function, $\Lambda(t) = \int_0^t \lambda(\tau)d\tau = \Lambda_0(t)e^{\mathbf{x}'\beta}$, and $\Lambda_0(t)$ is the baseline cumulative hazard function, and $R_0(t)$ is the baseline reliability function with $R_0(t) = \exp(-\Lambda_0(t))$. The failure function is the compliment of reliability function, i.e., $F(t) = 1 - R(t)$.

To obtain the GLM formula for censored failure time data, we need to (1) identify a variable and its distribution that belongs to the exponential family; (2) define the linear predictor; and (3) find a suitable link function. We will start from the interval censoring case and then discuss the right censoring case.

Suppose that all test units are inspected at times $t_1, t_2, \ldots$ until t_k, where $t_1 < t_2 < \cdots < t_k$. For a single test unit, one observes the following data set, (0, t_1, r_{i1}), (t_1, t_2, r_{i2}), ..., (t_{k-1}, t_k, r_{ik}), where r_{ij} is a binary variable indicating whether or not the test unit fails in the corresponding interval, (t_{j-1}, t_j). Let T_i be the failure time of test unit i, then

$$r_{ij} = \begin{cases} 1 & \text{when } t_{j-1} \le T_i < t_j \\ 0 & \text{otherwise} \end{cases}$$

As shown in Collett (2003), r_{ij} can be treated as independent binomial random variable with parameters m_{ij} and π_{ij}, where m_{ij} is the number of partially survived test unit at the time t_{j-1} and π_{ij} is the conditional probability of failure in the j^{th} interval; i.e.,

$$m_{ij} = \begin{cases} 1 & \text{when } T_i > t_{j-1} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\pi_{ij} = P(T_i \le t_j | T_i > t_{ij}).$$

The sample likelihood of r_{ij}'s is given by

$$L = \prod_{i=1}^{n} \left[\prod_{j=1}^{k} \pi_{ij}^{r_{ij}} (1 - \pi_{ij})^{m_{ij} - r_{ij}} \right] \tag{A.3}$$

From the definition of π_{ij}, we have

$$1 - \pi_{ij} = P(T_i \ge t_j \mid T_i \ge t_{j-1}) = \frac{R_i(t_j)}{R_i(t_{j-1})}. \tag{A.4}$$

Together with Eq. (A.2), we have

$$1-\pi_{ij}=\left(\frac{R_0(t_j)}{R_0(t_{j-1})}\right)^{\exp(\eta_i)}, \tag{A.5}$$

where η_i is a linear predictor, i.e., $\eta_i=\mathbf{x}_i'\beta$. Then, we have a complementary log-log link function as

$$\log[-\log(1-\pi_{ij})]=\eta_i+\log\left[\log\left(\frac{R_0(t_{j-1})}{R_0(t_j)}\right)\right]. \tag{A.6}$$

The second term on the right hand side of link function is an offset term, as it is irrelevant to stress factors $\mathbf{x}_i$. Thus, we have formulated a GLM for the ALT with interval censoring. For finding the weight matrix of an experimental design matrix and the U_c-optimal design, one may refer to T. Yang & Pan (2013).

In Section 2.2, we use an example adopted from Monroe et al. Monroe et al. (2010), where the linear predictor is derived from Peck's acceleration model with two stress factors – temperature and relative humidity. After factor standardization, the linear predictor is given by

$$\eta_i=-4.086x_{1i}-1.476x_{2i}+0.01x_{1i}x_{2i}, \tag{A.7}$$

where x_1 and x_2 may take values from 0 to 1.

