ABSTRACT

Title of Dissertation: ESTIMATING THE RELIABILITY OF A NEW CONSUMER PRODUCT USING USER SURVEY DATA AND RELIABILITY TEST DATA

Neda Shafiei, Doctor of Philosophy in Mechanical Engineering, 2022

Dissertation directed by: Professor Mohammad Modarres, Professor Jeffrey W. Herrmann, Department of Reliability Engineering

Because new products enter the market rapidly, estimating their reliability is challenging due to insufficient historical data. User survey data about similar devices (e.g., older versions of the new device) can be used as the prior information in a Bayesian analysis integrated with evidence in the form of product returns, reliability tests, and other reliability data sources to improve reliability estimation and test specification of the new product. User surveys are usually designed for purposes other than reliability estimation. Therefore, extracting reliability information from these surveys may be tricky or impossible. Even when possible, the extracted reliability information contains significant uncertainties.

This dissertation introduces the critical elements of a reliability-informed user survey and offers methods for collecting them. A generic and flexible mathematical approach is then proposed. This approach uses the survey and reliability test data of similar products, for example, an older

generation of the same product as prior knowledge. Then it combines them through a formal Bayesian analysis with the reliability test data to estimate the life distribution of the new product. The approach models continuous life distributions for products exposed to many damage-induced cycles. It proposes discrete life distribution models for products whose failures occur within several damaging cycles. The actual cycles for various applicable damaging stress profiles are converted into the equivalent (pseudo) cycles under a reference stress profile. When damage-induced cycles are estimated from user surveys, they may involve biases, as is the nature of most nontechnical users' responses. This bias is minimized using an approach based on the Kullback-Leibler divergence method. The survey data and other evidence from similar products are then combined with the test data of the new product to estimate the parameters of the reliability model of the new product.

The dissertation developed approaches to design reliability test specifications for a new product with unknown failure modes. The number of samples, stress levels, and the number of cycles for the accelerated life test are determined based on the manufacturer's requirements, including the desired warranty time, the desired reliability with some confidence level at the warranty time, and the maximum number of samples. The actual use conditions (i.e., actual stress profiles and usage cycles) are grouped using clustering techniques. The centers of clusters are then used to design frequency-accelerated or stress-accelerated reliability tests.

The application of the proposed reliability estimation approach and the test specification design approach is illustrated and used to validate the proposed algorithms using the simulated datasets for a hypothetical handheld electronic device with the failure mode of cracking caused by accidental drops.

The proposed approaches can adequately estimate the reliability model and design test specifications for a wide range of consumer products. These approaches require reliability data about an existing product that is similar to the new product, however.

ESTIMATING THE RELIABILITY OF A NEW CONSUMER PRODUCT USING USER SURVEY DATA AND RELIABILITY TEST DATA

by

Neda Shafiei

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Prof. Mohammad Modarres, Chair/Advisor
Prof. Jeffrey W. Herrmann, Co-advisor
Prof. Abhijit Dasgupta
Prof. Katrina M. Groth
Prof. Vasiliy Krivtsov

© Copyright by
Neda Shafiei
2022

Dedication

I dedicate this dissertation to my dear mother, Sadieh Parhizkar, my dear father, Mohammad Zahed Shafiei, and my loving husband, Hamid. They taught me the value of hard work and supported me throughout my graduate studies. Thank you from the bottom of my heart for being a source of motivation and encouraging me during this journey. We did this work together. I will never forget your kindness and appreciate all you have done for me.

Acknowledgments

I sincerely thank my advisor Prof. Mohammad Modarres and my co-advisor, Prof. Jeffrey W. Herrmann, for giving me the inspiration and guidance to accomplish this dissertation. This work could not have been done without your continuous support and kind advice. I am forever amazed at your deep knowledge, excellent ideas, patience, and great personality. Thank you for teaching me many helpful behavioral and technical lessons that I will carry throughout my life's journey.

I appreciate Amazon company's Lab 126 for funding this research and allowing me to expand my reliability engineering knowledge and familiarity with some of the real world's reliability-related industrial problems.

I want to thank Prof. Abhijit Dasgupta for supporting me throughout my Ph.D. study and being part of my advisory committee. Also, I would like to thank my other advisory committee members, Prof. Katrina M. Groth, Prof. Vasiliy Krivtsov, and the dean's representative, Prof. Mohamad Al-Sheikhly for their time and consideration.

Finally, thanks to my dear siblings Samira, Pedram, and Elham, and my best friends Mahsa, Behzad, Forough, Mani, Alireza, Hamed, Iman, Seda, Chien-Ming, and Nam Kyoung for their kindness, support, and encouragements through these years.

Table of Contents

Dedication ii
Acknowledgments iii
Table of Contents iv
List of Tables vi
List of Figures vii
List of Acronyms vii
Chapter 1: Chapter 1: Introduction 1
Chapter 2: Literature Review 8
Chapter 3: Designing a Reliability-Informed User Survey 22
3.1. Case Study 27
3.2. Conclusions 30
Chapter 4: Estimating the Reliability Model of a New Consumer Product Assuming a Continuous Life Distribution 31
4.1. Calculating the Equivalent Cycles of the Devices 33
4.2. Estimating the Bias about the Equivalent Cycles 36
4.3. Estimating the Parameters of the Reliability Model of the New Product ... 40
4.4. Case Study 42
4.4.1. The Simulated User Survey Dataset for the Similar Product 43
4.4.2. The Simulated Test Dataset of the Similar Product 49
4.4.3. The Simulated Test Dataset of the New Product 51
4.4.4. Reliability Analysis Using the Simulated User Survey and Test Datasets 51
4.4.5. Sensitivity and Uncertainty Analysis 55
4.4.6. Conclusions 62
Chapter 5: Estimating the Reliability Model of a New Consumer Product Assuming a Discrete Life Distribution 64
5.1. The Reliability Estimation Approach 68
5.2. Estimating The Parameters of The Reliability Model of The New Product 68
5.3. Case Study 71
5.3.1. The Simulated Datasets 71
5.3.2. Estimating the reliability model of the new product 74
5.4. Conclusion 83
Chapter 6: Designing Test Specification of a New Consumer Product 85
6.1. Summarize Use Stress Profiles 87
6.1.1. K-Means Clustering 87
6.1.2. Gaussian Mixture Model 88
6.1.3. SI-Cycle Graph 88
6.2. Table of Possible Stress Profiles and S.I. values 89
6.3. Design Test Specifications 89
6.3.1. Design Test Specifications for a Frequency-Accelerated Reliability Test 90
6.3.2. Design Test Specifications for a Stress-Accelerated Reliability Test 92
6.4. Case Study 94

6.4.1. Summarized Use Conditions 95
6.4.2. Design Test Specifications for a Frequency-Accelerated Reliability Test 104
.6.4.3 Design Test Specifications for a Stress-Accelerated Reliability Test .. 105
6.5. Conclusions 106
Chapter 7: Summary and Conclusions 108
7.1. Summary of contributions 108
7.2. Conclusions 109
7.3. Future Works 111
Appendix A 114
Glossary 126
Bibliography 127

List of Tables

Table 4-1 Scores of the stress adjustors 45
Table 6-1 Table of all possible stress profiles and their SI values........ 102
Table 6-2 The grouped stress profiles. 103
Table 6-3 The test conditions of the stress-accelerated test. 106

List of Figures

List of Acronyms

SI	Stress index
KL	Kullback_Leibler
MLE	Maximum likelihood estimate
KDE	Kernel density estimate
MCS	Monte Carlo Simulation
BIC	Bayesian information criterion
AIC	Akaike information criterion

Chapter 1: Introduction

This chapter briefly explains the motivation, objectives, research approach, conclusions, and dissertation outline. A road graph also shows different elements of the dissertation and their relationship.

1.1.Motivation

Assessing the reliability of a new consumer product is challenging because the reliability depends on numerous use conditions that are hard to predict and replicate in reliability laboratories. As a result, laboratory test data usually come from a narrower spectrum of use conditions than the actual use conditions and may not be fully representative of the product. Therefore, a reliability estimate solely based on laboratory test data may be inaccurate and have significant uncertainties. To some extent, using field reliability data of a similar product (such as the older generation of the new product) as prior information in Bayesian analysis can reduce the uncertainties.

The primary sources of field reliability data are warranty data and user survey data. Collecting warranty data is costly and time consuming. Also, warranty data represents the subpopulation of the failed or returned devices, which might have a smaller portion of right-censored units than the entire population and thus underestimate the reliability. Moreover, warranty data is not always available or may not contain the applicable use conditions. Even if the warranty data and the applicable use conditions are available, they cover a narrow period (i.e., the warranty time) and do not contain information about the entire useful life of the devices.

In contrast, collecting user survey data is quick and cost-effective. Also, user survey data can better than the warranty data represent the entire population of the devices. Moreover, a user survey can be designed to collect the applicable use conditions over the entire lifetime of the devices.

Designing such a user survey and using user survey data for estimating the reliability model of a product has not been thoroughly investigated in the past. This issue constitutes the first motivation of the present dissertation.

Consumer products are used under various use conditions. Each use condition partially damages the device and reduces its reliability. Previous studies measured accumulated damage in reliability tests and proposed empirical models that fitted the collected test data [1, 2]. Empirically-based physics of failure (damage-based) models are not practical for estimating reliability using user survey data. This is because users usually cannot adequately express and assess the true magnitude of damage as some damages are invisible or hard to approximate. For instance, users cannot identify or estimate the amount of damage (e.g., corrosion) on many everyday devices such as cell phones or dishwashers. However, users most likely know the conditions (i.e., stress events) that caused damage (e.g., corrosion). For instance, users know their geographical location, the number of times they spilled a liquid on their cell phones, the number of times they soaked their cell phones in a liquid, or the number of times they used their dishwashers in a typical month. Therefore, a stress-based model is needed to estimate a product's reliability using user survey data. A novel physics of failure model based on surveyed applied stresses can be developed considering all

applicable use conditions of a device. This issue constitutes the second motivation of the present dissertation.

1.2. Objectives

This dissertation describes research that supports three objectives as listed below.

1. The first objective is to design a reliability-informed user survey that collects essential data for estimating a consumer product's reliability model.
2. The second objective is to develop a stress-based reliability model that (1) considers all actual use conditions, (2) works with different reliability data sources such as user surveys and reliability tests, (3) removes bias from user survey responses, and (4) uses multiple and diverse data sources to reduce the uncertainties of reliability.
3. The third objective is to design test specifications for a new product with unknown failure modes using user survey data of a similar product.

1.3. Research Approach

Designing a reliability-informed user survey requires identifying what should be collected (i.e., critical elements) and how. This dissertation determines the critical elements based on the definition of reliability. A broad literature review is then conducted to understand how to collect the critical elements.

An approach is developed that estimates the new product's reliability model using the user survey data about the similar product and reliability test data about the similar and new product. The approach assumes products with many cycles to failure. Therefore, a continuous lifetime distribution is considered. Also, it is assumed that the

user survey data about the similar product is biased because the user responses are uncertain. The bias in user responses is measured by (1) calculating the KL-divergence distance between the lifetime distributions of the surveyed and tested devices from the similar product and (2) minimizing the distance using the gradient descent algorithm. Only 30% of the surveyed and tested similar devices are used to estimate the bias value. The remaining 70% is saved to build a prior joint distribution for the parameters of the new product's reliability model. The lifetimes of the remaining 70% of the surveyed devices are multiplied by the estimated bias value to remove bias. Then, a 3-step sequential Bayesian analysis estimates the reliability model's parameters. In the first step, the primary joint posterior distribution of the parameters is estimated using the Bayesian analysis and kernel density estimation (KDE) method. This step assumes weakly informative prior distributions and uses the bias-removed user survey data as the likelihood data. In the second step, the primary joint posterior distribution is used as the prior distribution, and the remaining 70% of the test data about the similar product builds the likelihood. Then, the intermediate joint posterior distribution is estimated. In the third step, the intermediate joint posterior distribution and the test data about the new product are used as the prior distribution and likelihood data. The final joint posterior distribution is then estimated. Finally, the mean reliability model and its uncertainty region are calculated.

A discrete life distribution estimates reliability when the new product has a few damage cycles to failure. Some discrete lifetime distributions may have a summation term that its upper bound is unknown. The unknown value is calculated using the MLE and gradient descent algorithm. Then, the KL-divergence method removes bias from

user responses, and the 3-step sequential Bayesian analysis estimates the reliability model’s parameters.

For designing reliability test specifications for the new product, first, the user survey data about the similar product is clustered through a clustering method (e.g., K-means, Gaussian mixture model (GMM), and SI-cycle graph). Then, a stress-accelerated and a frequency-accelerated test are designed. The frequency-accelerated test directly uses the clustered use conditions (i.e., clustered stresses and usage times). In contrast, the stress-accelerated test uses a stress-life model to increase the clustered stresses and decrease the clustered usage times such that the amount of damage at each cluster remains similar to the original clustered use conditions. The number of samples for both tests is determined using a binomial distribution and based on the manufacturer's constraints, including the maximum test duration, the maximum number of failures, and a reliability target with a specified confidence level at the desired time. The details of the above approaches will be discussed in Chapters 3-6.

1.4. Research Summary

A well-designed user survey reflects actual use conditions in the field and provides valuable information for supplementing a new product’s life testing and reliability analysis. Using a similar product's user survey data helps to establish appropriate prior information in a Bayesian reliability estimation. The proposed reliability estimation approach makes an accurate assessment of the reliability model of the new product due to each failure mode.

This dissertation designs a reliability estimation approach that applies to products with a few or many damage cycles to failure. It is shown that a discrete lifetime

distribution would perform better than a continuous distribution in estimating the reliability model of a product experiencing a failure mode only after a few damage cycles.

The user survey data and information about a similar product is a rich source of reliability data for designing test specifications of a new product. The designed test reveals the new product's most common failure modes. Taking mitigation actions against the observed failure modes improves the new product's reliability. The proposed methods in this dissertation are generic and can be applied to a wide range of consumer products such as electronics, appliances, automotive, and handheld devices.

1.5. Outline of the Dissertation

This dissertation divides into seven chapters. Chapter 1 (the current chapter) introduces the dissertation's overall structure and discusses the motivations, objectives, and research approaches. Chapter 2 reviews the previous survey studies and reliability estimation methods. Chapter 3 introduces the critical elements of a reliability-informed user survey and provides recommendations for designing the survey. Chapter 4 explains the reliability estimation approach applied to products with many damage cycles to failure. The approach considers a continuous lifetime distribution. The application of the approach is illustrated using simulated survey and test datasets. Chapter 5 proposes a method for estimating the reliability model of a new product that experiences only a few damage cycles to failure. Chapter 6 designs test specifications for a new product with unknown failure modes using user survey data of a similar product. Finally, Chapter 7 concludes the research and recommends possible future directions for this work. The outline of the dissertation is shown in Figure 1.1.

Figure 1.1 Outline of the dissertation.

Chapter 2: Literature Review

This chapter reviews relevant literature to the topic of this dissertation, including (1) user survey literature, (2) methods for estimating the reliability model of a product having a small or large number of damage cycles (i.e., the usage cycles that cause damage), (3) test planning studies. The chapter helps to understand the state of art reliability estimation approaches and the research gap.

2.2. User Survey Studies

A broad range of studies has been conducted using user survey research summarized in four groups. The first group focuses on data collection methods (e.g., email surveys, postal questionnaires, face-to-face interviews, and telephonic interviews). Taylor-Powell and Marshall [3] suggested considering the effect of the anticipated response rate, ability to follow up, speed of data collection, and availability of sampling frame in deciding the mode of delivery. Kelley et al. [4] investigated the advantages and disadvantages of postal, face-to-face, and telephone interviews. They found that the postal response rate was low (about 20%), depending on the content and length of the survey. Therefore, a large sample is needed for the postal survey. The

face-to-face interview had a higher response rate than the postal survey. However, it was costly and time-consuming. The telephone interview was quicker and cheaper than the face-to-face survey but had a higher level of refusal. Nayak and Narayan [5] investigated the strengths and weaknesses of online surveys. They found that preparing the questionnaire, collecting data, storing data, and visualizing data in online surveys were more straightforward than the other methods. Besides, running an online survey was cheaper and quicker. However, sampling, maintenance of confidentially, and ethical issues were problematic when the survey was online. Pratama [6] created an online-based user data collection application for mobile devices using the bootstrap platform. The application provided a solution for companies, researchers, and individuals who wanted to conduct a survey. Users created an account in the application and responded to the questionnaire. The user data was stored on a hosting server.

The second group of survey studies discusses designing a user survey. This group investigates how to ask questions in a survey and how to arrange them. Some common suggestions given in these studies are as follows [4, 7, 8, 9, 10, 3]:

1. Identify the purpose of the survey and the information that needs to be collected.
2. Use one or a few warm-up questions to prepare the respondents for the following questions in the survey.
3. Avoid unnecessary, lengthy, ambiguous, technical, and threatening questions.
4. Use shared vocabulary in the questions.
5. Ask one question at a time.

The third group of survey studies investigates the methods of evaluating survey validity and reliability [11, 12]. Validity refers to the extent to which a survey measures

what it is intended to measure. Deniz and Alsaffar [13] gave the same survey to two groups of respondents, one group known to have higher knowledge about the concept surveyed. They used the correlation between the two groups of responses as a metric that assessed the validity of the survey [13]. A survey is reliable if it produces the same result on repeated trials. Roberta and Twycross [14] evaluated survey reliability using test reliability, alternate-form reliability, and internal consistency reliability. With the test-retest method, the same questionnaire was carried out on the same respondents at different times, and the obtained scores were compared. A high correlation between the scores showed that the questionnaire was reliable. The alternate-form reliability measured the agreement between two or more research instruments, such as two differently worded surveys that measured the same attribute or construct. The internal consistency reliability measured the reliability of a survey by asking questions about the same thing in different ways.

The fourth group of survey studies focuses on uncertainty quantification and bias detection of the responses [15]. The uncertainties are either aleatory or epistemic. The aleatory uncertainties which are due to inherent variability or randomness in a system are irremovable and irreducible. The epistemic uncertainties are due to a lack of human knowledge and could be reduced by obtaining more information about systems [16, 17]. Survey data has epistemic uncertainties that arise from different sources. Studies have grouped the epistemic uncertainties in various ways [18, 19]. In one way, uncertainties are grouped into sampling errors and non-sampling errors. Sampling error results from using a sample from a population rather than analyzing the entire population. The sampling error will be reduced if more data is used. Non-sampling

errors arise from the design, data collection, and processing methods. Unlike the sampling errors, the uncertainties stemming from the non-sampling error increase with the sample size. The non-sampling errors are the specification, coverage, nonresponse, unit-level measurements, and processing errors [20, 21, 19].

Specification errors appear when survey questions cannot or do not perfectly measure the concept they plan to measure. The main reason is that people have different subjective views that cause miscommunication between the interviewee and the interviewer. The Multitrait-multimethod (MTMM) developed by Campbell and Fiske [22] can be used to understand if the survey questions could measure the concept they plan to measure. MTMM repeatedly measures different traits (constructs) using different methods and builds the MTMM matrix using the measurements of those methods. The MTMM matrix is used to understand the quality of the measures. It is expected to see a high correlation between similar traits and a weak correlation between dissimilar traits. If this expectation is not satisfied, the questionnaire will fail to measure what it is supposed to measure. In such a case, the sources of invalidity need to be recognized, and appropriate steps must be taken to improve the data quality [23, 24].

Coverage error appears when the samples are clustered and do not cover the entire population. To reduce coverage errors, the interviewees must be diverse (i.e., be selected from different age groups, genders, and geographical locations) [21].

Non-sampling errors are divided into total non-sampling errors and item non-sampling errors. The total non-sampling errors occur when the customers fail to respond to all questions, while the item non-sampling errors happen when the customers do not respond to several questions. Non-response errors can be decreased

through non-theoretical and theoretical approaches. Response rate could be non-theoretically increased using radio or television announcements, preparing the questionnaire in some common languages, and giving incentive material or financial [21, 19]. It is suggested to provide the incentives only if the customers fail to respond to the questionnaire after two- or three-times direct requests. The total non-response error could be theoretically decreased using weighting adjustment methods. Through the weighting adjustment methods, the weights of the respondents are increased to compensate for the non-respondents. Weighting adjustments overall, weighting class adjustment, population weighting adjustment, and sample weighting adjustment are some versions of weighting adjustment methods. The item non-sampling error could be decreased using imputation. Imputation is the process of assigning values to miss responses using the auxiliary information that is obtained from the sample units. Mean imputation, random imputation, hot-deck imputation, and regression imputation are some examples of imputation methods [25].

Unit-level measurement error results from possible differences between what is measured and the actual values for the sample units. This issue happens when interviewees want to please or hide information from the interviewees. In addition, when questionnaires are poorly designed, the interviewees misunderstand the questions and provide incorrect answers. The interviewees should be assured that telling the truth does not penalize them [19].

Processing errors are recording-, checking-, coding-, and survey data preparing-related errors. Making correct design decisions, trying different analytical methods, and selecting the highest accuracy method can reduce processing errors. When the non-

sampling errors are addressed, more data collection can further reduce the input data uncertainties [19].

Garthwaite et al. [23] used elicitation techniques to quantify uncertainties in surveyed people's responses. Elicitation is a process by which a person's knowledge and beliefs about one or several uncertain quantities are formulated into a (joint) probability distribution of the quantities. Elicitation is used to specify the prior distribution for the Bayesian analysis, which is then combined with the likelihood distribution (obtained from other data sources) to derive the posterior distribution or formulate uncertainties about inputs to a mathematical model or a decision problem.

Another topic that has gained attention in the survey studies is identifying and reducing inaccurate responses (biases). Studies show that people may rely on heuristics such as representativeness, availability, adjustment-and-anchoring when responding to the survey questions [23, 26]. These heuristics may create inaccurate responses in the survey data. Representativeness occurs when people answer the questions in the form of "what is the probability that objects A belongs to class B?" in answering this type of question, people typically evaluate the probability by the degree to which A is representative of B. At the same time, this approach leads to bias because similarity is not influenced by the factors that should affect judgments of probability [23, 26]. The availability issue occurs when people assess an event's frequency (or likelihood) by relying on the number of instances or occurrences that can be recalled. In some situations, people may be asked to estimate by adjusting a given initial value. Studies show that the adjustments are insufficient [23, 26].

A few studies used user surveys to collect product reliability information. Yang et al. [27] used a set of 22 checkboxes, open-answers, and ranking questions to gather information about the reliability issues of power electronic converters in the field, their most fragile components, the most applicable stresses, failure costs, and failure causes. To design the questions, they focused on the factors that affect power electronic system reliability, such as application category, operating range, stress level, duty time, and maintenance. For the checkbox questions, they allowed respondents to select more than one box and used the weighted average of the selected choices for analyzing the responses. Cochrane et al. [28] used a set of 55 questions to collect reliability information about the operation, maintenance, spare parts, availability, and performance of convertor stations worldwide. They, however, did not discuss the details of their survey design and data collection. Tollefson et al. [29] designed a user survey to collect information about the customer costs (monetary loss) associated with power interruptions, the number of failures experienced in the last two years, and the applicable stresses (e.g., the type of heating system (electrical, nonelectrical, or a combination)). They mailed the surveys to the users and, in total, got a 30% response rate. They used the collected data associated with money loss in industrial and commercial applications to quantify the average cost of electricity interruption from 2 seconds to 1 day.

In summary, the knowledge obtained from the previous survey studies is helpful but insufficient for designing and analyzing a reliability-informed user survey that can be used for estimating reliability and preparing test specifications. Assessing a product's reliability requires several types of essential data (called "critical elements")

to be collected in the user survey. This dissertation introduces the critical elements of a reliability-informed user survey.

2.2. Estimating Reliability Using a Parametric Distribution

Bayesian and MLE are the two most widely used methods for estimating reliability. The Bayesian approach has received increasing application and acceptance for estimating reliability when multiple data sources are available [30]. Pan [31] used the Bayesian method and a calibration factor to integrate field data and accelerated life test (ALT) data. Field data is the most reliable source of data to build the reliability models under the use condition. Collecting field data is often time-consuming. As a result, the existing field data is usually limited. A common approach to deal with this issue is to run a set of ALTs, build a stress-life model, and extrapolate the product's failure characteristics to the use level. The ALT, however, cannot perfectly simulate the accurate stress profile experienced by the units. Therefore, the stress-life model obtained from the ALT results is an approximation with many uncertainties. Pan [31] integrated field and ALT data to reduce the uncertainties. The proposed approach by Pan utilized a (known) calibration factor to bridge the reliability estimate from ALT to the actual reliability of the product in the field.

The Bayesian method is also used when abundant expert knowledge is available, but other reliability data sources (e.g., reliability test results or field data) are limited. Guo et al. [30] used a Bayesian modeling method (BMM) to estimate a system's reliability by integrating various sources of expert knowledge and reliability test data at subsystem and system levels. They used linear and geometric pooling methods to incorporate multiple experts' prior information. They used the BMM to update the

subsystem prior reliability distribution by integrating the system and subsystem prior reliability distributions. The BMM also evaluated the posterior reliability distribution by integrating the updated prior distribution and the reliability test data.

Wang et al. [32] used evolving, insufficient, and subjective datasets to predict system reliability. They integrated probability encoding methods with the Bayesian updating technique to elicit subjective data from users, model them with statistical distributions, and update the uncertainty distributions with evolving subjective datasets. However, they ignored the users' biases. They also did not consider the applicable stress profiles and used a binomial distribution to update the probability of failure. They illustrated their approach using conjugate distributions and did not address integrating dependent datasets through the Bayesian analysis.

Yuan et al. [33] combined historical degradation data and expert experience as the prior information and accelerated degradation test data as the observation in Bayesian analysis to estimate the reliability model of a product. Li et al. [1] developed a damage-based reliability estimation method that considered the effect of catastrophic and nonfatal failure. They used uncertainty analysis theory for multi-state systems and assumed that state transmission time is an uncertain variable. They estimated the reliability of a system by calculating the state probability and the states of the system. They verified the accuracy of their method using the Monte Carlo Simulation (MCS) method. The problem with the damage-based reliability methods is that users usually cannot adequately express the magnitude of damage as some damages are invisible or hard to assess accurately. Thus, the damage-based methods cannot be used for analyzing user survey data.

2..2.1. Application of Continuous and Discrete Distributions in Estimating the Reliability Model of a Product

Another topic that has gained attention in reliability engineering is the application of continuous and discrete distributions in estimating the reliability model of a product. A continuous distribution is applicable when the random variable representing life is continuous (e.g., expressed in calendar time), or the product undergoes many damage-accumulating usage cycles (e.g., hundreds of cycles of a device dropping on a hard floor). On the other hand, when the performance of a product is measured occasionally (e.g., every week or every month), or the product undergoes a few damage cycles before failing, a discrete distribution can more accurately estimate the life distribution or reliability model [34, 35].

