2011

Methods for planning repeated measures degradation tests

Brian Phillip Weaver
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd
Part of the Statistics and Probability Commons

Recommended Citation
Weaver, Brian Phillip, "Methods for planning repeated measures degradation tests" (2011). Graduate Theses and Dissertations. 11967. https://lib.dr.iastate.edu/etd/11967

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Methods for planning repeated measures degradation tests

by

Brian Phillip Weaver

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Program of Study Committee:
William Q. Meeker, Major Professor
Mervyn G. Marasinghe
Max Morris
Vivekananda Roy
Alyson Wilson

Iowa State University
Ames, Iowa
2011

Copyright © Brian Phillip Weaver, 2011. All rights reserved.
DEDICATION

I would like to dedicate this thesis to my wife Kathryn whose patience is infinite and to my daughters Emmeline and Maisie who helped keep me “in check.”
TABLE OF CONTENTS

LIST OF TABLES ................................................................. viii
LIST OF FIGURES ............................................................ xi
ACKNOWLEDGEMENTS ......................................................... xv
ABSTRACT ........................................................................... xvi

CHAPTER 1. General Introduction ........................................ 1
  1.1 Introduction ................................................................ 1
  1.2 Dissertation Organization ......................................... 3
  1.3 Literature Review .................................................... 4
  1.4 References ............................................................. 5

CHAPTER 2. Methods For Planning Repeated Measures Degradation Studies ........................................... 7
  2.1 Introduction ............................................................. 8
    2.1.1 Motivating Examples ...................................... 8
    2.1.2 Related Work ............................................... 9
    2.1.3 Overview .................................................. 10
  2.2 Repeated Measures Degradation Model ..................... 11
    2.2.1 Model and Data ............................................ 11
  2.3 Model Likelihood Function and Fisher Information Matrix ........................................... 12
    2.3.1 Likelihood ................................................ 12
    2.3.2 Variance Covariance Matrix .......................... 12
    2.3.3 Linear Mixed-Model Parameter Estimation .... 13
2.4 Estimating Quantiles of the Degradation Distribution ........................................ 14
  2.4.1 The Quantile of the Degradation Distribution .............................................. 14
  2.4.2 Standard Error for the Maximum Likelihood Estimator of the \( p \) Quantile .... 14
  2.4.3 Confidence Interval for the Degradation Distribution Quantiles .................. 15
2.5 Degradation Test Planning ................................................................. 16
  2.5.1 Simple Degradation Test Plans ................................................................. 16
  2.5.2 Degradation Test Plans with Differing Schedules ......................................... 18
  2.5.3 Using Simulation to Evaluate Test Plans ...................................................... 20
  2.5.4 Selecting a Test Plan Under a Cost Constraint ............................................. 21
2.6 Failure-Time Distribution ................................................................. 22
  2.6.1 Relationship Between Degradation and Failures .......................................... 23
  2.6.2 Estimating the \( p \) Quantile of the Failure-Time Distribution ..................... 25
  2.6.3 Standard Error for the Maximum Likelihood Estimator of the Failure-Time Quantile ................................................................. 25
  2.6.4 Degradation Test Planning Using \( t_p \) ......................................................... 25
2.7 Conclusions and Future Work ............................................................... 26
2.8 Appendix ............................................................................................. 28
  2.8.1 Derivation of the Information Matrix in Section 2.3.2 ................................... 28
  2.8.2 Forms of the Partial Derivatives in Section 2.4.2 ......................................... 29
  2.8.3 Derivation of \( t_p \) in Section 2.6.2 ............................................................... 29
  2.8.4 Forms of the Partial Derivatives in Section 2.6.3 ......................................... 31
2.9 References .......................................................................................... 31

CHAPTER 3. Methods For Planning Accelerated Repeated Measures Degradation Studies ................................................................. 33
3.1 Introduction .......................................................................................... 34
  3.1.1 Motivating Examples ..................................................................................... 34
  3.1.2 Related Work ............................................................................................... 36
  3.1.3 Overview .................................................................................................. 37
3.2 Degradation Models .................................................. 38
  3.2.1 Accelerated Repeated Measures Degradation Models ............ 38
  3.2.2 Approximate Variance-Covariance Matrix of the Maximum Likelihood Estimators ........................................... 39
3.3 Failure-Time Distribution for Degradation Models ................. 39
  3.3.1 Failure-Time Cumulative Distribution Function ................. 39
  3.3.2 Failure-Time Quantiles ........................................... 40
  3.3.3 Approximate Variance for the Maximum Likelihood Estimator of a Failure-Time Quantile ...................................... 41
3.4 Repeated Measures Accelerated Degradation Test Planning .......... 41
  3.4.1 Planning Information .............................................. 41
  3.4.2 Inspection Schedule Specification .............................. 41
  3.4.3 Accelerating Variable Plan Specification ......................... 42
3.5 Optimum RMADT ..................................................... 42
  3.5.1 Criterion for Choosing a Plan .................................... 42
  3.5.2 Checking the Initial Optimum Plan .............................. 43
3.6 Other RMADT Plans .................................................. 43
  3.6.1 Traditional Plans .................................................. 43
  3.6.2 Compromise RMADT ............................................... 44
3.7 Examples of Planning RMADT ........................................ 44
  3.7.1 Test Plan to Test the Increase of Resistance in Carbon-Film Resistors ....................................................... 44
  3.7.2 Test Plan to Assess Scar Width Growth .......................... 47
  3.7.3 Test Plan to Study LED Output Degradation ...................... 49
3.8 Monte Carlo Simulation To Evaluate a RMADT Plans .................. 56
3.9 Conclusions and Areas for Future Research ........................ 57
3.10 Appendix ............................................................... 59
  3.10.1 Derivation of the Information Matrix ......................... 59
  3.10.2 Derivation of $t_p$ in Section 3.3.2 ............................. 61
CHAPTER 4. Bayesian Methods For Planning Accelerated Repeated Measures Degradation Studies

4.1 Introduction ............................................. 65
  4.1.1 Related Work ................................... 66
  4.1.2 Motivating Example ............................. 66
  4.1.3 Overview ....................................... 67
4.2 Degradation Models ..................................... 67
  4.2.1 Accelerated Repeated Measures Degradation Models .... 67
  4.2.2 Model for Resistor Degradation .................... 69
  4.2.3 Reparameterization of the Model for Prior Distribution Specification ........................................ 69
4.3 Failure-Time Distribution for Degradation Models ............... 70
  4.3.1 Failure-Time Cumulative Distribution Function .......... 70
  4.3.2 Failure-Time Quantiles ............................ 70
4.4 Planning Bayesian Repeated Measures Accelerated Degradation Tests ...... 71
  4.4.1 Plan Specification ................................ 71
  4.4.2 Bayesian Planning Criterion ....................... 71
  4.4.3 The Prior Distribution ............................ 72
  4.4.4 General Equivalence Theory ....................... 73
  4.4.5 The Wynn Algorithm Used to Find Optimum Plans .......... 74
4.5 Numerical Example ...................................... 75
  4.5.1 Prior Specification ................................ 75
  4.5.2 Bayesian Optimum Test Plans ..................... 77
  4.5.3 Bayesian Compromise Test Plans ................... 79
4.6 Simulation Evaluation of a Bayesian RMADT Plan .................. 81
  4.6.1 Simulations Based on Fixed Model Parameters ............ 81
4.7 Conclusions and Future Work ................................ 82
4.8 Appendix ............................................... 84
4.8.1 Fisher Information Matrix ........................................... 84
4.8.2 Derivation of $t_p$ in Section 4.3.2 ................................. 86
4.9 References ..................................................................... 86

CHAPTER 5. General Conclusions ........................................... 89
5.1 General Discussion .......................................................... 89
5.2 Recommendation for Future Research ............................. 90
## LIST OF TABLES

| Table 2.1 | Asymptotic standard error SE\(_{\hat{c}_{0.10}}\) for three different degradation test plans. | 26 |
| Table 3.1 | Optimum, compromise, and traditional test plans as well as the large-sample approximate standard error under Plan 1 for the resistor example and assuming \(n = 29\) test units (as in the original experiment). | 46 |
| Table 3.2 | Optimum, compromise, the original plan from the study, and traditional test plans as well as the large-sample approximate standard error under Plan 2 for the resistor example and assuming \(n = 29\) test units (as in the original experiment). | 47 |
| Table 3.3 | Optimum, compromise, the original plan from the study, and traditional test plans as well as the large-sample approximate standard error for the metal wear example and assuming \(n = 12\) test units (as in the original experiment). | 49 |
| Table 3.4 | Optimum, compromise, the original plan from the study, and traditional test plans for the LED data. | 52 |
| Table 3.5 | Large-sample approximate standard error for the test plans given in Table 3.4 and assuming \(n = 180\) test units (as in the original experiment). | 53 |
| Table 3.6 | Optimum, compromise, the original plan from the study, and traditional test plans for the LED data where a constraint was put on the experimental region. | 55 |
| Table 3.7 | Large-sample approximate standard error for the test plans given in Table 3.6 and assuming \(n = 180\) test units (as in the original experiment). | 57 |
Table 3.8  Comparison of values of ASE(\(\hat{t}_p\)) and simulation method for different
RMADT plans and different sample sizes for the scar-width data. . . . 57

Table 4.1  Different combinations of prior specification. \(p_0(\vartheta)\) denotes the point-
mass prior, \(p_1(\vartheta)\) denotes the strongly informative prior for all param-
eters, \(p_2(\vartheta)\) denotes the weakly informative prior for all parameters,
\(p_3(\vartheta)\) denotes the strongly informative prior for only one parameter,
and \(p_4(\vartheta)\) denotes the diffuse prior. .............................. 73

Table 4.2  Prior information specified by independent marginal distributions for
the approximate point-mass prior distributions. ......................... 76

Table 4.3  Prior information specified by independent marginal distributions for
the strongly informative prior distributions. ............................ 76

Table 4.4  Prior information specified by independent marginal distributions for
strongly informative for only one parameter prior distribution. ...... 76

Table 4.5  Prior information specified by independent marginal distributions for
the weakly informative prior distributions. ............................. 77

Table 4.6  Optimum test plans for Cases A and B and the absolute value of the
Bayesian planning criterion in (4.9) evaluated at the prior information
for different sample sizes. .................................................. 79

Table 4.7  Optimum test plans for Cases C, D, and E and the absolute value
of the Bayesian test planning criterion in (4.9) evaluated at the prior
information for different sample sizes. .................................. 80

Table 4.8  Compromise test plans and optimum test plans for Case C and for
different sample sizes. Also present on this table is the absolute value
of the Bayesian test planning criterion in (4.9) evaluated at sample
size, and test plans where the expectation is with respect to the prior
distribution for test planning. ............................................. 80

Table 4.9  Summary statistics for the simulations of \(t_{0.10}\) at the use conditions for
Case B and for Case E. ....................................................... 83
Table 4.10  Summary statistics for the simulations of the posterior estimate of $\beta + \gamma x_U$ for Case B and for Case E.
LIST OF FIGURES

Figure 2.1 Simulated shelf-life degradation data for $n = 12$ units. .................. 9
Figure 2.2 Scatterplot of printhead migration data. The horizontal line indicates
point at which a failure is declared. ............................................. 10
Figure 2.3 Estimated 0.10 (solid line), 0.50 (dashed line), and 0.80 (dotted line)
quantiles of the degradation distribution at different points in time. . . 15
Figure 2.4 ML estimates of $d_{0.10}$ as a function of years in service, along with ap-
proximate 95% confidence intervals. The vertical lines indicate the
point where the data analysis was done after 5 years (a), after 10 years
(b), and after 20 years (c). ......................................................... 16
Figure 2.5 Contour plot of the asymptotic standard error $SE_{\hat{d}_{0.10}}$ as a function of
$n$ and $m$. For 2.5a the measurement error variability $\sigma^2$ is small and
for 2.5b the measurement error variability is large. ....................... 18
Figure 2.6 Comparison of asymptotic standard errors of $\hat{d}_{0.10}$ for 2 different test
plans with extrapolation out to 20 years. The point where the line
changes from solid to dashed or dotted is the time at which extrapola-
tion begins. ................................................................. 19
Figure 2.7 Simulations of ML estimates of the 0.10 quantile of the degradation
level using the planning values in Example 3. The longer thick line
represents the 0.10 quantile for the planning information. ............... 21
Figure 2.8  Asymptotic standard errors of the estimator of the 0.1 quantile of the degradation level for the 12(21) test plan using simulation (solid lines) and large sample approach (dashed lines) for comparing different test plans. ............................................. 22

Figure 2.9  Contour plot of cost (thick solid lines) and $SE_{\hat{d}_{0.10}}$ (dashed lines) for Example 6. The asterisk indicates the constrained optimum test plan. The cross-hatched region corresponds to the pairs $(n, m)$ that do not satisfy the constraint $SE_{\hat{d}_{0.10}} \leq 0.50$. The labels for the cost contours have been multiplied by $10^{-3}$ for readability. ......................... 23

Figure 2.10  Estimated failure-time distribution based on the printhead data. The dashed lines represent pointwise approximate 95% confidence intervals. 24

Figure 2.11  Contour plot of the asymptotic standard error $SE_{\hat{t}_{0.10}}$ as a function of $n$ and $m$. ................................................................. 27

Figure 2.12  Contour plot of cost (thick solid lines) and $SE_{\hat{t}_{0.10}}$ (dashed lines). The asterisk indicates the constrained optimum test plan. The cross-hatched region corresponds to the pairs $(n, m)$ that do not satisfy the constraint $SE_{\hat{t}_{0.1}} \leq 0.80$. The labels for the cost contours have been multiplied by $10^{-3}$ for readability. ......................... 28

Figure 3.1  Carbon-Resistor data from Shiomi and Yanagisawa (1979) plotted on the original scale and a square root of time scale. The horizontal line is the point for which a failure is declared (resistance = 230 ohms). . . 35

Figure 3.2  Scar width data of Meeker and Escobar (1998) plotted on the original scale and on a log-log scale. The horizontal line represents the point of degradation where a failure is declared (scar width = 50 microns). . . 36
Figure 3.3  LED output versus time in hours. The rows of the graph correspond to levels of current and the columns correspond to levels of junction temperature. The figure on the left (3.3a) presents the original data. The figure on the right (3.3b) presents the original data on a square root of time scale.

Figure 3.4  Plot of the directional derivative in (3.9), evaluated at the optimum plan in Section 3.5 versus lowest level of temperature.

Figure 3.5  The experimental region for the LED test plan where the four points correspond to the optimal test plan.

Figure 3.6  Plot of (3.9) versus ambient temperature and current evaluated at the optimum plan found in the LED example.

Figure 3.7  Constrained experimental region for the LED data. The area within the shape corresponds to acceptable testing conditions. The six dots represent the test points of the original study.

Figure 3.8  The constrained experimental region for the LED test plan where the four points correspond to the optimal test.

Figure 3.9  Plot of the directional derivative in (3.9) versus temperature and current evaluated at the optimum plan found in the LED example and under the constraint in Figure 3.7. Visually, the lines are the same for all levels of ambient temperature and are stacked on top of each other in this plot.

Figure 3.10 10000 simulated ML estimates of failure-time distribution quantiles at use conditions (50°C), for the different plans as in Table 3.3. The vertical line indicates the failure-time quantile at use conditions, evaluated at the planning information.

Figure 4.1  Carbon-Resistor data from Shiomi and Yanagisawa (1979) plotted on the original scale and a square root of time scale.
Figure 4.2  Plot of the directional derivative in (4.10) evaluated at the optimum test plan as a function of $x$ for Case A.  

Figure 4.3  Results of 5000 simulations of the posterior median of $t_{0.10}$ at the use conditions for Case B with diffuse prior information for inference (left plot) and for Case E with strongly informative prior information for inference (right plot). The black, vertical line in the plots corresponds to the true value of $t_{0.10}$ which is $0.19$.  

Figure 4.4  5000 simulations of the posterior estimate of $\beta + \gamma x_U$ for Case B (left plot) and for Case E (right plot).
ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those many individuals who introduced me to statistics, inspired me, and helped me to survive in this difficult field.

First and foremost, I would like to thank Dr. William Q. Meeker for his patience, willingness to work long distance, and extremely thoughtful advice. I will forever be indebted to him for the experience of being one of his fortunate students.

Besides my advisor, I would like to thank other Statisticians who have influenced me throughout my graduate career. I would like to thank Dr. Harry Martz for getting me “sold” on statistics. I would also like to thank Dr. Alyson Wilson for taking a chance on some math student from the University of New Mexico and becoming an influential mentor. I would next like to thank Dr. Max Morris for being an exceptional instructor, giving me help and advice on General Equivalence theory, and most importantly being extremely understanding and helpful during a very difficult time of my life.

I would also like to thank the other members of my committee: Dr. Mervyn Marasinghe and Dr. Vivekananda Roy for their patience and insightful advice on my dissertation research.

I must also mention the wonderful collaboration I have had between Dr. Luis Escobar from the Department of Experimental Statistics, Louisiana State University and Dr. Joanne Wendelberger from the Statistical Sciences Group with Los Alamos National Laboratory. This collaboration provided me the inspiration and early theoretical results that turned into a thesis.

Last but not least, I would like to thank my wife, daughters, friends, and parents (including in-laws) for their endless love, support, encouragement, and (sometimes) money.
ABSTRACT

The failure mechanism of an item often can be linked directly to some sort of degradation process. This degradation process eventually weakens the item which then induces a failure. As system components have become highly reliable, traditional life tests, where the response is time to failure, provide few or no failures during the life of a study. For such situations, degradation data can sometimes provide more information for assessing the item’s reliability. Repeated measures degradation is a form of degradation where the engineers are able to make multiple nondestructive measurements of the item’s level of degradation. For some items, however, the degradation rates at nominal use conditions are so low that no meaningful information can be extracted. Thus the engineers will use accelerating methods to increase the degradation rate. Before a test can be performed, the engineers need to know the number of items to test, the points of time to make the measurements, and at what values of the accelerating variable should the units be exposed in order to achieve the best estimation precision possible. In this thesis we study the test planning methods for designing repeated measures degradation and accelerated degradation tests. First, Chapter 2 provides methods for selecting the number of units and the number of measurements per unit for repeated measures degradation tests without acceleration. Selection of these testing parameters is based on the asymptotic standard error of an estimator of a function of the model parameters. These methods can also be used to assess how the estimation precision changes as a function of the number of units and measurements per items. Chapter 3 describes methods for planning repeated measures accelerated degradation tests (RMADTs) where the engineers need to know the accelerated conditions at which the items should be tested. Chapter 4 is similar to Chapter 3, but uses a Bayesian approach for planning RMADTs.
CHAPTER 1. General Introduction

1.1 Introduction

Manufacturers generally need to develop highly reliable products. To demonstrate reliability, engineers depend on different sorts of tests to estimate the failure-time distribution of the components and subsystems. Traditional life tests, where the response is time to failure, were originally used (and in some cases are still used). As products became more reliable, however, traditional life tests would produce few or no failures within a reasonable amount of time, making reliability assessment extremely difficult. An alternative to these traditional life tests is to make measurements of the degradation of item that are related to future failure. There are two different forms of repeated measures degradation data. In some applications, actual degradation can be observed with time. An example of this would be a crack growing with time on some component of an item. As the crack grows to a certain width, the item will fail. Another example would be the depth of the tread of an automobile tire. When the tread depth gets to a certain level (say 2/32 of an inch) the tire is deemed to be unsafe.

Another form of degradation is when the actual degradation process cannot be observed but measuring the product’s performance is possible. An example of this is the reduction of luminosity of a light-emitting diode (LED). When the luminosity has decreased to a certain level, the LED is considered to have failed. Degradation data can be used to assess product reliability even if a traditional life test of the product would produce very few or no failures. Moreover, observations of the physical degradation process or surrogate of this process may allow modeling of the underlying failure causing mechanism, providing more justification when extrapolations are needed (as in accelerated testing).

There are many applications for which degradation (both physical degradation and per-
formance based degradation) can be easily measured. In particular, it may be rather easy to make many measurements of the degradation over time without disrupting the degradation process or destroying the item to make the measurement (as opposed to destructive degradation). For example, the luminosity of an LED can be measured at different points of time and one would not expect this type of measurement to harm the LED or alter its degradation process. This type of degradation measurement is known as repeated measures degradation. These situations where one is able to make repeated measurements of degradation provide the motivation for the current work. A test with repeated measured degradation is called a “repeated measured degradation test” or in the presence of acceleration “repeated measures accelerated degradation test”.