In the right censoring case one observes the following data pair, (t_i, c_i), for each test unit, in which t_i is failure time or censoring time and c_i is a binary variable of whether or not a failure is observed. As shown in Aitkin & Clayton (1980), c_i can be treated as a Poisson distributed variable with mean such as

$$\mu_i=\Lambda_i(t_i)=\Lambda_0(t_i)e^{\eta_i}. \tag{A.8}$$

The link function is, therefore, a logarithm function as

$$\log\mu_i=\eta_i+\log\Lambda_0(t_i). \tag{A.9}$$

Again, the second term in the right hand side of the equation above is not related to stress factors, so it is an offset term. For the derivation of optimal test plans with exponential and Weibull failure time distributions, one may refer to Monroe et al. (2011) and Pan & Yang (2011).

APPENDIX B

SAS Code

```
1  libname tojmp 'D:\program files\SAS results';
2  /* this code is used to find U-optimal designs for interval ...
       censoring case with different number of intervals*/
3  /* this code is used to generate the Tables 3 to 6 in Interval ...
       censoring paper                                    */
4  proc iml;
5
6     /* define the U-optimality function */
7     start UOptimal(xx)
8
9        global (Weight, Var1,b0, b1, b2, b3, t, a,Phi,sum,k);
10
11       /* shape function used here to make a 100*2 size matrix*/
12       x=shape(xx,nrow(xx)*ncol(xx)/2,2);
13
14       /* F is the design matrix ...
                                                                    */
15       /* F is constructed by horizontal concatenating (pipe ...
            operator)                  */
16       /* First column is all 1's, then column x1, x2, x1*x1, ...
            x1*x2, and x2*x2 */
17       /* H is the (x1, x2) matrix ...
                                                                 */
18
19       F=j(nrow(x),1) || x[,1]  || x[,2]||x[,1]#x[,2];
20       G= F ;
21       H= x[,1] || x[,2];
22
23
24       /* b matrix is a column vector (denoted by //) of beta ...
            coefficients               */
25       b=b0//b1//b2//b3;
26
27       /* Codes below are used to find weight matrix ...
                                              */
28       a1=G*b;
29       a2=exp(a1);
30       a3=exp(-a2*t);
31       a4=(-a2*t)#(-a2*t);
32       a5=a4#a3;
33
34
35       Phi=a5/(1-a3);
36
37
38       W=diag(Phi);
39
40
41
42
43       /* This is the use condition location (x1,x2)=(1.758, ...
            3.159)                        */
44
```

```
45        use={1 1.758 3.159 5.553};
46
47
48        sum=0;
49
50        /* k is the number of intervals and in the paper it has ...
             values of 2, 5, 10, and 30 */
51        k=5;
52
53
54          do j=1 to k by 1;
55
56              a6=(a3##(j-1));
57              sum=sum+a6;
58          end;
59
60
61
62
63
64        /* calculate the prediction variance ...
                                                                  */
65        XWX=F`*(sum#W)*F;
66        Var=use*inv(XWX)*use`;
67        pred_var=Var;
68
69        /* Round function makes output easier to read and avoids ...
             scientific notation    */
70
71        Var1=round(Var,0.01);
72        Weight=round(Phi,0.01);
73        XWX1=round(XWX,0.01);
74
75        /* Return the value back to the nlpcg function which is a ...
             nonlinear optimization process */
76        return(Var);
77
78    finish;
79
80    /*to invoke the create statement the first time through ...
                                    */
81    first=1;
82
83    /* set the values of parameters ...
                                                                    */
84    do b0=0 to 0;
85       do b1=-4.086 to -4.086;
86          do b2=-1.476 to -1.476;
87             do b3=0.01 to 0.01;
88
89
90                      /* when change the value of k, change the ...
                           denominator of gap      */
91                      gap=30/5;
```

```
92                    a=1;
93                    t=gap**a;
94
95                  /* number of test units                          ...
                                          */
96                  Nd=100;
97
98                  /* ranuni is random univariate number generator ...
                       between 0 and 1         */
99                  /* 2*rand -1 generates a set of numbers between ...
                       -1 and +1               */
100                 /* 1:Nd*2 is a row vector, with transpose it is ...
                       a Nd*2 by 1 column vector */
101                 x0init = ranuni((1:(Nd*2))`);
102
103