The Weibull distribution is a well-known lifetime distribution that is widely used in reliability engineering for estimating the reliability model of a product. The Weibull distribution is available in continuous and discrete forms.

Two-parameters (2P) and 3-parameters (3P) Weibull distributions are the most common forms of continuous Weibull distribution. The continuous Weibull distribution for some certain parameter values is related to other distributions like exponential and Rayleigh distribution [36]. Some studies extended the continuous Weibull distribution. For instance, Dan et al. [37] proposed a 4-parameters continuous Weibull distribution and estimated the distribution of its parameters in a Bayesian analysis that assumed Gamma prior distributions. Wang and Zhou [38] proposed a 2P-

Weibull segmented model for failure modeling and used least-squares Hough transform algorithm to estimate its parameters.

Three types of discrete Weibull distributions have been proposed: (1) type I has a reliability function that mimics the reliability function of a continuous Weibull distribution, (2) type II has a hazard function that mimics the hazard function of a continuous Weibull distribution, and (3) the type III is more generic and does not follow any function of a continuous Weibull distribution [39].

Various extensions of discrete Weibull distributions have also been introduced, and different methods for estimating their parameters have been suggested. For instance, Jia et al. [40] proposed a discrete extended Weibull distribution and used MLE for estimating its parameters. Barbiero [34] proposed three methods, including the method of proportion, MLE, and the method of moments for estimating the parameters of a discrete Weibull type III distribution.

The above studies assumed that the product is used under a fixed stress level, and the number of use cycles (damage producing cycles) is equivalent to the number of times the stress is applied to the product. This dissertation argues that consumer products are used under varying stress levels, and the concept of equivalent cycles should be used to calculate the number of cumulatively damaging usage cycles under a reference stress condition. Moreover, neither of the above studies used user survey data to estimate a product's reliability model. This dissertation applies the concept of equivalent cycles and uses user survey data for estimating reliability.

2.2. Reliability Test Planning

The last topic that is reviewed in Chapter 2 is reliability test planning. Manufacturers assess the reliability of their products using laboratory tests to understand whether they meet or exceed their minimum reliability requirement. A reliability test follows specifications including sample size, test conditions (i.e., stress levels and their frequencies of occurrence), the analytical procedures for assessing reliability, and the acceptance criteria (e.g., numerical limits) that are used to evaluate the reliability of a product. The product is accepted for mass production if the sample's test results are within the acceptance criteria. Otherwise, a root cause analysis is needed to understand the reason for the poor test results [41, 42].

Yang [43] divided the reliability demonstration test methods into the bogey test (or zero-failure test), life test, and degradation test. In a bogey test, the sample size, test time, and level of stresses are predetermined; and the target reliability is achieved if no units fail during the test. Yang [43] argued that the quantitative reliability could not be estimated if any samples failed in a bogey test. Life test methods include sequential life tests (e.g., step-stress test) and conventional life tests (e.g., failure terminated test). In a life test, the reliability estimate is based on the number of failures. A degradation test is performed on a product with performance characteristics that degrade over time, which leads to failure. Reliability can be estimated by measuring the performance characteristics at different times during testing.

Yang [43] developed a degradation model that modeled the product's performance characteristics with the Weibull distribution and estimated the model's parameters using Maximum Likelihood Estimation (MLE). He also estimated the test cost as a

function of test time, number of samples, and cost of measurement. He estimated the number of samples and the test time by minimizing the cost function.

Gerokostopoulos et al. [44] overviewed the methods for determining the sample size for a reliability test. These methods were divided into (1) methods that use the theory of confidence interval (known as the estimation approach) and (2) methods that control the type I and type II errors (known as the risk control approach).

Gerokostopoulos et al. [46] determined the sample size using (1) the confidence interval theory and (2) the risk control approach. In the first method, they assumed a Weibull distribution and calculated the ratio between the reliability model's upper and lower bound as a function of the number of samples. Then, they estimated the number of samples by assuming a numeral ratio. The second method determined the number of samples by controlling type I risk (i.e., the probability that the product meets the reliability requirement but does not pass the demonstration test) and type II risk (i.e., the probability that the product does not meet the reliability requirement but passes the demonstration test).

In an ALT, the stresses are selected by (1) determining the most critical failure modes of the product, (2) finding the relevant stresses to the failure modes (i.e., agents of failure), and (3) selecting several stress levels that higher than the use stress such that the maximum stress remains below the destructive limit of the product. This strategy allows fitting a reliability model more confidently to the failure data. Gerokostopoulos et al. [44] discussed four methods for determining stress levels: the two levels statistically optimum plan, the three levels best standard plan, the three levels best compromise plan, and the three levels best equal expected number failing plan.

Some studies used immature test specifications that were arbitrarily selected without considering the actual use conditions. Yang et al. [45] designed an accelerated degradation test for predicting the reliability of smart electricity meters (SEMs) using a mix of analytical methods and arbitrary choices for stress levels. Chang et al. [46] converted field use data of automotive headlamps to a laboratory bench test specification using the theory of fatigue damage equivalence between time domain and frequency domain data but arbitrarily selected the number of samples.

A product may have several potential failure modes. However, only a few of them may be observed by users. This is because some failure modes require specific stress levels or usage cycles that are out of the range of the actual use conditions. Addressing those failures is time-consuming, expensive, and unnecessary. Planning a reliability test based on the actual use conditions of similar products in use avoids addressing unnecessary failure modes. To the best of the author's knowledge, such a reliability test planning has not been developed yet. This dissertation plans the reliability test based on the actual use conditions of similar products collected through a user survey.

Chapter 3: Designing a Reliability-Informed User Survey[1]

This chapter provides a guideline for designing a reliability-informed user survey that collects the necessary information for estimating the reliability model of a consumer product. Then, a reliability-informed user survey is designed for a handheld electronic product with the failure mode of "cracking" caused by accidental drops.

One who designs a reliability-informed user survey should ensure that the survey collects the critical elements needed for estimating the reliability model. The critical elements are determined based on the definition of reliability which is the probability that a product performs its intended function adequately for a specified time under a specified use condition [47]. Therefore, five critical elements should be collected to estimate reliability, as shown in Figure 3.1.

The first element (i.e., usage time) is how long or how much the product has been used in calendar time or the number of cycles. The second element describes the actual stresses that the product has undergone. Stress is referred to a usage profile that causes damage to the device. The applied stresses could be mechanical (e.g., drop, twist,

[1] This chapter is a modified version of the paper about designing a reliability-informed user survey [69]: N. Shafiei, J. W. Herrmann, A. Krive, G. Sethi and M. Modarres, "Designing Reliability-Informed User surveys," in *European Safety and Reliability (ESREL)*, Angres, 2021.

bend), electrical (e.g., battery overcharging or discharging), thermal (e.g., battery overheating), and chemical (e.g., corrosive attacks).

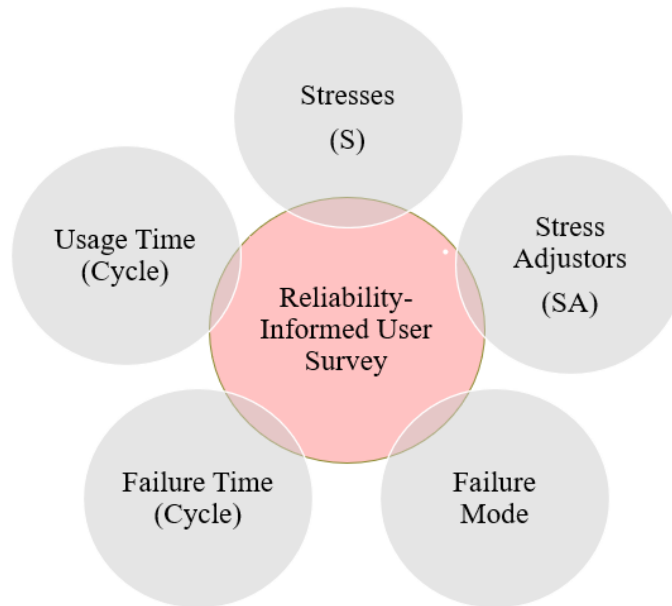


Figure 3.1 The critical elements of a reliability-informed user survey.

The third element describes stress adjustment factors which are the conditions or activities that increase or decrease the impact and damage from the applied stress or enhance stress absorption. For instance, when "drop" is the applicable stress of a device, the surface type on which the device is dropped, the human activity during a drop, and the height of drop are possible stress adjustors. The environmental conditions (e.g., temperature and humidity) are also stress adjustments. The fourth element describes the failure modes that have been observed. The fifth element describes when failures (if any) occurred. A product may have several failure modes and, consequently, several failure times.

The details of these five critical elements depend upon the product and the reliability analysis that will be performed. For instance, if the product is used continuously in a

well-understood environment (i.e., under a stress-controlled condition) until it suffers a catastrophic failure, the reliability analysis may be considered a simple time-to-failure model. Then, the fifth element should describe the total time the product operated until it failed. If no failure has been observed, the first critical element is used as the suspension time, and the device's failure mode is considered right-censored during data analysis. The other elements are not as essential because they should not vary.

Suppose the product is used occasionally in various circumstances and has numerous failure modes. In that case, however, the reliability analysis may consider multiple models for each failure mode. Each one may consider a Cox hazard model or another approach that considers multiple stress factors that influence the time to failure. In this case, many more details about the critical elements are required.

Regarding the first element, the survey may ask about the calendar time or the number of cycles. If the survey asks about the calendar time, it should ask for the time the product has been in use (not merely owned). If the survey asks about the number of cycles, it may be helpful to ask for two values: the time that the product has been in use and the frequency at which the product has been used during that time (e.g., the number of times per day, week, or month that the product is used).

It may be necessary to ask about combinations of the second and third elements (e.g., drops from different heights on different surfaces during different activities). Suppose the reliability analysis requires data about how many times different stresses and stress adjustors have occurred. In that case, it may be helpful to ask about the frequency of such events and then multiply it by the usage time (from the first element).

An appropriate survey should ask the users to build their applicable stress profile(s) (i.e., a combination of stresses and stress adjustors) rather than asking about the relevant stresses and stress adjustors in separate questions. This is desired because a device might have experienced several stress profiles. If those profiles' stresses and stress adjustors are collected using individual questions, their relationship will be lost (i.e., it will not be clear which sets of stress adjustors are relevant).

Although the process of designing the actual survey is outside of the scope of this study, two key questions are needed to collect the stress profiles and their frequency of use. The first question should request information about the use conditions (stress profiles). Offering predefined choices for the applicable stresses and stress adjustors simplifies the equation and let the user select relevant stresses and stress adjustors. For instance, in an impact-prone handheld device where “drop” is the applicable stress, the following simple question collects the applicable stress profile(s): “Please select the conditions under which you dropped your device in the last month.” The survey can offer predefined stress adjustors and multiple choices for each stress adjustor for this question, such as the following:

- Height of drop: (a) knee height, (b) waist height, (c) chest height, (d) head or higher height.
- Surface type: (a) soft surface (e.g., grass, thick carpet), (b) semi-soft surface (e.g., thin carpet, workout mats), (c) semi-hard surface (e.g., laminate, hardwood), (d) hard surface (e.g., asphalt, concrete).
- Activity type: (a) benign (e.g., standing, walking), (b) harsh (e.g., running, exercising, playing).

Then the user should select one item of each stress adjustor at a time. The approach explained above avoids using failure and reliability engineering jargon like stress adjustors and stress profiles in the body of the question.

The next class of questions should be designed to collect the frequency of occurrence of the stress profile. For instance, “please select the number of times the above condition occurred in the last month.” Again, the user may select the relevant number among several predefined ranges or enter a number in response. The two questions may be repeated to collect more stress profiles.

For the fourth element, the survey should precisely define the failure modes as the manner of failure. The loss of certain functions, cracks, scores, and scratches are examples of failure modes. The failure modes should be those that the product’s user can observe. Selection of failure modes follows by questions such as: “After using your device at the above condition, have you observed any problem in your device?” Providing multiple choices (i.e., some probable failure modes) for this question should help users understand the meaning of the “problem” and remember the observed failure modes.

For collecting the failure times, the survey might ask, “After how long (e.g., months) of use the device did you first observe this problem?”. Regarding the usage time, the survey might ask for two values in separate questions: (1) the calendar time (e.g., number of months) that the product has been in use and (2) the frequency at which a common stress profile is applied to the device (e.g., the duration or the number of times that a stress profile occurs in a typical month). The usage time is then calculated by multiplying the ownership times (in months) by the frequencies (in terms of 1/month).

Although the fifth element will usually be an essential factor, a difficulty can arise in cases where a failure does not prevent one from using the product. For example, one can use a phone or tablet even if the screen has some cracks. If the reliability analysis requires the time until the first crack, the survey should ask for this explicitly; the time the product has been in use will not suffice. If the product has multiple failure modes, the survey should ask about the times each failure mode occurred. It is also essential to ask which failure modes have not yet been observed if this censored data is needed for the reliability analysis.

Although it is easier for the person completing the survey to answer a question if the available responses are broad ranges (e.g., “1 to 5 times every day” or “2 to 3 months”), such imprecise data yield imprecise results from the reliability analysis. Thus, one would prefer to ask for a specific number (e.g., “How many times a day do you use this product?”). A challenge with that question is that the answer may vary over time (e.g., the product is used five times on some days and only once on other days). If that is sufficient, one can modify the question to ask about the average.

3.1. Case Study

Below is an example of a user survey that collects the five critical elements for estimating the reliability model of an electronic product that the user accidentally drops.

1. For how many months have you used your device?
2. Have you ever observed any problem with your device?

   Scratch   Crack   Battery does not hold the charge   Battery caught fire

Speakers malfunction I have not observed any problem Others (please specify)

3. After how many months of using the device did you observe this problem first?

4. Please select the condition where you drop your device in a typical month. It includes (a) the height of the drop, (b) the type of surface that the product hits when it is dropped, and (c) the type of activity that you are performing when the product is dropped.

(a) Height:

- Knee height
- Waist height
- Chest height
- Head or higher height

(b) Surface:

- Soft (e.g., grass, thick carpet)
- Semi-soft (e.g., thin carpet, exercising flooring)
- Semi-hard (e.g., laminate, hardwood)
- Hard (e.g., asphalt, concrete)

(c) Activity:

- Benign (e.g., inserting or removing the device from the case or ears, standing, walking)

- Harsh (e.g., running, exercising, playing)

5. How many times under the above condition do you drop your device in a typical month?

   More than once a day   once a day   1-6 times a week

   1-3 times in a month   less than once a month

6. Do you have another condition where you drop your device in a typical month? Yes   No

3.2. Conclusions

A well-designed user survey can provide valuable data for conducting a reliability analysis of a new product based on how users use the product in the field without waiting for returned products. Chapter 3 discussed the challenges of designing a reliability-informed user survey. A fundamental principle is that the survey should collect the data needed for reliability analysis. Because different products will need different types of reliability analysis, the user survey should be designed with the analysis in mind (instead of adopting a reliability analysis technique to deal with the survey data).

Five critical elements that a survey should collect were identified. These elements are the length of ownership time, failure modes, frequency and size of normal and accidental user-applied stresses, stress adjustors (i.e., the conditions or activities that increase or decrease the stress magnitude or absorption), and failure time/cycles.

A well-designed reliability-informed user survey requires collaboration among multiple groups in a product development organization, including reliability engineers, marketing analysts, and survey design experts. The user surveys containing information about the critical elements could be used to estimate the life distribution of products.

Chapter 4: Estimating the Reliability Model of a New Consumer Product Assuming a Continuous Life Distribution[2]

This chapter develops a novel mathematical approach for estimating the reliability model of a new product with a large number of damage cycles to failure using the combined user survey data and reliability test data of a similar product, and the reliability test data of the new product. The application of the approach is demonstrated using a simulated dataset for an electronic device with the failure mode of cracking caused by accidental drops.

Determining the reliability of a consumer product using user survey data is not straightforward. A product may experience various use-stress profiles. The product's damage depends on the applicable stress profile(s). Suppose only the calendar time (or

[2] This is a reproduction of the papers about estimating the reliability model of a consumer product using user survey data and reliability test data [62, 54]:

N. Shafiei, J. W. Herrmann, A. Krive and M. Modarres, "Estimating the Reliability of Consumer Electronics from User Survey Data and Test Data," in *Reliability and Maintainability Symposium (RAMS)*, Tuscon, USA, 2022.

N. Shafiei, J. W. Herrman, A. Krive, M. Nikiforov, G. Sethi and M. Modarres, "Analyzing Product Reliability Using User Survey Data and Reliability Test Data," *IEEE Transactions on Reliability,* 2022.

frequency of use in terms of (mean) number of cycles) is used for reliability assessment (i.e., a simple time-to-failure analysis). In that case, the effect of different stress profiles will be disregarded. Some studies have proposed damage-based approaches that estimated reliability by accumulating damage over the product's usage time [1, 2]. The problem with the damage-based methods is that users usually cannot adequately express the magnitude of damage as some damages are invisible or hard to assess accurately. For instance, users cannot estimate the crack length or the amount of corrosion on an internal component of their electronic devices. Instead of the magnitude of damage, the users can easily express the applied stress profiles. Although users may not know about the extent of the damage, they are likely to tell the approximate frequency and intensity of the stresses that caused the damage. For example, users can estimate how many times or the conditions under which their devices were dropped, overheated, underheated, or soaked over a recent time. This chapter proposes a stress-based reliability estimation approach as the applicable stress profiles can be collected from user surveys. The proposed approach is generic and can estimate the reliability of a wide range of consumer products such as consumer electronics, appliances, handheld devices, and automobile components.

The approach is divided into three steps:

1. Calculating the equivalent cycles of the devices.
2. Estimating the bias about the equivalent cycles.
3. Estimating the parameters of the reliability model.

Details about these steps are given in the following.

4.1. Calculating the Equivalent Cycles of the Devices

The concept of equivalent cycles is introduced to consider the effect of actual stress profiles in the reliability assessment. In their lifetimes, devices experience different stress profiles and the number (or mean number) of cycles or times the stress profiles are applied. The number of times a stress profile is repeated in a device's lifetime is referred to as a stress block that causes partial but cumulative damage that reduces the device's life. The number of times that the application of a reference stress block produces the same amount of damage as the actual stress blocks experienced is referred to as the equivalent (pseudo) time. Figure 4.1 schematically shows the concept of equivalent cycles. Assume a user drops his handheld electronic device various times under various stress profiles (that means the user drops the device under various stress blocks). The equivalent cycles are the number of cycles under the reference stress profile that causes the same damage as the actual cycles under actual stress profiles (i.e., actual stress blocks).

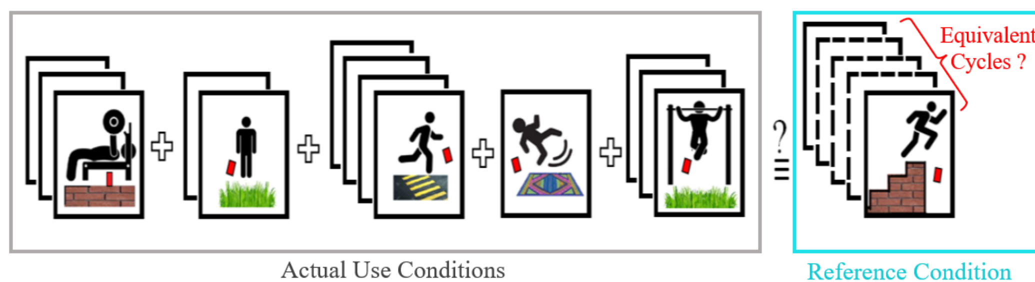


Figure 4.1. Actual cycles vs. equivalent cycles.

The process of calculating the equivalent cycles is shown in Figure 4.2. Converting the actual cycles into the equivalent cycles through the stress-life model of the product requires quantitative stresses and stress adjustors. Nevertheless, surveys usually collect

subjective data. Therefore, it is proposed to linearly or non-linearly score each actual usage cycle's stresses, s, and stress adjustors, s_a, on a numeric scale in the range, such as (0, 1) or (0, 100). The sensitivity of the reliability estimates to these scores will be discussed later in Chapter 4.6.

The s and s_a values should then be combined through mathematical forms such as linear, log-linear, multiplicative, additive, or quadratic stress-index (SI) models to find a unique SI value for each usage cycle, as shown in Eq. (4-1) to Eq. (4-6).

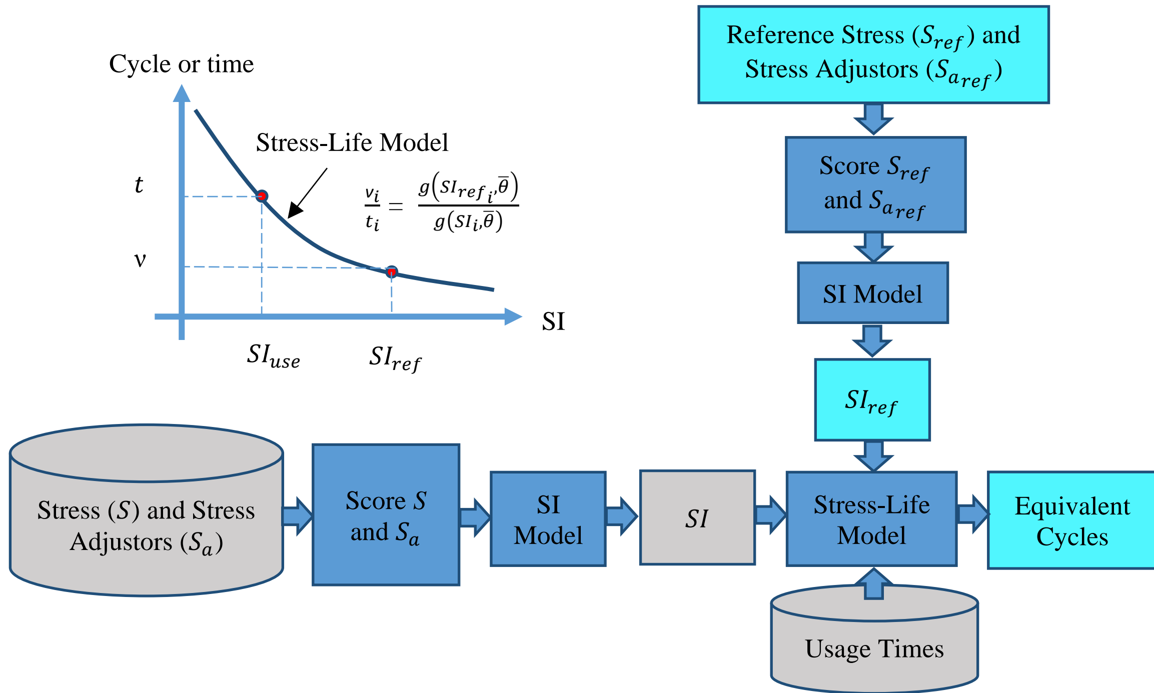


Figure 4.2. The process of calculating the equivalent cycles.

Let the S.I. function shown in Eq. (4-1) relates the dependent variable, SI_i, representing the stress index of the i-th product to the independent variables vector of applicable stresses, $\bar{s}_i$, the vector of applicable stress adjustors, $\bar{s}_{a_i}$, and the vector of stress weights, $\overline{w}_i$. For the additive S.I. model, the sum of the stresses and stress

adjustors’ weights is constant as shown in Eq (4-2), where c is a preselected value (e.g., 1 or 100), J and K are the numbers of stresses and stress adjustors in the model, respectively.

Equation (4-3) is the additive model that combines the stresses and stress adjustors. Alternatively, log-linear Eq. (4-4), multiplicative Eq. (4-5), or quadratic Eq. (4-6) models may be used. In Eq. (4-3) to Eq. (4-6), s_{ij} is the j-th stress value of the i-th device and w_{ij} is its corresponding weight, and $s_{a_{ik}}$ is the k-th stress adjustor value of the i-th device and $w_{a_{ik}}$ is its corresponding weight. A decision on the function's choice depends on the data's nature and can be made judgmentally or formally optimized. The sensitivity of the reliability model to this function is discussed in Chapter 4.6.

$$SI_i = f(\bar{s}_i, \bar{s}_{ai} | \bar{w}_i, \bar{w}_{a_i}) \quad (4\text{-}1)$$

$$\sum_{j=1}^{J} w_i + \sum_{k=1}^{K} w_{a_i} = \mathrm{c} \quad (4\text{-}2)$$

$$SI_i = \sum_{j=1}^{J} w_{ij} s_{ij} + \sum_{k=1}^{K} w_{a_{ik}} s_{a_{ik}} \quad (4\text{-}3)$$

$$SI_i = \sum_{j=1}^{J} w_{ij} log\ (s_{ij}) + \sum_{k=1}^{K} w_{a_{ik}} log\ (s_{a_{ik}}) \quad (4\text{-}4)$$

$$SI_i = \prod_{j=1}^{J} s_{ij} \times \prod_{k=1}^{K} s_{a_{ik}} \quad (4\text{-}5)$$

$$\mathrm{SI}_i = \sum_{j=1}^{J} w_{ij} {s_{ij}}^2 + \sum_{k=1}^{K} w_{a_{ik}} {s_{a_{ik}}}^2 \quad (4\text{-}6)$$

A stress-life model is then selected that converts the S.I. values into time (including cycles or the mean number of cycles). The general form of the stress-life model is shown in Eq. (4-7), where g is the stress-life function, SI_i is the stress index of the i-th product, and $\bar{\theta}$ is the vector of model parameters. Possible stress-life models are linear,

log-linear, exponential (e.g., Cox cumulative hazard model), and inverse power law (IPL). A combination of these models is also possible. Selecting a model depends on the nature of the data, the device, and the type of stresses applied.

$$t_i = g(SI_i, \bar{\theta}) \quad (4\text{-}7)$$

Afterward, the actual usage times under various stress profiles are converted to equivalent times under the reference stress profile. This is done by finding the ratio between the actual usage times and equivalent times. Thus, the equivalent times are obtained from Eq. (4-8). The calculated equivalent cycles at this step are parametric because they contain the unknown parameters of the stress-life model. These parameters are later estimated through our proposed approach.

$$v_i = t_i \frac{g\left(SI_{ref_i}, \bar{\theta}\right)}{g(SI_i, \bar{\theta})} \quad (4\text{-}8)$$

4.2. Estimating the Bias about the Equivalent Cycles

When the equivalent cycles of the surveyed devices are calculated, the results are biased (inaccurate) because users' responses to the applied stresses, stress adjustors, and the number of usage cycles are biased. To measure the amount of bias in user responses, the lifetime distribution of a similar product (e.g., an older version of the new product) is estimated the first time using its user survey data and the second time using its reliability test data. The reliability test data is assumed to represent true life because the test conditions are recorded in the laboratory. Thus, the difference between the two life distributions is because of the bias in the survey responses, and if the bias is removed, the distributions will become similar. It is assumed that if all equivalent

cycles in the life distribution of the surveyed devices are multiplied by ψ (i.e., the bias value), the shifted distribution will be close to the life distribution of the tested devices. This is schematically shown in Figure 4.3.