With limited resources (money, time, facilities, etc.), engineers need to carefully plan tests. When planning a test, engineers need to determine the points in time at which to make measurements (including the length of time the test will be performed), the accelerated conditions to which units will be exposed (in an accelerated test), and the proportion of units to be allocated to these accelerated conditions. Additionally, engineers need to specify the form of a degradation model, planning information for model parameters, a distribution form of the variability, the level of degradation for which a unit fails, the number of units to be tested, and a range of the accelerating variable (in an accelerated test). The planning information mentioned previously should be based on past experience with similar items or expert judgment. It may be that the engineers are uncertain about the planning information and instead provide a plausible range for these values.

It is the job of the statistician to work with the engineers during the test planning phase to help get the best accuracy and precision for estimation possible with the limited resources that are available. One approach that is commonly used in estimation is the method of maximum likelihood (ML) estimation. A useful criterion for choosing a degradation test plan is the minimization of the variance of the ML estimator of a function of interest (such as a quantile of the degradation distribution or the failure-time distribution). We achieve this minimization using the large-sample approximate variance of the ML estimator which is obtained by inverting...
the Fisher information matrix. If the engineers could assume that the model and planning information for the parameters were known without error, then one could find the optimum test plan that would minimize the asymptotic variance among all test plans. However, engineers never know the model parameters and alternate test plans called compromise test plans could be used that are robust to deviations from planning assumptions and still have good statistical properties.

The main purpose of this research is to study planning methods for both accelerated and unaccelerated repeated measures degradation tests. This research is motivated by several different applications. The methods that are developed here are general enough that they could be applied to many other situations.

1.2 Dissertation Organization

This dissertation consists of three main chapters following this general introduction, and ending with a general conclusion. Each of the main three chapters corresponds to a journal article. Chapter 2 presents methods for planning repeated measures degradation tests where all the units are tested at the same physical conditions (such as temperature, humidity, etc.). This chapter looks at how different testing factors affect estimation precision. These factors include the number of items to be tested, the number of measurements in time, different measurement schedules, and test planning under the constraint on the cost of a test. This chapter assumes that there is only one failure-causing mechanism and that the units degrade within a reasonable amount of time. Chapter 3 addresses the situation when the units will not produce any meaningful degradation at the nominal use conditions. An accelerating factor is used to accelerate the degradation of the items. This chapter looks at how to choose the levels of an accelerating variable(s) to be used in a test and the proportion of the units to be tested at those levels so that the best estimation precision is achieved. This chapter assumes that the planning information for the model parameters is exact. In reality, however, engineers never know the exact values of parameters and instead can only describe a plausible range of values. This information could be described as a joint prior distribution for the model
parameters. Chapter 4 presents Bayesian methods that incorporates this prior information into test planning.

1.3 Literature Review

Laird and Ware (1982), Jenrich and Schluchter (1986), and Lindstrom and Bates (1988) present the theoretical results for calculating the Fisher information matrix for the type of models this dissertation assumes. Gertsbakh and Kordonsky (1969) discuss repeated measures degradation for a simple linear model with random slope and intercept for assessing product reliability. They assumed zero correlation between the random slope and intercept term. They then present the failure time distribution under this degradation model which is known as the Bernstein distribution. Chapters 13 and 21 of Meeker and Escobar (1998), Chapter 7 of Tobias and Trindade (1995), Lu and Meeker (1993), and Meeker, Escobar, and Lu (1998) discuss derivation of the failure-time distribution that is induced by of the degradation model and present statistical methods for estimation of degradation model parameters based on repeated measures data. Yu and Tseng (1999) and Yu (2003) illustrates design methods for planning degradation tests under a cost constraint where the lifetime distribution is lognormal for the unaccelerated and accelerated cases, respectively. In particular, they focus on estimation precision of a $p$ quantile of the failure-time distribution based on the asymptotic distributions of their moment-based estimators. This differs from our work since our work is based on asymptotic distributions of maximum likelihood estimators and for the accelerated case Whittle's general equivalence theorem is used to verify optimality. Moreover, their assumed model is a special case of our more general model. Boulanger and Escobar (1994) discuss experimental design for repeated measures accelerated degradation tests where the amount of degradation over time levels off to a plateau.

Decision theoretic approaches for Bayesian experimental design have been presented by Raiffa and Schlaifer (1961) and Lindley (1972). Chaloner and Verdinelli (1995) give a useful review of Bayesian experimental design theory and methods. Chaloner and Larntz (1992) and Zhang and Meeker (2006) present Bayesian methods for planning accelerated life tests

1.4 References


CHAPTER 2. Methods For Planning Repeated Measures Degradation Studies

Brian P. Weaver and William Q. Meeker
Department of Statistics
Iowa State University
Ames, IA, 50010

Luis A. Escobar
Department of Experimental Statistics
Louisiana State University
Baton Rouge, LA, 70803

Joanne R. Wendelberger
Statistical Sciences Group
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Abstract

Repeated measures degradation studies are used to assess product or component reliability when there are few or even no failures expected during a study. Such studies are used to assess the shelf life of materials and products. We show how to evaluate the properties of proposed test plans needed to identify statistically efficient tests. We consider test plans for applications where parameters related to the degradation distribution or the lifetime distribution are to be estimated. We use the approximate large-sample variance-covariance matrix of the parameters of a mixed effects linear regression model for repeated measures degradation data to assess the effect of sample size (number of units and number of measurements within the units) on estimation precision of both degradation and failure-time distribution quantiles. We also illustrate the complementary use of simulation-based methods for evaluating and comparing test plans. These test-planning methods are illustrated with examples.
Keywords: Repeated Measures Planning, Aging and Degradation, Lifetime Distributions, Degradation Distributions

2.1 Introduction

2.1.1 Motivating Examples

Engineers often need to quantify the failure-time distribution of highly reliable items. Traditional life tests, where the response is time to failure, typically yield few or no failures. Instead, engineers can sometimes use methods that measure the degradation of an item, providing more information than the traditional life tests. One such method is to use non-destructive repeated measurements over time on the degradation of each item. Given a degradation model and a relationship between degradation and failure, a failure-time distribution can be established. Before the test is performed, however, the engineers need to decide how many items should be measured and how often should these measurements be made in order to achieve a certain level of precision.

This work is motivated by two different applications that we have encountered. The first application involved a long-term shelf-life study on chemical degradation of a certain compound in a particular environment. A total of 12 items were randomly selected from a much larger population of items in storage. The engineers would then make annual measurements of the concentration of the chemical compound in units of parts per million volume (ppmv). Because of the importance of the application the available data would be analyzed and a summary report would be prepared annually. Since the data were sensitive and not available for release, Figure 2.1 shows data that were simulated on a modified scale to mimic the original study. The question asked by the engineers was, “Given the pattern of the observations in Figure 2.1 (from a previous similar study), how should the next shelf-life study be performed?”

The second application involves a study involving inkjet printer heads. The engineers involved in this example were interested in performing a system reliability study for which the print heads were a component. The engineers wanted an estimate of the failure-time distribution where failure-time depends on the degradation level of the print head. Degradation
Figure 2.1 Simulated shelf-life degradation data for $n = 12$ units.

was defined to be the diffusion of an ink-related substance in the printheads. As time progresses, if this substance reaches a certain location in the printhead, a failure will soon follow.

In the experiment, measurements were taken periodically on a sample of 12 parts. At each inspection time, the parts were measured to determine how far this substance had moved (in millimeters) after a certain amount of time. Figure 2.2 shows a scatterplot of the print head degradation data. Again, the data were scaled to protect proprietary information. The first point in time (time point zero) is considered the point for which the printhead had been initially loaded with ink. According to the coordinate system used, failure will occur when the degradation level reaches 60 mm.

2.1.2 Related Work

This section reviews some of the literature on degradation test plans or related applications. Yu and Tseng (1999) discuss how the optimization of degradation plans under the constraints of total experimental costs and the assumption that the lifetime distribution is lognormal. Lu and Meeker (1993) derive an analytical form of the lifetime distribution under different models and assumptions on the model parameters. Lenth (2006), on his website, has several Java programs that allow design of experiments based on controlling power. Diggle, Heagerty, Liang,
and Zeger (2002) give sample size calculations for longitudinal models where the number of measurements per unit is specified. Boulanger and Escobar (1994) discuss experimental design for accelerated degradation tests where the amount of degradation over time levels off to a plateau. Vickers (2003) discusses how adding more measurements in a repeated measure study can affect the power of the test (i.e., the rate of detecting a difference caused by a treatment when a difference is truly present).

2.1.3 Overview

The remainder of this paper is organized as follows. Section 2.2 describes the linear degradation model used in our work. Section 2.3 gives the likelihood function and the asymptotic variance-covariance matrix of the maximum likelihood estimator, followed by a discussion of parameter estimation. Section 2.4 gives the degradation distribution quantile function and shows the use of the Fisher information matrix for inference on this function. Section 2.5 illustrates the use of the Fisher information matrix and simulation for degradation test planning and for comparing test plans. Section 2.6 describes test plans that focus on estimating quantities of the failure-time distribution induced by the degradation model. Section 2.7 gives
conclusions and describes possible areas for future related research.

2.2 Repeated Measures Degradation Model

2.2.1 Model and Data

Let $y_{ij}$ be the observed degradation at time $t_{ij}$ on unit $i$ where $i = 1, \ldots, n$ and $j = 1, \ldots, m_i$. The linear degradation random effects model is

$$y_{ij} = D_{ij} + \epsilon_{ij} \quad (2.1)$$

where the actual degradation path is

$$D_{ij} = b_{0i} + b_{1i} t_{ij}. \quad (2.2)$$

The intercept $b_{0i}$ and the slope $b_{1i}$ are modeled as random realizations from the bivariate-normal distribution $(b_0, b_1)^T \sim \text{BVN} (\beta, V)$, where the elements of $\beta = (\beta_0, \beta_1)^T$ are fixed terms representing the population’s mean intercept and slope and

$$V = \begin{pmatrix} \sigma_{b_0}^2 & \rho \sigma_{b_0} \sigma_{b_1} \\ \rho \sigma_{b_0} \sigma_{b_1} & \sigma_{b_1}^2 \end{pmatrix}$$

is the covariance matrix.

Collecting into $Y_i = (y_{i1}, \ldots, y_{im_i})^T$ the observations from unit $i$, an equivalent expression for the linear degradation model in (2.1) is

$$Y_i = X_i \beta + Z_i b_i^* + \epsilon_i, \quad (2.3)$$

where $b_i^* = (b^*_{0i}, b^*_{1i})^T$ is modeled as $(b^*_{0i}, b^*_{1i})^T \sim \text{BVN} (0, V)$, $X_i$ and $Z_i$ are matrices of explanatory variables defined by

$$X_i = Z_i = \begin{pmatrix} 1 & t_{i1} \\ \vdots & \vdots \\ 1 & t_{im_i} \end{pmatrix},$$

and $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{im_i})^T$. 
Assuming independence between $\epsilon_i$ and $b_i^*$ and that the components of $\epsilon_i$ are independent and jointly normal distributed, that is $\epsilon_i \sim \text{MVN}(0, \sigma^2 I_i)$ where $I_i$ is a $m_i \times m_i$ identity matrix, it follows that $Y_i \sim \text{MVN}(X_i\beta, \Sigma_i)$ with

$$\Sigma_i = \text{Var}(Y_i) = \text{Var}(X_i\beta + Z_i b_i + \epsilon_i) = Z_i V Z_i^T + \sigma^2 I_i.$$  \hspace{1cm} (2.4)

Notice that the independence assumption among the components of $\epsilon$ implies that the error terms are not autocorrelated which is a reasonable assumption when spacing between observations is not too small.

See Jenrich and Schluchter (1986) for more details.

### 2.3 Model Likelihood Function and Fisher Information Matrix

#### 2.3.1 Likelihood

Suppose that $y_1, \ldots, y_n$ are $n$ independent observations from $Y_1, \ldots, Y_n$, respectively. The log-likelihood for observational unit $i$ is

$$L_i = -\frac{1}{2} \log |\Sigma_i| - \frac{1}{2} (y_i - X_i\beta)^T \Sigma_i^{-1} (y_i - X_i\beta).$$  \hspace{1cm} (2.5)

The total log-likelihood for $n$ units is

$$\mathcal{L} = \sum_{i=1}^{n} L_i = -\frac{1}{2} \sum_{i=1}^{n} \log |\Sigma_i| - \frac{1}{2} \sum_{i=1}^{n} (y_i - X_i\beta)^T \Sigma_i^{-1} (y_i - X_i\beta).$$  \hspace{1cm} (2.6)

#### 2.3.2 Variance Covariance Matrix

Let $\theta = (\beta^T, \vartheta^T)^T$ be the parameter vector where $\vartheta = (\sigma_{b_0}, \sigma_{b_1}, \rho, \sigma)^T$. Recall that the Fisher information matrix is defined as $I(\theta) = -E[\partial^2 \mathcal{L}/\partial \theta^2]$. From large sample theory, the large-sample approximate covariance matrix of the maximum likelihood (ML) estimators is

$$\text{AVar}(\hat{\theta}) = [I(\theta)]^{-1}.$$  \hspace{1cm} (2.7)

$\text{AVar}(\hat{\theta})$ can be estimated by evaluating (2.7) at the ML estimates $\hat{\theta}$. We denote this estimate by $\hat{\text{AVar}}(\hat{\theta})$. The derivation of the information matrix is given in the appendix.
2.3.3 Linear Mixed-Model Parameter Estimation

Our focus is on test planning, but it is necessary to mention how mixed effects model parameters can be estimated. Laird and Ware (1982) discuss ML and restricted ML (REML) parameter estimation for a general class of mixed effects models, which includes our repeated measures model, using the EM algorithm. Jenrich and Schluchter (1986) derive the derivatives and second derivatives needed in a Newton-Rhapson algorithm for ML estimation of parameters for a general class of model that includes mixed effects models. Lindstrom and Bates (1988) extended the work of Laird and Ware (1982) and Jenrich and Schluchter (1986) and developed efficient algorithms for computation of both ML and REML estimates for mixed-effects models.

Faraway (2006) is a useful reference for methods to estimate model parameters using the R software package. In particular, he focuses on using the package “lme4” with the built-in function “lmer” to perform all the analysis and find the estimates of both the fixed and random effects. For more information on the package “lme4” and its functions, see Bates and Maechler (2009). Another model fitting function in R is “lme” found in the package “nlme.” This package contains functions that also allow for interval estimation of both fixed and random effects model parameters as well as best linear unbiased predictors of the response. We use this function to calculate ML estimates of the model parameters. See Pinheiro, Bates, Debroy, Sarkar and the R Core Team (2008) for additional information on “nlme” and its underlying functions. Although both lme and lmer both produce ML and REML estimates, there are some differences between the two packages. To name a few, lmer is a quicker function than lme and also handles crossed random effects, unlike lme which only handles nested random effects. lme is a more stable function and is easier for handling heteroscedasticity than lmer and provides p-values for significance of effects (though there is debate on the validity of these p-values). Finally, the SAS procedure “Proc Mixed” is widely used for the fitting and estimation of mixed effects models. Littell, Milliken, Stroup, Wolfinger, and Shabenberger (2006) give many examples of fitting mixed effects models in SAS as well as the theory involved.
2.4 Estimating Quantiles of the Degradation Distribution

2.4.1 The Quantile of the Degradation Distribution

From the model in Section 2.2.1, it follows that the degradation at time $t$ is given by $D = b_0 + b_1 t$. When $(b_0, b_1)^T$ has a bivariate normal distribution, $D$ is normally distributed with $E(D) = E(b_0 + b_1 t) = \beta_0 + \beta_1 t$ and $\text{Var}(D) = \text{Var}(b_0 + b_1 t) = \sigma_{b_0}^2 + \sigma_{b_1}^2 t^2 + 2t \rho \sigma_{b_0} \sigma_{b_1}$. The $p$ quantile of the degradation distribution at time $t$ is

$$d_p(t) = E(D) + \Phi^{-1}_{\text{nor}}(p) \sqrt{\text{Var}(D)}$$

where $\Phi^{-1}_{\text{nor}}(p)$ is the inverse standard normal cumulative distribution function. The ML estimate of $d_p$ can be computed by evaluating (2.8) at the ML estimates $\hat{\theta}$.

Example 1 Consider the simulated shelf-life data set in Figure 2.1. The data were simulated using (2.1) and the parameter values $\beta_0 = 8$, $\beta_1 = -0.2$, $\sigma_{b_0} = 0.3$, $\sigma_{b_1} = 0.04$, $\rho = 0.7$, and $\sigma = 0.3$ for $n = 12$ items and for a length of time of 20 years. The R function “lme” provides the ML estimates of these parameters as $\hat{\beta}_0 = 7.98$, $\hat{\beta}_1 = -0.19$, $\hat{\sigma}_{b_0} = 0.35$, $\hat{\sigma}_{b_1} = 0.05$, $\hat{\rho} = 0.9$, and $\hat{\sigma} = 0.28$. For given values of $p$ and $t$, the ML estimate of the degradation quantile is (2.8) evaluated at the ML estimates of $\theta$. This is illustrated in Figure 2.3 for $p = 0.10, 0.50$, and $0.80$ and at different points in time.

2.4.2 Standard Error for the Maximum Likelihood Estimator of the $p$ Quantile

This section deals with the estimation of the standard error of the ML estimator of the degradation quantile $d_p$ in (2.8). This quantile is a function of the parameters $\theta = (\beta_0, \beta_1, \sigma_{b_0}, \sigma_{b_1}, \rho, \sigma)^T$. Using the invariance property of ML estimators, the ML estimator $\hat{d}_p$ of $d_p$ is obtained by evaluating (2.8) at $\hat{\theta}$. The formula for the approximate standard error of $\hat{d}_p$ was derived using the delta method. Let $c$ be a vector with elements $c_i = \partial d_p / \partial \theta_i$, $i = 1, \ldots, 6$. Then by the delta method, the large-sample approximate variance of $\hat{d}_p$ is

$$\text{AVar}(\hat{d}_p) = c^T \text{AVar}(\hat{\theta}) c.$$
Figure 2.3 Estimated 0.10 (solid line), 0.50 (dashed line), and 0.80 (dotted line) quantiles of the degradation distribution at different points in time.

The standard error of $\hat{d}_p$ is $SE_{\hat{d}_p} = \sqrt{AVar(\hat{d}_p)}$ which is estimated by evaluating (2.9) at $\hat{\theta}$ giving $SE_{\hat{d}_p} = \sqrt{AVar(\hat{d}_p)}$. The explicit forms of the partial derivatives are given in the appendix.

2.4.3 Confidence Interval for the Degradation Distribution Quantiles

An approximate 100$(1 - \alpha)$ confidence interval for $d_p$ is

$$[\hat{d}_p - \bar{d}_p, \hat{d}_p + \bar{d}_p] = \hat{d}_p \pm z(1 - \alpha/2)SE_{\hat{d}_p}, \quad (2.10)$$

where $z(1 - \alpha/2)$ is the $1 - \alpha/2$ standard normal quantile.

**Example 2** Returning to Example 1, Figures 2.4a, 2.4b, and 2.4c show 95% confidence intervals for $d_{0.10}$ when extrapolations were made to estimate degradation based on data available at 5 years, 10 years, and 20 years respectively. As expected, the width of the confidence intervals decreases as more information becomes available.
2.5 Degradation Test Planning

This section describes planning methods for repeated measures degradation tests. Section 2.5.1 shows a simple, graphical approach for test planning that assumes all units have the same inspection schedule. Section 2.5.2 describes an approach that allows for different schedules for different units. Section 2.5.3 illustrates a simulation-based approach that complements the analytical evaluations and can also be applied to the more general settings. Finally Section 2.5.4 describes an approach for minimizing the cost of a test subject to a constraint on estimation precision.

2.5.1 Simple Degradation Test Plans

In a simple degradation test plan all units are measured using the same schedule. Evaluation of statistical test-plan properties help to determine the number of units to measure in the study and how many measurements should be made over time. We use the asymptotic standard error $\SE_{d_p}$ to quantify and compare the precision provided by alternative test plans. In particular, we obtain a contour plot of the $\SE_{d_p}$ values obtained over a grid of $n = 3, 4, \ldots, 10$ experimental units and $m = 3, 4, \ldots, 10$ measurements per unit. Test-plan decisions and recommendations...
are based on the actual values of $\text{SE}_{\hat{d}_p}$ calculated over the grid and the corresponding contour plot. In the following two examples, we use rather extreme values for the measurement error variability planning values to illustrate the strong effect that this parameter can have on degradation test plans.