104                 /* con is the constraint matrix ...
                                                          */
105                 /* shape is (value=-1, # of rows, # of cols ...
                                                 */
106                 /* double slash (//) operator is a vertical ...
                       concatenation               */
107                 /* con = {0 occurs Nd times, 1 occurs Nd times} ...
                                                 */
108                 /* first shape function is for lower bounds, ...
                       the second for upper bound */
109                 con     = shape(0,1,Nd*2) // shape(1,1,Nd*2);
110
111
112
113                 /* Nonlinear program using NLPCG                 ...
                                                */
114                 call nlpcg(rc,                  /* Return code ...
                                                */
115                           x0,                   /* Returned ...
                              optimum factors             */
116                           "UOptimal",           /* Function to ...
                              optimize                 */
117                           x0init,               /* Initial ...
                              value of factors            */
118                           0,                    /* Specify a ...
                              minimization              */
119                           con);                 /* Specify ...
                              constraints               */
120
121
122                 /* Round function makes output easier to read ...
                       and avoids scientific notation    */
123                 toout=round(shape(x0, nd,2),0.001) ||
124                       Weight ||
125                       repeat(Var1,nd,1) ||
126                       repeat(b0,nd,1) ||
127                       repeat(b1,nd,1) ||
128                       repeat(b2,nd,1) ||
```

```
129                          repeat(b3,nd,1) ||
130                          repeat(t,nd,1)  ||
131                          repeat(a,nd,1);
132
133                    if first then do;
134                       colnames={'x1' 'x2' 'weight' 'pred_var' 'b0' ...
                             'b1' 'b2' 'b3' 't' 'a'};
135                       create tojmp.newdat from ...
                             toout[colname=colnames];
136                    end;
137                    append from toout;
138                    first=0;
139
140                end;
141             end;
142          end;
143     end;
144
145
146
147 quit;
148 proc freq data=tojmp.newdat;
149    /* make the table called summary */
150    tables b0*b1*b2*b3*t*a*x1*x2*pred_var / noprint ...
          out=tojmp.summary(drop=percent);
151 run;
```

```
libname tojmp 'D:\program files\SAS results';
/* this code is used to find I-optimal designs for right ...
   censoring case                                          */
/* this code is used to generate the Tables 1 to 3 in Model ...
   Checking paper                                        */

proc iml;
/* define the Integration part of I-optimality  */
start norpdf2(u1) global(u2,c);

use=1||u1||u2||u1*u2;
pvs=use*c*use`;
return(pvs);
finish;




start marginal(v) global(yy,u2,c);

interval = 1.458 || yy;

u2 = v;

call quad(pm,"NORPDF2",interval);
return(pm);
finish;



start norcdf2(au,bu) global(yy,c);


yy = bu;

interval= 2.859 || au;


call quad(p,"MARGINAL",interval);


return(p);
finish;




   /* define the I-optimality function */
   start UOptimal(xx)

      global (Weight, Var1,b0, b1, b2, b3, t, a,Phi,sum,k, c, Var);

      /* shape function used here to make a 100*2 size matrix*/
```

```
52        x=shape(xx,nrow(xx)*ncol(xx)/2,2);
53
54        /* F is the design matrix ...
                                                                    */
55        /* F is constructed by horizontal concatenating (pipe ...
             operator)                  */
56        /* First column is all 1's, then column x1, x2, x1*x1, ...
             x1*x2, and x2*x2 */
57        /* H is the (x1, x2) matrix ...
                                                                  */
58        F=j(nrow(x),1) || x[,1]  || x[,2]||x[,1]#x[,2];
59        G= F ;
60        H= x[,1] || x[,2];
61
62
63        /* b matrix is a column vector (denoted by //) of beta ...
             coefficients             */
64        /* b can be changed based on the different models. e.g. for ...
             main effects only model,