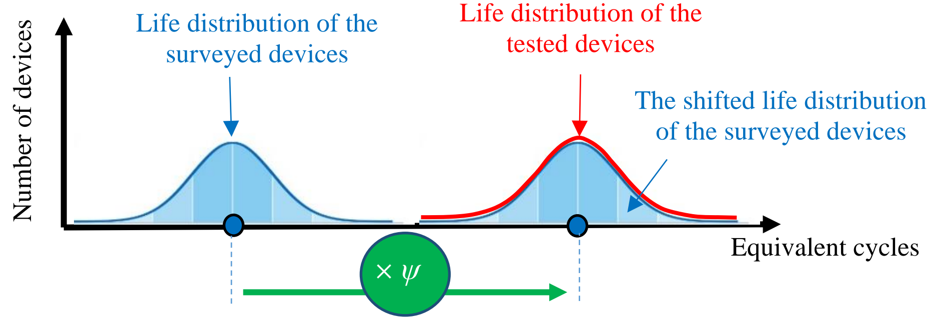


Figure 4.3. The process of removing bias from user responses.

Subchapter 4.2 proposes a mathematical approach to measure and reduce bias. First, the similar product's test data and survey data are randomly split into two parts (x % in part 1 and 100-x % in part 2), as shown in Figure 4.4. Then, the first part of the test data and the first part of the survey data are used to estimate the bias in the survey responses.

Since the selected stress-life model (i.e., $g(SI_i|\bar{\theta})$) in chapter 4.1 is uncertain, an error term (i.e., $\varepsilon(0|\sigma,\eta)$) in the form of a probability density function (PDF) with the zero mean and the standard deviation σ is added to the model. Therefore, the model output is the PDF of time-to-failure (i.e., life distribution). This is mathematically shown in Eq. (4-9), where $p(\nu, SI|\bar{\Omega})$ is the PDF of time, $\bar{\Omega}$ is the vector of time function parameters and error PDF parameters (i.e., $\bar{\Omega} = [\bar{\theta}, \sigma, \bar{\eta}, \bar{w}]$), $g(SI|\bar{\theta})$ is the stress-life model, and $\varepsilon(0|\sigma,\bar{\eta})$ is the error. To make the scatter about the origin, the error is changed to $\varepsilon(g(SI_{ref}|\bar{\theta})|\sigma,\eta)$ and the PDF of time is expressed by Eq. (4-10). The PDF has unknown parameters that should be estimated through the maximum

likelihood estimation (MLE) or Bayesian method. It is assumed that the PDF follows a known family of distributions such as Weibull, Normal, and exponential.

$$p(\nu, SI|\bar{\Omega}) = g(SI|\bar{\theta}) + \varepsilon(0|\sigma, \eta) \quad (4\text{-}9)$$

$$p(\nu, SI_{ref}|\bar{\Omega}) = \varepsilon(g(SI_{ref}|\bar{\theta})|\sigma, \eta) \quad (4\text{-}10)$$

Then, the likelihood function of the selected distribution is built, as shown in Eq. (4-11) and Eq. (4-12), where F_{ts} and F_{ss} are the number of failures of the tested and surveyed similar devices, C_{ts} and C_{ss} are the number of censoring among the tested and surveyed similar devices, $p(\nu_{ts_f}, SI_{ref}|\bar{\Omega}_{ts})$, and $p(\nu_{ss_f}, SI_{ref}|\bar{\Omega}_{ss})$ are the PDFs of the selected distribution for the tested and surveyed similar devices, $R(\nu_{ts_c}, SI_{ref}|\bar{\Omega}_{ts})$, and $R(\nu_{ss_c}, SI_{ref}|\bar{\Omega}_{ss})$ are their corresponding reliability functions, $\bar{\Omega}_{ts}$ and $\bar{\Omega}_{ss}$ are the vectors of time function parameters and error PDF parameters (i.e., $\bar{\Omega}_{ts} = [\bar{\theta}_{ts}, \sigma_{ts}, \bar{\eta}_{ts}, \bar{w}_{ts}]$ and $\bar{\Omega}_s = [\bar{\theta}_{ss}, \sigma_{ss}, \bar{\eta}_{ss}, \bar{w}_{ss}]$]), ν_{ts_f} and ν_{ss_f} are the equivalent failure times of the tested and surveyed similar devices, and ν_{ts_r} and ν_{ss_r} are the equivalent censoring times of the tested and surveyed similar devices, respectively. The point estimates of the parameters of the likelihood functions are inferred from the MLE or Bayesian analysis. In the Bayesian analysis, the point estimates are the mean values of the posterior distributions obtained from Eq. (4-13) and Eq. (4-14). The π_{ts_i} and π_{ss_i} in Eq. (4-13) and Eq. (4-14) show the posterior distributions of $\bar{\Omega}_t$ and $\bar{\Omega}_s$ for the i-th tested and surveyed similar devices, respectively.

$$L_{ts}=\prod_{f=1}^{F_t} p\left(\nu_{ts_f}, SI_{ref}\middle|\bar{\Omega}_{ts}\right) \prod_{c=1}^{C_t} R(\nu_{ts_c}, SI_{ref}|\bar{\Omega}_{ts}) \quad (4\text{-}11)$$

$$L_{ss}=\prod_{f=1}^{F_s} p\left(\nu_{ss_f}, SI_{ref}\middle|\bar{\Omega}_{ss}\right) \prod_{c=1}^{C_s} R(\nu_{ss_c}, SI_{ref}|\bar{\Omega}_{ss}) \quad (4\text{-}12)$$

$$\pi_{t_i}\left(\bar{\Omega}_{ts}\middle|\nu_{ts_i}, SI_{ref}\right) = \frac{L\left(\nu_{ts_i}, SI_{ref}\middle|\bar{\Omega}_{ts}\right)\pi_0(\bar{\Omega}_{ts})}{p(\nu_{t_i}, SI_{ref})} \quad (4\text{-}13)$$

$$\pi_{s_i}\left(\bar{\Omega}_{ss}\middle|\nu_{ss_i}, SI_{ref}\right) \frac{L\left(\nu_{ss_i}, SI_{ref}\middle|\bar{\Omega}_{ss}\right)\pi_0(\bar{\Omega}_{ss})}{p(\nu_{ss}, SI_{ref})} \quad (4\text{-}14)$$

It is assumed that the difference between the PDF distributions estimated using the test data and the PDF distributions estimated using the survey data is because of the bias in the survey responses. To account for the bias, the equivalent times, or mean cycles (ν) of the first part of the surveyed similar devices are multiplied by an unknown correction parameter ψ. The diversion between the tested and surveyed devices' PDFs is then used as an objective function to determine ψ. The KL divergence method, as shown in Eq. (4-15), is used to measure the distance between the PDF of the tested similar devices (roughly representing true failure behavior) and surveyed similar devices. The KL divergence is a distance or similarity measure between two probability distributions. The KL distance is then minimized through optimization techniques (e.g., gradient descent algorithm) to estimate ψ. The process of estimating the ψ value is shown at the top box in Figure 4.4.

$$\Delta = \int_0^\infty \pi_t\left(\nu\middle|SI_{ref}, \bar{\Omega}_{ts}\right).log\left[\frac{\pi_t\left(\nu\middle|SI_{ref}, \bar{\Omega}_{ts}\right)}{\pi_s\left(\psi\nu\middle|SI_{ref}, \bar{\Omega}_{ss}\right)}\right] d\nu,\ \nu > 0 \quad (4\text{-}15)$$

4.3. Estimating the Parameters of the Reliability Model of the New Product

The parameters of the reliability model of the new product are estimated using the second part of the survey and test data about the similar product and the entire test data about the new product. However, first, the bias in the survey responses should be removed from the survey data. The process of estimating the parameters of the reliability model is shown at the bottom of Figure 4.4.

First, the equivalent times of the second part of the surveyed and tested devices are calculated using the approach discussed in Chapter 4.1. To reduce the bias, the parametric equivalent times of the surveyed devices are multiplied by ψ, which now is a known value. The equivalent time is a function of one or multiple PDF model parameters which are yet unknown.

A 3-step sequential Bayesian analysis then estimates the parameters of the PDF model (and consequently the parameters of the reliability or unreliability model) of the new product. The sequential Bayesian analysis is composed of a sequence of Bayesian analyses. Each analysis uses one of the reliability data sources to build the likelihood function and considers the joint posterior distribution estimated in the previous step as the prior distribution of the current step. However, the first Bayesian analysis uses user-defined (informative or non-informative) priors. The joint posterior distributions of the parameters are estimated using the kernel density estimate (KDE) method, which is a non-parametric way to estimate the (joint) PDF of a (or multiple) random variable(s). Finally, the test data of the new product is used to build the likelihood function of the

last Bayesian analysis. It means that the reliability data of the old products sequentially contribute to constructing the prior distribution for the Bayesian analysis of the new product.

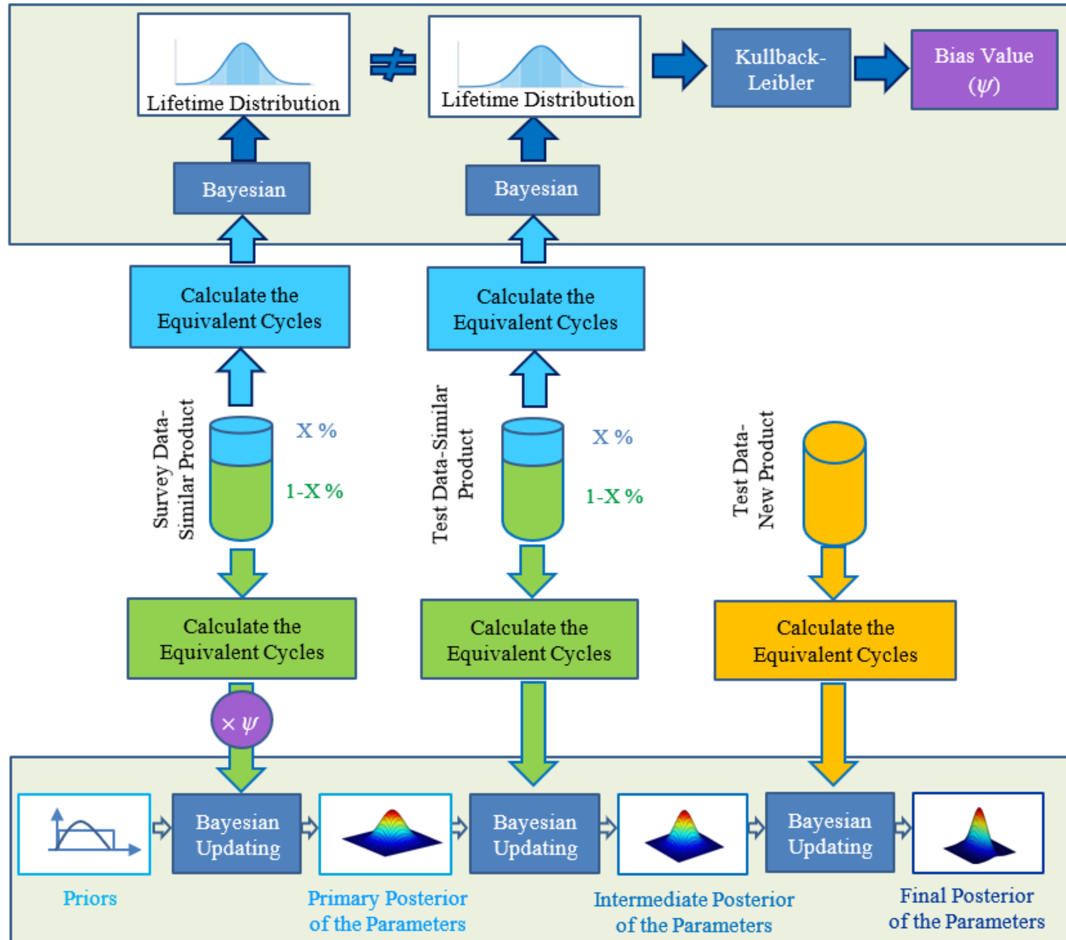


Figure 4.4. The Reliability estimation process.

The proposed approach considers a single failure mode. Suppose the reliability in the presence of multiple failure modes is sought. In this case, the approach is applied to each failure mode separately, and the total reliability is measured as the events that the system is reliable under all failure modes. Thus, one should multiply the reliability models of all failure modes and estimate the total reliability model. The process of

estimating the reliability model in the presence of multiple failure modes is shown in Figure 4.5.

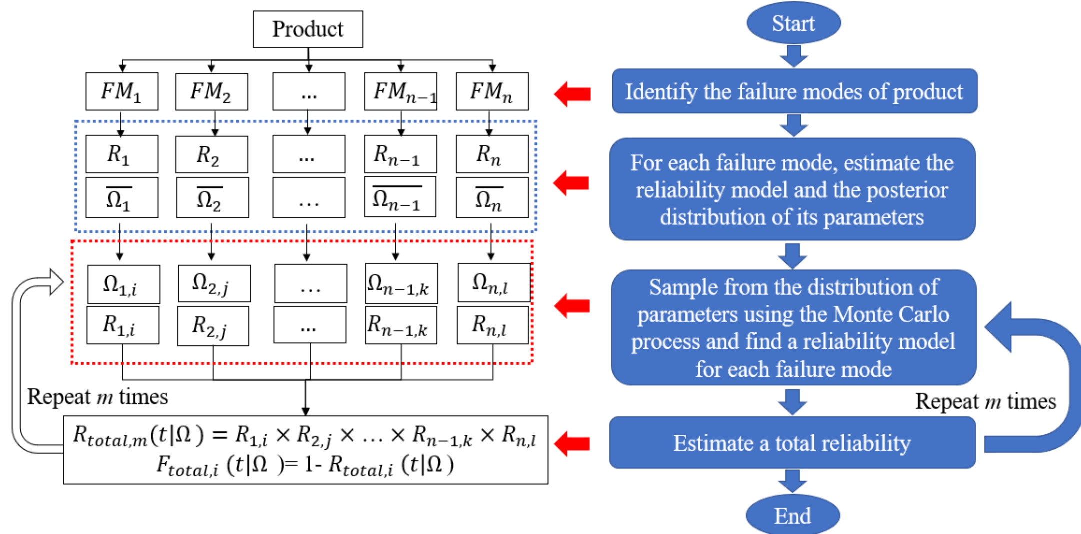


Figure 4.5. The process of estimating the reliability model in the presence of multiple failure modes.

4.4. Case Study

Chapter 4.4 illustrates the application of the proposed reliability estimation approach using a simulated ALT dataset for a new consumer electronic product with the failure mode of cracking caused by accidental drops, a simulated user survey dataset for a similar product (older version of the new product) with the same failure mode, and a simulated ALT dataset for the similar product. The stress for this product is “drop”, and the relevant stress adjustors are (a) the height of the drop, (b) the type of surface that the product hits when it is dropped, and (c) the type of activity that the user is performing when the product is dropped. Weather conditions are not considered significant stress adjustors for this device. Chapters 4.4.1 and 4.4.2 describe steps in simulating the user survey and test datasets for similar and new products.

4.4.1. The Simulated User Survey Dataset for the Similar Product

A user survey dataset was simulated for a hypothetical handheld product. It was assumed that this product was the older version of a new product. The dataset simulates the information typically collected in actual user surveys, which usually give multiple-choice questions instead of asking for precise values. Each (simulated) respondent chooses one or multiple stress profiles. Each stress profile contains a drop height, a surface type, and an activity type from discrete choices in the simulated survey. Each drop height is in the set {knee height, waist heigh, chest height, head or higher height}. Each surface type is in the set {soft (e.g., thick carpet / rug / mat, grass), semi-soft (e.g., thin carpet / rug / mat, exercise flooring), semi-hard (e.g., hardwood, laminate, packed earth), hard (e.g., concrete, asphalt, tile, brick)}. Each activity type is in the set {benign (e.g., standing, walking), harsh (e.g., running, playing sports)}.

The survey dataset was simulated as follows: In step 1, it was assumed that the user's age affected the characteristics of the drop (e.g., height) and generated a sample of 500 users, containing 24 %, 63 %, and 13 % of young (age $<$ 18 years), middle-aged (age in the range 18 to 64 years), and senior (age $>$ 64 years) users, respectively. These fractions are consistent with the U.S. population age groups [48]. The males and females were not distinguished because the mean height of males and females was used for each age group, assuming that the surface and activity types are the same for males and females.

In the second step, eight common stress profiles were assumed, including (1) knee height-semisoft surface-benign activity, (2) knee height -hard surface -harsh activity, (3) waist height -semihard surface benign activity, (4) waist height -hard surface -harsh

activity, (5) chest height -semisoft surface -benign activity, (6) chest height -semihard surface -harsh activity, (7) head height -soft surface -hard activity, and (8) head height -semihard surface -harsh activity. A user in the dataset dropped the device under one or more of these stress profiles during ownership.

In the third step, the stress adjustors between 0 and 10 were sampled from normal distributions, $\mathcal{N}$, with the means and standard deviations shown in Table *4-1*. Harsher stress adjustors had higher scores. The height scores in Table *4-1* were determined as follows. Because the middle-aged group had the highest heights, first, the heights of this group were scored. It was assumed that the mean score of the middle-aged group's knee, waist, chest, and head height were 2.50, 5.00, 7.50, and 9.60, respectively. Then, the mean heights of the other age groups were calculated by multiplying the mean heights of the middle-aged group by the ratio between the height of the intended age group and the middle-aged group. The ratio was determined using the published anthropometric data [49, 50, 51]. For instance, assuming that the ratio between the height of the young and middle-aged group is 0.936, the mean knee height of the young group becomes $2.50\times 0.936= 2.34$. The standard deviation of 0.10 was selected for the head height of the middle-aged group because it resulted in distribution with the upper bound of 9.90 (assuming a six-sigma region) that was close but less than 10. The standard deviation of the other age groups was arbitrarily selected as 0.20.

Table 4-1 Scores of the stress adjustors

Stress Adjustor	**Young**	**Middle -Aged**	**Senior**
Knee Height	$\mathcal{N}(2.34, 0.2)$	$\mathcal{N}(2.50, 0.2)$	$\mathcal{N}(2.45, 0.2)$
Waist Height	$\mathcal{N}(4.67, 0.2)$	$\mathcal{N}(5.00, 0.2)$	$\mathcal{N}(4.90, 0.2)$
Chest Height	$\mathcal{N}(7.01, 0.2)$	$\mathcal{N}(7.50, 0.2)$	$\mathcal{N}(7.35, 0.2)$
Head Height	$\mathcal{N}(8.97, 0.2)$	$\mathcal{N}(9.60, 0.1)$	$\mathcal{N}(9.40, 0.2)$
Soft Surface	$\mathcal{N}(2.50, 0.2)$	$\mathcal{N}(2.50, 0.2)$	$\mathcal{N}(2.50, 0.2)$
Semisoft Surf.	$\mathcal{N}(5.00, 0.2)$	$\mathcal{N}(5.00, 0.2)$	$\mathcal{N}(5.00, 0.2)$
Semihard Surf.	$\mathcal{N}(7.50, 0.2)$	$\mathcal{N}(7.50, 0.2)$	$\mathcal{N}(7.50, 0.2)$
Hard Surface	$\mathcal{N}(9.60, 0.1)$	$\mathcal{N}(9.60, 0.1)$	$\mathcal{N}(9.60, 0.1)$
Benign Act	$\mathcal{N}(5.00, 0.2)$	$\mathcal{N}(5.00, 0.2)$	$\mathcal{N}(5.00, 0.2)$
Harsh Act	$\mathcal{N}(9.60, 0.1)$	$\mathcal{N}(9.60, 0.1)$	$\mathcal{N}(9.60, 0.1)$

Normal distributions allowed us to scatter the scores and generate different but close stress adjustor values for users in the same age group. The distributions of the drop height scores varied among the age groups, but the other distributions were the same.

Step 4 randomly assigned properties including age, ownership time, stress profiles (i.e., drop height, surface type, and activity type during the drop), and stress adjustors' bias to the users. The process of assigning the properties is shown in Fig. 3. First, an age group was selected. For each user in the age group, an ownership time from $\mathcal{W}$ (1.3, 2) was randomly selected, where $\mathcal{W}$ is the Weibull distribution. The ownership times (in years) were rounded to their nearest integers. Also, a stress adjustor's bias was selected randomly from $\mathcal{N}$ (0.9, 0.05) for each user, where $\mathcal{N}$ represents the normal distribution. Then, a random stress profile from the eight common stress profiles was assumed for each year of ownership time. It was assumed that the values generated for the stress adjustors were accurate. As the survey data is biased, the accurate stress adjustors are multiplied by the random bias values associated with the

users. The bias values of the three stress adjustors were the same and constant during the ownership time but varied by the user.

A multiplicative S.I. model, as shown in Eq. (4-16), was then used to combine the stress adjustors and obtain an S.I. value (0< SI< 10^3) for each stress profile.

$$SI = \prod_{i=1}^{3} s_i \tag{4-16}$$

Step 5 selected a drop from the head or above height on a hard surface during a harsh activity as the reference stress profile. The S.I. value for the reference stress profile was SI_{ref}=10^3.

Step 6 assumed that the underlying stress-life model of a similar product was known, as shown in Eq. (4-17), where SI is the stress-index value and N illustrates the mean number of cycles. The drop numbers in the user survey dataset (still not assigned) were viewed as the mean number of drops since the user responses were assumed to express their average number of drops. Thus, as the mean value, the number of cycles (drops) and the number of equivalent cycles can be assumed to follow the continuous Weibull PDF model in Eq. (4-18), where $\mathcal{W}$ illustrates the Weibull function, ν_{ss} shows the distribution of the mean of equivalent cycles of the surveyed similar devices, α_{ss} is the scale parameter that follows the stress-life model, and β is the shape parameter. If the number of cycles is discretely modeled, the discrete PDFs (such as the discrete Weibull types I- III) may be used [34, 52].

$$N = 1800.SI^{-0.5} \quad (4\text{-}17)$$

$$\nu_{ss} = \mathcal{W}\ (\alpha_{ss} = 1800(10^3)^{-0.5} = 56.92,\ \beta{=}1.5) \quad (4\text{-}18)$$

In step 7, the (parametric) mean number of equivalent cycles under the reference stress profile was calculated for each stress profile of the users using (4-19), where $\nu_{ss_{i,y}}$ is the mean equivalent cycle, $d_{i,y}$ is the actual mean number of drops, and $SI_{i,y}$ is the stress-index value of the y-th stress profile of the i-th user. Equation (4-19) was obtained from Eq. (4-8), knowing that the mean number of cycles followed the IPL model. The number of drops in Eq. (4-19) has not been determined yet, so the numerical equivalent cycles are still unknown.

$$\nu_{ss_{i,y}} = d_{i,y}\left(\frac{SI_{i,y}}{10^3}\right)^{0.5} \quad (4\text{-}19)$$

Step 8 assigned numerical equivalent cycles and failure statuses (i.e., failed or right-censored) using the random censoring algorithm proposed in [16]. First, a random probability of failure was drawn from Uniform (0,1), and its corresponding cycles were calculated from Eq. (4-20), where pr_{ss_i} is the random probability and t_i is the corresponding cycle. Then a random cycle (i.e., c_i) was drawn from Uniform (0, C=100), where C is an arbitrary value that controls the number of failed and right-censored units. If $c_i > t_i$, the total number of equivalent cycles during the ownership time of the user was selected as t_i and the device was considered as failed. Otherwise, the total number of equivalent cycles was c_i and the device was right-censored. Two hundred fifty-nine failed and 241 right-censored devices were simulated using this approach.

$$pr_{ss_i} = 1 - e^{-(\frac{t_i}{56.92})^{1.5}} \quad (4\text{-}20)$$

Step 9 assigned the number of drops under the use stress profiles during each year of ownership. The total number of equivalent cycles calculated in step 8 was randomly divided into x parts, where x was the integer of the ownership time. The number of drops during each year of ownership was then calculated using Eq. (4-19).

Step 10 made the number of drops biased by multiplying them by the random values selected from $\mathcal{N}$ (0.85, 0.1), where $\mathcal{N}$ represents the normal distribution. The bias value varied among the users but was the same during the ownership time of the users. Also, the bias of the number of drops was different from the bias of the stress adjustors.

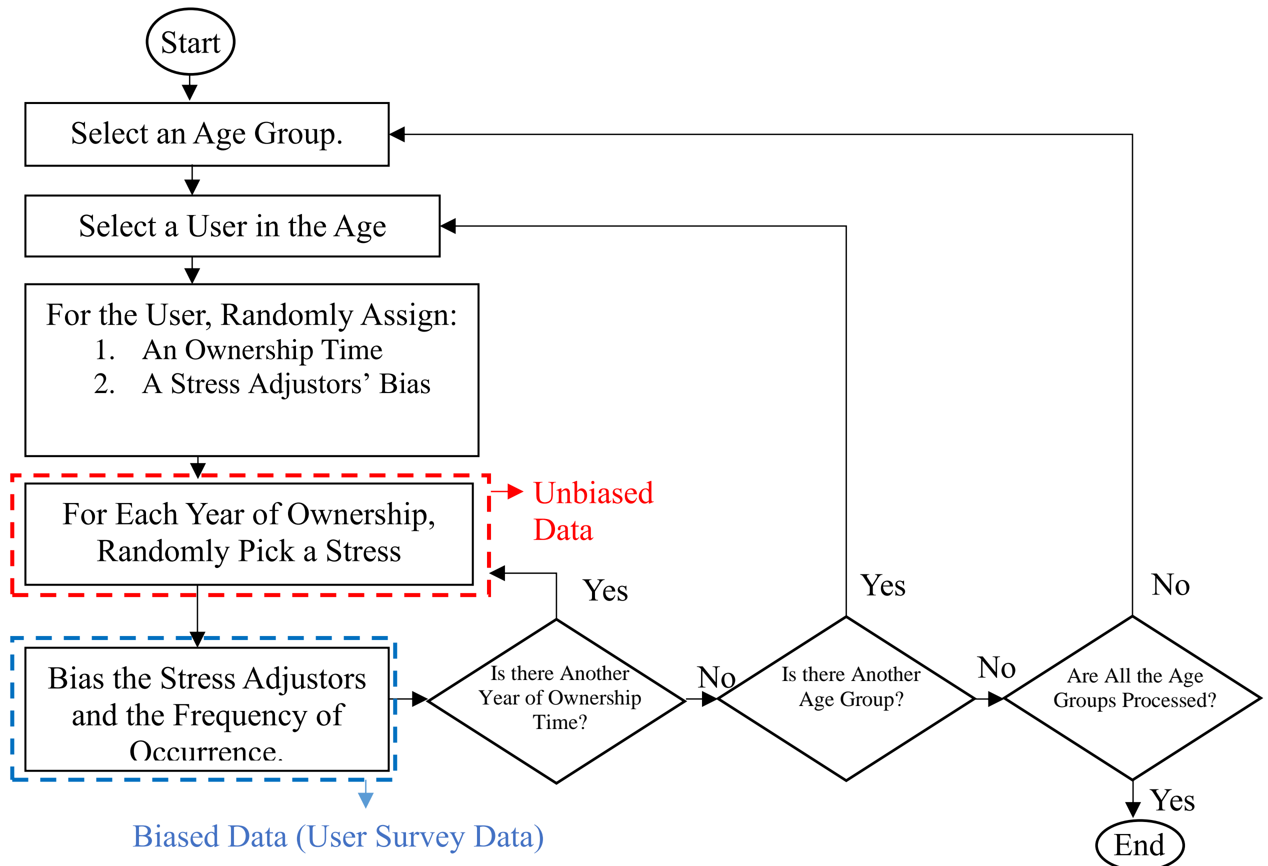


Figure 4.6 The process of assigning properties to users in the simulated user survey dataset.