**Example 3** Suppose that the objective is to assess the trade-off between the number of measurements per unit and the number of units being used in the study. The shelf-life study is expected to run for 20 years. The parameter values from Example 1 are used as the planning information and they are denoted by $\beta_0 = 8$, $\beta_1 = -0.2$, $\sigma_{\beta_0} = 0.3$, $\sigma_{\beta_1} = 0.04$, $\rho = 0.7$, and $\sigma = 0.3$, that is, a parameter value with an empty box in the superscript. Figure 2.5a shows the results for some simple test plans. This plot shows that for the proposed planning values, the smallest standard error that could be obtained is less than 0.46 (corresponding to $n = 10$ and $m = 10$). The plot suggests that a trade-off could be made by choosing a small number of units, say 6, and measuring them 7 times without losing much in terms of precision (in this region $\text{SE}_{\hat{d}_p} \approx 0.59$). Because the measurement error is relatively small in this example, increasing the number of measurements over time will not have a large effect on estimation precision.

In the next example, we use a much larger planning value for the measurement error variability to illustrate its effect on estimation precision.

**Example 4** Now suppose that the planning information value for $\sigma$ is increased to $\sigma = 3$. Figure 2.5b shows the asymptotic standard error $\text{SE}_{\hat{d}_{0.10}}$ for different combinations of $n$ and $m$ with the new planning information. In this case, the plot shows that a test should be chosen from the North–East region where $\text{SE}_{\hat{d}_{0.10}}$ is less than 1.5. In the South–West corner of the plot, however, $\text{SE}_{\hat{d}_{0.10}}$ reaches values larger than 5.5. In summary, to compensate for the large variability in measurements (i.e., large $\sigma$), the test plan requires more units and more measurements per unit to achieve a smaller standard error when compared to $\sigma = 0.3$ in Example 3.
2.5.2 Degradation Test Plans with Differing Schedules

The use of different inspection schedules for groups of units is motivated by two concerns of test planners:

- Inspections can be expensive and there can be substantial savings if some units are sampled less frequently.

- There was concern that the measurements could have an effect on the degradation process. Having groups of units on different inspection schedules can provide information to detect and model the effect of such changes, if they exist.

Example 5 Again consider a shelf-life study that is to be performed over a period of 20 years. Periodic evaluation of a sample of units is scheduled for 5, 10, and 20 years. Some questions of interest are “what is the current state of the units in the larger population of units from which the sample was taken” and “can the future state of the population of units be predicted?” Two different plans will be compared for this study, $\mathcal{N}(21), \mathcal{N}(11), \mathcal{N}(6)$ and $\mathcal{N}(21)$. This notation means that the first plan will use 12 units of which 4 units are measured 21 times (i.e., every
year starting at time 0), 4 are measured 11 times (i.e., every other year), and 4 units are measured 6 times (i.e., every 4 years). In the alternative plan, all 12 units will be measured every year. Notice that the first of these plans will involve 152 measurements and the second will have 252 measurements. Thus if the first plan gives adequate information, it would be preferred because it costs much less than the second plan. Figure 2.6a shows a comparison of these two plans. There is a large difference in the standard errors for plans 12(21) and 4(21),4(11),4(6) when extrapolation to 20 years is performed after 5 years of observations. After 10 years of observations, however, there is little difference between the two plans. These results suggest that the sampling plan 4(21),4(11),4(6), especially after 10 years of inspection could result in large savings in both time and money with little loss of precision.

If the number of units tested could be doubled from 12 to 24, estimation precision would be improved. Figure 6b shows the results from the two plans 24(21) and 8(21),8(11),8(6). It is easy to show that standard errors for the 24(21) test plan are the same as those from the 12(21) plan, divided by $\sqrt{2}$. This is only approximately so for the 8(21),8(11),8(6) plan with respect to the 4(21), 4(11), 4(6) plan.

(a) $n = 12$ units per test.  
(b) $n = 24$ units per test.

Figure 2.6 Comparison of asymptotic standard errors of $\hat{d}_{0.10}$ for 2 different test plans with extrapolation out to 20 years. The point where the line changes from solid to dashed or dotted is the time at which extrapolation begins.
2.5.3 Using Simulation to Evaluate Test Plans

This section describes a complementary simulation-based method for comparing test plans. Simulation provides visualization of sampling variability and insight into the test planning process. Simulation results, presented graphically, are particularly useful when communicating with engineers. In general, simulation methods for evaluating and comparing test plans are also useful in situations where the delta method might not provide a good approximation, when it is difficult to derive an analytical method, or when there is not enough time to derive an analytical method. The simulation algorithm for repeated measures degradation testing is as follows:

1. With a given test plan and planning values, simulate data vectors $Y_i^*$ from the model in (2.3) where $i = 1, \ldots, B$ and $B$ being a large number, say 10,000.

2. For each simulated data set $Y_i^*$, calculate the ML estimates of $\theta_i^*$, say $\hat{\theta}_i^*$.

3. Calculate functions of $\hat{\theta}_i^*$ that are of interest, say $g(\hat{\theta}_i^*)$ (e.g., $\hat{d}_{p,i}^*$).

4. Plot the estimates to illustrate the trial to trial variability.

5. Estimate the standard error of the components of $\hat{\theta}_i^*$ or $g(\hat{\theta}_i^*)$ by calculating the sample standard deviation of the simulated estimates.

Figure 2.7 shows an example of Step 3 for $\hat{d}_{0.10}^*$ in the above algorithm for the planning information given in Section 2.5.1. Figure 2.8 shows a comparison between the simulation method and the large sample method for the plan 12(21) when considering again the standard error for the 0.10 quantile of the degradation distribution. Notice that the two approaches agree with each other and are very similar in shape and numerical values. The simulation approach, however, indicates smaller standard errors than the large-sample approximation value for the 0.10 quantile for the case when extrapolation began after 5 years.
2.5.4 Selecting a Test Plan Under a Cost Constraint

This section describes the selection of a degradation test plan when there is a constraint on $\hat{SE}_{d_p}$ and there is a desire to minimize the cost of running the experiment. Suppose the cost of the experiment is

$$\text{cost}(n, m) = c_1 + c_2 n + c_3 nm,$$

where $c_1$ denotes the fixed cost of running the experiment, $c_2$ is the cost of testing an experimental unit, and $c_3$ is the cost of a measurement on an item. Although the approach is general, we will use the exact same measurement schedule for each unit and equally spaced inspections.

Let $\gamma$ denote the maximum acceptable value of $\hat{SE}_{d_p}$. Then we wish to find the values of $n$ and $m$, say $n^*$ and $m^*$ such that $\hat{SE}_{d_p} \leq \gamma$ and $\text{cost}(n^*, m^*) = \min_{n,m} \{\text{cost}(n, m)\}$.

Example 6  Consider a shelf-life study that is to be performed for 20 years with the planning information given in Example 3. The information after 10 years of observations will, however, be used to make important predictions at 20 years. This study has a limited budget and a test plan is to be chosen so that the cost of the study is to minimized subject to the constraint
The individual cost components of the study are $c_1 = $15,000, $c_2 = $1,500, and $c_3 = $75. Figure 2.9 shows the results of this optimization using the analytical methods. In the cross-hatched region $SE_{d_{10}} > 0.50$. The asterisk on the plot corresponds to the constrained optimum test plan. The plot indicates that $n = 11$ items should be measured at $m = 7$ equally-spaced times. For this test plan, $SE_{d_{10}} = 0.498$ and cost$(n,m) = $37,275.

### 2.6 Failure-Time Distribution

This section derives the failure-time distribution implied by the linear degradation model in (2.2) and a specification of the degradation level $D_t$ for failure. Chapter 13 of Meeker and Escobar (1998) provides a more general discussion of failure-time distributions that are implied by a degradation model.
2.6.1 Relationship Between Degradation and Failures

We assume a degradation process with soft failures. That is, the failure time for a unit is defined to be the time at which the degradation level reaches the specified degradation level $D_f$. Let $T$ define the random variable associated with the unit’s time to failure.

For a fixed $t$, $b_0 + b_1 t \sim \text{NOR}(\beta_0 + t \beta_1, \sigma_0^2 + t^2 \sigma_1^2 + 2t \rho \sigma_0 \sigma_1)$. First, consider the case of increasing degradation. In this case

$$\Pr(T \leq t) = F(t; \theta) = \Pr(b_0 + b_1 t \geq D_f) = 1 - \Phi_{\text{nor}}[\kappa(\theta)],$$

where $\kappa(\theta) = (D_f - \beta_0 - t \beta_1) / \sqrt{\sigma_0^2 + t^2 \sigma_1^2 + 2t \rho \sigma_0 \sigma_1}$ and $\Phi_{\text{nor}}$ is the cdf for a standard normal distribution. Similarly, if failure happens when the degradation level decreases to $D_f$, then

$$\Pr(T \leq t) = F(t; \theta) = \Pr(D = b_0 + b_1 t \leq D_f) = \Phi_{\text{nor}}[\kappa(\theta)].$$
When \( \rho = 0 \), \( F(t; \theta) \) is known as the Bernstein distribution (e.g., Lu and Meeker 1993, Gertsbakh and Kordonskiy 1969, and Ahmad and Sheikh 1984). The ML estimate of the failure-time distribution is \( F(t; \hat{\theta}) \) where \( \hat{\theta} \) is the ML of \( \theta \). Meeker and Escobar (1998), page 330, describes a numerical integration and a simulation based approach to evaluate the failure-time distribution for more complicated models where a closed form solution for the cdf \( F(t; \theta) \) does not exist.

**Example 7** The printhead of an inkjet cartridge is a component in a larger series system for a printer. Estimation of its lifetime distribution was needed to estimate the lifetime distribution for the entire system. As described in Section 2.1.1, the failure mechanism was diffusion of a failure-causing substance. The engineers defined a degradation level of \( D_t = 60 \) to be a failure. This degradation level is represented by the horizontal line in Figure 2.2. The ML estimates of the model parameters for the degradation model are

\[
\begin{align*}
\hat{\beta}_0 &= 11.22, \\
\hat{\beta}_1 &= 1.14, \\
\hat{\sigma}_{b_0} &= 0.45, \\
\hat{\sigma}_{b_1} &= 0.07, \\
\hat{\rho} &= -0.82, \text{ and } \hat{\sigma} = 2.6.
\end{align*}
\]  

(2.13)

*Figure 2.10* gives the ML estimate of the failure-time distribution for the print head degradation data.
2.6.2 Estimating the \( p \) Quantile of the Failure-Time Distribution

From (2.11) the \( p \) quantile of the failure-time distribution is

\[
t_p = -\left( k\sigma_{b_0 b_1} + h\beta_1 \right) \pm \sqrt{k^2\sigma_{b_0 b_1}^2 + k\sigma_{b_0}^2\beta_1^2 + h^2k\sigma_{b_1}^2 - k^2\sigma_{b_0}^2\sigma_{b_1}^2 + 2hk\beta_1\sigma_{b_0 b_1}}
\]

(2.14)

where \( \sigma_{b_0 b_1} = \rho\sigma_{b_0}\sigma_{b_1} \) is the covariance between \( b_0 \) and \( b_1 \), \( h = D_f - \beta_0 \) and \( k = [\Phi^{-1}_n(1-p)]^2 \) or \( k = [\Phi^{-1}_n(p)]^2 \) depending on whether a failure is declared when \( D \geq D_f \) or \( D \leq D_f \), respectively. The derivation of (2.14) is given in the appendix. If \( 0 < p < 0.5 \), \( t_p \) is the root where the radical is added. If \( 0.5 < p < 1 \), \( t_p \) is given by the root where the radical is subtracted.

2.6.3 Standard Error for the Maximum Likelihood Estimator of the Failure-Time Quantile

Let \( \mathbf{c} \) be the gradient vector with elements \( c_i = \partial t_p / \partial \theta_i, \ i = 1, \ldots, 6 \). Using the delta method, the large-sample approximate variance of \( \hat{t}_p \) is

\[
\text{AVar}(\hat{t}_p) = \mathbf{c}^T \text{AVar}(\hat{\theta}) \mathbf{c}.
\]

(2.15)

The standard error of \( \hat{t}_p \) is \( \text{SE}_{\hat{t}_p} = \sqrt{\text{AVar}(\hat{t}_p)} \) which is estimated by evaluating (2.15) at \( \hat{\theta} \) giving \( \text{SE}_{\hat{t}_p} = \sqrt{\text{AVar}(\hat{t}_p)} \). The explicit forms of the partial derivatives in \( \mathbf{c} \) are given in the appendix.

2.6.4 Degradation Test Planning Using \( t_p \)

This section applies the test planning techniques described in Sections 2.5.1 and 2.5.2 to \( t_p \). This work is motivated by the inkjet cartridge example. The engineers were interested in estimating \( t_{0.10} \), the time at which 10% of the items in the population would fail. They were interested in performing other degradation tests in the future on similar parts and wanted to know how many items should be measured and how many measurements should be made on each item. The ML estimates obtained in Example 7 will be used as the planning information (i.e., \( \beta_0^\square = 11.22, \ \beta_1^\square = 1.14, \ \sigma_{b_0}^\square = 0.45, \ \sigma_{b_1}^\square = 0.07, \ \rho^\square = -0.82, \) and \( \sigma^\square = 2.6 \)).
First, we consider the simple test plans described in Section 2.5.1 where each unit is measured the same number of times. Figure 2.11 shows several simple degradation test plans using the planning information given above. As expected the best plan is the 12(21) on the North–East corner for which $\text{SE}_{\hat{t}_{0.10}} \approx 0.83$.

A drawback to the plan 12(21) is that it might be expensive or time consuming to complete because it requires 252 measurements. Thus, it is of interest to entertain other degradation test plans that involve different measurement sequences on the units. For example, Table 2.1 shows the $\text{SE}_{\hat{t}_{0.10}}$ for the degradation plans 12(21) and 4(21),4(11),4(6), and 3(21), 3(11), 3(6), 3(3) respectively. Notice that there is not a large difference in the estimation precision for the three different plans. One could achieve savings for both time and money at the sacrifice of only a small amount of estimation precision if the plan 3(21), 3(11), 3(6), 3(3) (with 123 measurements) is selected over the plan 12(21).

<table>
<thead>
<tr>
<th>Plan</th>
<th>Number of Measurements</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>12(21)</td>
<td>252</td>
<td>0.83</td>
</tr>
<tr>
<td>4(21),4(11),4(6)</td>
<td>152</td>
<td>0.88</td>
</tr>
<tr>
<td>3(21),3(11),3(6),3(3)</td>
<td>123</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 2.1 Asymptotic standard error $\text{SE}_{\hat{t}_{0.10}}$ for three different degradation test plans.

As in Section 2.5.4, test planning to minimize cost under the constraint $\text{SE}_{\hat{t}_p} \leq \alpha$ can be applied to the failure-time distribution quantile. Consider again the cost structure from Example 6 but now with the constraint $\text{SE}_{\hat{t}_{0.1}} \leq 0.80$ for a test that is going to run for 50 hours. Figure 2.12 shows the constrained optimum plan to be $n = 16$ and $m = 8$. The cost associated with this test plan is $48,600 with $\text{SE}_{\hat{t}_{0.1}} = 0.791$.

### 2.7 Conclusions and Future Work

Nondestructive repeated measures degradation tests are useful in understanding the material or performance degradation of a product over time. It is important to plan these tests carefully in order to acquire the desired level of precision while working within resource con-
Figure 2.11 Contour plot of the asymptotic standard error $\hat{SE}_{t_{0.10}}$ as a function of $n$ and $m$.

Constraints (time, number of units, and number of measurements). The methodology presented in this paper can be extended to more complicated situations. The following list suggests future work:

- Extend to models with more complicated covariance structures such as autocorrelations which might be needed when one has smaller spacing between measurements.

- In some applications accelerated repeated measures degradation testing is needed (e.g., when using a regression model to describe the effect of temperature on degradation rates). For examples, see Chapter 21 of Meeker and Escobar (1998).

- Bayesian methods are often useful when there is prior knowledge (e.g., from physics of failure or previous experience with similar products). When such information is available, it should be incorporated into both the analysis and test planning.
Figure 2.12  Contour plot of cost (thick solid lines) and $SE_{t_{0.10}}$ (dashed lines). The asterisk indicates the constrained optimum test plan. The cross-hatched region corresponds to the pairs $(n, m)$ that do not satisfy the constraint $SE_{t_{0.1}} \leq 0.80$. The labels for the cost contours have been multiplied by $10^{-3}$ for readability.

2.8 Appendix

2.8.1 Derivation of the Information Matrix in Section 2.3.2

Using equation (4) of Jenrich and Schluchter (1986), it can be shown that, using our notation from Section 2.2, the Hessian Matrix, $H_i$, for unit $i$, is given by

$$
H_i = \begin{pmatrix}
H_{\beta\beta,i} & H_{\beta\phi,i} \\
H_{\phi\beta,i} & H_{\phi\phi,i}
\end{pmatrix} = \begin{pmatrix}
\frac{\partial^2 L_i}{\partial \beta \partial \beta} & \frac{\partial^2 L_i}{\partial \beta \partial \phi} \\
\frac{\partial^2 L_i}{\partial \phi \partial \beta} & \frac{\partial^2 L_i}{\partial \phi \partial \phi}
\end{pmatrix}.
$$

Then the information matrix can be expressed as

$$
I_i(\theta) = \begin{pmatrix}
X_i^T \Sigma_i^{-1} X_i & 0 \\
0 & M_i
\end{pmatrix},
$$

where $M_i$ is a $4 \times 4$ symmetric matrix with elements

$$
M_{jk}^i = \frac{1}{2} \text{tr}(\Sigma_i^{-1} \Sigma_j \Sigma_i^{-1} \Sigma_k), \ j = 1, \ldots, 4; \ k = 1, \ldots, 4,
$$
and

\[ \dot{\Sigma}_{ij} = \frac{\partial \Sigma_i}{\partial \phi_j}, \quad j = 1, \ldots, 4. \]

From equation (2.4), it follows that

\[ \dot{\Sigma}_{i1} = \frac{\partial \Sigma_i}{\partial \sigma_{b0}} = Z_i \begin{pmatrix} 2\sigma_{b0} & \rho \sigma_{b1} \\ \rho \sigma_{b1} & 0 \end{pmatrix} Z_i^T, \quad \dot{\Sigma}_{i2} = \frac{\partial \Sigma_i}{\partial \sigma_{b1}} = Z_i \begin{pmatrix} 0 & \rho \sigma_{b0} \\ \rho \sigma_{b0} & 2\sigma_{b1} \end{pmatrix} Z_i^T, \]

\[ \dot{\Sigma}_{i3} = \frac{\partial \Sigma_i}{\partial \rho} = Z_i \begin{pmatrix} 0 & \sigma_{b1} \sigma_{b0} \\ \sigma_{b1} \sigma_{b0} & 0 \end{pmatrix} Z_i^T, \quad \dot{\Sigma}_{i4} = \frac{\partial \Sigma_i}{\partial \sigma} = 2\sigma_i I_i. \]

Then the information matrix for all \( n \) units is

\[ I(\theta) = \sum_{i=1}^{n} I_i(\theta). \]

### 2.8.2 Forms of the Partial Derivatives in Section 2.4.2

The individual elements of \( c \) in (2.9) are

\[ \frac{\partial d_p}{\partial \beta_0} = 1, \quad \frac{\partial d_p}{\partial \beta_1} = t, \quad \frac{\partial d_p}{\partial \sigma_{b0}} = \zeta(2\sigma_{b0} + 2t\rho\sigma_{b1}), \]

\[ \frac{\partial d_p}{\partial \sigma_{b1}} = \zeta(2t^2 \sigma_{b1} + 2t\rho\sigma_{b0}), \quad \frac{\partial d_p}{\partial \rho} = \zeta(2t\sigma_{b0}\sigma_{b1}), \quad \text{and} \quad \frac{\partial d_p}{\partial \sigma} = 0, \]

where

\[ \zeta = \frac{\Phi_{\text{nor}}^{-1}(p)}{2 \sqrt{\sigma_{b0}^2 + \sigma_{b1}^2 t^2 + 2t\rho\sigma_{b0}\sigma_{b1}}}. \]