65           b=b0//b1//b2
66        */
67        b=b0//b1//b2//b3;
68
69        /* Codes below are used to find weight matrix ...
                                                 */
70        a1=G*b;
71        a2=exp(a1);
72        Phi= 1-exp(-(t**a)*a2) ;
73        W=diag(Phi);
74
75
76
77
78
79        XWX=F`*W*F;
80        c=inv(XWX);
81
82
83
84
85
86
87       /* Calculate the average prediction ovet the whole use ...
            region                    */
88       p = norcdf2(3.459,2.058)/0.36;
89
90
91
92
93        Var=p;
94        pred_var=Var;
95
96        /* Round function makes output easier to read and avoids ...
             scientific notation    */
```

```
97        Var1=round(Var,0.01);
98        Weight=round(Phi,0.01);
99
100
101       /* Return the value back to the nlpcg function which is a ...
             nonlinear optimization process */
102       return(Var);
103
104    finish;
105
106
107    /*to invoke the create statement the first time through ...
                                    */
108    first=1;
109    do b0=0 to 0;
110       do b1=-4.086 to -4.086;
111          do b2=-1.476 to -1.476;
112             do b3=0.01 to 0.01;
113
114
115                   /* Cenroing time                          ...
                                                    */
116                   t=30;
117                   a=1;
118
119                /* number of test units                        ...
                                          */
120                Nd=100;
121
122                /* ranuni is random univariate number generator ...
                      between 0 and 1         */
123                /* 2*rand -1 generates a set of numbers between ...
                      -1 and +1               */
124                /* 1:Nd*2 is a row vector, with transpose it is ...
                      a Nd*2 by 1 column vector */
125                x0init = ranuni((1:(Nd*2))`);
126
127                /* con is the constraint matrix ...
                                                              */
128                /* shape is (value=-1, # of rows, # of cols ...
                                                 */
129                /* double slash (//) operator is a vertical ...
                      concatenation              */
130                /* con = {0 occurs Nd times, 1 occurs Nd times} ...
                                          */
131                /* first shape function is for lower bounds, ...
                      the second for upper bound */
132                con     = shape(0,1,Nd*2) // shape(1,1,Nd*2);
133
134
135                /* Nonlinear program using NLPCG              ...
                                             */
136                call nlpcg(rc,                 /* Return code ...
                                          */
```

```
137                                 x0,                   /* Returned ...
                                       optimum factors            */
138                                 "UOptimal",           /* Function to ...
                                       optimize                 */
139                                 x0init,               /* Initial ...
                                       value of factors           */
140                                 0,                    /* Specify a ...
                                       minimization               */
141                                 con);                 /* Specify ...
                                       constraints                */
142                       print x0;
143                       print c;
144                       print Var;
145
146
147                       /* Round function makes output easier to read ...
                             and avoids scientific notation    */
148                       toout=round(shape(x0, nd,2),0.001) ||
149                             Weight ||
150                             repeat(Var1,nd,1) ||
151                             repeat(b0,nd,1) ||
152                             repeat(b1,nd,1) ||
153                             repeat(b2,nd,1) ||
154                             repeat(b3,nd,1) ||
155                             repeat(t,nd,1)  ||
156                             repeat(a,nd,1);
157
158                       if first then do;
159                          colnames={'x1' 'x2' 'weight' 'pred_var' 'b0' ...
                                'b1' 'b2' 'b3' 't' 'a'};
160                          create tojmp.newdat from ...
                                toout[colname=colnames];