4.4.2. The Simulated Test Dataset of the Similar Product

Also, a failure-terminated test dataset was generated for the similar product. This dataset, which resembles the information collected in a laboratory during the reliability test, includes the drop height, surface type, activity type, and the number of drops until failure or censoring.

In the first step, it was assumed that the number of failed and right-censored devices were $n_{fs} = 80$, and $n_{cs} = 20$, respectively. The total number of tested devices was thus n_{ts}= 100 units. The test was stopped when the number of failures reached 80 units.

The second step assumed that the reliability test was performed at three conditions: (1) a drop from 0.5 m on a hard surface (e.g., concrete, asphalt, tile, brick) during a harsh activity (i.e., an action with an initial velocity or acceleration), (2) a drop from 1 m on a hard surface during a harsh activity, and (3) a drop from 2 m on a hard surface during a harsh activity. It was assumed that 0.5 m, 1 m, and 2 m were equivalent to a user's knee height, waist height, and head or higher height in the middle-aged group of the simulated user survey dataset. Identical scores were given to the quantitative heights in the test dataset as the subjective heights in the survey dataset. The scores for surface and activity in the test dataset were the same as those in the user survey dataset. As shown in Eq. (4-16), a multiplicative S.I. model was then used to combine the stress adjustors.

The third step defined the reference stress profile as a drop from 2 m on a hard surface during a harsh activity. This profile is equivalent to the reference stress profile selected

for the user survey dataset. The SI value for the reference stress profile was $SI_{ref} = 10^3$.

The fourth step assumed that the relationship between the S.I. value and the number of usage cycles was described by the same IPL stress-life model used for the user survey dataset, as shown in Eq. (4-17). Also, it is assumed that the equivalent cycles (ν_{ts}) of the tested similar devices follow the same Weibull distribution used for the user survey dataset, as shown in Eq. (4-18).

It was assumed that the failures occurred at different usage cycles due to the randomness in the manufacturing process and material properties. In step 5, the random number of usage cycles was generated. First, a random probability of failure (pr_i) was assigned to each failed device using Uniform (0, 1). Second, the numbers of equivalent cycles under the reference stress profile were calculated using Eq. (4-21). Third, the number of usage cycles was calculated using Eq. (4-22), where d_i is the number of drops of the i-th device, SI_i is the stress-index value of the stress profile under which the i-th device was tested, and ν_{ts_i} is the number of equivalent cycles for the i-th device.

$$\nu_{ts_i} = \exp\left[\frac{1}{1.5} Ln\left(Ln\left(\frac{1}{1-pr_i}\right)\right) + \ln(56.92)\right] \quad (4\text{-}21)$$

$$d_i = \nu_{ts_i}\left(\frac{SI_i}{10^3}\right)^{-0.5} \quad (4\text{-}22)$$

Steps 6 to 8 determine the number of drops and failure status of the right-censored devices. First, step 6 used the highest equivalent cycle of the failed devices as the censoring equivalent cycle. Next, in step 7, the numbers of drops of the 20 right-

censored units were estimated using Eq. (4-22). Finally, in step 8, the status of "right-censored" was considered for the censored units.

4.4.3. The Simulated Test Dataset of the New Product

It was assumed that the new product was the latest version of the similar product and its failure mode (i.e., cracking) was delayed. It means that the distribution of the equivalent cycles of the new devices had the same shape parameter as the similar devices, but its scale parameter was larger. Therefore, through the same algorithm used for simulating the test dataset of the similar product, a dataset with the shape parameter of β=1.5 and the scale parameter of $\alpha_{tn} = 2200(10^3)^{-0.5} = 69.57$ was generated. The distribution of the equivalent cycles of the new devices is shown in Eq. (4-23), where $\mathcal{W}$ represents the Weibull function, ν_{nt} is the equivalent cycles, α_{tn} is the scale parameter, and β is the shape parameter. The number of failed and right-censored in this dataset are $n_{fn} = 80$, and $n_{cn} = 20$, respectively.

$$\nu_{tn} = \mathcal{W}\ (\alpha_{tn} = 2200(10^3)^{-0.5} = 69.57, \beta=1.5) \qquad (4\text{-}23)$$

4.4.4. Reliability Analysis Using the Simulated User Survey and Test Datasets

The reliability model of the hypothetical new product was estimated using the process discussed in Chapters 4.1 To 4.3. First, the scored stress adjustors were combined through the multiplicative S.I. model. An IPL stress-life model with unknown parameters was then assumed, as shown in Eq. (4-24), where N illustrates the mean number of cycles, SI is the stress-index value, and A and n are the model's

unknown parameters. Next, the IPL model was used to find the relationship between the mean number of cycles under the use stress profile and the mean number of equivalent cycles under a reference stress profile, as shown in Eq. (4-25), where ν_i is the mean number of equivalent cycles under the reference stress profile, d_i is the mean number of drops under the actual use profile, SI_i is the stress-index value of the actual use profile, SI_{ref} is the reference stress index, and n is an unknown parameter. Finally, the reference stress profile was assumed as a drop from the head or higher height (2 m) on a hard surface during a harsh activity ($SI_{ref} = 10^3$). The mean equivalent cycles of the surveyed and tested devices were then estimated using Eq. (4-25).

$$N = A.SI^{-n} \tag{4-24}$$

$$\nu_i = d_i\left(\frac{SI_i}{SI_{ref}}\right)^n \tag{4-25}$$

It was assumed that the mean equivalent cycles followed a Weibull distribution model, as shown in Eq. (4-26). In Eq. (4-26) $\mathcal{W}$ illustrated the Weibull model, α is the scale parameter that wathe IPL stress-life model replaced, A and n are the parameters of the stress-life models, and β is the shape parameter. Then, 30% of the surveyed devices and 30% of the tested similar devices were separately used in two MLE analyses to estimate the parameters of their Weibull distribution models. The point estimates of the parameters for the survey dataset were A_s=1300, n_s=0.5, and β_s = 1.48, and for the test dataset were A_t=1799.99, n_t=0.5, β_t=1.4.

$$\nu \sim \mathcal{W}\left(\alpha = A(SI_{ref})^{-n}, \beta\right) \tag{4-26}$$

It was assumed that the difference between the Weibull distributions (i.e., point estimates of the parameters) of the survey dataset and test dataset was because of the bias in the survey data. If the bias were removed, the user survey and test data would follow the same life distribution model. The mean equivalent cycles of the surveyed similar devices were thus multiplied by an unknown value (bias parameter shown by ψ) to correct this bias. The ψ was estimated by minimizing the KL divergence value between the life distribution of the surveyed similar devices and the life distribution of the tested similar devices, as shown in Eq. (4-27), where π shows the life distribution of the tested similar devices, f is the life distribution of the surveyed similar devices, SI_{ref} is the reference stress index, α_t and β_t are the scale parameter and the shape parameter of the life distribution of the tested devices, α_s and β_s are the scale parameter and the shape parameter of the life distribution of the surveyed devices, A_t , A_s, n_t and n_s are the parameters of the stress-life model, and ν is the mean number of cycles which is a positive value. This analysis resulted in ψ = 0.82. then the mean equivalent cycles of the remaining 70% of the surveyed devices were multiplied by 0.82 to correct the bias.

$$\Delta = \int_0^{\infty} \pi\left(\nu \middle| \alpha_t = A_t SI_{ref}{}^{-n_t}, \beta_t\right) . \log\left[\frac{\pi(\nu|\alpha_t = A_t SI_{ref}{}^{-n_t}, \beta_t)}{f(\Psi\nu|\alpha_s = A_s SI_{ref}{}^{-n_s}, \beta_s)}\right] d\nu, \nu > 0 \quad (4\text{-}27)$$

The bias-corrected user survey data, the remaining 70% of the test data about the similar product, and the test data about the new product were used in a 3-steps sequential Bayesian analysis to estimate the parameters of the life distribution model. The first step assumed Normal prior distributions for the reliability model's parameters.

The mean values of the Normal distributions were approximated using the MLE method, and their standard deviations (std) were half-normal distributions with hyper priors (i.e., std) equal to 1/10th of the MLE estimates. The bias-corrected survey data of similar devices were also used as the likelihood data (observation). Next, the primary joint posterior distribution was estimated using the KDE method. It was used as the prior distribution of the second Bayesian analysis. In the second Bayesian analysis, the remaining 70 % of the test data of the similar device was used as the likelihood data. The intermediate joint posterior distribution estimated in this step was used as the prior distribution of the third Bayesian analysis. In the third Bayesian analysis, the test data about the new product was used to build the likelihood function, and the final joint posterior distribution of the parameters was estimated. Figure 4.7 shows the parameters' primary, intermediate, and final posterior distributions. Although the joint distribution of the parameters was estimated, for illustration purposes in Figure 4.7, the distributions of the parameters were plotted independently. One thousand realizations from the estimated final joint posterior distribution of the parameters were then generated and used to quantify the uncertainty of the estimated reliability.

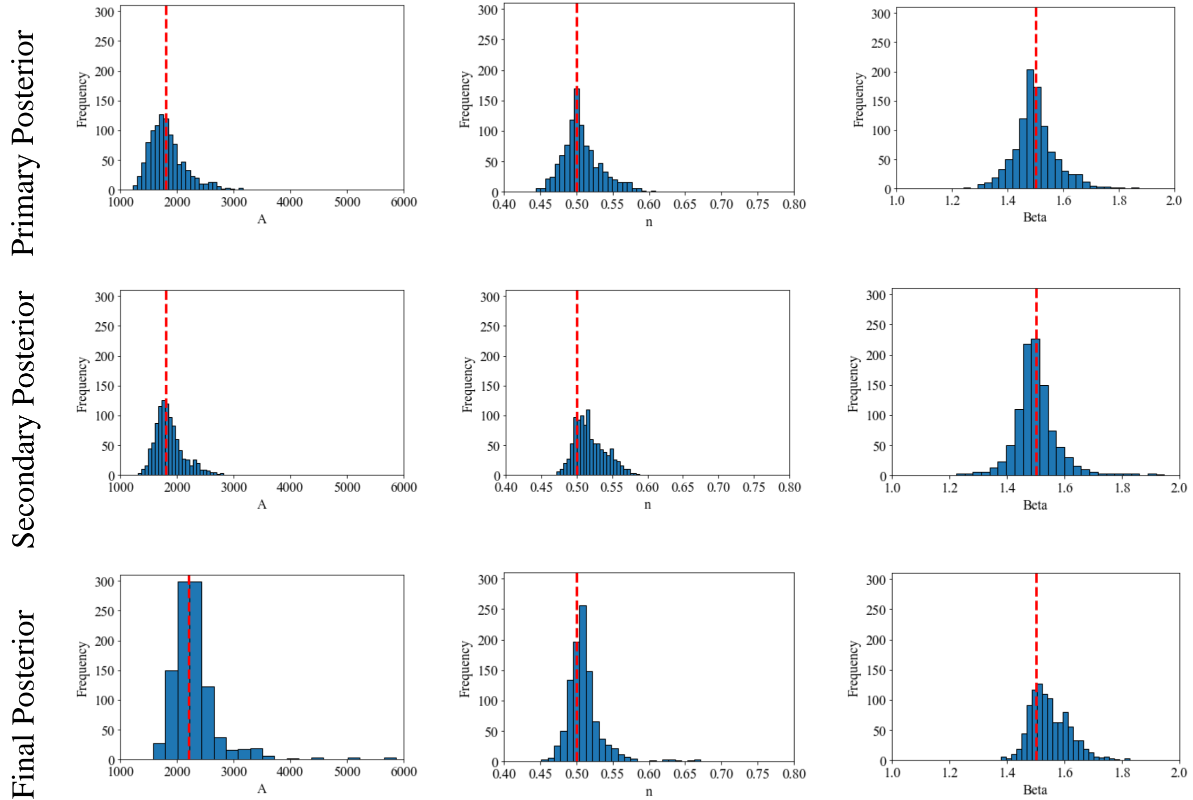


Figure 4.7. Primary, intermediate, and final posterior distributions of the parameters of the reliability model.

Figure 4.8 illustrates the estimated unreliability models (in blue) and the true unreliability model (in black). The true model is the original model from which the test data of the new product was generated. As shown in Figure 4.8, the true model is within the uncertainty region of the estimated unreliability models. The model in Figure 4.8 is called the "main model" and is used in Subchapter 4.4.5.

4.4.5. Sensitivity and Uncertainty Analysis

In estimating the reliability model in this example, the stress adjustors were arbitrarily scored, an S.I. model was arbitrarily selected, and a reference stress profile

was arbitrarily selected. Subchapter 4.4.5 describes the sensitivity of the results to changes in these parameters. Three analyses were performed: (1) the linear scores used in Subchapters 4.4.1 to 4.4.3 were replaced by the non-linear (squared) scores (2) an additive S.I. model replaced the multiplicative S.I. model used in Subchapter 4.4.1 to 4.4.3, and (3) the reference stress profile used in Subchapter 4.4.1 to 4.4.3 was replaced by a drop from the head or above height (2 m) on a hard surface during a benign activity.

To compare the results of the three analyses with the main model, the parameters of their stress-life model were selected such that their true Weibull life distribution models remained the same as the true model in Figure 4.8. The proposed approach (that was used to find Figure 4.8) was then used to estimate the reliability models of the three cases. Figure 4.9 to Figure 4.11 show the uncertainty regions of the estimated reliability models in blue and the true reliability models in black. In all cases, the true model is within the uncertainty region of the estimated reliability models.

The uncertainties of the estimated reliability models stem from two sources: (1) the aleatory uncertainties that result from inherent variability (e.g., the natural variability in materials) that are irremovable and irreducible, and (2) the epistemic uncertainties due to insufficient human knowledge, analyzing a sample rather than the population, and processing errors (e.g., recording, coding, and data preparing-related errors, model selection). The epistemic uncertainties can be reduced by obtaining more information about the product [17, 16].

To compare the uncertainty regions of Figure 4.8 and Figure 4.9 to Figure 4.11, the difference between the maximum % failure and minimum % failure at every 20

equivalent cycles (0 < equivalent cycles < 400) was used as the comparison metric. The differences are shown in Figure 4.12. The curves labeled “non-linear scoring”, “additive S.I. model”, and “reference stress profile” in Figure 4.12 are the curves related to Figure 4.9 to Figure 4.11, respectively. According to Figure 4.12, the main model (i.e., the reliability model with linear scoring) has less uncertainty than the model with non-linear scoring. As the uncertainty due to linear and non-linear scorings depends on the nature of the product and its applicable stresses and stress adjustors, the reliability analyzer may try both linear and non-linear scorings and select the one that shows less uncertainty for the intended application.

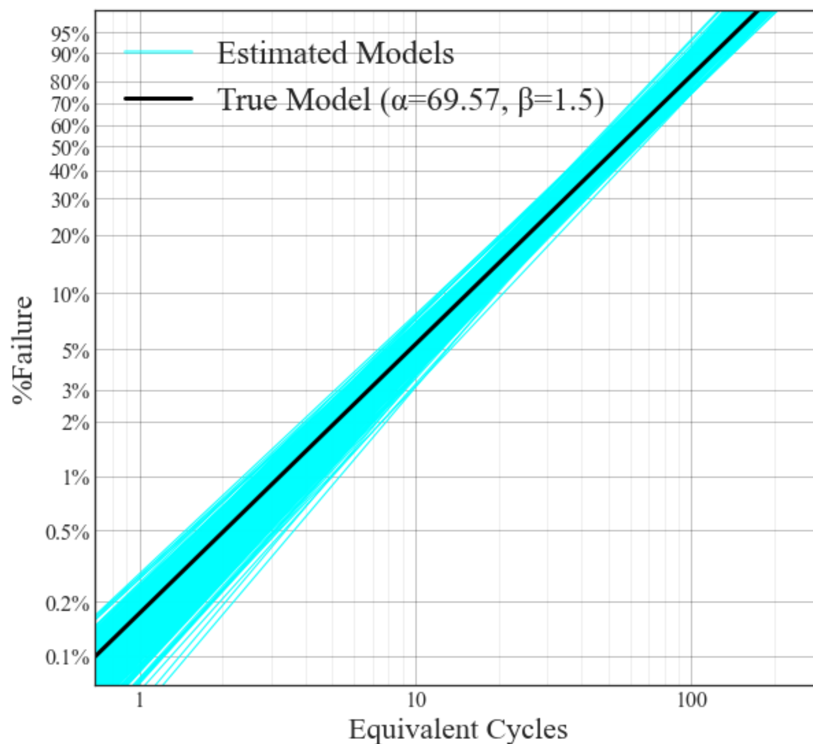


Figure 4.8. The estimated reliability models and the true reliability model.

Figure 4.12 illustrates that when a multiplicative S.I. model is used (i.e., the main model), the uncertainty region of the reliability models is narrower than the additive S.I. model. This is because the additive model has several unknown parameters (i.e., the weights of the applicable stresses and stress adjustors), and their estimates introduce additional uncertainties in the reliability model.

The predicted reliability model is less sensitive to the selected reference stress profile than the S.I. model and scorings, as shown in Figure 4.12. As the model with linear scoring and multiplicative S.I. model shows less uncertainty than the models in Figure 4.9 to Figure 4.11, it was used as the most representative model in this study.

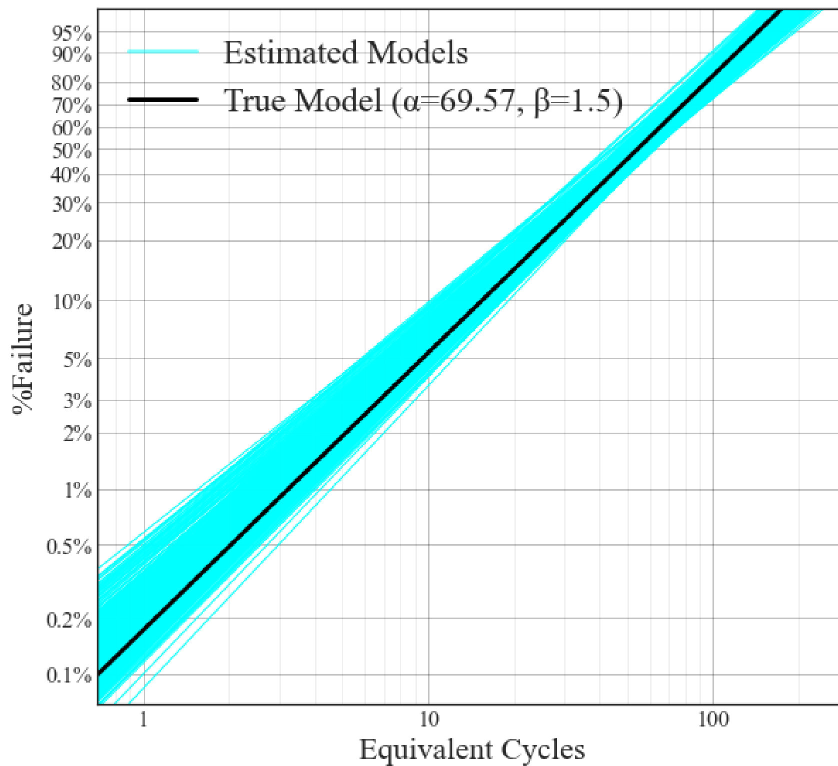


Figure 4.9. The estimated reliability models using non-linear scores for the stress adjustors.

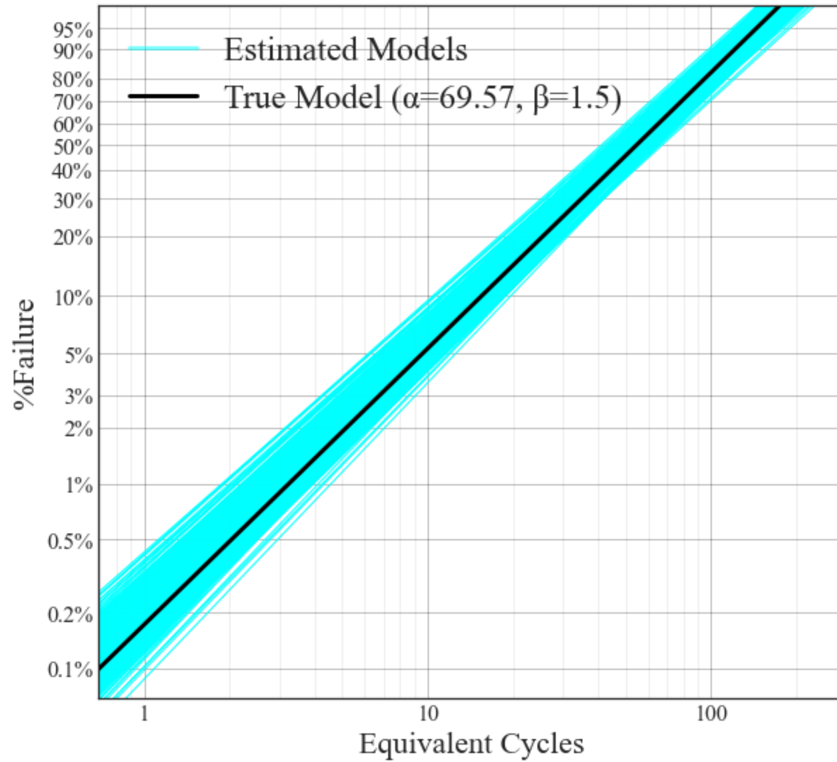


Figure 4.10. The estimated reliability models using the additive stress-index model.

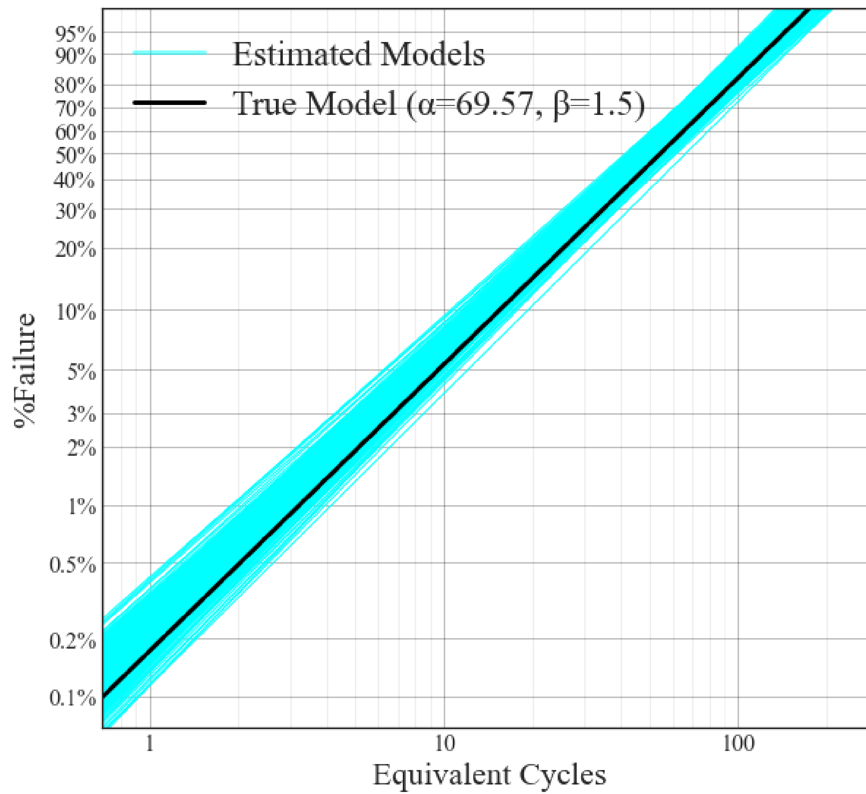


Figure 4.11. The estimated reliability models using a drop from the head or higher (2 m) height on a hard surface during a benign activity as the reference stress profile.

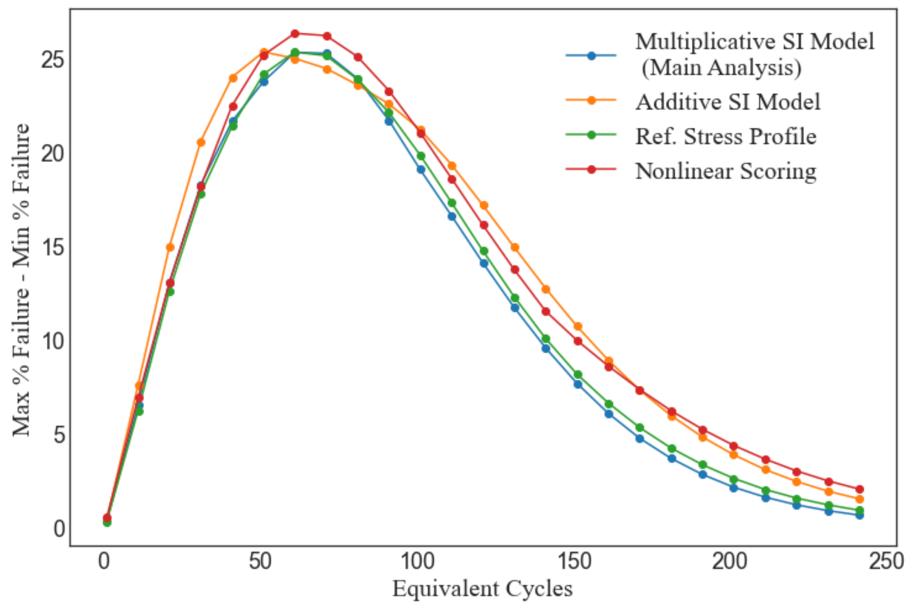


Figure 4.12. Comparing the uncertainty regions of the models with different parameter selections using the difference between maximum and minimum % failure as the comparison criterion.