### 2.8.3 Derivation of \( t_p \) in Section 2.6.2

Let \( F \) denote the CDF of the random variable \( T \), corresponding to the first failure definition in Section 2.6, i.e., failure occurs when \( b_0 + b_1 t \geq D_t \).

\[
F(t_p) = 1 - \Phi_{\text{nor}} \left( \frac{D_t - \beta_0 - t_p \beta_1}{\sqrt{\sigma_{b0}^2 + t_p^2 \sigma_{b1}^2 + 2t_p \rho \sigma_{b0} \sigma_{b1}}} \right) = p
\]

\[
\Phi_{\text{nor}} \left( \frac{D_t - \beta_0 - t_p \beta_1}{\sqrt{\sigma_{b0}^2 + t_p^2 \sigma_{b1}^2 + 2t_p \rho \sigma_{b0} \sigma_{b1}}} \right) = 1 - p
\]

\[
\frac{D_t - \beta_0 - t_p \beta_1}{\sqrt{\sigma_{b0}^2 + t_p^2 \sigma_{b1}^2 + 2t_p \rho \sigma_{b0} \sigma_{b1}}} = \Phi_{\text{nor}}^{-1}(1-p)
\]

\[
\frac{(D_t - \beta_0 - t_p \beta_1)^2}{\sigma_{b0}^2 + t_p^2 \sigma_{b1}^2 + 2t_p \rho \sigma_{b0} \sigma_{b1}} = \left[ \Phi_{\text{nor}}^{-1}(1-p) \right]^2.
\]

(2.16)
Let \( k = \left[ \Phi_{\text{nor}}^{-1}(1 - p) \right]^2 \), \( h = D_1 - \beta_0 \), and \( l = k\sigma_{b_0}^2 \). Then

\[
k \left( \sigma_{b_0}^2 + t_p^2 \sigma_{b_1}^2 + 2t_p \sigma_{b_0} \sigma_{b_1} \right) = (h - t_p \beta_1)^2
\]

\[
l + t_p^2 k \sigma_{b_1}^2 + 2t_p k \sigma_{b_0} \sigma_{b_1} = h^2 - 2h \beta_1 t_p + \beta_1^2 t_p^2
\]

\[
t_p^2 (k \sigma_{b_1}^2 - \beta_1^2) + 2t_p (k \sigma_{b_0} + h \beta_1) + (l - h^2) = 0.
\]

Let \( a = (k \sigma_{b_1}^2 - \beta_1^2) \), \( b = 2(k \sigma_{b_0} + h \beta_1) \), \( c = (l - h^2) \). Then this equation is of the form:

\[
at_p^2 + bt_p + c = 0 \quad (2.17)
\]

with the following solutions for \( t_p \):

\[
t_p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]

\[
= \frac{-2(k \sigma_{b_0} + h \beta_1) \pm \sqrt{4(k \sigma_{b_0} + h \beta_1)^2 - 4 \left( k \sigma_{b_1}^2 - \beta_1^2 \right) (l - h^2)}}{2 \left( k \sigma_{b_1}^2 - \beta_1^2 \right)}
\]

\[
= \frac{-k \sigma_{b_0} + h \beta_1 \pm \sqrt{(k \sigma_{b_0} + h \beta_1)^2 - \left( k \sigma_{b_1}^2 - \beta_1^2 \right) (l - h^2)}}{k \sigma_{b_1}^2 - \beta_1^2}
\]

\[
= \frac{-k \sigma_{b_0} + h \beta_1 \pm \sqrt{k^2 \sigma_{b_0}^2 + k \sigma_{b_0} \beta_1^2 + h^2 k \sigma_{b_1}^2 - k^2 \sigma_{b_0}^2 \sigma_{b_1}^2 + 2hk \beta_1 \sigma_{b_0} \sigma_{b_1}}}{k \sigma_{b_1}^2 - \beta_1^2}.
\]

The derivation is similar when the failure definition is \( b_0 + b_1 t \leq D_1 \), using

\[
F(t_p) = \Phi_{\text{nor}} \left( \frac{D_1 - \beta_0 - t_p \beta_1}{\sqrt{\sigma_{b_0}^2 + t_p^2 \sigma_{b_1}^2 + 2t_p \sigma_{b_0} \sigma_{b_1}}} \right).
\]
2.8.4 Forms of the Partial Derivatives in Section 2.6.3

Let
\[ \psi = \sqrt{k \beta_1^2 \sigma_{b_0}^2 + k \sigma_{b_1}^2 (\beta_0 - D_f)^2 - k^2 \sigma_{b_0}^2 \sigma_{b_1}^2 + k^2 \rho^2 \sigma_{b_0}^2 \sigma_{b_1}^2 - 2k \rho \beta_1 \sigma_{b_0} \sigma_{b_1} (\beta_0 - D_f)}. \]

Then,
\[ \frac{\partial}{\partial \beta_0} t_p = \frac{\beta_1}{k \sigma_{b_1}^2 - \beta_1^2} \pm \frac{k \sigma_{b_1}^2 [\beta_0 - D_f] - k \rho \beta_1 \sigma_{b_0} \sigma_{b_1}}{(k \sigma_{b_1}^2 - \beta_1^2) \psi}. \]
\[ \frac{\partial}{\partial \beta_1} t_p = \frac{2 \beta_1}{(\beta_1^2 - k \sigma_{b_1}^2)^2} \left( \beta_0 \beta_1 - \beta_1 D_f \pm \psi - k \rho \sigma_{b_0} \sigma_{b_1} \right) \]
\[ - \frac{1}{\beta_1^2 - k \sigma_{b_1}^2} \left[ \beta_0 - D_f \pm \frac{k \beta_1 \sigma_{b_0}^2 - k \rho \sigma_{b_0} \sigma_{b_1} (\beta_0 - D_f)}{\psi} \right]. \]
\[ \frac{\partial}{\partial \sigma_{b_0}} t_p = \frac{1}{\beta_1^2 - k \sigma_{b_1}^2} \left[ \sigma_{b_0} k \rho \pm \frac{\sigma_{b_0} \sigma_{b_1} k^2 - \sigma_{b_0} \sigma_{b_1} k \beta_1^2 - \sigma_{b_0} \sigma_{b_1} k^2 \rho^2 + \sigma_{b_1} k \beta_1 (\beta_0 - D_f)}{\psi} \right]. \]
\[ \frac{\partial}{\partial \sigma_{b_1}} t_p = \frac{1}{\beta_1^2 - k \sigma_{b_1}^2} \left[ \sigma_{b_0} k \rho \pm \frac{\sigma_{b_0} k (\beta_0 - D_f)^2 - \sigma_{b_0} \sigma_{b_1} k^2 + \sigma_{b_0} \sigma_{b_1} k^2 \rho^2 - \sigma_{b_0} k \beta_1 (\beta_0 - D_f)}{\psi} \right] \]
\[ - \frac{2 \sigma_{b_1} k}{(\beta_1^2 - k \sigma_{b_1}^2)^2} \left( \beta_0 \beta_1 - \beta_1 D_f - \sigma_{b_0} \sigma_{b_1} k \rho - \psi \right). \]
\[ \frac{\partial}{\partial \rho} t_p = \frac{1}{\beta_1^2 - k \sigma_{b_1}^2} \left[ k \sigma_{b_0} \sigma_{b_1} \pm \frac{k^2 \rho \sigma_{b_0} \sigma_{b_1}^2 - k \beta_1 \sigma_{b_0} \sigma_{b_1} (\beta_0 - D_f)}{\psi} \right]. \]
\[ \frac{\partial}{\partial \sigma} t_p = 0. \]

For all cases, excluding the partial derivative with respect to \( \rho \), replace \( \pm \) with “+” if \( 0 < p < 0.5 \) and with “−” if \( 0.5 \leq p \leq 1 \). For the partial derivative with respect to \( \rho \), replace \( \pm \) with “−” if \( 0 < p < 0.5 \) and with “+” if \( 0.5 \leq p \leq 1 \).

2.9 References


CHAPTER 3. Methods For Planning Accelerated Repeated Measures Degradation Studies

Brian P. Weaver and William Q. Meeker
Department of Statistics
Iowa State University
Ames, IA, 50010

Abstract

Accelerated repeated measures degradation tests can be sometimes used to assess product or component reliability when one would expect few or even no failures during a study. Such tests are used to estimate the lifetime distributions of highly reliable items. This paper describes methods for selecting an accelerated repeated measures degradation test plan when the (possibly transformed) degradation is linear in (possibly transformed) time. To find optimum test plans, we use a criterion based on the estimation precision of the quantile of the failure-time distribution at use conditions. We also discuss how to find compromise test plans that satisfy practical constraints. We use the general equivalence theorem to verify that a test plan is globally optimum. The resulting optimized plans are also evaluated using simulation and compared with other test plans.

Keywords: Nondestructive Degradation, Mixed Effects Linear Models, Accelerated Degradation Testing, General Equivalence Theorem
3.1 Introduction

3.1.1 Motivating Examples

Engineers often need to quantify the lifetime distribution of highly reliable items. Traditional life tests where the response is time to failure typically yield few or no failures, even with acceleration. Instead engineers can, in some applications, use methods that measure the degradation of an item, providing more information than traditional life tests. In some situations, however, the degradation rate is so low that noticeable degradation will not be observed during the test. To address this issue, engineers will expose the items to accelerated conditions, such as higher use rate, voltage, temperature, or humidity. Under these accelerated conditions, measurements of degradation are made and the relationship between the accelerating variables and the degradation rate can be modeled and unknown parameters in the model can be estimated. This relationship can then be used to extrapolate to estimate lifetime at normal use conditions. When planning a repeated measures accelerated degradation test (RMADT), the engineers need to specify the levels of the accelerating variable at which test units will be exposed and the number of units that should be allocated to each level.

This work is motivated by several different examples. The first example is based on an experiment performed by Shiomi and Yanagisawa (1979). The engineers measured the resistance of carbon-film resistors at particular points in time. At the beginning of the experiment, the resistance values varied between 215.92 and 224.7 ohms. The resistors were exposed to three different levels of temperature (83°C, 133°C, and 173°C) in order to accelerate the degradation. They defined degradation to be the amount of increase in resistance over time and we suppose that failure occurs when resistance exceeds 230 ohms. Figure 3.1a presents the original data and Figure 3.1b shows the data plotted versus the square root of time (i.e., on a square root axis). We suppose that the engineers wanted to estimate a quantile of the failure-time distribution at the nominal use conditions (50°C).

The next example is the sliding metal wear application found in Chapter 21 of Meeker and Escobar (1998). The experiment was conducted to test the wear resistance of a particular
metal alloy. The engineers applied a range of different weights (in grams) to a piece of metal which was then slid over another piece of metal. The engineers then measured the widths of cracks (in microns) that formed on the metal piece. The engineers defined a failure to be a crack width of 50 microns. The purpose of the experiment was to study the effect of weight on wear rate and to gain a better understanding of the wear mechanism. Figure 3.2 presents the original data (Figure 3.2a) as well as the data plotted on a log-log axis (Figure 3.2b).

The final example is a test that ran 30 light emitting diodes (LEDs) at each of six different combinations of temperature and electrical current, as described in Pascual, Meeker, and Escobar (2006). The engineers discovered that the LEDs tested at the combination of highest temperature and current had an unexpected failure mechanism that is different than the one at other test conditions. Unless the failure mode can be eliminated by redesigning the LED device, this region of the experimental region will be excluded from future tests. The engineers measured the output of the LEDs (in Lumens) and this was then converted to relative change in light output. Figure 3.3 presents the original data (Figure 3.3a) on a linear scale and the relative change with square root of time (Figure 3.3b).
3.1.2 Related Work

Figure 3.3  LED output versus time in hours. The rows of the graph correspond to levels of current and the columns correspond to levels of junction temperature. The figure on the left (3.3a) presents the original data. The figure on the right (3.3b) presents the original data on a square root of time scale.

3.1.3 Overview

The rest of the paper is organized as follows. Section 3.2 introduces the assumed general form of the degradation model and the corresponding Fisher information matrix. Section 3.3 discusses the relationship between the degradation model and the lifetime distribution, the failure-time distribution quantile, and the large-sample approximate variance of the maximum likelihood estimator for the failure-time distribution quantile. Section 3.4 describes RMADTs. Section 3.5 discusses the criterion for finding test plans and optimum test plans. Section 3.6 discusses other types of RMADT plans. Section 3.7 goes through different examples of RMADT planning. Section 3.8 shows how to use simulation to compare different types of test plans. Section 3.9 provides concluding remarks and areas for future research.
3.2 Degradation Models

3.2.1 Accelerated Repeated Measures Degradation Models

The actual degradation level for a typical observational unit at time $t$ and accelerating variable(s) is denoted by $D = D(\tau, x_1, x_2, \theta)$ where $\tau = h_t(t)$ is a monotone increasing transformation of time, $x_1$ and $x_2$ are vectors of (possibly transformed) accelerating variables, and $\theta$ is the unknown parameter vector.

All of the models used in the examples are special cases of the following generic model. The (possibly transformed) observed degradation $Y$ at some (possibly transformed) time point $t$ for a fixed value(s) of the (possibly transformed) accelerating variables $x_1$ and $x_2$ is

$$Y = h_d(D) + \varepsilon = \mu(\tau, x_1, x_2) + \varepsilon = x_1 \gamma_1' + x_2 \gamma_2' \tau + b_0 + b_1 \tau + \varepsilon = x_1 \gamma_1' + x_2 \gamma_2' \tau + zb' + \varepsilon,$$

so that $\mu(\tau, x_1, x_2) = h_d(D)$ is a location parameter for the distribution of $Y$ that depends on the unknown parameters $\theta$, $z' = (1, \tau)'$, $\varepsilon \sim iid N(0, \sigma^2)$ random variable that describes the within unit variation, and $h_d$ is a monotone increasing or decreasing transformation of $D$. In (3.1), the term $x_1 \gamma_1'$ describes how the intercept or initial value of degradation changes as a function of the accelerating variable(s), $x_2 \gamma_2'$ describes how the degradation slope or rate changes as a function of the accelerating variable(s), and $zb'$ describes the unit-to-unit variability in the degradation intercepts and slopes. Note that $x_1$ and $x_2$ will sometimes be exactly the same vectors (but in general they do not have to be).

We assume that the variability in the linear regression parameters $b$ can be described by a bivariate normal distribution

$$(b_0, b_1)' \sim BVN(\beta, V)$$
where \( \beta = (\beta_0, \beta_1)^T \) is the mean vector and

\[
V = \begin{pmatrix}
\sigma_{b0}^2 & \rho \sigma_{b0} \sigma_{b1} \\
\rho \sigma_{b0} \sigma_{b1} & \sigma_{b1}^2
\end{pmatrix}
\]

is the covariance matrix. We further assume that \((b_0, b_1)'\) is independent of \(\varepsilon\). This assumption of independence between \((b_0, b_1)'\) and \(\varepsilon\) implies that different units are independent of each other.

### 3.2.2 Approximate Variance-Covariance Matrix of the Maximum Likelihood Estimators

Let \( \theta = (\gamma_1, \gamma_2, \beta, \vartheta)' \) be the \(r\)-element parameter vector where \( \vartheta = (\sigma_{b0}, \sigma_{b1}, \rho, \sigma)' \). From large-sample theory, the approximate covariance matrix of the maximum likelihood (ML) estimators is

\[
\text{AVar}(\hat{\theta}) = [\mathcal{I}(\theta)]^{-1}
\]

(3.2)

where \( \mathcal{I}(\theta) = -E(\partial^2 \mathcal{L} / \partial \theta^2) \) is the Fisher Information matrix and \( \mathcal{L} \) is the total loglikelihood. The derivation of the information matrix is given in the appendix. \( \text{AVar}(\hat{\theta}) \) can be estimated by evaluating (3.2) at the ML estimates \( \hat{\theta} \). We denote this estimator by \( \hat{\text{AVar}}(\hat{\theta}) \).

### 3.3 Failure-Time Distribution for Degradation Models

This section describes the relationship between the degradation model and the induced failure-time model. If a degradation model and definition of failure are given, a failure-time distribution is implied (e.g., Chapter 13 of Meeker and Escobar 1998).

#### 3.3.1 Failure-Time Cumulative Distribution Function

This section shows how to derive the failure-time distribution from the degradation model in (3.1). A degradation process with soft failures is assumed. A soft failure occurs when the degradation level of a unit, \( D \), reaches a pre-specified degradation level \( D_f \). This is equivalent to \( h_d(D) \geq \mu_f \) for increasing \( D \) or \( h_d(D) \leq \mu_f \) for decreasing \( D \) where \( \mu_f = h_d(D_f) \). The
failure time for a unit is the time at which it reaches the degradation level $D_f$. Let $T$ define
the random variable associated with the item’s time to failure. If $b_0 + b_1 \tau \sim N(\beta_0 + \tau \beta_1, \sigma_b^2 + \tau^2 \sigma_{b_1}^2 + 2 \tau \rho \sigma_{b_0} \sigma_{b_1}) = N(z\beta', \sigma_{b_0}^2 + \tau^2 \sigma_{b_1}^2 + 2 \tau \rho \sigma_{b_0} \sigma_{b_1})$, then for increasing $D$

$Pr(T \leq t) = Pr(D \geq D_t) = Pr(h_\tau(D) \geq \mu_t) = Pr(b_0 + b_1 \tau \geq \mu_t - x_1 \gamma'_1 - x_2 \gamma'_2) = 1 - Pr(b_0 + b_1 \tau \leq \mu_t - x_1 \gamma'_1 - x_2 \gamma'_2) = 1 - \Phi_{nor}(\kappa) \quad (3.3)$

where

$$\kappa = \frac{\mu_t - x_1 \gamma'_1 - x_2 \gamma'_2 - z\beta'}{\sqrt{\sigma_{b_0}^2 + \tau^2 \sigma_{b_1}^2 + 2 \tau \rho \sigma_{b_0} \sigma_{b_1}}}$$

and $\Phi_{nor}$ is the CDF of the standard normal distribution.

Similarly, for decreasing $D$, a failure happens when $D \leq D_f$ and

$F(t; x) = Pr(D \leq D_t) = \Phi_{nor}(\kappa). \quad (3.4)$

It should be noted that the failure time is a function of the true, unobserved degradation but
note our modeling for estimation of parameters is based on the observed, possibly transformed,
degradation. $F_T(t; x)$ can be estimated by evaluating (3.3) or (3.4) at the ML estimates of
$\theta$. Meeker and Escobar (1998), page 330, describes a numerical integration and a simulation
based approach for more complicated models where a closed form solution does not exist.

### 3.3.2 Failure-Time Quantiles

From (3.3) or (3.4) the $p$ quantile of the failure-time distribution, denoted by $t_p$, is $t_p = h_t^{-1}(\tau_p)$ where

$$\tau_p = -\left[\frac{k \rho \sigma_{b_0} \sigma_{b_1} + (\mu_t - x_1 \gamma'_1 - \beta_0) (x_2 \gamma'_2 + \beta_1)}{[k \sigma_{b_1}^2 - (x_2 \gamma'_2 + \beta_1)^2]} \right] \pm \sqrt{\psi}. \quad (3.5)$$

Here $k = \left[\Phi_{nor}^{-1}(1 - p)\right]^2$ or $k = \left[\Phi_{nor}^{-1}(p)\right]^2$ depending on whether a failure is declared when
$D \geq D_t$ or $D \leq D_t$, respectively, $k \sigma_{b_1}^2 \neq (x_2 \gamma'_2 + \beta_1)^2$, and

$$\psi = \left[k \rho \sigma_{b_0} \sigma_{b_1} + (\mu_t - x_1 \gamma'_1 - \beta_0) (x_2 \gamma'_2 + \beta_1)\right]^2 - \left[k \sigma_{b_1}^2 - (x_2 \gamma'_2 + \beta_1)^2\right] \left[k \sigma_{b_0}^2 - (\mu_t - x_1 \gamma'_1 - \beta_0)^2\right].$$
The derivation of (3.5) is given in the appendix. If $0 < p < 0.5$, then $t_p$ is the root where the radical in (3.5) is added. If $0.5 < p < 1$, $t_p$ is given by the root where the radical in (3.5) is subtracted.