161                       end;
162                       append from toout;
163                       first=0;
164
165                    end;
166                 end;
167              end;
168        end;
169
170
171
172   quit;
173   proc freq data=tojmp.newdat;
174      /* make the table called summary */
175      tables b0*b1*b2*b3*t*a*x1*x2*pred_var / noprint ...
            out=tojmp.summary(drop=percent);
176   run;
177
178   proc print data=tojmp.summary;
179   run;
```

```
1
2  libname tojmp 'D:\program files\SAS results';
3
4  /* this code is used to find Dual objective designs with D- and ...
       Ds-optimalities                                    */
5  /* this code is used to generate the Tables 8 to 10 in Model ...
       Checking paper                                     */
6  proc iml;
7
8
9     /* define the dual objective optimality function */
10    start DOptimal(xx)
11
12       global (Weight, DCompound, b0, b1, b2, b3, b4, b5, t , a, ...
             DsEff, Deff);
13
14       /* shape function used here to make a 100*2 size matrix*/
15       x=shape(xx,nrow(xx)*ncol(xx)/2,2);
16
17       /* F is the design matrix ...
                                                                  */
18       /* F is constructed by horizontal concatenating (pipe ...
             operator)                  */
19       /* First column is all 1's, then column x1, x2, x1*x1, ...
             x1*x2, and x2*x2 */
20       /* H is the (x1, x2) matrix ...
                                                                 */
21       F=j(nrow(x),1) || x[,1] || x[,2]||x[,1]#x[,2] ||x[,1]#x[,1] ...
             ||x[,2]#x[,2];
22       G= F;
23       H= x[,1] || x[,2];
24
25       /* b matrix is a column vector (denoted by //) of beta ...
             coefficients              */
26       b=b0//b1//b2//b3//b4//b5;
27
28       /* Codes below are used to find weight matrix ...
                                                 */
29       a1=G*b;
30       a2=exp(a1);
31       Phi= 1-exp(-(t**a)*a2);
32       W=diag(Phi);
33
34
35       /* Computing the determinant of the Fisher Information ...
             Matrix                    */
36       XWX=F`*W*F;
37       D=det(XWX);
38
39
40       /* set Ds-optimality part for the dual objective design ...
                                   */
41       D11=det(XWX[1:4,1:4]);
```

```
42        r=4;
43        s=2;
44
45        /* K is the value to decide the weight of D-optimal design ...
             for the simpler model*/
46        k=0.8;
47
48
49        /* Dual objective function ...
                                                                      */
50        DCompound=(6*k-r)/r/s*log(D11)+(1-k)/s*log(D);
51
52        Ds=D/D11;
53
54
55        /* Calculate the Ds and D efficiencies ...
                                                        */
56        DsEff=(D/D11/0.90295)**(1/s);
57
58        Deff=(det(D11)/24438.53)**(1/r);
59
60
61
62        /* Round function makes output easier to read and avoids ...
             scientific notation    */
63        D1=round(D,0.01);
64        Weight=round(Phi,0.01);
65        XWX1=round(XWX,0.01);
66
67        /* Return the value back to the nlpcg function which is a ...
             nonlinear optimization process */
68        return(DCompound);
69
70     finish;
71
72
73     /*to invoke the create statement the first time through ...
                                    */
74     first=1;
75     do b0=0 to 0 by -1;
76        do b1=-4.086 to -4.086;
77           do b2=-1.476 to -1.476;
78              do b3=0.01 to 0.01;
79              b4=0.03;
80              b5=0.03;
81
82
83                    /* Cenroing time                          ...
                                                      */
84                    t=30;
85                    a=1;
86
87
88                  /* number of test units                        ...
```

```
                                        */
Nd=100;

/* ranuni is random univariate number generator ...
   between 0 and 1         */
/* 2*rand -1 generates a set of numbers between ...
   -1 and +1               */
/* 1:Nd*2 is a row vector, with transpose it is ...
   a Nd*2 by 1 column vector */
x0init = ranuni((1:(Nd*2))`);

/* con is the constraint matrix ...
                                                    */
/* shape is (value=-1, # of rows, # of cols ...
                                   */
/* double slash (//) operator is a vertical ...
   concatenation               */
/* con = {0 occurs Nd times, 1 occurs Nd times} ...
                                  */