This study used a computer with Intel(R) Core (TM) i7-9700K CPU and 32 GB RAM. The computational time for the multiplicative S.I model, non-linear scoring, and reference stress profile analyses was about 17 minutes, and for the additive S.I. model analysis was about 35 minutes.

The reliability model of the new product was also analyzed through a simple Bayesian analysis that used the ALT data of the new product as the observation and did not consider the historical data of the similar product. Figure 4.13 compares the reliability models estimated by the ALT data of the new product with the main model (i.e., the reliability models estimated using the ALT data of the new device, ALT data of the similar device, and unbiased survey data of the similar device in a sequential Bayesian analysis). The figure illustrates that the main model has less uncertainty.

The estimated reliability model in Figure 4.8 can be used to understand the % failure (and consequently reliability) and its associated uncertainty at a given time. For

instance, we can find the % failure with 90% confidence after one year of warranty or four years when the product becomes obsolete, assuming that one year is equivalent to 20 cycles under the reference stress profile), as shown in Figure 4.14, or we can find the time-to-failure at given reliability (e.g., 95% reliability).

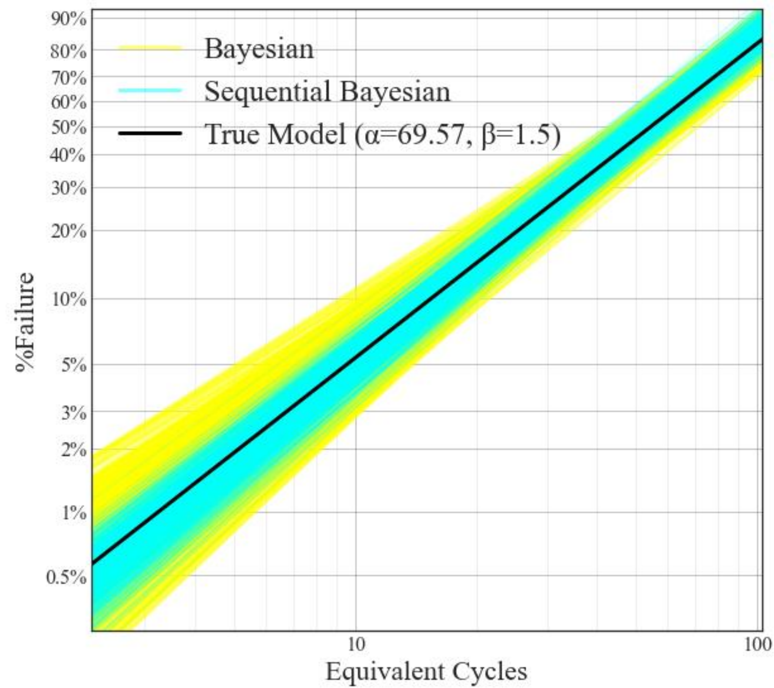


Figure 4.13. The reliability models estimated by the ALT data of the new product in Bayesian analysis and the reliability models estimated by the ALT data of the new device and the historical data of the similar device in sequential Bayesian analysis.

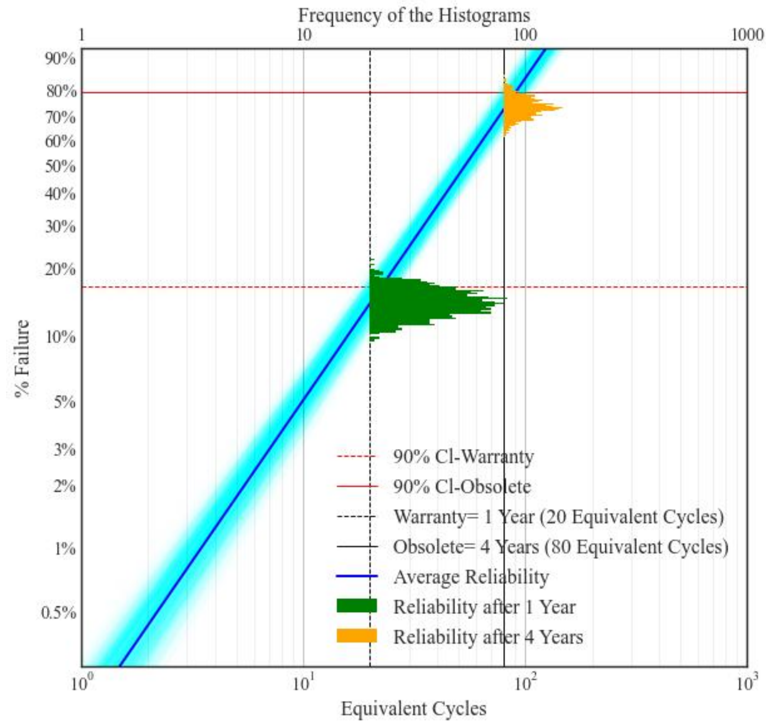


Figure 4.14. Distribution of % failure after 1 and 4 years, estimated using the main model.

4.4.6. Conclusions

Chapter 4 discussed designing and applying a user survey that collects critical data for estimating the reliability model of a consumer product. A stress-based reliability estimation approach was developed that approximated the reliability model of a new product using the reliability test data of a new product and the user survey data and reliability test data from an older version of the same or similar product. The bias in the user survey responses was estimated and removed using the Kullback-Leibler divergence method. The joint posterior distribution of the new product's reliability model parameters was then estimated in sequential Bayesian analysis.

A case study simulated the reliability data of a new hypothetical product and a similar product from the known reliability models. It was shown that the true reliability model of the new product was within the uncertainty region of the estimated reliability models of the new product.

Chapter 4 showed that a user survey could be a cost-effective and quick way to collect field data to estimate a product's reliability. Besides using the user survey data of a similar product as appropriate prior information in a Bayesian reliability estimation, the paper showed that the proposed approach could make an accurate assessment of the reliability of the new product due to a single failure mode. The proposed approach is generic and can estimate the reliability model of a wide range of consumer products. If the product experiences a small number of cycles to failure in its lifetime (e.g., a laptop that may break after a few drops), discrete forms of life distributions for estimating the reliability model would be more appropriate. The application of the discrete life models is shown in Chapter 5.

Chapter 5: Estimating the Reliability Model of a New Consumer Product Assuming a Discrete Life Distribution3

This chapter develops a novel mathematical approach for estimating a new product's reliability model with few damage cycles to failure. The application of the proposed approach is shown, and the results are validated using the simulated datasets for an electronic device with the failure mode of cracking caused by accidental drops.

The reliability of a product is the probability that a product performs its intended function adequately for a specified time under specified use conditions [47]. The reliability of a product, thus, depends on its usage time, applied stresses (i.e., agents that cause damage to the product), and stress adjustors (i.e., conditions that increase or decrease the stress magnitude or absorption).

---

[3] This chapter is a reproduction of the papers about applying user survey data and accelerated test data to estimate reliability of new consumer products using a discrete life distribution model [71, 70]:

[52] N. Shafiei, J. W. Herrman and M. Modarres, "Applying User Surveys and Accelerated Tests Data to Estimate Reliability of New Consumer Products Using a Discrete Life Distribution Model," *IEEE Access,* 2022 June 29.

[53] N. Shafiei, J. W. Herrmann and M. Modarres, "Estimating the Reliability Model of a New Consumer Product Assuming a Discrete Life Distribution," in *RAMS (under review)*, Florida, 2022.

Some previous studies used damage-based reliability estimation approaches to consider the effect of various stress and stress adjustors in reliability analysis [33, 53, 1]. These methods apply to reliability test or sensor data that are costly to collect.

Our previous studies [54, 55] developed a stress-based reliability estimation approach that utilized user survey data as a cost-effective source for estimating reliability.

That approach assumed a continuous lifetime distribution and, through Bayesian analysis, estimated the reliability model of a consumer product. In addition, the concept of the equivalent cycle was defined to consider the effect of various stresses and stress adjustors.

A continuous distribution is applicable when the random variable representing life is continuous (e.g., expressed in calendar time), or the product undergoes many damage-accumulating usage cycles (e.g., hundreds of cycles of a device dropping on a hard floor). When the performance of a product is measured occasionally (e.g., every week or every month), or the product undergoes a few damage cycles, a discrete distribution can more accurately estimate the reliability model [34, 35].

This chapter extends the approach of the previous chapter for discrete lifetime distributions. Some discrete life distributions have a summation term whose upper bound can be expressed as an unknown equivalent cycle. Such upper bounds make reliability estimation through a Bayesian or Maximum Likelihood Estimation (MLE) method (i.e., a method for estimating the parameters of a parametric distribution) very difficult. The discrete Weibull distribution type III has this issue. Ideally, a method for selecting the best distribution should be used in reliability analysis. However, to demonstrate the proposed generic approach in this chapter, the case study considers the

discrete Weibull type III distribution as a complicated discrete distribution with the summation term.

The Weibull distribution is a well-known lifetime distribution that is widely used in reliability engineering. The Weibull distribution is available in continuous and discrete forms. Three types of discrete Weibull distributions have been proposed: (1) type I has a reliability function that mimics the reliability function of a continuous Weibull distribution, (2) type II has a hazard function that mimics the hazard function of a continuous Weibull distribution, and (3) type III is more generic and does not follow any function of a continuous Weibull distribution [39].

Various extensions of discrete Weibull distributions have also been introduced, and different methods for estimating their parameters have been suggested. For instance, Jia et al. [40] proposed a discrete extended Weibull distribution and used MLE for estimating its parameters. Barbiero [34] proposed three methods, including the method of proportion, MLE, and the method of moments for estimating the parameters of a discrete Weibull type III distribution. The above studies assumed that the product is used under a fixed stress level, and the number of damage cycles is equivalent to the number of times the stress is applied to the product. On the other hand, the approach presented in this chapter assumes that consumer products are used under varying stress levels, and the concept of equivalent cycles is used to calculate the number of cumulatively damaging usage cycles under a reference stress condition. As discussed earlier, when using the idea of equivalent cycles, the upper bound of the summation term becomes an unknown value that is hard to estimate.

This chapter considers a consumer product that is used under various stress conditions, and its failure time is expected to occur within a few applied stress cycles (loads). Thus, the discrete forms of lifetime distributions are proposed for estimating the reliability model. The reliability estimation approach developed in our previous studies [6, 7] is generic and can be used with a discrete lifetime distribution, but the Bayesian method alone may not be enough to estimate the reliability model's parameters, and further elaboration is needed.

The reliability model of a discrete distribution (such as the Weibull Type III) might have a summation term with a time variable as an upper bound [34, 39]. In the reliability estimation approach proposed in our previous studies [54, 55], this upper bound becomes an equivalent cycle, which is a function of one or more unknown parameters. This makes calculating the summation term and consequently estimating the parameters difficult. Therefore, this study suggests initializing the upper bound and updating it using MLE and a gradient descent algorithm until the upper bound has converged. The upper bound is then fixed, and Bayesian analysis is used to estimate the parameters of the reliability model as discussed in previous studies.

This chapter also presents a case study that shows the application of the proposed approach using simulated survey and test datasets for a consumer product with the failure mode of cracking caused by accidental drops. It is assumed that the equivalent cycles of the devices follow a discrete Weibull distribution type III. Then, the parameters of the reliability model are estimated using the proposed approach. The final reliability model of the product is compared with the reliability model estimated in the previous study using the continuous Weibull distribution [54].

This chapter is organized as follows: Subchapter 4.1 explains the reliability estimation approach assuming a discrete lifetime distribution. Subchapter 4.2 presents the case study. Finally, Subchapter 4.3 concludes this chapter.

5.1. The Reliability Estimation Approach

This subchapter describes the proposed approach for estimating the reliability model of a new consumer product with a few life cycles using reliability test data of the new product coupled with user survey data and reliability test data of a similar product (e.g., the older versions of the new product). The current approach extends the approach proposed in Chapter 4 by assuming a discrete lifetime distribution [55]. The outline of the approach is shown in Figure 4.4. The approach has three steps: (1) calculating the equivalent cycles, (2) removing bias from the survey responses, and (3) estimating the parameters of the reliability model of the new product using a sequential Bayesian approach. The first two steps were discussed in Subchapter 4.1 and Subchapter 4.2. Chapter 5.2 describes the third step.

5.2. Estimating The Parameters of The Reliability Model of The New Product

The new product's reliability model parameters are estimated using a sequential Bayesian analysis [10]. The analysis involves three steps, as shown in the box at the bottom of Figure 4.4. All steps assume a parametric discrete life distribution for the equivalent cycles.

In the first step, non-informative or weakly-informative prior distributions are selected for the parameters of the reliability model. Next, the equivalent cycles of the

second portion of the surveyed devices are multiplied by ψ, which now is a known value. This removes the bias from the equivalent cycles. The adjusted (bias-removed) equivalent cycles are then used as observations (i.e., likelihood data) in the first Bayesian analysis. The primary joint posterior distribution of the parameters is then estimated.

In the second step, the primary joint posterior distribution is used as the prior distribution for the second Bayesian analysis. Then, the equivalent cycles of the second portion of the tested similar devices are used as observations, and the intermediate joint posterior distribution is estimated.

In the third step, the intermediate posterior distribution is used as the prior distribution, and the equivalent cycles of the new device are used as the observation. Finally, the final joint posterior distribution of the parameters is estimated in the Bayesian analysis and used to approximate the reliability model of the new product.

When assuming continuous life distributions, Bayesian analysis is straightforward. However, for some discrete life distributions, the analysis is complicated. For instance, in the probability mass function (pmf) of a discrete Weibull distribution type III, shown in (5-6) [53], the upper limit of the summation is a time value, which, in our approach, becomes an equivalent cycle that depends upon the unknown parameters of the stress-life model, which are unknown. This makes calculating the summation term and completing the Bayesian analysis difficult. In the following, a method is proposed to deal with this difficulty.

$$f(t) = \left(1 - e^{-c(t+1)^{\beta}}\right) e^{-c \sum_{j=1}^{t} j^{\beta}} \tag{5-6}$$

The proposed method utilizes a gradient descent algorithm to fix the upper bound of the summation term. Gradient descent is an optimization technique used to estimate the parameters of a linear model by minimizing its cost function [56]. The cost function measures the difference between a variable's predicted and true value, and these differences are used to update the estimates in each iteration. For example, consider a linear model with the slope of W and the intercept of b (i.e., $y(x) = W.x + b$). The gradient descent estimates for W and b are shown by Eq. (5-7) and Eq. (5-8), respectively, where l indicates the number of iterations in the gradient descent technique, α is the learning rate, x_i and $y(x_i)$ are the variables of the model, y_{t_i} is the true value of $y(x_i)$, and W and b are the parameters [56].

$$W_{l+1} = W_l + \alpha . \frac{\sum_{i=1}^{m}(y(x_i) - y_{t_i})x_i}{m} \tag{5-7}$$

$$b_{l+1} = b_l + \alpha . \frac{\sum_{i=1}^{m}(y(x_i) - y_{t_i})}{2m} \tag{5-8}$$

The approach linearizes the equivalent cycle equation, initializes the linearized equation parameters, and updates the parameters using Eq. (5-7) and Eq. (5-8). Each iteration calculates the upper-bound of the summation term of Eq. (5-6) using the latest gradient descent estimates and, through the MLE method, estimates the true value of $y(x_i)$ (i.e., y_{t_i}). This process continues until the upper bound has converged. Then the upper bound is fixed, and the other parameters of Eq. (5-6) are estimated in the sequential Bayesian analysis. The outline of the process of estimating the upper bound is shown in Figure 5.1.

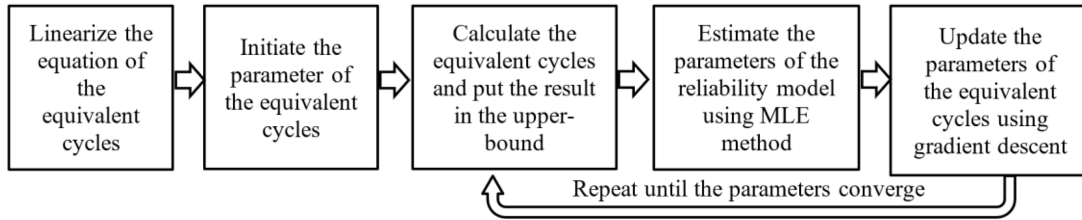


Figure 5.1. The process of estimating the upper bound of the summation term in the discrete life distribution.

5.3. Case Study

This subchapter shows the application of the proposed reliability estimation approach using simulated reliability data for an electronic product (e.g., a laptop) with the failure mode of cracking caused by a few (e.g., 10 to 20) accidental drops. The approach for simulating the test data of the new and old products and the survey data of the old product is explained. Then, the proposed reliability estimation approach is used to estimate the reliability model of the new product.

5.3.1. The Simulated Datasets

An accelerated life test (ALT) dataset for the older version of the new product. This dataset has 100 samples (80 failed and 20 were right-censored). It includes information about the applicable stress (i.e., drop), stress adjustors (i.e., drop height, surface type, and the severity of drop), and the number of cycles to failure or right-censoring. This simulated ALT dataset is based on the following assumptions:

1. During the reliability test, samples were dropped from 0.5 m, 1 m, 1.5 m, and 2 m heights on a hard surface during strenuous or harsh activity.

2. Stress and stress adjustors were scored and combined through a multiplicative SI model. The scores for heights were 25, 50, 75, and 100, for the hard surface was 100 and for harsh and strenuous activity was 100.

3. The stress-life model was a known inverse power law (IPL), as shown in Eq. (5-9) [57], where N illustrates the number of cycles, SI is the stress-index value, A is 285, and n is 0.5.

$$N = A\,SI^{-n} \tag{5-9}$$

4. The reference stress profile was a drop from 2 m on a hard surface during a harsh activity.

5. The life distribution was a known discrete Weibull distribution type III, as shown in Eq. (5-6), where c=0.05 for the old product, c=0.01 for the new product, and β=0.5.

Figure 5.2 shows the process of simulating the ALT dataset for the old product. The same approach was used for the new product, but it was assumed that the new product was a better version of the old product and that its failure had been terminated (i.e., has greater cycles to failure).

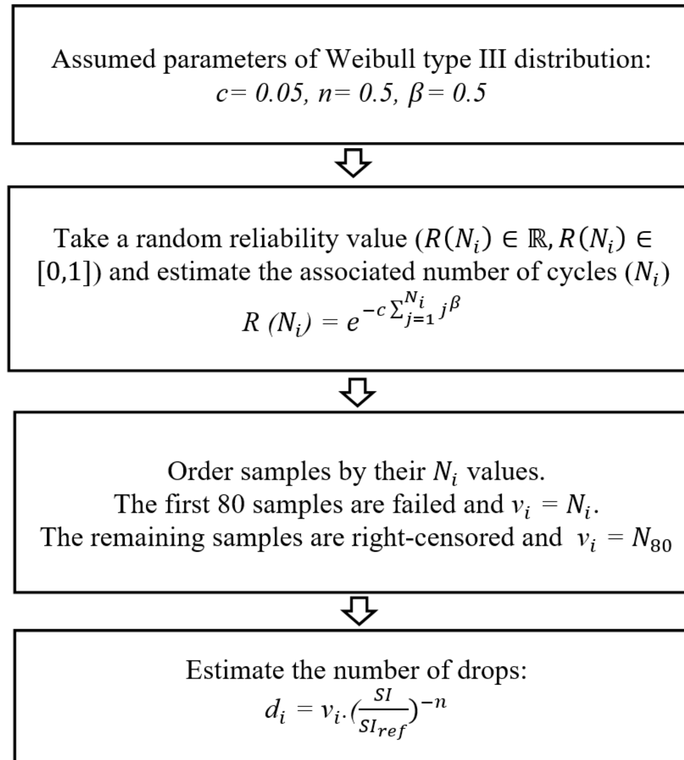


Figure 5.2. The process of simulating ALT data.

Also, a user survey dataset was simulated for the old product. The survey dataset considered 500 users from different age groups with different heights, activities, and biases in the frequency of drops and the applicable stress profiles. This dataset is based on the following assumptions:

1. User dropped their devices under one or several stress profiles. For example, 1) knee height-semisoft surface- benign activity, 2) knee height- hard surface- harsh activity, 3) waist height- semihard surface- benign activity, 4) waist height- hard surface- harsh activity, 5) chest height- semisoft surface- benign activity, 6) chest height- semihard surface- harsh activity, 7) head height- soft surface- hard activity, 8) head height- semihard surface- harsh activity.

2. Stress and stress adjustors were scored and combined through a multiplicative SI model. The scores for knee, waist, chest, and head or higher heights were 25, 50, 75, and 100 for soft, semi-soft, semi-hard, and hard surfaces were 25, 50, 75, 100, and for benign and harsh activity were 50 and 100, respectively.

3. The stress-life model was a known inverse power law (IPL), as shown in (9), where N illustrates the number of cycles and SI is the stress-index value, A is 285, and n is 0.5.

4. The reference stress profile was a drop from the head or higher height on a hard surface during a harsh activity.

5. The survey data was initially generated from the discrete Weibull life distribution shown in Eq. (5-6). Then, the SI value and number of drops in the dataset were multiplied by random values taken from a N (0.80, 0.1) distribution to bias the dataset. Taking random values from the normal distribution made it possible to assign different bias values to the users' responses (for the stress profile (SI value) and the number of drops). This bias value of each user remained constant during the ownership time. The process of simulating the user survey data is shown in Figure 5.3.

5.3.2. Estimating the reliability model of the new product

This subchapter assumes that the old and new products' stress-life model, life distribution, and reliability model are unknown. Therefore, the simulated datasets and the developed reliability estimation approach were used to estimate the reliability model of the new product. As discussed earlier, estimating the reliability model of the new product has three steps.

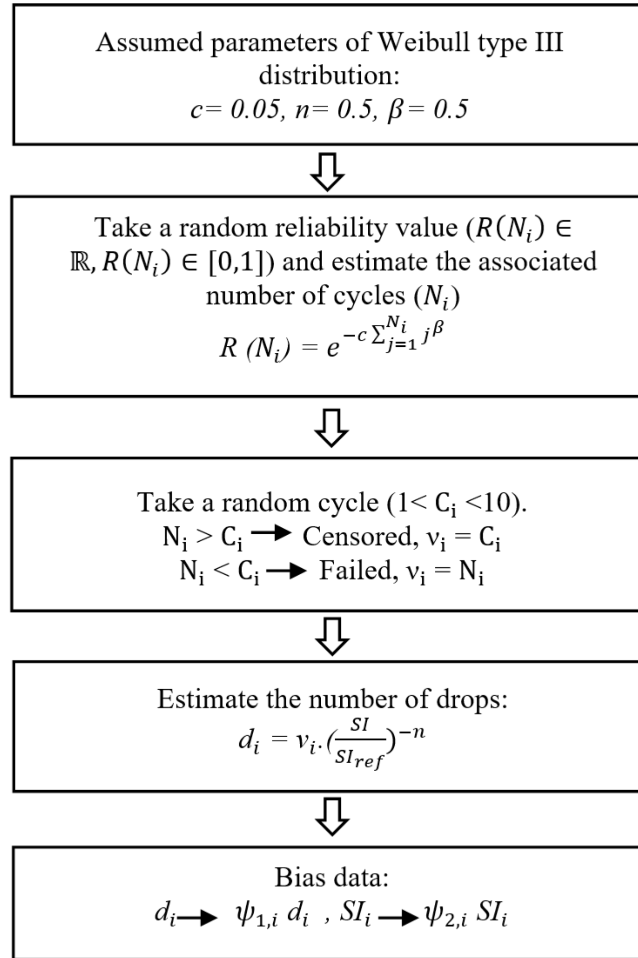


Figure 5.3. The process of simulating user survey data.

In the first step, the equivalent cycles of all surveyed and tested devices were calculated parametrically. For this purpose, the stress adjustors of the users' stress profiles were linearly scored and combined through a multiplicative SI model to calculate the SI_{use} values of the devices. Also, the reference stress profile was selected as a drop from the head or higher height on a hard surface during a harsh activity. Its SI value was calculated through the multiplicative SI model (SI_{ref}=1000). Because drop was mechanical stress (leading to damage), an IPL stress-life model was assumed. The ratio between the number of equivalent drops and the actual number of drops was

calculated. This resulted in Eq. (5-10), where ν_i is the equivalent cycle of the i-th device, d_i is the actual number of drops of the i-th device, $SI_{i_{use}}$ is the SI value of the i-th device, and SI_{ref} is the SI value of the reference stress profile equal to 1000.

$$\nu_i = d_i.(\frac{SI_{i_{use}}}{SI_{ref}})^n \tag{5-10}$$

In the second step, the bias from users' responses was removed. The survey and the test dataset of the similar product were split into two parts (30% in the first part and 70% in the second part [58]) and used the first portion of the two datasets to estimate the survey data bias. The parameters of the life distribution (i.e., the distribution of the equivalent cycles) of the tested and surveyed devices were estimated in an MLE analysis. The estimated parameters for the tested similar devices were β_{ts}= 0.43, c_{ts}= 0.046, n_{ts}= 0.50 and for the surveyed, similar devices were β_{ss}= 0.5, c_{ss}= 0.071, n_{ss}= 0.49. All equivalent cycles of the surveyed devices were then multiplied by an unknown number (i.e., λ) to shift their life distribution. The distance between the shifted life distribution of the surveyed devices and the life distribution of the tested devices was then calculated using the KL divergence method, Eq. (5-5), and minimized through the gradient descent algorithm. This resulted in the bias value of λ =1.32, which was close to the inverse of the average bias originally introduced to the survey dataset (ψ =1.397). According to Eq. (5-12) and because the bias values for d_i and $SI_{i_{use}}$ were randomly selected from $\mathcal{N}$ (0.80, 0.1), the original mean bias value was $0.80 \times 0.80^{0.5}$, which is approximately equal to 0.716, and its inverse is 1.397. To validate the estimated λ value, all equivalent cycles of the surveyed devices were multiplied by 1.32. The original value of n (n = 0.5) Was used to estimate the numerical value of the equivalent

cycles. Then, the reliability of the devices in the bias-reduced surveyed dataset was estimated using the Kaplan-Meier method and was compared with the true reliability model. As shown in Figure 5.4, the bias-reduced survey dataset follows the true reliability model (i.e., the solid black line). This provides evidence that the KL divergence method can remove the bias from users' responses. The estimated bias value was then multiplied by the equivalent cycles of the second portion of the surveyed devices to remove its bias.