### 3.3.3 Approximate Variance for the Maximum Likelihood Estimator of a Failure-Time Quantile

Let $a$ be a vector with elements $a_i = \partial t_p / \partial \theta_i$, $i = 1, \ldots, r$. Then by the delta method, the large-sample approximate variance of $\hat{t}_p$ is

$$\text{AVar}(\hat{t}_p) = a' \text{AVar}(\hat{\theta}) a. \quad (3.6)$$

$\text{AVar}(\hat{t}_p)$ can be estimated by evaluating (3.6) at $\hat{\theta}$.

### 3.4 Repeated Measures Accelerated Degradation Test Planning

#### 3.4.1 Planning Information

RMADT planning requires specification of the form of (3.1), information (planning values) for the unknown model parameters, a distribution for the variability, the failure-defining critical degradation level $D_f$, the points in time at which degradation measurements will be made, and a range of the accelerating variables allowed for the test.

#### 3.4.2 Inspection Schedule Specification

Part of RMADT plan specification is the inspection schedule, i.e., the points in time that a measurement will be made on a unit. In this paper we choose not to optimize in the time dimension. The reason being typically when one optimizes in this dimension, the resulting optimum inspection schedule puts all measurements at the beginning and the end of the experiment. We do not view this as being very practical in real applications. Therefore, for our examples we will use the original inspections schedules as used by the researchers.
3.4.3 Accelerating Variable Plan Specification

Let \( x_i = (x_{1,i}', x_{2,i}')' \) be a factor level combination of the (possibly transformed) accelerated conditions for unit \( i \) and let \( \pi_i \) denote the proportion of units that are to be tested at \( x_i \). An RMADT plan will specify the levels of the accelerating variable to test and the proportion of units to be tested at those levels. A test plan with \( K \) levels is denoted by

\[
\eta = \begin{bmatrix}
  x_1, & \pi_1 \\
  x_2, & \pi_2 \\
  \vdots & \vdots \\
  x_K, & \pi_K
\end{bmatrix},
\]

where \( \sum_{i=1}^{K} \pi_i = 1 \).

3.5 Optimum RMADT

3.5.1 Criterion for Choosing a Plan

The purpose of conducting an RMADT is often to be able to estimate \( t_p \) at the use conditions. We use (3.6), the large-sample approximate variance of the ML estimator of \( t_p \), as our test selection criterion.

Consider a situation where \( n \) available units are to be tested at \( K \) different points in the experimental space \( \mathbf{x} \). Based on the assumption of independence across units, the Fisher information matrix can be expressed as

\[
\mathcal{I}(\theta, \pi) = \sum_{i=1}^{n} \mathcal{I}(\theta, \pi_i) = \sum_{i=1}^{K} n_i \mathcal{I}(\theta, \mathbf{x}_i) = n \sum_{i=1}^{K} \pi_i \mathcal{I}(\theta, \mathbf{x}_i)
\]

(3.7)

where \( n_i, \pi_i, \) and \( \mathcal{I}(\theta, \mathbf{x}_i) \) are the number of units, proportion of units, and the Fisher Information matrix at the \( i \)-th specified condition, respectively. Taking (3.7) and replacing it in (3.6) gives

\[
\text{Avar}(\hat{t}_p) = a' \mathcal{I}^{-1}(\theta, \eta) a = \frac{1}{n} a' \left[ \sum_{i=1}^{K} \pi_i \mathcal{I}(\theta, \mathbf{x}_i) \right]^{-1} a \propto a' \left[ \sum_{i=1}^{K} \pi_i \mathcal{I}(\theta, \mathbf{x}_i) \right]^{-1} a
\]

(3.8)

and (3.8) will be used as the optimization criteria for the selection of test plans.
3.5.2 Checking the Initial Optimum Plan

We can use a numerical optimization algorithm to minimize (3.8) for given test-planning inputs. Whittle’s (1973) general equivalence theorem (GET) can be used to check for global optimality of test plans. The GET can also be used to check whether an optimum test plan is unique or not.

For an optimality criterion $\Psi(\cdot)$, the directional derivative, $\Lambda$, at $\eta$ and in the direction of another test plan $\nu$ is defined as

$$
\Lambda(\eta, \nu) = \lim_{\delta \to 0^+} \frac{\Psi[(1 - \delta)\eta + \delta\nu] - \Psi(\eta)}{\delta}.
$$

For a criterion similar to (3.8), Shi, Escobar, and Meeker (2009) give the following expression as the directional derivative

$$
\Lambda(\eta, \nu) = a'[I(\eta)]^{-1}I(\nu)[I(\eta)]^{-1}a - a'[I(\eta)]^{-1}a
$$

(3.9)

and this expression applies also to our optimization problem. Shi, Escobar, Meeker (2009) in their appendix also show how (3.9) satisfies all the requirements needed to apply the GET and the same requirements are satisfied for our model. Therefore, Whittle’s GET can be used to determine if a test plan is optimum.

According to the GET, a test plan $\eta^*$ will be optimum if and only if

$$
\sup_{x} \Lambda_\phi(\eta^*, \eta_x) = 0
$$

for each singular test plan $\eta_x$ (i.e., a test plan where all units are tested at one factor-level combination $x$), and the set of conditions $x^*_i$ in the optimal plan $\eta^*$ are a subset of the conditions satisfying $\Lambda_\phi(\eta^*, \eta_x) = 0$.

3.6 Other RMADT Plans

3.6.1 Traditional Plans

A traditional test plan is one that uses equally spaced levels of the accelerating variable and equal allocation of units to those levels. In testing situations that require extrapolation, as in RMADT, traditional plans may not be statistically efficient, yielding less precise estimates.
3.6.2 Compromise RMADT

Optimum test plans give the smallest value of the large-sample approximate variance of the ML estimator of the failure-time quantile at use conditions. These test plans, however, may not be practical or robust to deviations from model assumptions and other inputs. Compromise test plans use three levels of the accelerating variable. These plans are more robust to deviations from specified inputs but still have good statistical properties.

For our compromise test plans, we impose the constraint that some proportion of test units is to be tested at a middle level of the accelerating variables (i.e., halfway between the lowest and highest level). This prevents the test plan from degenerating to an optimum test plan.

3.7 Examples of Planning RMADT

This section will give specific examples of planning RMADTs. For each example the planning information will be given, an optimum test will be found and verified, and the optimum test will be compared to other RMADTs.

3.7.1 Test Plan to Test the Increase of Resistance in Carbon-Film Resistors

This first example is based on the resistance data in Figure 3.1b. The square root transformation on time was chosen to make the sample paths approximately linear. Let $y_{ijk}$ be the observed degradation for unit $i$, $i = 1, \ldots, 19$, at time point $j$, $j = 1, \ldots, 5$ for level $k$ of the accelerating variable, $k = 1, 2, 3$, and let $x_k$ denote the Arrhenius transformation of temperature where $x = -11605/(\text{Temp}^\circ C + 273.15)$. The assumed linear accelerated degradation random effects model is

$$y_{ijk} = \gamma x_k \tau_{ij} + b_0 + b_1 \tau_{ij} + \varepsilon_{ijk}, \quad (3.10)$$

where $\tau = \sqrt{\text{Time}}$. Notice that (3.10) is a special case of (3.1) with $x_1 = 0$, $x_2 = x$ and $\gamma_2 = \gamma$. The coefficient -11605 in the Arrhenius transformation is used so that the coefficient $\gamma$ has the commonly used units eV/K where eV denotes electron volts and K denotes Kelvin.

The above acceleration model adequately describes the underlying failure mechanism only within a certain range of $x$, say $[x_U, x_H]$ where $x_U$ is value of the accelerating value at use
conditions, but testing will be limited to the interval \( [x_L, x_H] \). Testing beyond \( x_H \) would cause the model to break down (e.g., by generating new failure modes) and testing below \( x_L \) would not provide any meaningful degradation information because the degradation level would be too small. Generally, \( x_L > x_U \). Thus for practical RMADTs, the experimental region is between \( x_L \) and \( x_H \).

For this example, we will consider two different types of test plans. The first will be a test plan that tests some units at the nominal use conditions. These tests will be called Plan 1. The second type of plan will only test units within the practical experimental region \( [x_L, x_H] \). This test will be called Plan 2. The reason engineers may want to test some units at the nominal use conditions is to check for differences in behavior of the degradation at accelerated conditions (e.g., analytical chemical measurements could be used to assure that increased levels of the accelerating variables have not changed the nature of the chemical reactions).

The rest of the planning information is as follows:

- \( D_t = 230 \) ohms.
- All units will have the same measurement schedule with the inspections being made after 0, 452, 1030, 4341, and 8084 hours.
- \( 173°C \) is the maximum allowable testing temperature.
- \( 83°C \) is the minimum testing temperature.
- The use temperature is \( 50°C \).
- The test will run for 8084 hours.
- The proportion of units tested at the nominal conditions, denoted by \( \pi_U \), is 5%.

The goal is to develop a test plan that will evaluate the failure-time distribution of lower-cost resistors that are otherwise believed to be similar to those in the previous test. The test plan properties will depend on the unknown model parameters \( \theta \). The planning values for \( \theta \) are \( \beta_0 = 218.4, \beta_1 = 0.53, \gamma = 0.016, \sigma_{b_0} = 2.181, \sigma_{b_1} = 0.00038, \rho = 0.628, \) and \( \sigma = 0.59 \) (these are based on the ML estimates from the original data).
Next we find an optimum test plan for this example. From the optimum design theory, we generally expect that an optimum test plan will spread the units out to the boundary of the experimental region. This leads us to consider a two-point optimum plan where a proportion \( \pi_L \) of the units should be at \( x_L \) (the lowest allowable accelerating conditions of the test) and \( 1 - \pi_L \) of the units should be at \( x_H \) (the highest allowable accelerating conditions of the test). Then (3.8) reduces to

\[
a' \left[ \pi_U I_U(\theta, x_U) + \pi_L I_L(\theta, x_L) + (1 - \pi_U - \pi_L) I_H(\theta, x_H) \right]^{-1} a
\]

for tests under Plan 1 and for tests under Plan 2

\[
a' \left[ \pi_L I_L(\theta, x_L) + (1 - \pi_L) I_H(\theta, x_H) \right]^{-1} a.
\]

Table 3.1 gives the optimum RMADT under Plan 1 and Table 3.2 gives the optimum RMADT under Plan 2 for the planning information given above.

<table>
<thead>
<tr>
<th>Plan</th>
<th>Temperature°C</th>
<th>Allocation</th>
<th>( \sqrt{\text{AVar}(t_p)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum</td>
<td>(50, 83, 173)</td>
<td>(.05, .711, .239)</td>
<td>21808</td>
</tr>
<tr>
<td>Compromise</td>
<td>(50, 83, 111.5, 173)</td>
<td>(0.05, 0.625, 0.1, 0.225)</td>
<td>22162</td>
</tr>
<tr>
<td>Original</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Traditional</td>
<td>(50, 83, 113, 143, 173)</td>
<td>(.05, .2375, .2375, .2375, .2375)</td>
<td>24189</td>
</tr>
</tbody>
</table>

Table 3.1 Optimum, compromise, and traditional test plans as well as the large-sample approximate standard error under Plan 1 for the resistor example and assuming \( n = 29 \) test units (as in the original experiment).

In order to use the GET when units are constrained to be tested at the use conditions, a modification to the theorem must be made, as suggested by Ying and Meeker (2010). Instead of considering singular tests where all the units are tested at a single level of temperature, the alternative plan must also put the same proportion of units at use conditions and the remaining proportion at a single test condition. Figure 3.4 shows (3.9) at \( \eta^*_1 \) as a function of \( x_L \). Notice that at the end points of the test design space, the directional derivative reaches its maximum and is zero showing that this design is indeed optimum. The corresponding figure for \( \eta^*_2 \) is similar and is omitted to avoid redundancy.
Plan 2
Temperature°C Allocation $\sqrt{\text{AVar}(t_p)}$

<table>
<thead>
<tr>
<th>Plan</th>
<th>Temperature°C</th>
<th>Allocation</th>
<th>$\sqrt{\text{AVar}(t_p)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum</td>
<td>(83, 173)</td>
<td>(.748, .252)</td>
<td>22561</td>
</tr>
<tr>
<td>Compromise</td>
<td>(83, 111.5, 173)</td>
<td>(.662, .1, .238)</td>
<td>22853</td>
</tr>
<tr>
<td>Original</td>
<td>(83, 133, 173)</td>
<td>(.33, .33, .33)</td>
<td>26289</td>
</tr>
<tr>
<td>Traditional</td>
<td>(83, 113, 143, 173)</td>
<td>(.25, .25, .25, .25)</td>
<td>26578</td>
</tr>
</tbody>
</table>

Table 3.2 Optimum, compromise, the original plan from the study, and traditional test plans as well as the large-sample approximate standard error under Plan 2 for the resistor example and assuming $n = 29$ test units (as in the original experiment).

We next consider other RMADT test plans, beginning with a traditional test plan. Recall that a traditional test plan places an equal proportion of units to equally spaced levels of the accelerating variable. Traditional test plans for Plans 1 and 2 for this example are given in Tables 3.1 and 3.2.

A compromise between the optimum and traditional test plans will select a third, middle level in between the lowest and the highest levels of $x$ in the optimum test plan. For this example, the midpoint is $x_M = -30.17$ (the Arrhenius transformation of 111.5°C). Constraints must be put on $\pi_M$; otherwise the compromise test plan will degenerate to a two-point optimum plan. The constraint is $\pi_M \in (0.10, 1)$. The compromise test plan is found by selecting the values of $x_L, \pi_L,$ and $\pi_M$, subject to the constraints just mentioned, that minimize (3.14). For this example, the compromise test plan for both Plans 1 and 2 are also given in Tables 3.1 and 3.2, respectively.

### 3.7.2 Test Plan to Assess Scar Width Growth

This example is based on the scar-width data presented in Figure 3.2b. Let $y_{ijk} = \log(D_{ij})$ be the observed, transformed degradation for unit $i, i = 1, \ldots, 12,$ at time point $t_{i,j}, j = 1, \ldots, 8$ and for accelerating level $k, k = 1, 2, 3$, where $D$ was defined to be crack width in microns. A model that can describe these data is

$$y_{ijk} = \gamma_1 x_{1i} + \gamma_2 x_{2i} t_{ij} + b_0 + b_1 t_{ij} + \varepsilon_{ijk}$$

(3.13)
where $\tau_{ij} = \log(t_{ij})$, $x_{1i} = x_{2i} = \log(\text{weight}_i)$ where weight is in grams, and $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

Notice that (3.13) is a special case of (3.1) with $x_1 = x_1$, $\gamma_1 = \gamma_1$, $x_2 = x_2$, and $\gamma_2 = \gamma_2$.

The rest of the planning information is as follows:

- $D_f = 50$ microns ($y_f = \log(50) = 3.912023$).
- All units will have the same measurement schedule where inspections were made after 2, 5, 10, 20, 50, 100, 200, and 500 cycles.
- 100g is the maximum allowable testing weight.
- 10g is the lowest allowable testing weight.
- The use weight is 5g.
- The test will run for 500 cycles.

The goal is to develop a test plan that will evaluate the wear resistance of new lower cost but equally durable metal plates that are similar to those that were tested previously. The
test plan will depend on the unknown model parameters $\theta$. The planning values for $\theta$ are 
$\beta_0^\Box = 0.823, \beta_1^\Box = 0.183, \gamma_1^\Box = 0.018, \gamma_2^\Box = 0.00014, \sigma_{b_0}^\Box = 0.194, \sigma_{b_1}^\Box = 0.019, \rho^\Box = -0.812, \text{ and } \sigma^\Box = 0.048 \text{ (these are based on the ML estimates from the original data).}$

Next we find an optimum test plan for this example. As in Section 3.7.1, the optimum test is expected to spread the units out to the boundary of the experimental region. This leads us to consider a two-point optimum plan where $\pi_L$ of the units should be at $x_L$ (the lowest allowable accelerating conditions of the test) and $1 - \pi_L$ of the units at $x_H$ (the highest allowable accelerating conditions of the test). This again reduces (3.8) to 
\[
a'[\pi_L I_L(\theta, x_L) + (1 - \pi_L) I_H(\theta, x_H)]^{-1} a.
\] (3.14)

For the planning information given for this example, the optimum plan is given in Table 3.3.  Also present in this table is a compromise test using the same structure and constraints as in the resistors example, a traditional test plan, and a plan representing the original study. We verified optimality numerically by using the GET but omit the plot as it is similar to Figure 3.4.

<table>
<thead>
<tr>
<th>Plan</th>
<th>Weight in grams</th>
<th>Allocation</th>
<th>$\sqrt{\text{AVar}(t_p)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum</td>
<td>(10, 100)</td>
<td>(.95, .05)</td>
<td>1796554</td>
</tr>
<tr>
<td>Compromise</td>
<td>(10, 55, 100)</td>
<td>(.855, .1, .045)</td>
<td>1838556</td>
</tr>
<tr>
<td>Original</td>
<td>(10, 50, 100)</td>
<td>(.33, .33, .33)</td>
<td>2090628</td>
</tr>
<tr>
<td>Traditional</td>
<td>(10, 40, 70, 100)</td>
<td>(.25, .25, .25, .25)</td>
<td>2273572</td>
</tr>
</tbody>
</table>

Table 3.3  Optimum, compromise, the original plan from the study, and traditional test plans as well as the large-sample approximate standard error for the metal wear example and assuming $n = 12$ test units (as in the original experiment).

### 3.7.3 Test Plan to Study LED Output Degradation

This example is based on the LED data displayed in Figure 3.3a. The original study consisted of modeling the degradation of LEDs as a function of junction temperature and current. Ambient temperature, however, is the appropriate scale for test planning. Temperature increase is proportional to power dissipated in the form of heat in the devices. This result and
Ohm’s law imply that $\Delta T \propto (\text{Current})^2$ where $\Delta T$ is change in ambient temperature in degrees C and current is in mA. That is

\[
\text{Junction}^\circ\text{C} = \text{Ambient}^\circ\text{C} + \delta(\text{Current})^2
\]

Based on the junction temperature and current values given in the data set, we can infer that for the particular LED devices that were tested, $\delta = 0.2333$. This gives approximate values of ambient temperature in the original experiment to be 40°C, 73°C, and 93°C. For this example, test planning is in terms of the ambient temperature in the test chambers but modeling the relationship between the degradation rate and temperature is done in terms of the junction temperature.

Let $y_{ijk}$ be the observed degradation for unit $i$, $i = 1, \ldots, 180$, at time point $t_j$, $j = 1, \ldots, 10$, and for the $k$-th accelerating level factor combination, $k = 1, \ldots, 6$, where $D$ was defined to be the relative change in LED output. The model that we use to describe this situation is

\[
y_{ijk} = \beta_0 + \gamma_1 x_{1k} \tau_j + \gamma_2 x_{2k} \tau_j + b \tau_j + \varepsilon_{ijk}
\]

(3.15)

where $\tau = \sqrt{\text{time}}$, $x_{1k} = -11605/(\text{Junction}^\circ\text{C} + 273.15)$ is the Arrhenius transformation of junction temperature, $\text{Junction}^\circ\text{C} = \text{Ambient}^\circ\text{C} + \delta(\text{Current})^2$, $x_{2k} = \log(\text{Current})$, $b \sim N(\beta_1, \sigma_b^2)$, and $\varepsilon_{ijk} \sim N(0, \sigma^2)$. Notice that (3.15) is a special case of (3.1) with $x_1 = 0$, $\gamma_1 = 0$, $\mathbf{x}_2 = (x_1, x_2)$, $\sigma_{b0}^2 = \rho = 0$, and $\gamma_2 = (\gamma_1, \gamma_2)$.

The rest of the planning information is as follows:

- $D_f = 0.6$ relative change.
- All units will have the same measurement schedule where inspections were made after 138, 234, 330, 402, 498, 570, 666, 738, 834, and 906 hours.
- 93°C and 40mA are the maximum value of ambient temperature and current values allowed.
- 61°C and 30mA are the minimum allowable ambient temperature and current values.
• In the original experiment, a second failure mechanism was induced by the combination of 93°C and 40mA and in future testing we need to avoid testing at combinations of high temperature and current.

• The nominal use conditions are 25°C ambient temperature and 20mA current.

• The test will run for 906 hours.