/* first shape function is for lower bounds, ...
   the second for upper bound */
con     = shape(0,1,Nd*2) // shape(1,1,Nd*2);

/* Nonlinear program using NLPCG                ...
                                    */
call nlpcg(rc,                  /* Return code ...
                                 */
           x0,                  /* Returned ...
              optimum factors             */
           "DOptimal",          /* Function to ...
              optimize                 */
           x0init,              /* Initial ...
              value of factors            */
           1,                   /* Specify a ...
              maximization               */
           con);                /* Specify ...
              constraints                 */


/* Round function makes output easier to read ...
   and avoids scientific notation    */
toout=round(shape(x0, Nd,2),0.01) ||
      Weight ||
      repeat(DCompound,Nd,1) ||
      repeat(b0,Nd,1) ||
      repeat(b1,Nd,1) ||
      repeat(b2,Nd,1) ||
      repeat(b3,Nd,1) ||
      repeat(b4,Nd,1) ||
      repeat(b5,Nd,1) ||
      repeat(t,Nd,1)||
      repeat(a,Nd,1)||
      repeat(DsEff,Nd,1)||
      repeat(Deff,Nd,1);
```

```
126
127                     if first then do;
128                        colnames={'x1' 'x2' 'weight' 'Dcompound' ...
                              'b0' 'b1' 'b2' 'b3' 'b4' 'b5' 't' 'a' ...
                              'DsEff' 'Deff'};
129                        create tojmp.newdat from ...
                              toout[colname=colnames];
130                     end;
131                     append from toout;
132                     first=0;
133
134                 end;
135              end;
136           end;
137          end;
138
139
140
141  quit;
142  proc freq data=tojmp.newdat;
143     /* make the table called summary */
144     tables b0*b1*b2*b3*b4*b5*t*a*x1*x2*Dcompound*DsEff*Deff / ...
           noprint out=tojmp.summary(drop=percent);
145  run;
146
147  proc print data=tojmp.summary;
148  run;
```

```
1
2  libname tojmp 'D:\program files\SAS results';
3  /* this code is used to find I-optimal designs for interval ...
      censoring case with different number of intervals*/
4  /* this code is used to generate the Table 1 in design evaluation ...
      paper                                  */
5
6  proc iml;
7  /* Do the integration part */
8  start norpdf2(u1) global(u2,c);
9
10 use=1||u1||u2||u1*u2;
11 pvs=use*c*use`;
12 return(pvs);
13 finish;
14
15
16
17
18 start marginal(v) global(yy,u2,c);
19
20 interval = 1.491 || yy;
21
22 u2 = v;
23
24 call quad(pm,"NORPDF2",interval);
25 return(pm);
26 finish;
27
28
29
30 start norcdf2(au,bu) global(yy,c);
31
32
33 yy = bu;
34
35 interval= 2.713 || au;
36
37 call quad(p,"MARGINAL",interval);
38
39
40 return(p);
41 finish;
42
43
44
45
46    /* define the U-optimality function */
47    start UOptimal(xx)
48
49       global (Weight, Var1,b0, b1, b2, b3, t, a,Phi,sum,k, c, Var);
50
51       /* shape function used here to make a 100*2 size matrix*/
```

```
52      x=shape(xx,nrow(xx)*ncol(xx)/2,2);
53
54      /* F is the design matrix ...
                                                                    */
55      /* F is constructed by horizontal concatenating (pipe ...
           operator)                  */
56      /* First column is all 1's, then column x1, x2, x1*x1, ...
           x1*x2, and x2*x2 */
57      /* H is the (x1, x2) matrix ...
                                                                  */
58      F=j(nrow(x),1) || x[,1]  || x[,2]||x[,1]#x[,2];
59      G= F ;
60      H= x[,1] || x[,2];
61
62
63      /* b matrix is a column vector (denoted by //) of beta ...
           coefficients              */
64      b=b0//b1//b2//b3;
65
66      /* Codes below are used to find weight matrix ...
                                              */
67      a1=G*b;
68      a2=exp(a1);
69      a3=exp(-a2*t);
70      a4=(-a2*t)#(-a2*t);
71      a5=a4#a3;
72
73
74
75
76      Phi=a5/(1-a3);
77
78
79      W=diag(Phi);
80
81      sum=0;
82      k=2;
83
84
85        do j=1 to k by 1;
86
87            a6=(a3##(j-1));
88            sum=sum+a6;
89        end;
90
91
92
93
94
95
96      XWX=F`*(sum#W)*F;
97      c=inv(XWX);
98
99
```

```
100
101
102
103         /* calculate average prediction variance in use regoin ...
                                          */
104         p = norcdf2(3.714,2.048)/0.557;
105
106
107
108         /* calculate the prediction variance ...
                                                              */