In the third step, the second portion of the survey data (bias-reduced), the second portion of the similar device's test data, and the new device's entire test data were used in a sequential Bayesian analysis to estimate the reliability model of the new product. It was assumed that the equivalent cycles of all datasets followed discrete Weibull distribution type III. The likelihood of the discrete Weibull distribution type III for an incomplete dataset (i.e., a dataset containing failed and right-censored devices) is shown in Eq. (5-11), where c and β are the distribution parameters, ν_i represents the equivalent cycles of the i-th device, r is the number of failed devices, and m is the total number of devices.

$$L(c, \beta; T) = \prod_{i=1}^{r} e^{-c\sum_{j=1}^{\nu_i} j^{\beta}}[1\text{-}e^{-c(\nu_i+1)^{\beta}}].\ \prod_{i=r+1}^{m} e^{-c\sum_{j=1}^{\nu_i} j^{\beta}} \quad (5\text{-}11)$$

As discussed earlier, the upper bound of the summation term in the likelihood function was unknown and needed to be estimated before performing the Bayesian analysis. The ν_i in the upper-bound of the summation and ν_i in the exponential term of the likelihood function were distinguished; the ν_i in the upper-bound was replaced by ${\nu_g}_i^l$, where subscript g stands for guess, l shows the number of iterations, and i represents the device's number. The process of estimating ${\nu_g}_i^l$ is shown in Figure 5.5.

This process had five steps. First, the equation of the equivalent cycles was linearized. Second, an initial guess for the parameter of the equivalent cycle was made. This parameter is shown by ${n_g}^l$ in Figure 5.5.

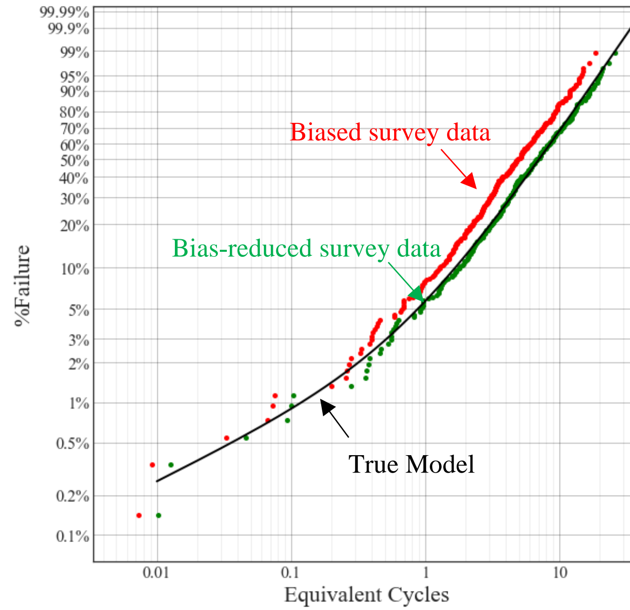


Figure 5.4. True reliability model, biased survey data, and bias-reduced survey data.

Third, the equivalent cycles of devices were calculated using Eq. (5-12), where d_i shows the actual number of drops for the i-th devices, SI_i is the SI value for the i-th device, SI_{ref} is the reference SI value, and ${n_g}^l$ is the estimated parameter of the model after l iterations. Fourth, ${v_g}_i^l$ was put in the upper bound of the summation term and the other parameters of the likelihood function were estimated through the MLE method. It was assumed that the MLE estimates were the true values of the parameters.

In the fifth step, Δn_l was calculated through the gradient descent algorithm, a learning rate of 0.1 was selected, and the value of ${n_g}^l$ was updated. The new value of ${n_g}^l$ was plugged into Eq. (5-12) again and the process of updating ${n_g}^l$ was continued until

convergence. Then, the value of n and consequently the upper bound was fixed, and sequential Bayesian analysis was performed to estimate the joint distribution of the parameters of the reliability model of the new product.

$$v_{g_i}^{\ l} = round(d_i . (\frac{SI_i}{SI_{ref}})^{n_g^{\ l}}) \text{ (5-12)}$$

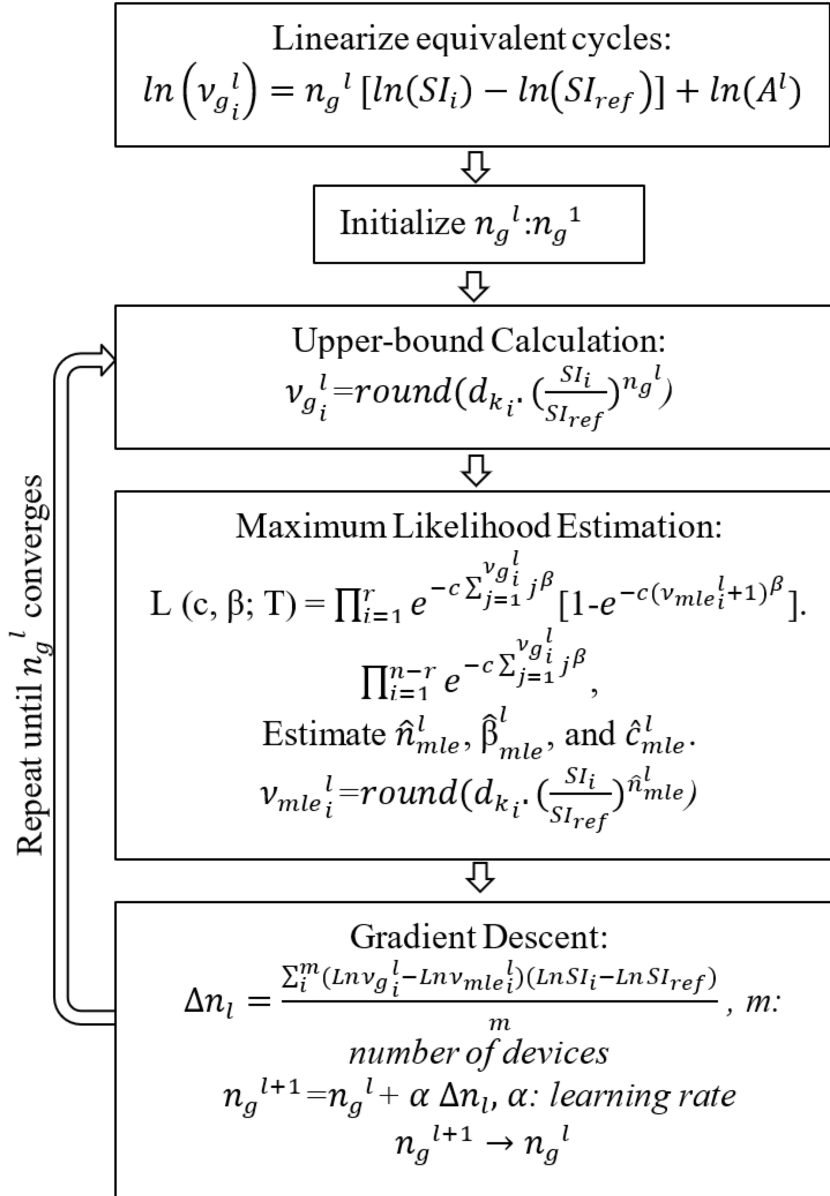


Figure 5.5. The process of estimating the upper bound of the summation term of the likelihood function of the discrete Weibull type III.

The sequential Bayesian analysis consisted of three regular Bayesian analyses. The first Bayesian analysis assumed weakly informative prior distributions for the parameters of the reliability model. These distributions were Normal distributions with the mean values estimated by the MLE method and standard deviations that were half-normal distributions with a mean of zero and standard deviation of 10% of the MLE estimates. The second part of the bias-reduced survey data was used as the observation in the first Bayesian analysis, and the primary posterior distributions of the parameters were estimated. The kernel density estimation (KDE) method was used for non-parametrically estimating the joint distribution of the parameters [59, 60]. In the second Bayesian analysis, the primary joint posterior distribution of the parameters was used as a prior distribution. The second portion of the test data is about a similar product used as the likelihood data. The intermediate posterior distributions of the parameters were estimated, and the corresponding joint distribution was approximated using the KDE method. Finally, the third Bayesian analysis used the intermediate joint posterior distribution as the prior distribution and the test data of the new product as observation and estimated the final posterior distributions of the parameters. Figure 5.6 shows the estimated reliability, mean, and true reliability models. The true reliability model is within the uncertainty region of the estimated models, and it is very close to the mean reliability model. Therefore, the proposed reliability estimation approach could adequately estimate the true reliability model.

The sensitivity of the estimated reliability model to the selected reference stress profile was evaluated. First, the reference stress profile was changed to a drop from a person's head or higher height on a hard surface during a benign (regular) activity.

Then the new product's reliability model was estimated. No noticeable change was observed in the estimated reliability model.

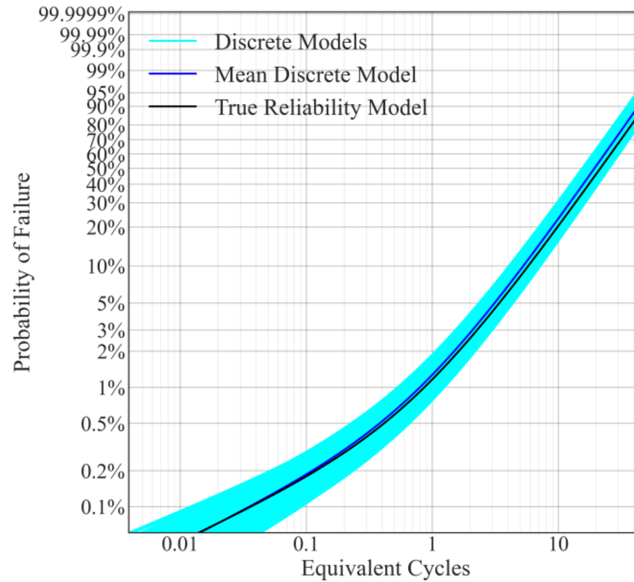


Figure 5.6. The estimated reliability models (discrete distribution analysis) and the True reliability model.

Also, the sensitivity of the reliability model to the selected SI model was evaluated. An additive SI model replaced the multiplicative SI model, and the proposed approach estimated the reliability model. The true model still was reasonably inside the region of the estimated reliability models. However, the maximum difference between the true reliability model and the mean reliability model over time was about 0.1 higher than the maximum difference obtained using the multiplicative SI model. This was because the additive SI model had three more parameters (the weights of the three stress adjustors) than the multiplicative SI model, which resulted in more uncertainties.

Also, the reliability model of the new product was estimated using the previous approach, which assumes a continuous Weibull distribution [54]. Figure 5.7 shows the

result of the continuous analysis, which used the same survey and test datasets. The difference between the continuous mean reliability model and the true reliability model and the difference between the discrete mean reliability model and the true reliability model over equivalent cycles are shown in Figure 5.8. As shown in Figure 5.8, the maximum difference between the discrete mean reliability model and the true model is 2.5%. This difference is more than three times lower than the maximum difference between the continuous mean and true reliability models. Therefore, the discrete life distribution model resulted in a better reliability estimate and is a better model for products with few cycles to failure.

This study assumed a single failure mode for the hypothetical product. In the presence of multiple failure modes, the proposed approach should be used to estimate the reliability model of each failure mode. Then, the total reliability model is measured as the events that the system is reliable under all failure modes. Thus, one should multiply the reliability models of all failure modes and estimate the total reliability model.

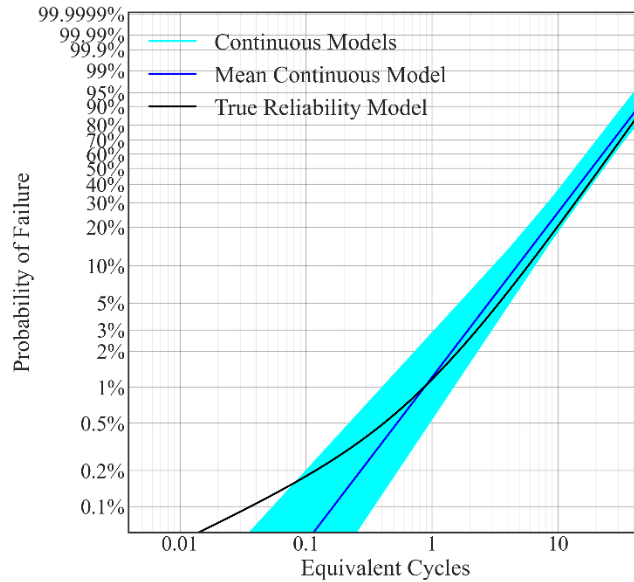


Figure 5.7. The estimated reliability models (continuous distribution analysis) and the True reliability model.

5.4. Conclusion

Previous studies used the concept of equivalent cycles. They estimated the reliability model of a new consumer product with large cycles to failure through a continuous lifetime distribution model using user survey data and reliability test data. This chapter introduced and applied the concept of equivalent cycles and developed a mathematical method to estimate the reliability model of a new product with a few cycles to failure through a discrete lifetime distribution model using user survey data and reliability test data.

This chapter showed that a user survey could be a cost-effective and quick way to collect field data for estimating the reliability model of a product having a few cycles to failure. Besides, it was shown that using a discrete life distribution can more accurately estimate the reliability model of a product having a few discrete cycles to failure than using a continuous distribution. However, the discrete distribution analysis

is computationally more involved because it requires additional steps related to the gradient descent algorithm and MLE.

The case study assumed a single failure mode. In the presence of multiple failure modes, one should multiply the reliability models of all failure modes and estimate the total reliability model. However, this case was not investigated in this study. This chapter suggested using x% of the survey and test data of a similar product for estimating the bias value and the remaining 1-x% for building prior distributions for the parameters of the reliability model of the new product. However, determining the value of x is out of the scope of this study. The value should be determined through an optimization analysis in the future. The proposed approach is generic and can estimate the reliability model of a wide range of consumer products such as laptops and monitors.

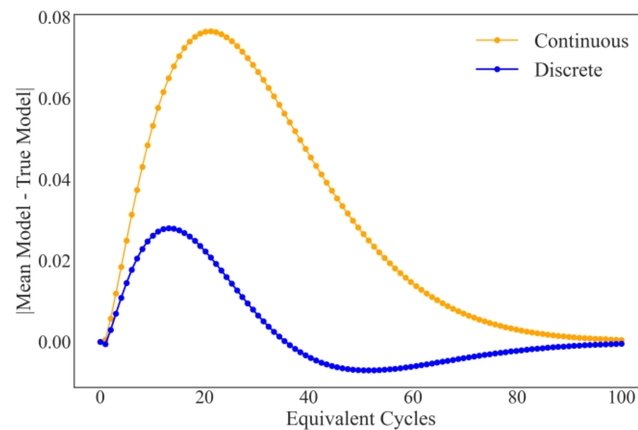


Figure 5.8. The Difference between the continuous mean reliability model and the true reliability model, and the difference between the discrete mean reliability model and the true reliability model over equivalent cycles.

Chapter 6: Designing Test Specification of a New Consumer Product

This chapter describes a procedure for designing test specifications for a new product with unknown failure modes. The proposed test design method is an extension of the bogey test [43], but this approach can estimate the reliability and its confidence intervals even if failures occur. The approach determines the sample size, stress levels, and their frequencies for reliability testing based on the target reliability, confidence level, manufacturer's constraints, including test duration and maximum sample size, and the actual use conditions of similar products collected by user survey. The similarity between the reliability of the new product and similar products depends on the extent to which the new product maintains a similar structure, material(s), and components [61].

The proposed test design approach uses some of the rules explained in our previous works like scoring the stress and stress adjustors and combining them through an additive or multiplicative stress-index (S.I.) model [62]. Although data about known failure modes was used in the previous chapters, the approach proposed here does not

consider a specific failure mode for the product. However, the stress, stress adjustors, and frequencies of drops are similar to the previous chapters.

This study designed two test specifications: (1) a frequency-accelerated test and (2) a stress-accelerated test. For both tests, the number of samples is determined by assuming a binomial distribution; and the use stress profiles of similar products are grouped through a multiplicative S.I. model as proposed in [62]. The multiplicative S.I. model multiplies the quantitative stress and stress adjustors of a stress profile together and delivers an S.I. value for each applicable stress profile. The S.I. values and the use frequencies of similar products are grouped using clustering methods. For the frequency-accelerated test, the grouped use frequencies are converted to the test frequencies and are applied to the samples during the test. In the stress-accelerated test, the grouped use conditions are replaced by some accelerated stress profiles, and their frequencies are determined using a known stress-life model. The accelerated stress profiles are applied to the samples during the test.

The rest of this chapter proceeds as follows. Chapter 6.1 summarizes the use conditions (S.I. values and their frequencies) using three clustering methods. Chapter 6.2 describes creating a table of all possible stress profiles and their S.I. values. Chapter 6.3 presents the method for designing the test specification for a frequency-accelerated and a stress-accelerated test. Chapter 6.4 illustrates the application of the approach using a simulated dataset for an electronic device that users accidentally drop. Chapter 6.5 concludes the chapter.

6.1. Summarize Use Stress Profiles

Our proposed approach utilizes the use conditions (i.e., the way owners are expected to use the product that leads to possible damaging stresses) of similar devices collected by user surveys to determine the test specification. As users may have many different use conditions, summarizing them into a small number of use condition groups will simplify the reliability test plan. Three clustering methods for this step are considered: (1) K-means clustering, (2) Gaussian mixture model (GMM), and (3) SI-cycle graph to group and summarize the use conditions.

The input data is a set of data points, where each data point has two values: (1) an S.I. (stress-index) value for one stress profile and (2) the frequency of occurrence (how frequently a user's device experienced that stress profile). The S.I. value is obtained through an additive or a multiplicative S.I. model which combines all scored stresses and stress adjustors of a stress profile [62]. The clustering approach yields a set of clusters, and the centroids of the clusters are taken as the grouped use conditions. (That is, each cluster has one grouped use condition.)

6.1.1. K-Means Clustering

K-means is a clustering method that allows finding groups of similar use conditions. K-means is computationally very efficient compared to the other clustering algorithm, but it does not have any mechanism to handle the uncertainties [63, 64]. The K-means algorithm performs clustering as follows. It first specifies K centroids and initializes their coordinates randomly. Then, it calculates the distance between the data points and the centroids to assign the data points to their nearest centroids. Finally, it updates the

coordinates of each centroid to the mean of the data points in the centroid's cluster. The elbow graph which plots the distortion (i.e., the average of the squared distances from the cluster centers) or inertia (i.e., the sum of squared distances of samples to their closest cluster center) versus the possible number of clusters is then used to assign K [63]. The centroids (geometric means) of the K clusters are the grouped use conditions.

6.1.2. Gaussian Mixture Model

The GMM is a clustering technique that uses a probabilistic assignment of data points to clusters and unlike the K-means algorithm considers uncertainties in clustering. The GMM algorithm performs clustering as follows. First, it specifies K multivariate Gaussian models (clusters) and randomly initializes their means and variances. Then, it calculates each data point's probability density function (PDF) using the existing Gaussian models and assigns the data point to the cluster with the highest PDF value. Finally, it updates the mean and variance of each cluster to the mean and variance of all data points assigned to that cluster. The trend of the Akaike information criterion (AIC) or Bayesian information criterion (BIC) over the number of clusters are then used to determine the number of groups, K, representing the number of multivariate models in the GMM. The optimum K is on the elbow of the graph. The centroids of the K clusters are known as the grouped use conditions.

6.1.3. SI-Cycle Graph

The SI-cycle graph is a two-dimensional graph that shows the S.I. values on one axis and the frequencies of the S.I. values (i.e., frequencies of the stress profiles) on the other axis. The area on the plot is divided into N equal elements. Each element contains

some data points. The number and shape of the elements are updated based on the optimized number of clusters (K) obtained by the K-means and GMM algorithm. The N-K elements with the least number of data points on the graph (scarcely occupied elements) are combined with their nearest neighbors. The nearest neighbor is the element with the closest boundary to the data points of the scarcely occupied element. This combination reduces the number of elements to K. The centroids of the K clusters are known as the grouped use conditions.

6.2. Table of Possible Stress Profiles and S.I. values

Each cluster's centroid is associated with an S.I. value and the frequency of its occurrence. These are called S.I. values and frequencies the "group S.I. values" and "grouped frequencies." The next step in our approach creates a table of all possible stress profiles and their corresponding S.I. values. Then, each group S.I. value is compared with the entries in the table, and the stress profile associated with the next higher S.I. value in the table is known as the "grouped stress profile." Thus, this step "translates" each group S.I. value to an appropriate stress profile that can be used to specify conditions for the reliability test.

6.3. Design Test Specifications

This subchapter designs test specifications for a new product with unknown failure modes. A frequency-accelerated and a stress-accelerated reliability test are proposed. Details about the tests and approaches for assigning their specification are discussed in subchapters 6.4.1 and 6.4.2.

6.3.1. Design Test Specifications for a Frequency-Accelerated Reliability Test

A frequency-accelerated reliability test applies the group usage stress profiles to the device but accelerates the group use frequencies. The test specification is determined based on the manufacturer's reliability requirements, including the desired warranty time, reliability level, confidence level, maximum number of units for the test, and test duration. The use conditions of similar devices are also needed to determine the test stress profiles and their frequencies of occurrence.

The procedure of designing test specifications for a frequency-accelerated test is shown in Figure 6.1. The various use conditions are grouped into K conditions using the clustering methods discussed in subchapter 6.1 and their associated stress profiles are estimated using the table of possible stress profiles and S.I. values introduced in subchapter 6.2. If the manufacturer desires to run the test under a smaller number of stress profiles than K, it may replace some of the stress profiles with the harsher profiles in the list of K profiles. This results in a rigorous reliability estimate. When the grouped stress profiles are determined, their corresponding frequencies of occurrence are multiplied by the ratio between the usage time window that the use frequency was calculated from it (e.g., 1 year) and the test duration (e.g., 1 week) to determine the frequencies of the stress profiles during the test (i.e., test frequencies). This is mathematically shown in Eq. (6-1).

$$\text{Test frequency}\left(\frac{\text{time(or cycle)}}{\text{test duration}}\right) = \qquad \text{Eq. (6-1)}$$

$$\text{usage frequency} \times \left(\frac{\text{time (or cycle)}}{\text{usage time window}}\right) \times \frac{\text{usage time window}}{\text{test duration}}$$

The number of test samples is determined using Eq. (6-2), where m is the number of samples, R_l is the lower-bound reliability determined by the manufacturer, 1-α is the confidence level, and l is the desired maximum number of failures when the test is complete [65].

$$1\text{-}\alpha=\sum_{i=0}^{l} \frac{m!}{i!(m-i)!}(1-R_l)^i \, R_l{}^{(m-i)} \qquad \text{Eq. (6-2)}$$

The reliability test is performed on m samples under the K (or smaller) stress profiles with the test frequencies. The test outcome is the number of failed (f) and right-censored (r) samples. If the number of failures is more significant than l, the actual reliability is less than the target reliability of the manufacturer. The point estimate of reliability is obtained from Eq. (6-3) where $\hat{R}$ is the point estimate of reliability, f is the number of failed samples when the test is complete, and m is the number of tested samples. The lower-bound reliability is calculated using the regularized incomplete Beta function, as shown in Eq. (6-4), where R_l is the lower-bound reliability, I_R is the regularized incomplete Beta function, m is the total number of tested samples, f is the number of failed samples, and $1-\alpha$ is the confidence level.

$$\hat{R} = 1 - \frac{f}{m} \qquad \text{Eq. (6-3)}$$

$$R_l = I_R(m-f,\ f+1) \le \alpha \qquad \text{Eq. (6-4)}$$

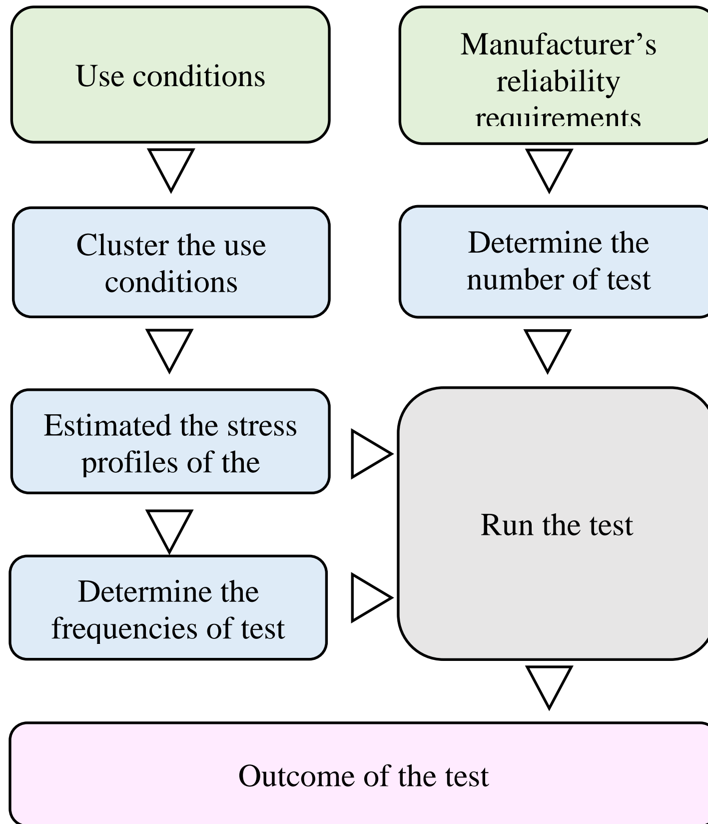


Figure 6.1 procedure of designing test specification for a frequency-accelerated test.

6.3.2. Design Test Specifications for a Stress-Accelerated Reliability Test

The stress-accelerated reliability test is performed at stress levels higher than the use stress levels. This test requires an additional input compared to the frequency-accelerated test which is the underlying stress-life model of the class of products. To run the test, the manufacturer selects some harsher stress profiles (i.e., accelerated stress profiles) than the grouped stress profiles. The stress-life model is then used to convert the frequency of the grouped stress profiles into the frequency of the accelerated stress profiles.

The procedure of designing test specifications for a stress-accelerated test is shown in Figure 6.2. Similar to the frequency-accelerated test, the use conditions are grouped using the clustering methods discussed in Subchapter 6.1. The centroids of the clusters show the group S.I. values and the grouped frequencies. The grouped stress profiles are determined using the table of possible stress profiles and S.I. values introduced in Subchapter 6.2. The manufacturer then selects several accelerated stress profiles which are harsher than the grouped stress profiles. The grouped frequencies are converted into the equivalent frequencies (frequencies of the accelerated stress profiles) using the underlying stress-life model, as shown in Eq. (6-5), where P is the cumulative density function (CDF), t_s is the grouped frequency, SI_s is the S.I. value of the grouped stress profile, v_a is the accelerated frequency, and SI_a is the S.I. value of the accelerated stress profile. The accelerated frequencies are then converted into the test frequencies using Eq. (6-6).

$$P(t_s, SI_s) = P(v_a, SI_a) \qquad \text{Eq. (6-5)}$$

$$\text{Test frequency}\left(\frac{\text{time(or cycle)}}{\text{test duration}}\right) \qquad \text{Eq. (6-6)}$$

$$= \text{accelerated frequency} \times \left(\frac{\text{time (or cycle)}}{\text{accelerated time window}}\right)$$

$$\times \frac{\text{accelerated time window}}{\text{test duration}}$$

The number of samples for the stress-accelerated test (m) is determined using Eq. (6-2). Then, the stress-accelerated test is performed on m samples with their associated test frequencies under the accelerated stress profiles. The outcome of the test is the

number of failed (f) and right-censored (r) units. The point estimate reliability and the lower-bound reliability are estimated using Eq. (6-3) and Eq. (6-4).