We suppose that the goal is to develop a test plan that will evaluate new LEDs that have properties similar to those in the original experiment. Again the test plan properties will depend on the unknown model parameters $\theta$. The planning values for $\theta$ are $\beta_0 = 1.09$, $\beta_1 = -0.0093$, $\gamma_1 = -0.0002$, $\gamma_2 = -0.0013$, $\sigma_\beta = 0.0003$, and $\sigma_\sigma = 0.005$ (these values are based on the ML estimates from the original data).

We consider two separate situations for this example. The first situation considers using the entire experimental region as in the original study, which we will call Situation 1. This would be a legitimate case if the engineers were able to remove the extra failure mechanism induced by the larger levels of current and ambient temperature. The second situation will consider using only the portion of the experimental region that does not invoke the second failure mechanism. This situation will be called Situation 2.

### 3.7.3.1 Situation 1

Here we find an optimum RMADT plan for the entire experimental region. Figure 3.5 shows the optimum plan obtained using the Wynn algorithm (Wynn 1970). The optimum test plan is given in Table 3.4 along with the value of the large-sample approximate standard error for $n = 180$ test units as was in the original study. To verify optimality, Figure 3.6 shows (3.9) for the continuous optimum test plan given in Table 3.4 as a function of current and ambient temperature. In Figure 3.6, a specific curve corresponds to a constant value of ambient temperature and changes as a function of current. This plot shows that ambient temperature does not appear to be as influential as current. This comes as no surprise since by looking again at a single row in Figure 3.3a (corresponding to a fixed current) and moving
Plan (Ambient Temperature, Current) Proportion Allocation

<table>
<thead>
<tr>
<th>Plan</th>
<th>(Ambient Temperature, Current)</th>
<th>Proportion Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum</td>
<td>(40, 30), (93, 30), (40, 40), (93, 40)</td>
<td>.488, .143, .123, 0.246</td>
</tr>
<tr>
<td>Compromise</td>
<td>(40, 30), (93, 30), (66.5, 35), (40, 40), (93, 40)</td>
<td>.522, .049, .100, .247, .082</td>
</tr>
<tr>
<td>Original</td>
<td>See Figure 3.7</td>
<td>1/6, 1/6, ..., 1/6</td>
</tr>
<tr>
<td>Traditional</td>
<td>(40, 30),(40, 35), ...,(93, 40)</td>
<td>1/9, 1/9, ..., 1/9</td>
</tr>
</tbody>
</table>

Table 3.4 Optimum, compromise, the original plan from the study, and traditional test plans for the LED data.

across columns (corresponding to changing temperature), we see that the ambient temperature has very little, if any, noticeable affect on the degradation slopes. Now by looking at a single column and moving across rows, we see that current has a strong effect on the degradation slopes. This is also evident in the directional derivative plot since the two different lines are directly on top of each other. Notice also that at the points in the continuous optimum test plan, the value of the directional derivative is zero, satisfying the GET. If one were to reverse the situation and put temperature on the x-axis, the plot would be very similar to Figure 3.6.

Figure 3.5 The experimental region for the LED test plan where the four points correspond to the optimal test plan.
Figure 3.6  Plot of (3.9) versus ambient temperature and current evaluated at the optimum plan found in the LED example.

Table 3.4 also gives a compromise test plan where a constraint is set that 10% of the units should be tested at (66.5°C, 35mA), a traditional test plan, and a test plan that is the same as the original test. Table 3.5 also gives the large-sample approximate standard error for the 0.10 failure-time quantile for the different test plans.

<table>
<thead>
<tr>
<th>Plan</th>
<th>$\sqrt{\text{Var}(t_p)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum</td>
<td>110.15</td>
</tr>
<tr>
<td>Compromise</td>
<td>115.65</td>
</tr>
<tr>
<td>Original</td>
<td>134.09</td>
</tr>
<tr>
<td>Traditional</td>
<td>152.72</td>
</tr>
</tbody>
</table>

Table 3.5  Large-sample approximate standard error for the test plans given in Table 3.4 and assuming $n = 180$ test units (as in the original experiment).
3.7.3.2 Situation 2

This situation will require putting a constraint on the experimental region to avoid inducing the extra failure mechanism that was discovered at the highest combination of ambient temperature and current in the original test. As mentioned previously, the factor-level combination 93°C ambient temperature and 40 mA of current caused a new failure mechanism. Suppose that the engineers are able to determine that part of the experimental region that caused the second failure mechanism. Figure 3.7 shows the new constrained experimental region.

![Figure 3.7 Constrained experimental region for the LED data. The area within the shape corresponds to acceptable testing conditions. The six dots represent the test points of the original study.](image)

We first consider continuous optimization (i.e., optimization of the proportion of units to be allocated to each combination of stresses). Again based on use of the Wynn algorithm, the continuous optimized test plan is given in Table 3.6. Table 3.6 also considers a traditional test plan, a compromise test plan with constraints similar to Situation 1 and a test plan that is similar to the original study. Table 3.7 gives the value of the large-sample approximate standard error for the different test plans in Table 3.6 and assuming \( n = 180 \) test units as was
in the original study.

<table>
<thead>
<tr>
<th>Plan</th>
<th>(Ambient Temperature, Current)</th>
<th>Proportion Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum</td>
<td>(40, 30), (93, 30), (40, 40), (73, 40)</td>
<td>.533, .098, .105, .264</td>
</tr>
<tr>
<td>Compromise</td>
<td>(40, 30), (93, 30), (66.5, 35), (40, 40), (73, 40)</td>
<td>.522, .048, .10, .204, .126</td>
</tr>
<tr>
<td>Original</td>
<td>See Figure 3.7</td>
<td>1/6, 1/6, ..., 1/6</td>
</tr>
<tr>
<td>Traditional</td>
<td>(40, 30), (40, 35), ..., (73, 40)</td>
<td>1/9, 1/9, ..., 1/9</td>
</tr>
</tbody>
</table>

Table 3.6 Optimum, compromise, the original plan from the study, and traditional test plans for the LED data where a constraint was put on the experimental region.

Figures 3.8 and 3.9 show the optimum test plan and the directional derivative for this test plan. As in Situation 1, it appears ambient temperature has little effect on degradation of the LEDs to the effect of current.

Figure 3.8 The constrained experimental region for the LED test plan where the four points correspond to the optimal test.
Figure 3.9  Plot of the directional derivative in (3.9) versus temperature and current evaluated at the optimum plan found in the LED example and under the constraint in Figure 3.7. Visually, the lines are the same for all levels of ambient temperature and are stacked on top of each other in this plot.

3.8 Monte Carlo Simulation To Evaluate a RMADT Plans

Simulation is a powerful tool to help visualize the variability involved in the test-planning and data analysis processes and allows evaluation of the properties of accelerated test plans without having to use large-sample approximations. As such, simulation and analytical evaluation are complimentary tools for test planning. Additionally, simulation can be used to evaluate the adequacy of large-sample approximations. We illustrate some test-planning simulation methods using the scar-width example.

For each RMADT test plan listed in Table 3.3, a simulation trial consisted of first 12 units and then 1200 units allocated to the accelerating levels based on the different test plan and the planning information. Data sets are then simulated for which ML estimates are calculated and used to estimate the lifetime quantile. The simulation was repeated 10,000 times for each of the four test plans.
Table 3.7  Large-sample approximate standard error for the test plans given in Table 3.6 and assuming $n = 180$ test units (as in the original experiment).

Figure 3.10 shows histograms of the 0.10 quantile of the failure-time distribution at use conditions for the different plans listed in Table 3.3. The plot shows the increased variability in the original and the traditional test plans. The vertical line in Figure 3.10 is the actual failure-time quantile according to the planning information. Table 3.8 shows the sample standard deviation computed from the simulated values in Figure 3.10 as well as the large-sample approximate standard errors for the $n = 12$ and $n = 1200$ cases. Notice that as the sample sizes increase, the large-sample and the simulation-based values are beginning to agree. For $n = 240$ (results not shown here) the two approaches produce similar values.

Table 3.8  Comparison of values of $\text{ASE}(\hat{t}_p)$ and simulation method for different RMADT plans and different sample sizes for the scar-width data.

3.9 Conclusions and Areas for Future Research

Nondestructive repeated measures degradation tests are useful for studying material or performance degradation of a product over time. It is important to plan these tests properly in order to achieve the desired level of precision while working within resource (time, number...
Figure 3.10  10000 simulated ML estimates of failure-time distribution quantiles at use conditions (50°C), for the different plans as in Table 3.3. The vertical line indicates the failure-time quantile at use conditions, evaluated at the planning information.

of units, and number of measurements) constraints. The methodology presented in this paper can be extended to more complicated situations. The following extensions are areas for future research:

- It would be possible to extend the work in this paper to handle models that cannot be transformed to linear, such as the model used in Meeker, Escobar, and Lu (1998).
- Allow more complicated covariance structures of the within-unit errors (such as autocorrelations that might be expected with smaller spacing between measurements).
- Extend to models where the accelerating variables can also influence the variability in the degradation slopes and intercepts.
- In our motivating applications, like most accelerated tests, the length of the tests were fixed according to a development schedule. It would be possible to extend our work to
test planning where optimization is performed over both the accelerating variable levels and length of the test. See Yu and Tseng (2003) for some work in this direction.

- Extend to competing risks models where there are multiple failure-causing degradation mechanisms.
- One could use Bayesian methodologies for test planning so that prior knowledge about the model parameters can be incorporated into both planning and analysis, similar to what has been done in Zhang and Meeker (2006).

3.10 Appendix

3.10.1 Derivation of the Information Matrix

For the derivation of the Fisher information matrix, we make (3.1) a little more concrete. Suppose there are a total of \( n \) items to be tested and item \( i \) will be measured \( m_i \) times at factor-level combination \( j \) of the (possibly transformed) levels of the accelerating variables \( x_j = (x_{1j}, x_{2j}) \). Then (3.1) can be rewritten as

\[
Y_{ijk} = x_{1j} \gamma' + x_{2j} \tau_{ik} \gamma' + z_i b' + \varepsilon_{ik},
\]

for \( i = 1, \ldots, n \), \( j = 1, \ldots, K \), and \( k = 1, \ldots, m_i \).

Collecting into \( Y_i = (Y_{ij1}, \ldots, Y_{ijm_i})^T \) the observations from unit \( i \), an equivalent expression for the linear degradation model in (3.1) is

\[
Y_i = X_i \gamma + Z_i b_i + \varepsilon_i
\]

where \( \gamma = (\gamma_1, \gamma_2) \), \( (b_0, b_1)^T \sim BVN(\beta, V) \), \( X_i \) and \( Z_i \) are matrices of explanatory variables

\[
X_i = \begin{pmatrix}
x_{1j} & \tau_{i1} x_{2j} \\
\vdots & \vdots \\
x_{1j} & \tau_{im_i} x_{2j}
\end{pmatrix}, \quad Z_i = \begin{pmatrix}
1 & t_{i1} \\
\vdots & \vdots \\
1 & t_{im_i}
\end{pmatrix},
\]

and \( \varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{im_i})^T \).
We assume independence between \( \varepsilon_i \) and \( b_i \) and that the components of \( \varepsilon_i \) are independent and jointly normal distributed (a reasonable assumption when spacing between measurements is not too small), expressed as, \( \varepsilon_i \sim \text{MVN}(0, \sigma^2 I_i) \) where \( I_i \) is an \( m_i \times m_i \) identity matrix. It follows that \( Y_i \sim \text{MVN}(X_i \gamma + Z_i \beta, \Sigma_i) \) where

\[
\Sigma_i = \text{Var}(X_i \gamma + Z_i \beta + \varepsilon_i) = Z_i V Z_i^T + \sigma^2 I_i. \quad (3.17)
\]

The log-likelihood for observational unit \( i \) is

\[
\mathcal{L}_i = -\frac{1}{2} \log \det(\Sigma_i) - \frac{1}{2} (Y_i - X_i \gamma - Z_i \beta)^T \Sigma_i^{-1} (Y_i - X_i \gamma - Z_i \beta).
\]

The total log-likelihood for \( n \) units is

\[
\mathcal{L} = \sum_{i=1}^{n} \mathcal{L}_i = -\frac{1}{2} \sum_{i=1}^{n} \log \det(\Sigma_i) - \frac{1}{2} \sum_{i=1}^{n} (Y_i - X_i \gamma - Z_i \beta)^T \Sigma_i^{-1} (Y_i - X_i \gamma - Z_i \beta).
\]

To save space, collect into \( \beta^* = (\gamma, \beta) \), the fixed effects model parameters, and collecting into \( \vartheta = (\sigma_{b_0}^2, \sigma_{b_1}^2, \rho, \sigma^2) \), the model variance components. Using Equation (4) of Jenrich and Schluchter (1986), it can be shown that the Hessian Matrix, \( H_i \), for unit \( i \) and its expected value are given by

\[
H_i = \begin{pmatrix}
H_{\beta^* \beta^*}, i & H_{\beta^* \vartheta}, i \\
H_{\vartheta \beta^*}, i & H_{\vartheta \vartheta}, i
\end{pmatrix} = \begin{pmatrix}
\frac{\partial^2 \mathcal{L}_i}{\partial \beta^* \partial \beta^*} & \frac{\partial^2 \mathcal{L}_i}{\partial \beta^* \partial \vartheta} \\
\frac{\partial^2 \mathcal{L}_i}{\partial \vartheta \partial \beta^*} & \frac{\partial^2 \mathcal{L}_i}{\partial \vartheta \partial \vartheta}
\end{pmatrix},
\]

where

\[
E(H_i) = I_i(\theta) = \begin{pmatrix}
X_i^T \Sigma_i^{-1} X_i & 0 \\
0 & M_i
\end{pmatrix}, \quad (3.18)
\]

where \( M_i \) is a \( 4 \times 4 \) symmetric matrix with elements

\[
M_{jk}^i = \frac{1}{2} \text{tr}(\Sigma_i^{-1} \Sigma_{ij} \Sigma_i^{-1} \Sigma_{ik}), \quad j, k = 1, \ldots, 4,
\]

and

\[
\Sigma_{ij} = \frac{\partial \Sigma_i}{\partial \vartheta_j}, \quad j = 1, \ldots, 4.
\]
From (3.17), it follows that

\[
\dot{\Sigma}_{i1} = \frac{\partial \Sigma_i}{\partial \sigma_{b0}} = Z_i \begin{pmatrix} 2\sigma_{b0} & \rho \sigma_{b1} \\ \rho \sigma_{b1} & 0 \end{pmatrix} Z'_i,
\]

\[
\dot{\Sigma}_{i2} = \frac{\partial \Sigma_i}{\partial \sigma_{b2}} = Z_i \begin{pmatrix} 0 & \rho \sigma_{b0} \\ \rho \sigma_{b0} & 2\sigma_{b1} \end{pmatrix} Z'_i,
\]

\[
\dot{\Sigma}_{i3} = \frac{\partial \Sigma_i}{\partial \rho} = Z_i \begin{pmatrix} 0 & \sigma_{b1} \sigma_{b0} \\ \sigma_{b1} \sigma_{b0} & 0 \end{pmatrix} Z'_i,
\]

\[
\dot{\Sigma}_{i4} = \frac{\partial \Sigma_i}{\partial \sigma} = 2\sigma I_i.
\]

Then the information matrix for all \( n \) units is

\[
I(\theta) = \sum_{i=1}^{n} I_i(\theta).
\]

### 3.10.2 Derivation of \( t_p \) in Section 3.3.2

Suppose that a failure occurs at the first point in time where \( D > D_f \). Let \( F \) denote the CDF of the random variable \( T \) and let \( \sigma_{b0b1} = \rho \sigma_{b0} \sigma_{b1} \). Then

\[
F(t_p) = 1 - \Phi_{\text{nor}} \left( \frac{y_t - x_1 \gamma'_1 - x_2 r \gamma'_2 - z \beta'}{\sqrt{\sigma_{b0}^2 + \tau_p^2 \sigma_{b1}^2 + 2\tau_p \sigma_{b0b1}}} \right) = p
\]

\[
\Phi_{\text{nor}} \left( \frac{y_t - \beta_0 - x_1 \gamma'_1 - (\beta_1 + x_2 \gamma'_2) \tau_p}{\sqrt{\sigma_{b0}^2 + \tau_p^2 \sigma_{b1}^2 + 2\tau_p \sigma_{b0b1}}} \right) = 1 - p
\]

\[
y_t - \beta_0 - x_1 \gamma'_1 - (\beta_1 + x_2 \gamma'_2) \tau_p = \Phi_{\text{nor}}^{-1} (1 - p)
\]

\[
\frac{(y_t - \beta_0 - x_1 \gamma'_1 - (\beta_1 + x_2 \gamma'_2) \tau_p)^2}{\sigma_{b0}^2 + \tau_p^2 \sigma_{b1}^2 + 2\tau_p \sigma_{b0b1}} = \left[ \Phi_{\text{nor}}^{-1} (1 - p) \right]^2
\]

\[
k \left( \sigma_{b0}^2 + \tau_p^2 \sigma_{b1}^2 + 2\tau_p \sigma_{b0b1} \right) = (h - r \tau_p)^2
\]

\[
k \sigma_{b0}^2 \tau_p^2 + 2\sigma_{b0b1} k \tau_p + k \sigma_{b0}^2 = h^2 - 2hr \tau_p + r^2 \tau_p^2
\]

\[
\tau_p^2 (k \sigma_{b1}^2 - r^2) + 2\tau_p (k \sigma_{b0b1} + hr) + (k \sigma_{b0}^2 - h^2) = 0
\]

where \( k = \left[ \Phi_{\text{nor}}^{-1} (1 - p) \right]^2 \), \( h = y_t - \beta_0 - x_1 \gamma'_1 \), and \( r = \beta_1 + x_2 \gamma'_2 \).
Let \( a = (k\sigma^2_{b_1} - r^2) \), \( b = 2(k\sigma_{b_0b_1} + hr) \), \( c = (k\sigma^2_{b_0} - h^2) \). This equation is of the form:

\[
a\tau_p^2 + b\tau_p + c = 0 \tag{3.20}
\]

with solutions for \( \tau_p \)

\[
\tau_p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]

\[
= -\frac{2(k\sigma_{b_0b_1} + hr) \pm \sqrt{4(k\sigma_{b_0b_1} + hr)^2 - 4(k\sigma^2_{b_1} - r^2)(k\sigma^2_{b_0} - h^2)}}{2(k\sigma^2_{b_1} - r^2)}
\]

\[
= -\frac{(k\sigma_{b_0b_1} + hr) \pm \sqrt{(k\sigma_{b_0b_1} + hr)^2 - (k\sigma^2_{b_1} - r^2)(k\sigma^2_{b_0} - h^2)}}{k\sigma^2_{b_1} - r^2}.
\]

The derivation when failure is defined as the first time at which \( D < D_f \) is similar, starting with

\[
F(t_p) = \Phi_{norm} \left( \frac{y_t - x_1\gamma'_1 - x_2\tau_p\gamma'_2 - z\beta'}{\sqrt{\sigma^2_{b_0} + \tau^2_p\sigma^2_{b_1} + 2\tau_p\sigma_{b_0b_1}}} \right).
\]

### 3.11 References


CHAPTER 4. Bayesian Methods For Planning Accelerated Repeated Measures Degradation Studies

Brian P. Weaver and William Q. Meeker
Department of Statistics
Iowa State University
Ames, IA, 50010

Abstract

Accelerated repeated measures degradation tests are often used to assess product or component reliability when there would be few or even no failures during a traditional life test. Such tests are used to estimate the failure-time distributions of highly reliable items in applications where it is possible to take repeated measures of some appropriate degradation measure. When engineers have valid prior information about failure mechanisms, it is important that such information be used in inference and test planning. Bayesian methods provide a vehicle for doing this. This paper describes methods for selecting a Bayesian repeated measures accelerated degradation test plan when the degradation and acceleration model is linear in the parameters. A Bayesian criterion based on estimation precision of the lifetime quantile at use conditions is selected for finding optimum test plans. We use a large-sample approximation for the posterior distribution to simplify the planning criterion. The general equivalence theorem is used to check for global optimality of the optimum test plan. We also discuss how to find a compromise test plan that satisfies practical constraints while still providing good statistical properties.