109         Var=p;
110         pred_var=Var;
111
112         /* Round function makes output easier to read and avoids ...
               scientific notation    */
113         Var1=round(Var,0.01);
114         Weight=round(Phi,0.01);
115
116
117         /* Return the value back to the nlpcg function which is a ...
               nonlinear optimization process */
118         return(Var);
119
120      finish;
121
122      first=1;
123      /* set the values of parameters ...
                                                                 */
124      do b0=0 to 0;
125         do b1=-4.086 to -4.086;
126            do b2=-1.476 to -1.476;
127               do b3=0.01 to 0.01;
128
129
130                     /* when change the value of k, change the ...
                           denominator of gap      */
131                     gap=30/2;
132                     a=1;
133                     t=gap**a;
134
135                  /* number of test units                        ...
                                              */
136                  Nd=100;
137
138                  /* ranuni is random univariate number generator ...
                        between 0 and 1         */
139                  /* 2*rand -1 generates a set of numbers between ...
                        -1 and +1               */
140                  /* 1:Nd*2 is a row vector, with transpose it is ...
                        a Nd*2 by 1 column vector */
141                  x0init = ranuni((1:(Nd*2))`);
142
143                  /* con is the constraint matrix ...
```

```
                                                                    */
144                  /* shape is (value=-1, # of rows, # of cols ...
                                                                   */
145                  /* double slash (//) operator is a vertical ...
                        concatenation                              */
146                  /* con = {0 occurs Nd times, 1 occurs Nd times} ...
                                                                  */
147                  /* first shape function is for lower bounds, ...
                        the second for upper bound */
148                  con     = shape(0,1,Nd*2) // shape(1,1,Nd*2);
149
150                  /* Nonlinear program using NLPCG                ...
                                                                   */
151                  call nlpcg(rc,                    /* Return code ...
                                                                  */
152                             x0,                    /* Returned ...
                                optimum factors                 */
153                             "UOptimal",            /* Function to ...
                                optimize                     */
154                             x0init,                /* Initial ...
                                value of factors                */
155                             0,                     /* Specify a ...
                                minimization                 */
156                             con);                  /* Specify ...
                                constraints                     */
157                  print x0;
158                  print c;
159                  print Var;
160
161                  /* Round function makes output easier to read ...
                        and avoids scientific notation    */
162                  toout=round(shape(x0, nd,2),0.001) ||
163                        Weight ||
164                        repeat(Var1,nd,1) ||
165                        repeat(b0,nd,1) ||
166                        repeat(b1,nd,1) ||
167                        repeat(b2,nd,1) ||
168                        repeat(b3,nd,1) ||
169                        repeat(t,nd,1)  ||
170                        repeat(a,nd,1);
171
172                  if first then do;
173                     colnames={'x1' 'x2' 'weight' 'pred_var' 'b0' ...
                           'b1' 'b2' 'b3' 't' 'a'};
174                     create tojmp.newdat from ...
                           toout[colname=colnames];
175                  end;
176                  append from toout;
177                  first=0;
178
179               end;
180            end;
181         end;
182      end;
```

```
183
184
185
186 quit;
187 proc freq data=tojmp.newdat;
188    /* make the table called summary */
189    tables b0*b1*b2*b3*t*a*x1*x2*pred_var / noprint ...
          out=tojmp.summary(drop=percent);
190 run;
```