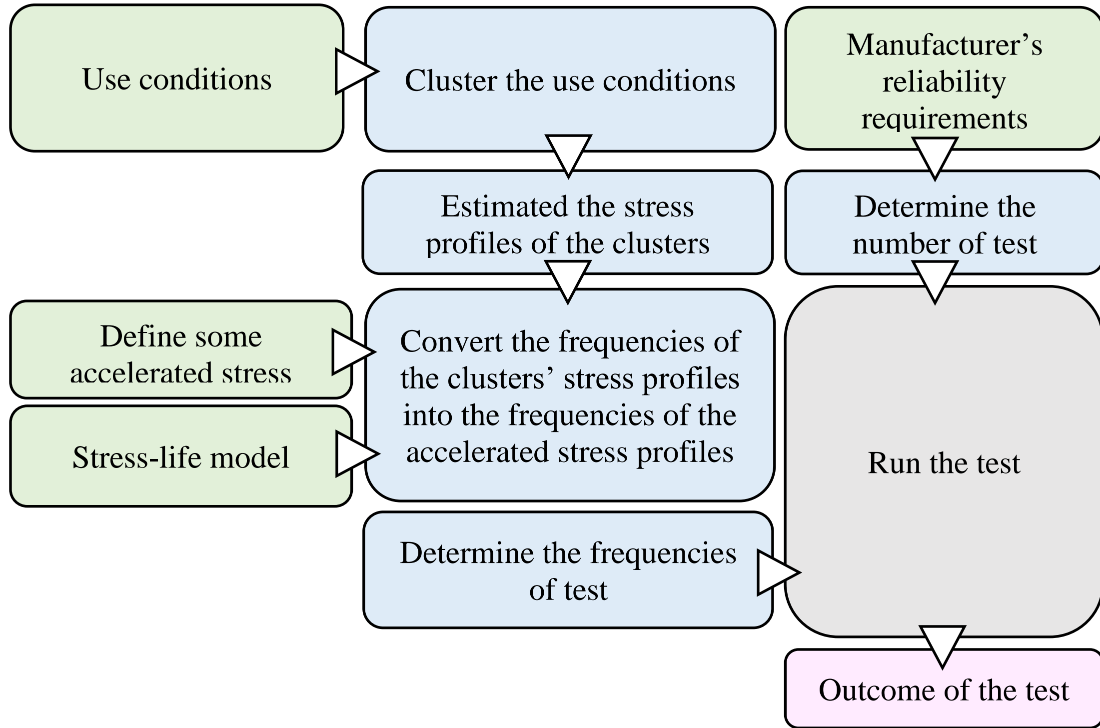


Figure 6.2 procedure of designing test specification for a stress-accelerated test.

6.4. Case Study

This subchapter illustrates the application of the proposed approach using a simulated user survey dataset for an electronic device that users accidentally drop. For this product, the applied stress is a drop during use. Three stress adjustors that can describe the relevant quantitative stress levels are (a) the drop height, (b) the type of surface on which the device dropped, and (c) the type of user's activity when the drop occurred. Weather conditions are not considered a significant stress adjustor for this product [62]. The process of simulating the user survey dataset is explained in detail in

our previous study [62]. The dataset contains 1000 users (24% young, 63% middle-aged, and 13% senior users). The fraction of age groups and the height of users in the groups are consistent with the US population data [48, 49, 50, 51]. The users dropped their devices under various use conditions (i.e., stress profiles) many times during their ownership. The ownership times were randomly drawn from a discrete distribution that contained 400 ownership times of 1 year, 300 ownership times of 2 years, 200 ownership times of 3 years, and 100 ownership times of 4 years.

The qualitative drop heights in the user survey were knee height, waist height, chest height, and head or higher height; the qualitative surface types were soft, semi-soft, semi-hard, and hard surface; and the qualitative activities were benign and harsh activity. The stress adjustors were scored between 0 and 100 using the method explained in [62]. The quantitative stress adjustors were combined through a multiplicative S.I. model, as shown in Eq. (6-7), to estimate an S.I. value for each stress profile.

$$SI = \prod_{i=1}^{3} s_i \qquad \text{Eq. (6-7)}$$

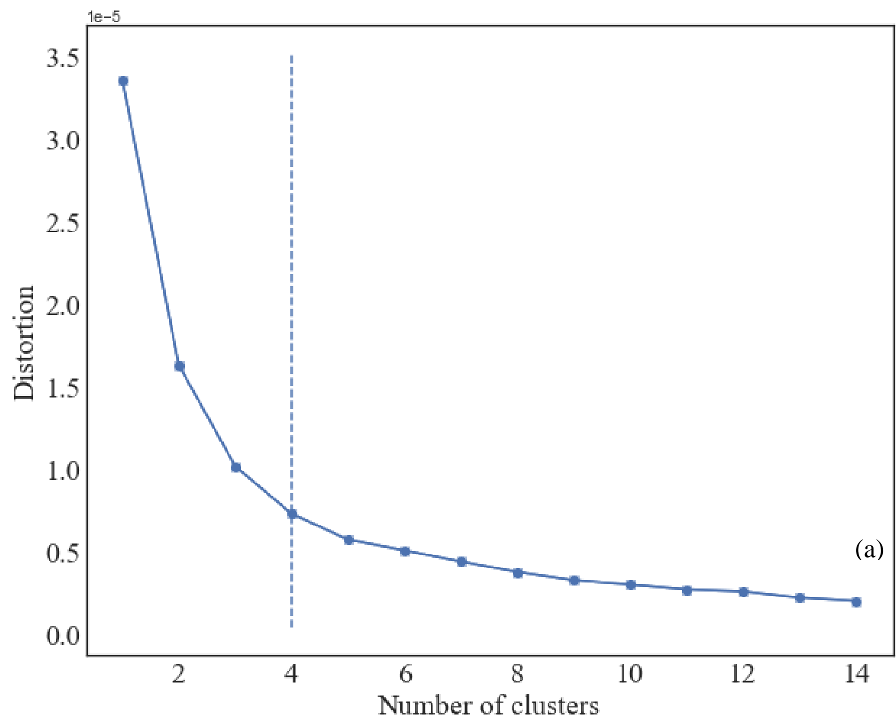
6.4.1. Summarized Use Conditions

The use conditions were grouped using K-means clustering, GMM, and SI-cycle graph. For K-means clustering, the elbow graph, as shown in Figure 6.3 (a), along with the Kneedle algorithm [66] was used to determine the best number of clusters. The elbow graph shows the trend of distortion versus the number of clusters. The best number of clusters is at the elbow graph's knee point, and the Kneedle algorithm

estimates the location of the knee point. Using this method, the best number of clusters was estimated as 4. Then, using the K-means algorithm the use conditions were divided into 4 clusters, as shown in Figure *6.3* (b). The centroids of the clusters (i.e., the geometric mean of the S.I. values and frequencies) represent the grouped use conditions.

For GMM, the BIC (or AIC) trend versus the number of clusters, as shown in Figure 6.4 (a), along with the Kneedle algorithm, was used to determine the best number of clusters. This analysis resulted in 4 clusters. The GMM was then used to divide the data into 4 clusters, as shown in Figure *6.4* (b). Each cluster in Figure *6.4* (b) has three shaded parts which show the six-sigma region of the cluster's Gaussian mixture distribution.

For the SI-cycle graph, first, the data were arbitrarily divided into 9 identical regions, as shown in Figure 6.5 (a). This resulted in four scarcely occupied regions (i.e., regions 2, 3, 6, and 9 in Figure 6.5 (a)). These regions were combined with their nearest neighbors and the number of regions was reduced from 9 to 4, as shown in Figure 6.5 (b). The final number of clusters is consistent with the number of clusters for the K-means and GMM algorithm.



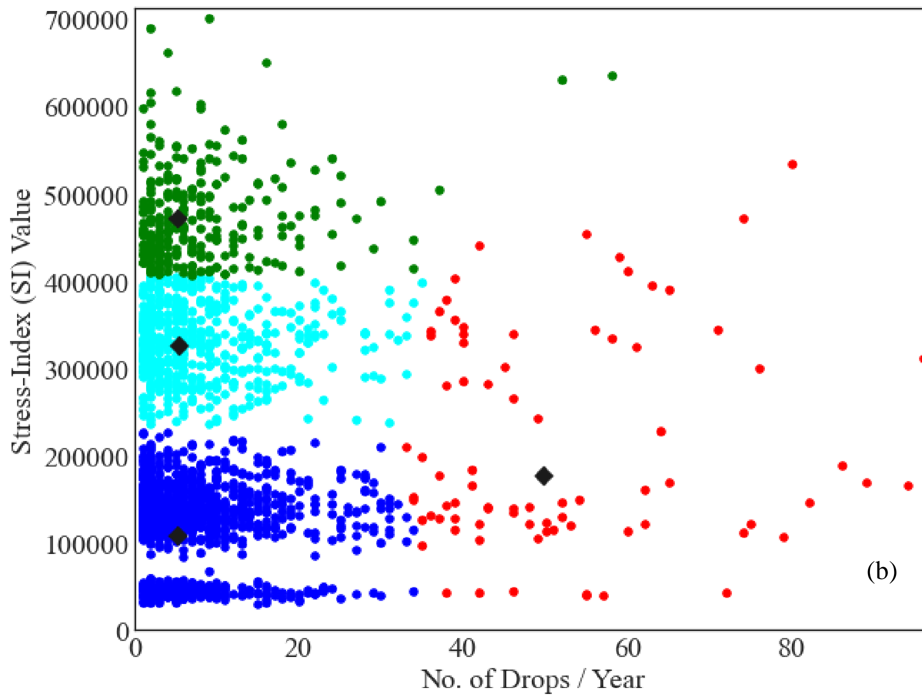
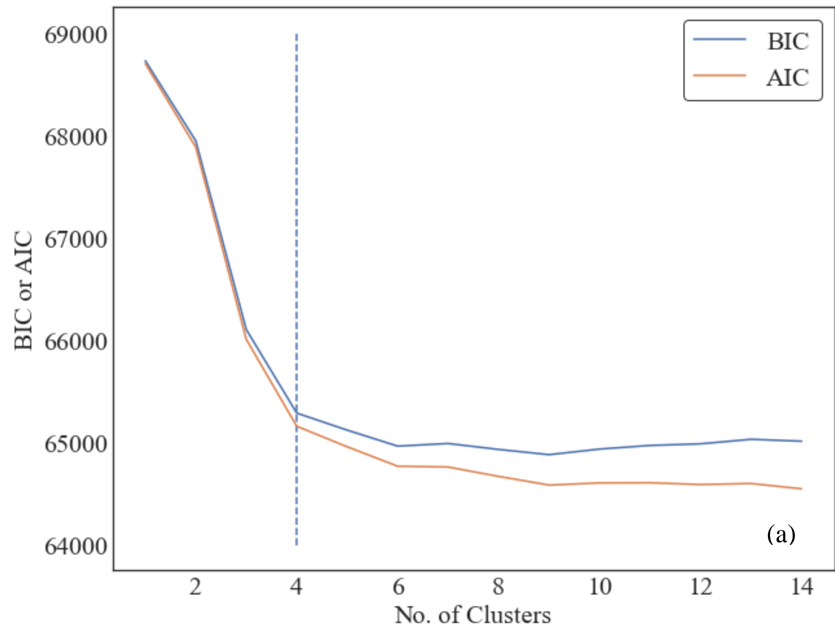


Figure 6.3 (a) the elbow graph, (b) the result of clustering using the K-means algorithm.



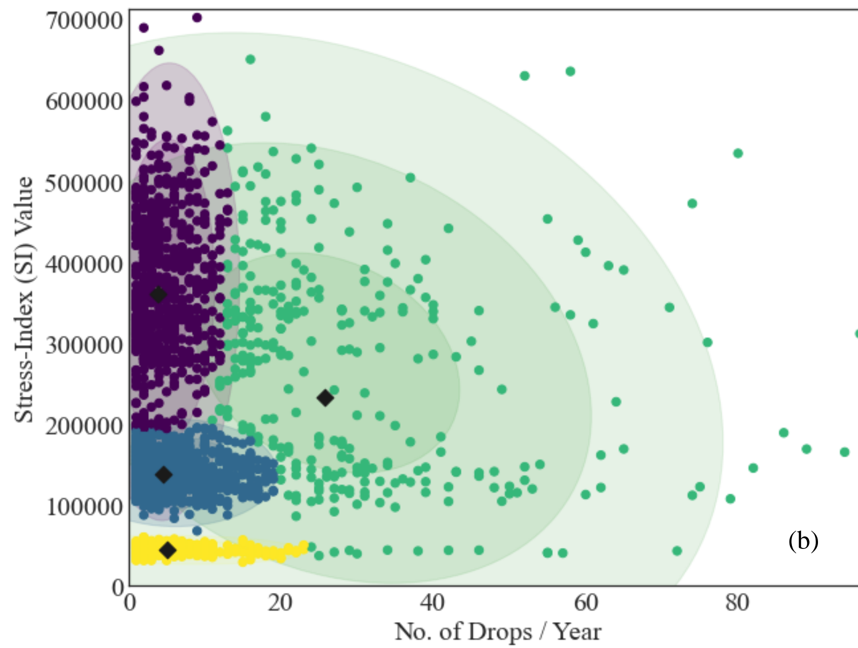
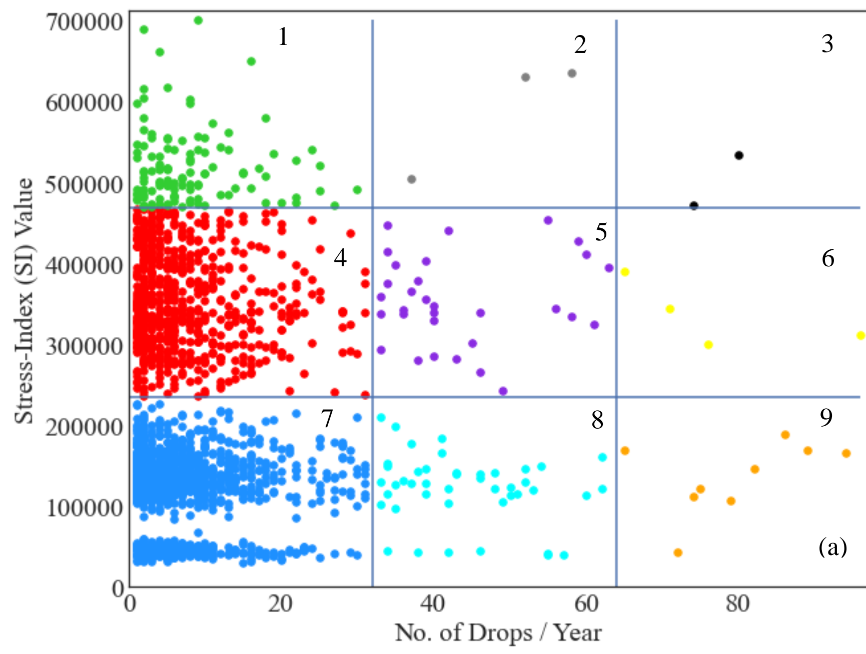


Figure 6.4 (a) trend of BIC and AIC vs. the number of clusters, (b) the result of clustering using the GMM algorithm.



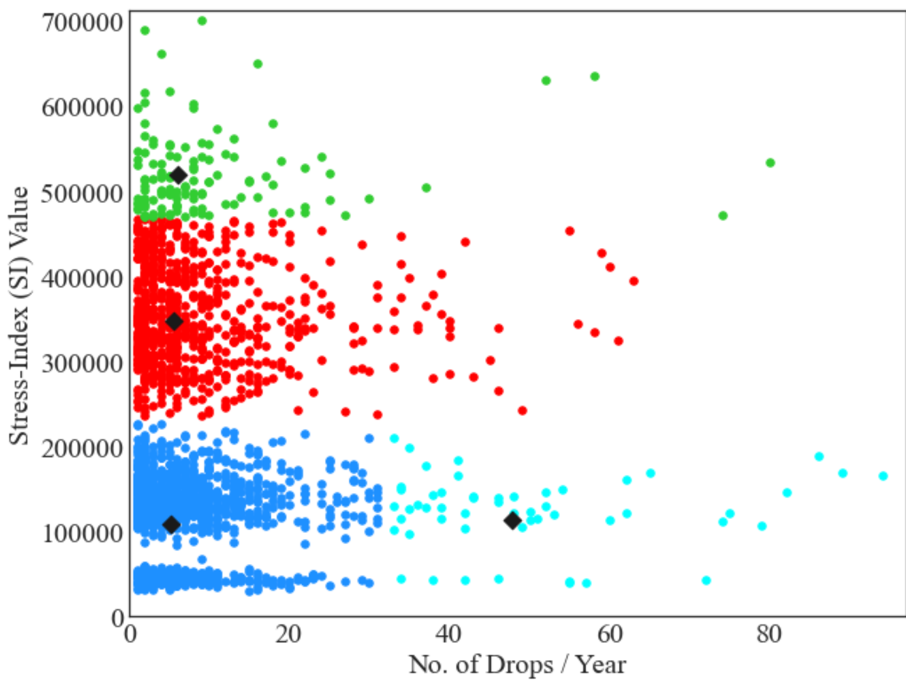


Figure 6.5. (a) SI-cycle graph with arbitrary divisions, (b) SI-cycle graph with 4 clusters.

Figure 6.6 compares the grouped use conditions (the centroids of the clusters) obtained by the three methods. The centroids were put into 4 groups which are shown

by black ellipses in the figure. In 3 out of the four groups (i.e., groups 1, 2, and 3), the centroids estimated by the GMM had the lowest S.I. values. Besides, in all four groups, the GMM resulted in the smallest number of drops per year. Therefore, the grouped use conditions obtained by GMM are optimistic. In 3 out of the four groups (i.e., groups 1, 2, and 3), the centroids estimated by the SI-cycle graph had the highest S.I. values and number of drops per year. Thus, the SI-cycle graph results in pessimistic grouped use conditions. The grouped use conditions obtained by K-means are moderate because the S.I. value or/and the number of drops estimated by K-means clustering are usually between the values estimated by the other two methods.

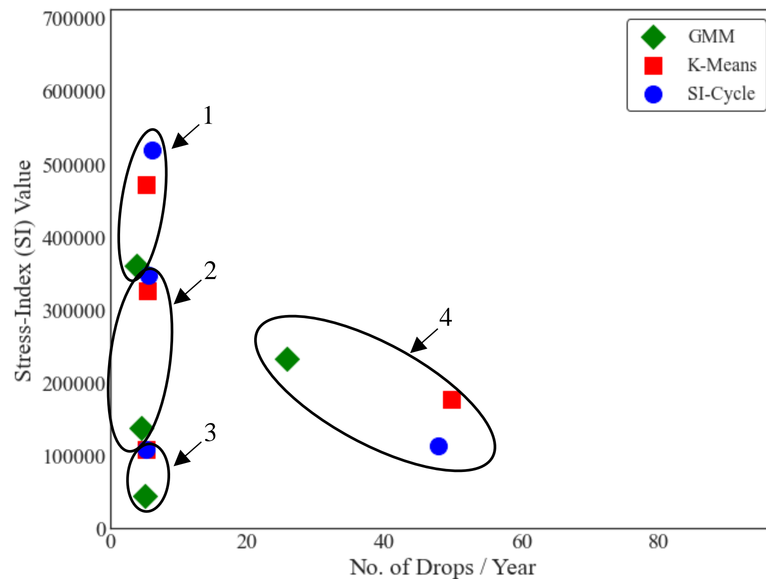


Figure 6.6 The grouped use conditions estimated by K-means clustering, GMM, and SI-cycle graph.

The pessimistic grouped use conditions obtained by the SI-cycle graph were used to infer the group use stress profiles because they resulted in the most rigorous reliability estimate. In order to infer the grouped stress profiles, the table of all possible stress profiles and their S.I. values was built, as shown in Table 6-1. Because there were four height choices (i.e., knee, waist, chest, and head or higher), four surface choices

(i.e., soft, semi-soft, semi-hard, and hard), and two activity choices (i.e., benign, and harsh) in the user survey, in total there were 32 different possible combinations of them ($\binom{4}{1} \times \binom{4}{1} \times \binom{2}{1} = 32$). Each combination is a possible stress profile that a user in the field may observe. The list of all 32 combinations is shown in Table 6-1. To calculate the S.I. values in the table, the scores for knee, waist, chest, and head (or higher) height were assumed as 25, 50, 75, 100, for soft, semi-soft, semi-hard, and hard surfaces were assumed as 25, 50, 75, 100, and for benign and harsh activity were assumed as 50 and 100. The height scores are consistent with the scores used for the middle-aged group in the user survey dataset. As the middle-aged group has the highest scores in the dataset, these scores result in a rigorous reliability estimate. The scores associated with the stress adjustors of the 32 stress profiles were combined through a multiplicative S.I model, Eq. (6-7), to find the S.I. values in Table 6-1. Table 6-1 shows some stress profiles with similar S.I. values. It is assumed that the device is equally damaged under the stress profiles with the same S.I. value.

Table 6-1 Table of all possible stress profiles and their SI values.

Stress Profile	**S.I. Value**	**Stress Profile**	**S.I. Value**
Knee- soft- benign	31,250	Waist- hard- benign	250,000
Knee- semisoft- benign	62,500	Waist- semisoft- harsh	250,000
Knee- soft- harsh	62,500	Head- semisoft- benign	250,000
Waist- soft- benign	62,500	Head- soft- harsh	250,000
Knee- semihard- benign	93,750	Chest- semihard- benign	281,250
Chest- soft- benign	93,750	Chest- hard- benign	375,000
Knee- hard- benign	125,000	Chest- semisoft- harsh	375,000
Knee- semisoft- harsh	125,000	Head- semihard- benign	375,000
Waist- semisoft- benign	125,000	Waist- semihard- harsh	375,500
Waist- soft- harsh	125,000	Head- hard- benign	500,000
Head- soft- benign	125,000	Head- semisoft- harsh	500,000
Knee- semihard- harsh	187,500	Waist- hard- harsh	510,000
Waist- semihard- benign	187,500	Chest- semihard- harsh	562,500
Chest- semisoft- benign	187,500	Chest- hard- harsh	750,000
Chest- soft- harsh	187,500	Head (or higher)- semihard- harsh	750,000
Knee- hard- harsh	250,000	Head (or higher)- hard- harsh	1,000,000

Each group S.I. value was compared with all S.I. values in Table *6-1* and selected the stress profile with the next higher S.I. value as the grouped stress profile. For instance, in Figure 6.7, the next higher S.I. value to the S.I. value of centroid 1 was 562,500 which belonged to a drop from chest height on a semihard surface during a harsh activity. The other grouped stress profiles were estimated using the same scenario and were listed in Table 6-2.

Table 6-2 The grouped stress profiles.

Centroid No. in Figure 6.7	Grouped stress profile	Grouped Frequency of Use
1	Chest- semihard- harsh	6 drops in 1 year
2	Chest- hard- benign Chest- semisoft- harsh Head- semihard- benign Waist- semihard- harsh	6 drops in 1 year
3	Knee- hard- benign Knee- semisoft- harsh Waist- semisoft- benign Waist- soft- harsh Head- soft- benign	6 drops in 1 year
4	Knee- hard- benign Knee- semisoft- harsh Waist- semisoft- benign Waist- soft- harsh Head- soft- benign	49 drops in 1 year

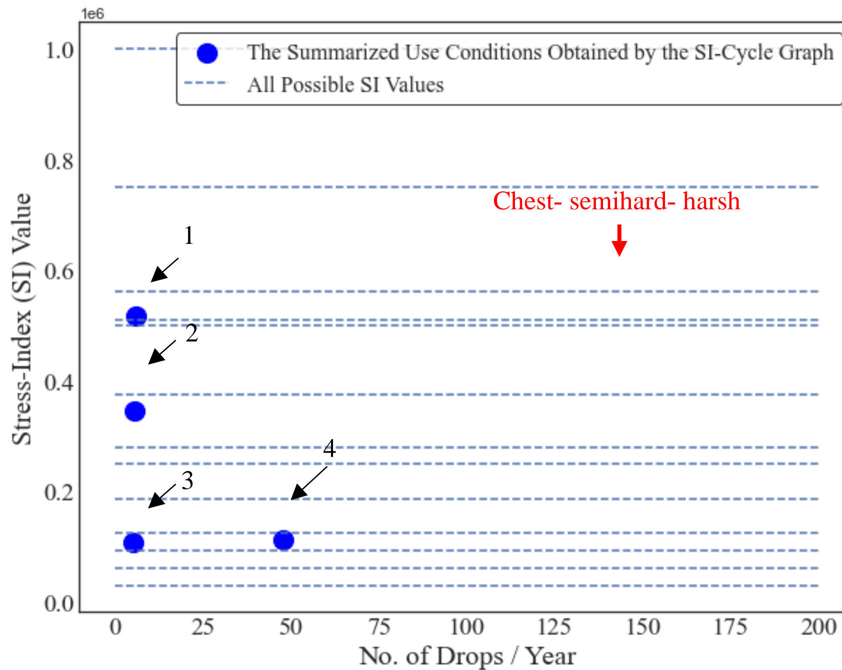


Figure 6.7 The grouped use conditions and all possible S.I. values.

In the cases where the next higher S.I. value belongs to several stress profiles, each can be selected as the grouped stress profile because it is assumed that the device is

equally damaged under all those stress profiles. The grouped stress profiles along with the grouped frequencies are used in Subchapter 6.4.2 and Subchapter 6.4.3 to design the test specification for a frequency-accelerated and a stress-accelerated test.

6.4.2. Design Test Specifications for a Frequency-Accelerated Reliability Test

A frequency-accelerated test requires three elements which are (1) the group use stress profiles, (2) the test frequencies, and (3) the number of test samples (or the allowed number of failures if the manufacturer has decided the maximum number of samples). The first element was determined in Subchapter 6.4.1 and listed in Table 6-2. The test frequencies were calculated using Eq. (6-1) and by assuming the test duration of 1 week. These frequencies are 6, 6, 6, and 49 drops in one week for the grouped stress profile 1 to 4 listed in Table 6-2, respectively.