Keywords: Nondestructive Degradation, Mixed Effects Linear Models, Accelerated Degradation Testing, General Equivalence Theory, Bayesian Experimental Design
4.1 Introduction

Manufacturers often need to demonstrate the reliability of devices used in their products. One approach to demonstrating reliability is to estimate the failure-time distribution of the device. Traditionally, the failure-time distribution could be estimated using life tests or accelerated life tests, where the response is time to failure. With highly reliable devices, however, traditional life tests typically yield few or no failures. Instead, engineers can sometimes use methods that measure the degradation of an item nondestructively by making degradation measurements over time, providing more information than the traditional life tests. In some situations, the degradation rate is so low that any noticeable degradation will not be observed during the life of the test. To address this issue, engineers will expose devices to accelerated usage conditions, such as higher temperature, stress, or humidity. Under these conditions, measurements of the degradation are made and a model describing the relationship between acceleration and degradation is fit to the data. This relationship can then be used to extrapolate to normal use conditions. Because resources (such as time and money) are limited, properly planning a test beforehand is important. One way of doing this is constructing optimum tests, for a given criteria such as optimum estimation of a lifetime quantile at use conditions, that can be used to help guide the construction of good test plans.

In some situations there is information about the degradation process beforehand, either in the form of past degradation tests, knowledge from the physics of failure, or from other expert judgement. Bayesian design methods use this past information in the form of prior information for model parameters. The information is used both for inference and test planning. The primary motivation of this paper is to use this prior information to construct tests that provide good estimation of the lifetime quantile of a device at use conditions. Bayesian test plans are a formal method to incorporate prior information about the model parameters into an analysis and also to develop plans with good statistical properties.
4.1.1 Related Work


4.1.2 Motivating Example

This work is motivated by a number of applications described in Weaver and Meeker (2011). In this paper we focus on an experiment that was performed by Shiomi and Yanagisawa (1979) that involved carbon-film resistors. The engineers measured the resistance of carbon-film resistors at particular points in time. The resistors were run at three different levels of temperature ($83^\circ C$, $133^\circ C$, and $173^\circ C$) in order to accelerate the degradation. They defined degradation to be the relative change in resistance over time. Figure 4.1a presents the original data and Figure 4.1b shows the data plotted versus the square root of time (i.e., on a square root axis). We suppose that the engineers plan to conduct a similar experiment on a resistor using a different formulation and want to estimate a quantile of the failure-time distribution at the nominal use conditions ($50^\circ C$) and that a failure would occur when the relative change has increased to 10%. For two reasons we are considering a different form of the data than that used in Weaver and Meeker (2011). First, it is quite common for failure to be defined in terms of relative change from some baseline measurement. Second, by using relative change
as the response we have reduced the number of parameters in our degradation model because
now the intercept is a known constant (not random) and is defined to be 0.

![Figure 4.1](image)

**Figure 4.1** Carbon-Resistor data from Shiomi and Yanagisawa (1979) plotted on the original scale and a square root of time scale.

### 4.1.3 Overview

The remainder of this paper is organized as follows. Section 4.2 describes the accelerated
degradation model used and related quantities. Section 4.3 derives the lifetime distribution and
the $p$ quantile of this distribution. Section 4.4 gives the methods needed to create Bayesian test
plans. Section 4.5 illustrates finding Bayesian test plans for the carbon-film resistor example.
Section 4.6 uses a simulation technique to compare different Bayesian test plans. Finally,
Section 4.7 gives concluding remarks and outlines areas of related future work.

### 4.2 Degradation Models

#### 4.2.1 Accelerated Repeated Measures Degradation Models

The actual degradation level for an observational unit at time $t$ and accelerating variable(s)
is denoted by $D = D(\tau, x_1, x_2, \theta)$ where $\tau = h_\theta(t)$ is a monotone increasing transformation of
time, $x_1$ and $x_2$ are vectors of (possibly transformed) accelerating variables, and $\theta$ is the unknown parameter vector.

The model used in the example and other applications that we have encountered are special cases of the following generic model. The (possibly transformed) observed degradation $Y$ at some (possibly transformed) time point $t$ for a fixed value(s) of the (possibly transformed) accelerating variables $x_1$ and $x_2$ is

$$Y = h_d(D) + \varepsilon$$

$$= \mu(\tau, x_1, x_2) + \varepsilon$$

$$= x_1\gamma_1' + x_2\tau\gamma_2' + b_0 + b_1\tau + \varepsilon$$

$$= x_1\gamma_1' + x_2\tau\gamma_2' + zb' + \varepsilon,$$

(4.1)

so that $\mu(\tau, x_1, x_2) = h_d(D)$ is a location parameter for the distribution of $Y$ that depends on the unknown parameters $\theta$, $z = (1, \tau)$, $\varepsilon \overset{iid}{\sim} N(0, \sigma^2)$ is a random variable that describes the within unit variation, and $h_d$ is a monotone increasing or decreasing transformation of $D$. In (4.1), the term $x_1\gamma_1'$ describes how the intercept or initial value of degradation changes with the accelerating variable(s), $x_2\gamma_2'$ describes how the degradation slope or rate changes with the accelerating variable(s), and $zb'$ describes the variability in the degradation slope and intercept across units. Note that in some applications, $x_1$ and $x_2$ will be the same (for example see the scar width application in Weaver and Meeker 2011).

We assume that the variability in the linear regression parameters $b$ can be described by a bivariate normal distribution

$$(b_0, b_1)' \sim BVN(\beta, V)$$

where $\beta = (\beta_0, \beta_1)'$ is the mean vector representing the slope and intercept common to all units and

$$V = \begin{pmatrix} \sigma_{b_0}^2 & \rho\sigma_{b_0}\sigma_{b_1} \\ \rho\sigma_{b_0}\sigma_{b_1} & \sigma_{b_1}^2 \end{pmatrix}$$

is the covariance matrix. We further assume that $(b_0, b_1)'$ is independent of $\varepsilon$. This implies that different units are independent of each other.
4.2.2 Model for Resistor Degradation

For the particular resistor application introduced in Section 4.1.2, the following model will be used. Let $y_{ijk}$ be the observed degradation for unit $i$, $i = 1, \ldots, 19$, at time point $j$, $j = 1, \ldots, 5$ for level $k$ of the accelerating variable, $k = 1, 2, 3$, and let $x_k$ denote the Arrhenius transformation of temperature, i.e., $x = -11605/(\text{Temp}^\circ C + 273.15)$. The assumed linear accelerated degradation random effects model is

$$y_{ijk} = \beta \tau_{ij} + \gamma x_k \tau_{ij} + b_i \tau_{ij} + \varepsilon_{ijk}$$

(4.2)

where $\tau_{ij}$ is the square root of time. Notice that (4.2) is a special case of (4.1) with $h_d(D) = D$ (i.e., the identity function), $x_1 = 0$, $x_2 = x$, and $\rho = \sigma_{b0} = 0$. We will assume that $b \overset{iid}{\sim} \mathcal{N}(0, \sigma_b^2)$ describes unit-to-unit variability in the slopes and is independent of $\varepsilon$.

The above acceleration model adequately describes the underlying failure mechanism only within a certain range of $x$, say $[x_U, x_H]$ where $x_U$ is the value of the accelerating variable at use conditions and $x_H$ is the highest value. For practical RMADTs, the interval will be restricted to $[x_L, x_H]$ where testing beyond $x_H$ will cause the model to break down due to possible changes in the failure mechanism and testing below $x_L$ will not provide any meaningful degradation information. Generally, $x_L > x_U$.

4.2.3 Reparameterization of the Model for Prior Distribution Specification

For the numerical calculations involved, we reparameterize (4.2) to form an alternative set of stable parameters, as described in Ross (1990). This reparameterization is used to break strong correlations between pairs of estimators and to speed convergence of algorithms used for estimation (both ML and MCMC).

Let $\bar{x}$ denote the mean of the accelerating variable. Then (4.2) can be reparameterized as

$$Y = \beta^* \tau + \gamma \tau (x - \bar{x}) + b \tau + \varepsilon$$

where $\beta^* = \beta + \gamma \bar{x}$ is the degradation slope for units tested at the average accelerating variable.
4.3 Failure-Time Distribution for Degradation Models

This section describes the relationship between the degradation model and the induced failure-time model. If a degradation model and definition of failure are given, a failure-time distribution is implied (e.g., Chapter 13 of Meeker and Escobar 1998).

4.3.1 Failure-Time Cumulative Distribution Function

This section shows how to derive the failure-time distribution from the degradation model in (4.2). A degradation process with soft failures is assumed. That is, a unit “fails” when it reaches a pre-specified degradation level $D_f$. This is equivalent to $D \geq D_f$ (for increasing $D$) or $D \leq D_f$ (for decreasing $D$). The failure time for a unit is the time at which it reaches the degradation level $D_f$. Let $T$ define the random variable associated with the item’s time to failure. If $b\tau \sim N(0, \sigma_b^2 \tau^2)$, then

$$F(t; x) = \Pr(T \leq t) = \Pr(D \geq D_f) = \Pr(b\tau \geq D_f - \beta \tau - \gamma \tau x) = 1 - \Phi_{nor}(\kappa)$$

(4.3)

where

$$\kappa = \frac{D_f - \beta \tau - \gamma \tau x}{\sigma_b \tau}$$

and $\Phi_{nor}$ is the CDF of the standard normal distribution.

Similarly, if a failure happens when $D \leq D_f$, then

$$F(t; x) = \Phi_{nor}(\kappa).$$

(4.4)

Meeker and Escobar (1998), page 330, describes a numerical integration and simulation based approach for more complicated models where a closed form solution does not exist.

4.3.2 Failure-Time Quantiles

From (4.3) or (4.4) the $p$ quantile, denoted as $t_p$, is $t_p = h_t^{-1}(\tau_p)$ where

$$\tau_p = \frac{D_f}{\zeta \sigma_b + (\beta + \gamma x)}$$

(4.5)
where \( \zeta = \Phi_{\text{nor}}^{-1}(1-p) \) or \( \zeta = \Phi_{\text{nor}}^{-1}(p) \) depending on whether a failure is declared when \( D \geq D_f \) or \( D \leq D_f \), respectively. The derivation of (4.5) is given in the appendix.

### 4.4 Planning Bayesian Repeated Measures Accelerated Degradation Tests

#### 4.4.1 Plan Specification

Let \( x_{(i)} = (x_{1,i}, x_{2,i}) \) be a factor level combination of (possibly transformed) accelerated conditions for unit \( i \) and let \( \pi_i \) denote the proportion of units that are to be tested at \( x_{(i)} \). An RMADT plan will specify the levels of the accelerating variable to test and the proportion of units to be tested at those levels. A test plan with \( K \) levels is denoted by

\[
\eta = \begin{bmatrix}
  x_{(1)}, & \pi_1 \\
  x_{(2)}, & \pi_2 \\
  \vdots & \vdots \\
  x_{(K)}, & \pi_K 
\end{bmatrix},
\]

where \( \sum_{i=1}^{K} \pi_i = 1 \).

#### 4.4.2 Bayesian Planning Criterion

We follow the general approach used in Zhang and Meeker (2006) to define our objective function. The purpose of RMADT is estimation of a lifetime quantile \( t_p \). Let the utility function be minus the posterior variance denoted by \( \text{var}_{\theta|Y,\eta}(t_p) \). The posterior variance depends on the unobserved data, so an expectation is taken over all possible data

\[
C(\eta) = -\int_Y \text{var}_{\theta|Y,\eta}(t_p)p(Y|\eta)dY.
\]  

(4.6)

Let \( g(\theta) = t_p, c_i(\theta) = \partial g(\theta)/\partial \theta_i \) for \( i = 1, \ldots, 4 \), and \( c(\theta) = (c_1(\theta), \ldots, c_4(\theta))' \). Chaloner and Verdinelli (1995) give as an approximation to the posterior variance

\[
\text{var}_{\theta|Y,\eta}(t_p) \approx c(\hat{\theta})'[S^{-1} + \hat{I}(\hat{\theta}, \eta)]^{-1}c(\theta)
\]

(4.7)

where \( S^{-1} \) is the prior precision matrix for \( \theta \), and \( \hat{I}(\theta, \eta) \) is the Fisher information matrix for the test plan \( \eta \) evaluated at the maximum likelihood estimate \( \hat{\theta} \) where the Fisher information
matrix is defined as $\mathbf{I}(\theta) = - \mathbf{E}(\partial^2 \mathbf{L}/\partial \theta^2)$ whose explicit form is given in the appendix.

Notice that (4.7) depends on the data only through the maximum likelihood estimator $\hat{\theta}$, so the integration of (4.6) can be done with respect to the predictive distribution of $\hat{\theta}$, denoted as $p(\hat{\theta})$ where

$$p(\hat{\theta}) = \int p(\hat{\theta}|\theta)p(\theta)d(\theta).$$

So, (4.6) can be approximated as

$$C(\eta) \approx - \int \mathbf{c}(\hat{\theta})'[\mathbf{S}^{-1} + \mathbf{I}(\theta, \eta)]^{-1}\mathbf{c}(\hat{\theta})d(p(\hat{\theta})).$$

The distribution $p(\hat{\theta}|\theta)$ is not tractable, however. Let $p(\theta|\vartheta)$ be the appropriate joint prior distribution for $\theta$ with known hyperparameters $\vartheta$. As the sample size increases, $\hat{\theta}$ converges in distribution to $\theta$. That is $p(\hat{\theta}|\theta)$ converges to $p(\theta|\vartheta)$ and (4.6) can further be approximated as

$$C(\eta) \approx - \int \mathbf{c}(\theta)'[\mathbf{S}^{-1} + \mathbf{I}(\theta, \eta)]^{-1}\mathbf{c}(\theta)d(p(\theta|\vartheta)).$$

In (4.9), the elements of the vector $\mathbf{c}$ are

$$\frac{\partial t_p}{\partial \beta} = - \frac{D_f(\zeta \sigma_b + \beta + \gamma x)^2}{(\zeta \sigma_b + \beta + \gamma x)^2}, \quad \frac{\partial t_p}{\partial \gamma} = - \frac{D_f x(\zeta \sigma_b + \beta + \gamma x)^2}{(\zeta \sigma_b + \beta + \gamma x)^2},$$

$$\frac{\partial t_p}{\partial \sigma_b \zeta} = - \frac{D_f}{(\zeta \sigma_b + \beta + \gamma x)^2}, \quad \frac{\partial t_p}{\partial \sigma} = 0.$$

### 4.4.3 The Prior Distribution

Prior information is generally elicited from engineers or other experts and is usually based on past experience and past data. In Bayesian experimental design there are two forms of prior information to consider:

1. The prior information used to design the experiment,
2. The prior information used in inference.

Tsutakawa (1972) and Etzioni and Kadane (1993) discuss the need to use different prior distributions for inferences and planning. They mention that inference about the parameters happens much later than the actual experimental design. They also argue that the risk and
criteria for those designing and conducting the experiment could be completely different than those using the results of inference from the experiment and so different prior information would be needed.

In (4.9), the prior information used to design the experiment is represented by $p(\theta|\vartheta)$. Generally the prior distribution used to design the experiment must be an informative prior. This is because test planning criteria are highly sensitive to this prior information. The prior used in inference is represented by the precision matrix $S^{-1}$. To specify a diffuse prior for inference, $S^{-1}$ is specified as a matrix of zeros.

We will explore several different combinations of prior distributions for Bayesian RMADT planning. In particular, we will use point-mass priors $p_0(\vartheta)$, a strongly informative prior for all parameters $p_1(\vartheta)$, a weakly informative prior for all parameters $p_2(\vartheta)$, a strongly informative prior for only one parameter $p_3(\vartheta)$, and a noninformative prior $p_4(\vartheta)$. Table 4.1 gives the different combinations of prior information to be used in our test planning examples. As mentioned in Section 4.4.3, the prior information for the test planning will always be informative.

<table>
<thead>
<tr>
<th>Cases</th>
<th>Test Planning</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$p_0(\vartheta)$</td>
<td>$p_4(\vartheta)$</td>
</tr>
<tr>
<td>B</td>
<td>$p_1(\vartheta)$</td>
<td>$p_4(\vartheta)$</td>
</tr>
<tr>
<td>C</td>
<td>$p_1(\vartheta)$</td>
<td>$p_2(\vartheta)$</td>
</tr>
<tr>
<td>D</td>
<td>$p_1(\vartheta)$</td>
<td>$p_3(\vartheta)$</td>
</tr>
<tr>
<td>E</td>
<td>$p_1(\vartheta)$</td>
<td>$p_1(\vartheta)$</td>
</tr>
</tbody>
</table>

Table 4.1 Different combinations of prior specification. $p_0(\vartheta)$ denotes the point-mass prior, $p_1(\vartheta)$ denotes the strongly informative prior for all parameters, $p_2(\vartheta)$ denotes the weakly informative prior for all parameters, $p_3(\vartheta)$ denotes the strongly informative prior for only one parameter, and $p_4(\vartheta)$ denotes the diffuse prior.

4.4.4 General Equivalence Theory

The appendix of Zhang and Meeker (2006) provides a review for the general equivalence theorem (GET). The GET depends on the directional derivative of (4.9). For this criterion,
Zhang and Meeker (2006) give the directional derivative as

$$
\Lambda(\eta, \xi) = \int c(\theta)'V(\theta, \eta)V^{-1}(\theta, \eta_\xi)V(\theta, \eta)c(\theta)d(p(\theta|\vartheta)) + C(\eta)
$$

(4.10)

where $V(\theta, \eta) = [S^{-1} + \mathcal{I}(\theta, \eta)]^{-1}$ and $\xi$ is an alternative plan from the design space (the set of all possible designs). Equation (4.10) will be used in the Wynn algorithm (outlined in Section 4.4.5) to find an optimum test plan and to also verify that it is optimum.

Zhang and Meeker (2006) discuss in their appendix how (4.10) satisfies all the requirements needed to apply the GET. Therefore, Whittle’s GET can be used to determine if a test plan is optimum or not. According to the GET, a test plan $\eta^*$ will be optimum if and only if for each singular plan in the design space $\eta_x$ (i.e., a test plan where all units are tested at one point in the test space, the set of all possible accelerating factor level combinations), the value of (4.10) evaluated at $\eta^*$ and $\eta_x$ is less than or equal to 0 and (4.10) obtains a maximum value of 0 and these values are achieved at the points in the design space that are also in the test plan $\eta^*$.

4.4.5 The Wynn Algorithm Used to Find Optimum Plans

We use the Wynn algorithm to find optimum plans. The outline of the algorithm (described more completely in Wynn 1972) is as follows:

1. Specify $p(\theta|\vartheta)$, the prior distribution for test planning.

2. Determine the range of accelerating values to which the items will be exposed. We call this range the test space and it is denoted by $\mathcal{X}$. For the resistor data, $\mathcal{X} = [83^\circ C, 173^\circ C]$.

3. Define a grid of values over $\mathcal{X}$. Randomly and uniformly pick $n$ points from this grid and assign equal weights to each point ($1/n$). This represents the initial candidate design $\eta$.

4. Create a set of singular test plans $\mathcal{H}$ by assigning a singular test plan to each value of $x$ in the grid of values over $\mathcal{X}$.

5. For each singular test plan $\eta_x \in \mathcal{H}$, calculate (4.10) evaluated at $\eta$ and pointing in the direction of $\eta_x$.

6. Save the largest value of (4.10) and the corresponding $x$; call these $\Lambda_{\text{max}}$ and $x_{\text{max}}$. 
7. If the directional derivative $\Lambda_{\text{max}}$ is less than a specified stopping value, then stop the algorithm. Otherwise, set $n = n + 1$, add $x_{\text{max}}$ to $\eta$ with weight $1/n$, down-weight other points in the design by the factor $1 - 1/n$, and return to step 4.

4.5 Numerical Example

This section presents a numerical example to illustrate the planning methods described in Section 4.4. This example is based on the experiment done by Shiomi and Yanagisawa (1979), described in Section 4.1.2.

4.5.1 Prior Specification

Prior distributions for the model parameters can be obtained from previous experiments and data on a similar product or from expert opinion. For this example, the degradation is increasing with time so we expect $\beta$ to be a positive value. $\gamma$ can be thought of as an effective activation energy, relating degradation rate to temperature. Engineers often have good information about this parameter from previous experience with particular failure mechanisms. The parameters $\gamma$, $\sigma_b$, and $\sigma$ are all positive and their prior distributions should put all mass on the positive real line. For $\beta$ we will use a normal prior distribution (equivalent to assigning $\exp(\beta)$ a lognormal distribution), $\gamma$, $\sigma_b$, and $\sigma$ will be assigned lognormal distributions. These prior distributions will be independent of each other.

Point-mass Prior Distribution $p_0(\vartheta)$:

Weaver and Meeker (2011) present non-Bayesian methods for finding RMADT plans. The locally optimum test plans developed in that paper require specifying values for the model parameters. These specific values are called “planning values.” This planning information is equivalent to setting a point-mass prior at those specified values. A point-mass prior is viewed to be highly informative that is centered around the planning information and will be approximated by normal distributions with small variances. The approximate point-mass priors are given in Table 4.2.

Strongly Informative Prior Distribution $p_1(\vartheta)$:
Table 4.2 Prior information specified by independent marginal distributions for the approximate point-mass prior distributions.

Table 4.3 summarizes the independent prior information for the model parameters, their distributions, and their hyperparameters.

Table 4.3 Prior information specified by independent marginal distributions for the strongly informative prior distributions.

**Strongly Informative for Only One Parameter Prior Distribution \( p_3(\vartheta) \):**

Table 4.4 summarizes the independent prior information for the model parameters, their distributions, and their hyperparameters.

**Weakly Informative Prior Distribution \( p_2(\vartheta) \):**
Table 4.5 summarizes the independent prior information for the model parameters, their distributions, and their hyperparameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior Specification</th>
<th>Hyperparameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01 quantile</td>
<td>0.99 quantile</td>
</tr>
<tr>
<td>( \beta )</td>
<td>-0.044</td>
<td>0.444</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>0.0001</td>
<td>0.01</td>
</tr>
<tr>
<td>( \sigma_b )</td>
<td>0.01</td>
<td>0.5</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.10</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>standard deviation</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.2</td>
<td>0.105</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>( \sigma_b )</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.35</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Table 4.5 Prior information specified by independent marginal distributions for the weakly informative prior distributions.

4.5.2 Bayesian Optimum Test Plans

For all the cases below, the Wynn algorithm, as described in Section 4.4.5, combined with the appropriate prior information was used to find the optimum designs. Also, optimality for each case was verified numerically using GET and only a plot for Case A is supplied since all the other plots looked very similar.

Case A

Non-Bayesian designs can be thought of as special cases of Bayesian designs where the prior distributions for the design assigned to the parameters have all their mass at a single point in the parameter space corresponding to the planning values required for a locally-optimum design problem and a noninformative prior is used for the inference. Weaver and Meeker (2011) present non-Bayesian approaches to planning RMADTs. The Bayes optimum plan is given in Table 4.6 along with the absolute value of (4.9) for the different sample sizes \( n = 29 \) and \( n = 2900 \). For this table and others, we introduce the notation \( \pi_L \) to denote the proportion of units that are to be tested at the lower level of temperature 83°C and \( \pi_H \) for the proportion of units allocated to the higher level of temperature 173°C.

The GET was used to numerically verify optimality of \( \eta^*_A \). Figure shows (4.10) evaluated at \( \eta^*_A \) and as a function of \( x \). Notice that at the endpoints, the maximum value of (4.10) is 0 and \( \eta^*_A \) is optimum by the GET. The Wynn algorithm was also used for the non-Bayesian
methodologies and an optimum plan as described in Weaver and Meeker (2011). This test plan is very similar to the plan found in Case A. Details and properties of the non-Bayesian optimum test can be found in Weaver and Meeker (2011).

Case B

A Bayesian optimum plan for Case B is given in Table 4.6. Notice that the test plan for Case B is very similar to Case A and the non-Bayesian approach which corresponds to one of the general conclusions presented by Chaloner and Larntz (1989). They concluded that prior distributions over a small region of the parameters space will produce optimum designs very similar to those obtained from non-Bayesian methods. Note, as expected, $|C(\eta)|$ is inversely proportional to $n$.

Cases C, D, and E

Cases C, D, and E are different from Cases A and B in that they introduce prior information for estimation of the parameters. With this additional prior information, one cannot factor out the sample size in (4.9), because $n$ affect only $I(\theta, \eta)$ and not the prior distribution. So to calculate optimum test plans for Cases C, D, and E, we will consider two
Table 4.6 Optimum test plans for Cases A and B and the absolute value of the Bayesian planning criterion in (4.9) evaluated at the prior information for different sample sizes.

| Case | $\pi_L$ | $\pi_H$ | $|C(\eta)|$ for $n = 29$ | $|C(\eta)|$ for $n = 2900$ |
|------|---------|---------|-------------------------|-------------------------|
| A    | 0.750   | 0.250   | $4.47 \times 10^{-4}$   | $4.47 \times 10^{-6}$   |
| B    | 0.749   | 0.251   | $7.3 \times 10^{-4}$    | $7.3 \times 10^{-6}$    |

different sample size, $n = 29$ and $n = 2900$. When the sample size is large, the posterior distribution will be less dependent on the prior information for inference. When the sample size is small, however, the posterior and the design will be more strongly influenced by the prior information for inference.

Table 4.7 shows the Bayesian optimum plans for the three different cases using informative prior information for inference and the two different sample sizes along with the absolute value of (4.9). Notice that for the smaller sample size, a larger portion of the units are to be tested at $83^\circ C$. This also corresponds with another general conclusion from Chaloner and Larntz (1989). They also concluded that as the amount of prior information increases, the proportion of points to be tested at the highest level of the temperature decreases. For our example, as the prior information increased for $\gamma$, we see that the number of units to be tested at the highest value of the accelerating variable decreases. Intuitively, this is saying if we go to the extreme and state we know $\gamma$ with 100% certainty, then we would not need to test at points that would improve estimation of $\gamma$. In contrast, for Cases A and B because there is no prior information for estimation, both optimum plans want to put more of the units across the design space (since $\gamma$ needs to have its estimation improved). As the sample size increases, however, the prior information for estimation has less of an influence and the optimum test plans found become more similar to those found in Cases A and B.

4.5.3 Bayesian Compromise Test Plans

Bayesian optimum test plans give the smallest value of the large-sample approximate variance of the expected marginal posterior distribution of the failure-time quantile at use con-
ditions. These test plans, however, may not be practical or robust to deviations from model assumptions because there are only two levels of the accelerating variable being tested. This section describes compromise test plans that are more robust to deviations from specified inputs and that still have good statistical properties.

We impose the constraint that some proportion of items is to be tested at a middle level of the accelerating variables (i.e., halfway between the lowest and highest level, in our example \( x_M = 128^\circ C \)). This prevents the test plan from degenerating to an optimum test plan. Notice that the Fisher Information matrix can be rewritten as

\[
I(\theta, \eta) = n[\pi_L I(\theta, x_L) + \pi_M I(\theta, x_M) + (1 - \pi_L - \pi_M) I(\theta, x_H)] \tag{4.11}
\]

which will then be evaluated in (4.9) to find the compromise plans.

Table 4.8 gives the compromise test plans for Case C and for the sample sizes \( n = 29 \) and \( n = 2900 \) as well as the absolute value of (4.9).

<table>
<thead>
<tr>
<th>Test</th>
<th>( n = 29 )</th>
<th></th>
<th>( n = 2900 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \pi_L )  ( \pi_M )  ( \pi_H )</td>
<td>(</td>
<td>C(\eta)</td>
</tr>
<tr>
<td>Compromise</td>
<td>0.85 0.1 0.05 0.00039</td>
<td></td>
<td>0.7 0.1 0.2 7.66 \times 10^{-6}</td>
</tr>
<tr>
<td>Optimum</td>
<td>0.947 0 0.053 0.00038</td>
<td></td>
<td>0.754 0 0.246 7.03 \times 10^{-6}</td>
</tr>
</tbody>
</table>

Table 4.8 Compromise test plans and optimum test plans for Case C and for different sample sizes. Also present on this table is the absolute value of the Bayesian test planning criterion in (4.9) evaluated at sample size, and test plans where the expectation is with respect to the prior distribution for test planning.
4.6 Simulation Evaluation of a Bayesian RMADT Plan

This section describes simulation methods for planning Bayesian RMADTs. Simulation is useful to help illustrate the potential variability in the estimation after a test plan has been selected. This is particularly useful to engineers as it helps visualize to them the possible range of estimated values they can expect for a given test plan. Simulation can also be used to determine how well the large-sample approximation are performing. For evaluation of test plans, one can repeatedly simulate data from one set of model parameters which can then be used to see how estimates behave for a specific plan.

4.6.1 Simulations Based on Fixed Model Parameters

A simulation study was conducted of the optimum test plans found for Case B and Case E where the true model parameters were selected to be the mean of the prior distribution. These values are \( \theta = (\beta = 0.2, \gamma = 0.0002, \sigma_b = 0.033, \sigma = 0.335) \) where the open box denotes planning information. Since this procedure is simulating an estimation technique, the prior information for estimation should be used (as opposed to the prior information for test planning). A total of 5000 data sets were simulated and for each data set the posterior median of \( t_{0.10} \) was calculated at use conditions. For the estimation, we use the histogram prior approach as outlined in Albert (2007). Note that \( t_{0.10} \) is a function of \( \theta \) and the posterior distribution of \( \theta \) is proportional to \( \text{likelihood}(\text{data}|\theta) \times p(\theta) \). Figure 4.3 shows the 5000 estimates of \( t_{0.10} \) at use conditions and Table 4.9 shows some summary statistics of the estimates. The black, vertical line in the plots corresponds to the true value of \( t_{0.10} \) which is 0.19. There are a couple of features of this plotted that should be mentioned. Notice that for Case B, the distribution is more spread out and seems to have a larger bias for \( t_{0.10} \). In contrast, the distribution for Case E has less spread than Figure 4.3a and also has a smaller bias. This comes as no surprise since the prior information used for Case E is highly informative and centered around the true values of the parameters.

Since one of the main purposes of accelerated degradation testing is to estimate the relationship between the rate of degradation and an accelerating variable, engineers might also be
Figure 4.3 Results of 5000 simulations of the posterior median of $t_{0.10}$ at the use conditions for Case B with diffuse prior information for inference (left plot) and for Case E with strongly informative prior information for inference (right plot). The black, vertical line in the plots corresponds to the true value of $t_{0.10}$ which is 0.19.

interested in estimating this relationship at the use conditions, i.e., $\beta + \gamma x_U$. Figures 4.4a and 4.4b show 5000 simulations of $\beta + \gamma x_U$ for Cases B and E, respectively. The black vertical line corresponds to the true value of $\beta + \gamma x_U$ which is 0.187. We again see similar features as those mentioned in Figures 4.3a and 4.3b. Notice also that for Case B, there is a significant amount of posterior mass for lower values of $\beta + \gamma x_U$, in particular $\Pr(\beta + \gamma x_U \leq 0.17) = 0.06$. In contrast, the posterior probability of this event under Case E is 0.003. Again, this comes as no surprise since the prior information used for Case E is highly informative.

4.7 Conclusions and Future Work

When prior information is available, planning an RMADT that incorporates this information is useful for making more precise inference about a failure-time distribution quantile. We presented Bayesian test planning methods for RMADTs. We use a Bayesian criterion function that is based on estimation for the quantile of the failure-time distribution. A large-sample approximation was employed to simplify the posterior distribution. The GET was then used
Table 4.9  Summary statistics for the simulations of $t_{0.10}$ at the use conditions for Case B and for Case E.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Case B</th>
<th>Case E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Median</td>
<td>0.200</td>
<td>0.194</td>
</tr>
<tr>
<td>Sample Standard Deviation</td>
<td>0.018</td>
<td>0.013</td>
</tr>
<tr>
<td>Sample 2.5% Quantile</td>
<td>0.167</td>
<td>0.171</td>
</tr>
<tr>
<td>Sample 97.5% Quantile</td>
<td>0.239</td>
<td>0.220</td>
</tr>
</tbody>
</table>

Figure 4.4  5000 simulations of the posterior estimate of $\beta + \gamma x_U$ for Case B (left plot) and for Case E (right plot).

to verify the optimality of proposed tests.

The methods in this paper can be extended to more complicated models that are of the form of (4.1). These could include models where there are multiple accelerating variables that affect the rate of degradation or that also affect the intercept (such as on a log scale). Examples of such models are found in Weaver and Meeker (2011). As the number of parameters increases, however, the computational time will increase as well.
Statistics | Case B | Case E  
--- | --- | ---  
Sample Median | 0.183 | 0.185  
Sample Standard Deviation | 0.008 | 0.005  
Sample 2.5% Quantile | 0.167 | 0.175  
Sample 97.5% Quantile | 0.198 | 0.196  

Table 4.10 Summary statistics for the simulations of the posterior estimate of $\beta + \gamma x_u$ for Case B and for Case E.

4.8 Appendix

4.8.1 Fisher Information Matrix

The model in Equation (4.2) will be used for the derivation of the Fisher information matrix. Collecting into $Y_i = (Y_{i1}, \ldots, Y_{i5})'$ the observations from unit $i$, an equivalent expression for the linear degradation model in (4.2) is

$$Y_i = X_i \beta' + Z_i b_i' + \epsilon_i$$  \hspace{1cm} (4.12)

where $\beta = (\beta, \gamma)$, $X_i$ and $Z_i$ are matrices of explanatory variables

$$X_i = \begin{pmatrix} \tau_{i1} & \tau_{i1} x_j \\ \vdots & \vdots \\ \tau_{i5} & \tau_{i5} x_j \end{pmatrix}, \quad Z_i = \begin{pmatrix} \tau_{i1} \\ \vdots \\ \tau_{im_i} \end{pmatrix}$$

and $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{i5})'$.

We assume independence between $\epsilon_i$ and $b_i$ and that the components of $\epsilon_i$ are independent and jointly normal distributed (a reasonable assumption when spacing between measurements is not too small), expressed as, $\epsilon_i \sim \text{MVN}(0, \sigma^2 I_i)$ where $I_i$ is an $5 \times 5$ identity matrix. It follows that $Y_i \sim \text{MVN}(X_i \beta, \Sigma_i)$ where

$$\Sigma_i = \text{Var}(X_i \beta' + Z_i b_i' + \epsilon_i) = \sigma_b^2 Z_i Z_i' + \sigma^2 I_i.$$  \hspace{1cm} (4.13)

The log-likelihood for observational unit $i$ is

$$L_i = -\frac{1}{2} \log \det(\Sigma_i) - \frac{1}{2} (Y_i - X_i \beta')' \Sigma_i^{-1} (Y_i - X_i \beta - Z_i \beta).$$
The total log-likelihood for \( n \) units is

\[
\mathcal{L} = \sum_{i=1}^{n} \mathcal{L}_i = -\frac{1}{2} \sum_{i=1}^{n} \log \left[ \det (\Sigma_i) \right] - \frac{1}{2} \sum_{i=1}^{n} (Y_i - X_i\beta')' \Sigma_i^{-1}(Y_i - X_i\beta').
\]

Let \( \phi = (\sigma_b, \sigma) \) and using (4) of Jenrich and Schluchter (1986), it can be shown that the Hessian Matrix, \( H_i \), for unit \( i \) and its expected value, \( E(H_i) \), are given by

\[
H_i = \begin{pmatrix}
H_{\beta\beta,i} & H_{\beta\phi,i} \\
H_{\phi\beta,i} & H_{\phi\phi,i}
\end{pmatrix} = \begin{pmatrix}
\frac{\partial^2 \mathcal{L}_i}{\partial \beta \partial \beta} & \frac{\partial^2 \mathcal{L}_i}{\partial \beta \partial \phi} \\
\frac{\partial^2 \mathcal{L}_i}{\partial \phi \partial \beta} & \frac{\partial^2 \mathcal{L}_i}{\partial \phi \partial \phi}
\end{pmatrix},
\]

\[
E(H_i) = \mathcal{I}_i(\theta) = \begin{pmatrix}
X_i' \Sigma_i^{-1} X_i & 0 \\
0 & M_i
\end{pmatrix}, \tag{4.14}
\]

where \( M_i \) is a \( 2 \times 2 \) symmetric matrix with elements

\[
M_{jk}^i = \frac{1}{2} \text{tr}(\Sigma_i^{-1}\Sigma_i^{-1} \Sigma_i^{-1} \Sigma_i^{-1} \Sigma_i^{-1} \Sigma_i^{-1} \Sigma_i^{-1} \Sigma_i^{-1}), \quad j = 1, 2; \quad k = 1, 2,
\]

and

\[
\Sigma_{ij} = \frac{\partial \Sigma_i}{\partial \phi_j}, \quad j = 1, 2.
\]

From (4.13), it follows that

\[
\Sigma_{i1} = \frac{\partial \Sigma_i}{\partial \sigma_b} = 2\sigma_b Z_i Z_i',
\]

\[
\Sigma_{i2} = \frac{\partial \Sigma_i}{\partial \sigma} = 2\sigma I_i.
\]

Then the information matrix for all \( n \) units is

\[
\mathcal{I}(\theta) = \sum_{i=1}^{n} \mathcal{I}_i(\theta).
\]
4.8.2 Derivation of $t_p$ in Section 4.3.2

Suppose that a failure occurs at the first point in time where $D > D_f$. Let $F$ denote the CDF of the random variable $T$. Using (4.3),

$$p = 1 - \Phi_{\text{nor}} \left( \frac{D_f - \beta \tau_p - \gamma \tau_p x}{\sigma_b \tau_p} \right)$$

$$\zeta = \frac{D_f - \beta \tau_p - \gamma \tau_p x}{\sigma_b \tau_p}$$

$$D_f = \tau_p (\zeta \sigma_b + \beta + \gamma x)$$

$$\tau_p = \frac{D_f}{\zeta \sigma_b + \beta + \gamma x}$$

where $\zeta = \Phi_{\text{nor}}^{-1} (1 - p)$. The derivation is similar when the failure definition is $D \leq D_f$, using

$$p = \Phi_{\text{nor}} \left( \frac{D_f - \beta \tau_p - \gamma \tau_p x}{\sigma_b \tau_p} \right).$$

4.9 References


CHAPTER 5. General Conclusions

5.1 General Discussion

Repeated measures degradation and accelerated degradation tests are extremely useful when many measurements can be made with time for assessing the reliability of an item, especially when few or no failures are expected during the life of a test. Planning good repeated measures degradation test (RMDTs) and RMADTs can be highly beneficial to industry. In Chapter 2, we described methods for planning RMDTs. We first introduced the assumed degradation distribution and in particular, a quantile of the degradation distribution. We then illustrate how the Fisher information matrix can be used for planning tests whose interest is in estimation of the degradation distribution quantile. We introduced simple test plans and test plans with differing measurement schedules and how these can be used to assess tradeoffs between the number of items to be tested and the measurement scheme for these items. We also discuss how simulation can be used for test planning purposes and compare this to the large-sample based methods. We then derived the failure-time distribution induced by the degradation model as well as the failure-time distribution quantile. We then illustrated how the methods presented earlier with the degradation distribution quantile can be applied to the failure-time distribution quantile.

In Chapter 3, we described methods for finding unconstrained and constrained optimum test plans for RMADTs for a specified test optimization criterion, in our case, we minimized the large-sample approximate variance of the maximum likelihood estimator of a failure-time distribution quantile. The General Equivalence theorem (GET) was used to verify optimality of the proposed test plans. Then we introduced a compromise test plan which is more robust to misspecification of unknown model inputs while still obtaining good statistical properties.
We then compared the different types of test plans using simulation and analytical methods. The generality of these methods was illustrated by applying them to three different real life examples.

In Chapter 4, we described Bayesian methods to planning RMADTs. We used a Bayesian criterion based on estimation of the failure-time distribution quantile at nominal use conditions for finding optimum test plans. The Bayesian criterion was simplified by using a large-sample approximation to the posterior distribution. Again, GET was used to numerically verify a proposed test plan was globally optimum. We also illustrated how prior information for estimation influences the selection of optimum test plans by varying the amount of information for the model parameters. We then introduced a Bayesian compromise test plan as an alternative to the Bayesian optimum plan that is more robust to misspecification of unknown model inputs while still obtaining good statistical properties.

5.2 Recommendation for Future Research

The methodology that has been presented in this dissertation is general enough that it can be extended to other important situations that suggest new areas for future research. Some of these situations are:

1. Extend to models with more complicated covariance structures such as autocorrelations which might be needed when one has smaller spacing between measurements.

2. In our motivating examples for RMADTS, the length of the tests were fixed according to a development schedule. It would be possible to extend our work to test planning where optimization is performed over both the accelerating variable and length of the test.

3. Extend to applications where there are multiple failure mechanisms (competing risk models).

4. Extend to models where acceleration effects the variance structure as well as the mean structure.