It was assumed that the manufacturer wanted to achieve at least 95% reliability with 90% confidence after one year of warranty, and the maximum number of test samples was 100. The allowed number of failures was estimated as two samples by substituting these values in Eq. (6-2). If the number of failures after completing the frequency-accelerated test is less than 2, the product meets the target reliability. Otherwise, a root cause analysis is needed to understand and resolve the reason for out-of-specification. For instance, if the number of failures is 3, the point estimate of reliability is 97% and the lower-bound reliability is 93.44% which is less than the minimum desired reliability of the manufacturer. Therefore, a root cause analysis should be conducted to understand the reason for out-of-specification and appropriate actions should be performed to improve the product’s reliability.

6.4.3. Design Test Specifications for a Stress-Accelerated Reliability Test

A stress-accelerated test requires three elements which are (1) the accelerated stress profiles, (2) the equivalent frequencies (i.e., the frequencies of the accelerated stress profiles), and (3) the number of test samples (or the allowed number of failures if the manufacturer has decided the maximum number of samples). The accelerated stress profiles are harsher than the grouped stress profiles and are decided by the manufacturer. For instance, the accelerated stress profiles for this case study can be the profiles shown in

Table 6-3. These accelerated stress profiles have at least one harsher stress adjustor than their corresponding grouped stress profiles listed in Table 6-2.

The next step is calculating the equivalent frequencies for the accelerated stress profiles such that the accelerated test conditions cause the same amount of damage as the grouped use conditions. The underlying stress-life model of the class of products is needed to calculate the equivalent frequencies. An inverse power law (IPL) stress-life model with known parameters was assumed. This model is shown in Eq. (6-8), where *N* and *SI* represent the frequency of drops and the S.I. value, respectively. Equation (6-9) was obtained by the ratio between the frequency of an accelerated stress profile and the frequency of a grouped stress profile where ν and SI_a are the equivalent frequency and the S.I. value of the accelerated stress profile, and d and SI_a are the frequency and the S.I. value of the grouped stress profile. The SI_a and SI_s were calculated using Eq. (6-7). These frequencies are smaller than the grouped frequencies and thus reduce the testing time. The test frequencies were estimated using Eq. (6-66) and are listed in

Table 6-3. The allowed number of failures, the point estimate reliability, and the lower-bound reliability for the stress-accelerated test are estimated using the same equations, and the same scenario explained for the frequency-accelerated test.

$$N = 300. SI^{-1.2} \quad \text{Eq. (6-8)}$$

$$\nu = d. (\frac{SI_s}{SI_a})^{1.2} \quad \text{Eq. (6-9)}$$

Table 6-3 The test conditions of the stress-accelerated test.

Centroid No. in Figure 6.7 and Table 6-2	Accelerated Stress Profile	Test Frequency
1	Head- hard- harsh	3 drops in 1 week
2	Head- hard- benign	4 drops in 1 week
3	Chest- hard- harsh	1 drop in 1 week
4	Chest- hard- benign	13 drops in 1 week

6.5. Conclusions

This chapter showed that the reliability test specification of a new product with unknown failure modes could be designed based on the usage conditions of similar products collected using a reliability-informed user survey. A test specification based on the user data reveals the failure modes observed in the field. The user survey is a cost-effective and quick way to collect the use conditions. A frequency-accelerated and a stress-accelerated test were proposed. The test stress profiles, their frequencies, and the allowed number of failures were determined based on the manufacturer's reliability requirements including the warranty time, desired reliability, confidence level, test duration, and the maximum number of samples. The various use conditions collected

by the user survey data were grouped through the K-means clustering, GMM, and SI-cycle graph and were used to determine the test frequencies and the test stress profiles. The stress-accelerated test required an additional input compared to the frequency-accelerated test but delivered a shorter testing time. Designing test specifications was illustrated using a simulated dataset for an electronic device that users in the field accidentally dropped. Our case study showed that the SI-cycle graph resulted in the most pessimistic grouped use conditions and thus delivered the most rigorous test specification and reliability estimate.

Chapter 7: Summary and Conclusions

Chapter 7 summarizes the dissertation and research contributions and recommends future works.

7.1. Summary of contributions

1. This dissertation identified the critical elements of a reliability-informed user survey and offered recommendations for designing the survey.
2. The concept of “equivalent damage cycles” was introduced. It considered the effect of all applicable damage cycles in reliability analysis and reduced the uncertainty of the estimated reliability.
3. This research introduced a novel formal approach for applying the user survey data in a reliability estimation analysis. The approach estimated the reliability model of a new product (with no field experience) using user survey data and reliability test data of a similar product (e.g., an older version of the new product) and the reliability test data of the new product. A continuous lifetime distribution was assumed for products with large damage cycles to failure. The actual number of cycles under various use conditions was converted into the equivalent number of

cycles under a reference use condition. Then, the bias in user responses was reduced by (1) measuring the difference between the lifetime distributions of surveyed and tested devices of the similar product through the KL divergence method, (2) minimizing the difference using an optimization technique such as gradient descent algorithm to estimate the bias value, and (3) multiplying all equivalent cycles of the surveyed devices by the bias value. Then, the approach estimated the reliability model's parameters through a sequential Bayesian analysis.

4. The research extended survey and test-based reliability estimation for products with several damage cycles to failure. The extension relied on discrete forms of parametric distributions to describe such a product's lifetime distribution.
5. The introduction of a novel approach for designing test specifications for a new product with unknown failure modes was proposed.

7.2. Conclusions

1. This dissertation showed that user survey data could be a cost-effective and quick way to collect field data for estimating the reliability model of a consumer product.
2. Users usually cannot describe the damage on their devices as some damages are hidden or hard to approximate; however, users roughly know the conditions (stresses) that caused the damage. Therefore, a stress-based model is needed to estimate the reliability model of a consumer product using user survey data.
3. Stresses, stress adjustors, usage times, failure times, and failure modes were determined as the critical elements of a reliability-informed user survey.

4. The biases in user responses can be reduced or removed through the KL divergence method.
5. This study showed that using user survey data of a similar product as prior information in a Bayesian analysis could reduce the uncertainty of the estimated reliability of the new product.
6. A continuous distribution should estimate the reliability of a product with large cycles to failure. In contrast, a discrete distribution estimates the reliability model of a product with a few cycles to failure.
7. Designing a reliability test plan based on user survey data of a similar product in use reveals the most common failure modes of the new product. It saves time and money by disregarding the failure modes that do not happen during the actual use conditions.
8. The developed reliability estimation approach was applied to the real test and user survey datasets collected by Amazon lab 126 for some electronic handheld devices. The results of the simulated datasets were consistent with the results of the real datasets. Due to confidentiality reasons, this dissertation only presented the results of the simulated datasets.
9. The developed reliability estimation approach is applicable when the reliability data about a similar product (with a small variation from the new product) is available. The reliability data about the products with drastic variations from the new product cannot be used as prior information.

7.3. Future Works

This subchapter proposes several future works to expand this research.

1. This dissertation estimated the reliability model of a new product and its uncertainty using the reliability test data of the new product and the user survey data and reliability test data of a similar product. However, the sample size was not optimized in this study. Increasing the sample size reduces the uncertainty of the estimated reliability but increases the manufacturer's cost associated with testing more samples and surveying more users. Finding the best trade-off between the sample size and accuracy regarding the manufacturer's requirement could be a possible research direction to improve the current study. Previous studies that determined the samples size based on the manufacturer's requirement (e.g., the ratio or the difference between the upper-bound and lower-bound of the reliability model at a given time like the warranty time) and found the best trade-off between the sample size and accuracy could be a start point [44, 67].
2. In Chapter 5, a test and a user survey dataset were simulated for a product with a few damage cycles to failure. The reliability model was then estimated using a continuous and a discrete lifetime distribution. It was shown that the discrete lifetime distribution better estimated the reliability model of the product. A future study can simulate datasets with many damage cycles to failure and compare the reliability models estimated by a discrete and a continuous lifetime distribution.

3. Another potential future work can be preparing a lookup table that quantifies the amount of bias in user survey responses for different classes of products and different types of stresses. Many manufacturers can use the lookup table because the bias in user responses is independent of the manufacturer. Previous studies showed that the bias in user responses depends on the event's importance and the occurrence frequency. For instance, a user may accurately recall the events when she/he spilled a liquid on her/his cell phone in the past month. In another world, the user may accurately remember the number of spills, the type of liquid (e.g., water, coffee, soda), and the rough temperature of the liquid. However, the user may not accurately remember the number of times she/he pressed the volume button of the cell phone or the applicable pressures. The bias about these events does not depend on the device's manufacturer. Therefore, research can be performed to classify different products and stresses into several levels. Then, a generic lookup table that illustrates the bias values of different stress levels and products can be prepared using real-world reliability test datasets, user survey datasets, and the proposed bias estimation approach.
4. This study used continuous Weibull distribution and discrete type III Weibull distribution to illustrate the application of the proposed reliability estimation approach. Other lifetime distributions such as Normal, lognormal, and exponential distribution can be used to demonstrate the approach's generality.
5. This study selected the scoring and stress-life models through a sensitivity analysis. However, model selection methods such as BIC and AIC can be used

as comparison metrics to select the best combination of the scoring, lifetime distribution, and stress-life models. These methods use the model's log-likelihood, the number of parameters in the model, and the data size to score various model combinations and select the combination that best fits the data.

6. This study assumed that the similar product is an older version of the new product. However, sometimes an older version of the new product does not exist. Determining a similarity index that describes how much a new product is similar to the existing products allows using the reliability data of other products to build a (weighted) prior distribution for Bayesian analysis of the new product.
7. The case studies used normal distributions for simulating variation in the (scored) stress adjustors. However, using normal distributions rarely may generate random scores outside of the valid score range. Using beta distribution with parameters that are close to a symmetric bell shape model avoids this problem.

Appendix A

This appendix presents the examples of the python codes developed to obtain the results of this dissertation.

Appendix A.1: Python Code for Estimating the Reliability Model Using Continuous Distribution

This section presents an example of the python code developed to obtain the results provided in Chapter 4 of this dissertation.

```
# Import libraries:
import warnings
import pandas as pd
import arviz as az
import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np
import pymc3 as pm
import theano.tensor as tt
from pymc3 import Model, Normal, Slice, sample
from pymc3.distributions import Interpolated
from scipy import stats
from theano.compile.ops import as_op
from google.colab import files
uploaded = files.upload()
import io
# Import Data:
DF= pd.read_csv(io.BytesIO(uploaded['Survey1_1_15_22(1800-1.5-0.5)_multiplicative.csv']))
df1 = pd.read_csv(io.BytesIO(uploaded['Test1(A1800_B1.5_n0.5)_1_13_2022.csv']))
df2 = pd.read_csv(io.BytesIO(uploaded['Test2(A2200_B1.5_n0.5)_1_13_2022.csv']))
DF1= pd.read_csv(io.BytesIO(uploaded['DF1.csv']))
DF2 = pd.read_csv(io.BytesIO(uploaded['DF2.csv']))
DF3 = pd.read_csv(io.BytesIO(uploaded['DF3.csv']))
DF4 = pd.read_csv(io.BytesIO(uploaded['DF4.csv']))

# Define the kernel density function:
def from_posterior (param, samples):
    smin, smax = np.min(samples), np.max(samples)
    width = smax - smin
    x = np.linspace(smin, smax, 100)
```

```
    y = stats.gaussian_kde(samples)(x)
    x = np.concatenate([[x[0] - 3 * width], x, [x[-1] + 3 * width]])
    y = np.concatenate([[0], y, [0]])
    return Interpolated(param, x, y)

#Collect first half of the survey data about the similar device
scale_value=1
ns1 = round (DF1['biased drops'],2) # ns1 is the no. of drops
SIs=round (DF1['SIs']/scale_value,2) # SIs is the SI value of the first portion of the surveyed
devices

# Build the log-likelihood using the first portion of surveyed devices:
def logp (SIs, ns1):
    summ1 = 0
    for i in range(0, len(DF1)):
        SI=1000
        print(i)
        F=DF1['failure'][i] # Collect the failure status (healthy/failed) of the first portion of surveyed.
devices
        nu=(ns1[i])*(SIs[i]/SI)**n # Calculate the equivalent cycles.
        PDF = (B*nu**(B-1))/(A*SI**-n)**B # PDF is used to build the log-likelihood function.
        R = np.exp(-(nu/(A*SI**-n))**B) # R is used to build the log-likelihood function.
        logLik = (np.log ((PDF**F)*R)) # Log-likelihood function
        summ1 += logLik # Sum of the logLik.
    return (summ1)

# Bayesian estimation using the first portion of the surveyed devices:
with pm.Model() as model_ss1:

    MuB = 1.49
    SigmaB= pm.HalfNormal ("SigmaB", 0.15)
    B = pm.Normal ('B', mu=MuB, sigma=SigmaB)
    MuA = 1299.82
    SigmaA= pm.HalfNormal ("SigmaA", 130)
    A = pm.Normal('A', mu=MuA, sigma=SigmaA)
    Mun =0.5
    Sigman= pm.HalfNormal("Sigman", 0.05)
    n = pm.Normal('n', mu=Mun, sigma=Sigman)
    y = pm.DensityDist ('y', logp, observed={'SIs': SIs.values.astype (int), 'ns1': ns1.values.astype
(int)})
    trace_ss1 = pm.sample (1000, tune=1000, chains = 2, target_accept=0.99)
    print ('The code is running')
    Bi = pm.summary (trace_ss1, var_names=['B'])['mean'][0]
    Ai = pm.summary (trace_ss1, var_names=['A'])['mean'][0]
    ni = pm.summary (trace_ss1, var_names=['n'])['mean'][0]
```

```
    az.plot_trace (trace_ss1, var_names=['B','A','n'])
    print(str('[B, A, n]=[')+str(Bi)+str(',')+str(Ai)+str(',')+str(ni)+str(']'))

#Collect the first portion of test data about the similar product:
SI=1000
scale_value=1
nts1 = round (DF3['No. of drops_test_SI1'],2) # number of drops
SIts1=round (DF3['st1_SI1']*DF3['st2_SI1']*DF3['st3_SI1']/scale_value,2) # SI value

# Build the log-likelihood using the first portion of test data about the similar product:
def logp2(SIts1, nts1, SI):
    summ1 = 0
    for i in range (0, len(DF3)):
        print(i)
        F1=DF3['failure'][i]
        nu=nts1[i]*(SIts1[i]/SI)**n
        PDF = (B*nu**(B-1))/(A*SI**-n)**B
        R = np.exp(-(nu/(A*SI**-n))**B)
        logLik = (np.log ((PDF**F1)*R))
        summ1 += logLik
    return(summ1)

# Bayesian estimation using the first portion of test data about the similar product:
with pm.Model() as model_ts1:
    MuB =1.6
    SigmaB= pm.HalfNormal("SigmaB", 0.16)
    B = pm.Normal('B', mu=MuB, sigma=SigmaB) # Prior distribution of beta
    MuA =1789.99
    SigmaA= pm.HalfNormal("SigmaA", 179)
    A = pm.Normal('A', mu=MuA, sigma=SigmaA) # Prior distribution of A
    Mun = 0.52
    Sigman= pm.HalfNormal("Sigman", 0.052)
    n = pm.Normal('n', mu=Mun, sigma=Sigman) # Prior distribution of n
    y = pm.DensityDist('y', logp2, observed={ 'SI': SI,'SIts1': SIts1.values.astype(int), 'nts1': nts1.va
lues.astype(int)})
    trace_ts1 = pm.sample(1000, tune=1000, target_accept=0.99,chains = 2)
    print ('The code is running')
    Bi = pm.summary(trace_ts1, var_names=['B'])['mean'][0]
    Ai = pm.summary(trace_ts1, var_names=['A'])['mean'][0]
    ni = pm.summary(trace_ts1, var_names=['n'])['mean'][0]
    az.plot_trace(trace_ts1, var_names=['B','A','n'])
    print (str('[B, A, n]=[')+str(Bi)+str(',')+str(Ai)+str(',')+str(ni)+str(']'))


#%% Kullback-Leibler Analysis:
```

```
###################################################################
 X is the number of cycles of the tested units. Therefore, the equivalent cycles of the surveyed unit
s shoud be taken to the power of (or multiplied by) 1/Sai to follow the life distribution of the teste
d units.
###################################################################
# Estimate the bias value through the KL divergence analysis:
from scipy.optimize import minimize
import tensorflow.compat.v1 as tf
tf.disable_v2_behavior()
font=14
plt.style.use('seaborn-white')
SI=1000
learning_rate = 0.001;
epochs = 100
Sai=tf.Variable(np.ones(1))
[betas, As, ns]=[1.57,1299.99,0.51]; # These parameters were estimated using the first portion of
the surveyed devices.
alphas=As*(SI**-ns) # Calculate the scale parameter of the Weibull distribution using the survey
data
[betat,At,nt]=[1.4,1799.99,0.5]; # These parameters were estimated using the first portion of the
tested devices from the similar product.
alphat=At*(SI**-nt) # Calculate the scale parameter of the Weibull distribution using the test data
Time=300
x=np.arange(1,Time,(Time-1)/(m))
Sai = tf.Variable(0.5, trainable=True)
p=(betat*(x**(betat-1))/(alphat**betat))*np.exp(-(x/alphat)**betat).reshape(1, -1)
qqq=(betas*((x*Sai)**(betas-1))/(alphas**betas))*tf.exp(-((x*Sai)/alphas)**betas)
q=tf.reshape (qqq, [1, m])
kl_divergence = p * tf.log(p / q) # Calculate the KL distance
loss = kl_divergence
optimizer = tf.train.GradientDescentOptimizer(learning_rate).minimize(abs(kl_divergence)) #
Minimize the KL distance through the gradient descent algorithm to estimate the bias value.
with tf.Session () as sess:
   sess.run(tf.global_variables_initializer())
   for i in range (500):
      print(sess.run ([Sai])) # This line prints the estimated bias value at each iteration
      sess.run (optimizer)


#Removing bias from the second portion of the survey data and building the log-likelihood
function using the unbiased survey data:
SI=1000
scale_value=1
ns11 = DF2['biased drops']
SIs11=DF2['SIs']/scale_value
```

```
def logp1(SIs11, ns11, SI):
    summ1 = 0
    for i in range (0, len(DF2)):
        print(i)
        F=DF2['failure'][i]
        nu=ns11[i]*(SIs11[i]/SI)**n # Calculate the equivalent cycles using the original users' data.
        nu=nu*1.351 # Remove bias from the equivalent cycles by multiplying the equivalent cycles
by the estimated bias value.
        PDF = (B*nu**(B-1))/(A*SI**-n)**B
        R = np.exp(-(nu/(A*SI**-n))**B)
        logLik = (np.log ((PDF**F)*R))
        summ1 += logLik # Sum of the log-likelihood function
    return(summ1)

# Estimating the primary posterior distribution (first step of the sequential Bayesian analysis):
with pm.Model() as model_ss2:
    MuB = 1.49
    SigmaB= pm.HalfNormal("SigmaB", 0.15)
    B = pm.Normal('B', mu=MuB, sigma=SigmaB) # Prior distribution of beta
    MuA = 1299.82
    SigmaA= pm.HalfNormal("SigmaA", 130)
    A = pm.Normal('A', mu=MuA, sigma=SigmaA) # Prior distribution of A
    Mun =0.5
    Sigman= pm.HalfNormal("Sigman", 0.05)
    n = pm.Normal('n', mu=Mun, sigma=Sigman) # Prior distribution of n
    y = pm.DensityDist('y', logp1, observed={ 'SI': SI,'SIs11': SIs11.values.astype(int), 'ns11': ns11.
values.astype(int)})
    trace_ss2 = pm.sample(1000, tune=2000, chains = 2, target_accept=0.99)
    print ('The code is running')
    Bi = pm.summary(trace_ss2, var_names=['B'])['mean'][0] # Primary posterior distribution of
beta
    Ai = pm.summary(trace_ss2, var_names=['A'])['mean'][0] # Primary posterior distribution of A
    ni = pm.summary(trace_ss2, var_names=['n'])['mean'][0] # Primary posterior distribution of n
    az.plot_trace(trace_ss2, var_names=['B','A','n'])
    print(Bi, Ai, ni)


#Collect the second Part of the test data about the similar product
nts2 = round (DF4['No. of drops_test_SI1'],2)
SIts2=round (DF4['st1_SI1']*DF4['st2_SI1']*DF4['st3_SI1']/scale_value,2)


# Build the log-likelihood using the second portion of the test data about the similar product:
def logp3(SIts2, nts2):
    summ1 = 0
```

```
    for i in range(0, len(DF4)):
        print(i)
        SI=1000
        F11=DF4['failure'][i]
        nu=nts2[i]*(SIts2[i]/SI)**n
        PDF = (B*nu**(B-1))/(A*SI**-n)**B
        R = np.exp(-(nu/(A*SI**-n))**B)
        logLik = (np.log ((PDF**F11)*R))
        summ1 += logLik
    return(summ1)
traces = [trace_ss2]
model = Model ()

# Estimating the intermediate posterior distribution using the second portion of test data about the
similar product as the likelihood data (second step of the sequential Bayesian analysis):
with model:
    B = from_posterior ("B", trace_ss2["B"]) # Priors of B, A, and n are the posterior
distributions of the previous step.
    A = from_posterior ("A", trace_ss2["A"])
    n = from_posterior ("n", trace_ss2["n"])
    y = pm.DensityDist ('y', logp3, observed={'SIts2': SIts2.values.astype
(int), 'nts2': nts2.values.astype (int)})
    trace_ts2 = pm.sample (1000, tune=3000, chains = 2, target_accept =
0.99, start = {'A': 1800, 'n': 0.5, 'B': 1.5})
    traces.append(trace_ts2)
    az.plot_trace(trace_ts2, var_names=['B','A','n'])
    Bii = pm.summary(trace_ts2, var_names=['B'])['mean'][0] # Intermediate posterior distribution
of beta
    Aii = pm.summary(trace_ts2, var_names=['A'])['mean'][0] # Intermediate posterior distribution
of A
    nii = pm.summary(trace_ts2, var_names=['n'])['mean'][0] # Intermediate posterior distribution
of n
    print(Bii, Aii, nii)


# Collect test data about the new device:
scale_value=1
SI=1000
ntn = round(df2['No. of drops_test_SI1'],2)
SItn=round(df2['st1_SI1']*df2['st2_SI1']*df2['st3_SI1']/scale_value,2)


# Build the log-likelihood using the test data about the new product:
def logp4(SItn,ntn):
    # Resave the initial parameter guesses
```

```
    summ1 = 0
    for i in range(0,len(df2)):
        print(i)
        SI=1000
        F11=df2['failure'][i]
        nu=ntn[i]*(SItn[i]/SI)**n
        PDF = (B*nu**(B-1))/(A*SI**-n)**B
        R = np.exp(-(nu/(A*SI**-n))**B)
        logLik = np.log ((PDF**F11)*R)
        summ1 += logLik
    return(summ1)
traces = [trace_ts2]
model = Model()

# Estimating the final posterior distribution using the test data about the new product as the
likelihood data (third step of the sequential Bayesian analysis):
with model:
    # Priors are posteriors from previous iteration
    B = from_posterior("B", trace_ts2["B"])
    A = from_posterior("A", trace_ts2["A"])
    n = from_posterior("n", trace_ts2["n"])
    y = pm.DensityDist('y', logp4, observed={ 'SItn': SItn.values.astype(int), 'ntn': ntn.values.astype
(int)})
    trace_tn = pm.sample(1000, tune=3000, target_accept=0.999, chains = 2)
    traces.append(trace_ts2)
    az.plot_trace(trace_tn, var_names=['B','A','n'])
    Bii = pm.summary(trace_tn, var_names=['B'])['mean'][0] #Final posterior distribution of beta
    Aii = pm.summary(trace_tn, var_names=['A'])['mean'][0] #Final posterior distribution of A
    nii = pm.summary(trace_tn, var_names=['n'])['mean'][0] #Final posterior distribution of n
    print(Bii, Aii, nii)
```

A.2: Python Code for Estimating the Reliability Model Using Discrete Distribution

This section presents an example of the python code developed to obtain the results provided in Chapter 5 of this dissertation.

```
# Import libraries:
import warnings
import pandas as pd
import arviz as az
import matplotlib as mpl
```

```python
import matplotlib.pyplot as plt
import numpy as np
import pymc3 as pm
import theano.tensor as tt
from pymc3 import Model, Normal, Slice, sample
from pymc3.distributions import Interpolated
from scipy import stats
from theano.compile.ops import as_op
from google.colab import files
uploaded = files.upload()
import io
from scipy.optimize import minimize
import math



# Import data:
DF= pd.read_csv (io.BytesIO(uploaded['Survey_4_4_22(C0.05-B0.5_n0.5).csv']))
df1 = pd.read_csv (io.BytesIO(uploaded['Test1_4_4_2022(c0.05).csv']))
df2 = pd.read_csv (io.BytesIO(uploaded['Test2_4_4_2022(c0.025).csv']))
DF1= pd.read_csv (io.BytesIO(uploaded['DF1_4_4.csv']))
DF2 = pd.read_csv (io.BytesIO(uploaded['DF2_4_4.csv']))
DF3 = pd.read_csv (io.BytesIO(uploaded['DF3_4_4.csv']))
DF4 = pd.read_csv (io.BytesIO(uploaded['DF4_4_4.csv']))


#Define the kernel density function:
def from_posterior (param, samples):
    smin, smax = np.min(samples), np.max(samples)
    width = smax - smin
    x = np.linspace(smin, smax, 100)
    y = stats.gaussian_kde(samples)(x)
    x = np.concatenate([[x[0] - 3 * width], x, [x[-1] + 3 * width]])
    y = np.concatenate([[0], y, [0]])
    return Interpolated(param, x, y)

SIs11=DF2['SIs']
ns11 = DF2['biased drops']
SIref = 1000

Delta_n = []
N0 = []
Fun0 = []
ni = 0.05
lam = 0.1
```

```
delta_n = 0
iterations = 10
F=DF2['failure']

# Define the -log-likelihood function:
def regressLL(params):
    b = params [0]
    c = params [1]
    n = params [2]
    summ1 = 0.
    summ2 = 0.
    Nu0 = [None]*len(DF2)
    for i in range (0, len(DF2)):
        Nu0[i] = ((ns11[i]*(SIs11[i]/SIref)**n0)).astype(int)
        for j in range (1, Nu0[i]):
            summ1 += j**b
        summ2 += ((np.log(1-np.exp(-c*((((ns11[i]*(SIs11[i]/SIref)**n))+1)**b)))))*F[i]
    logLik = -c*summ1+summ2
    return(-logLik)

# Gradient descent algorithm to estimate n:
for k in range (1, iterations):
    n0 = ni+lam*delta_n
    start_pos = [0.5, 0.5, 0.6]
    bnds = [(0.4, 0.8), (0.4, 0.8), (0.5, 0.7)]
    results = minimize (regressLL, start_pos,bounds=bnds)
    ni = results.x[2]
    print(ni)
    delta_n = 0.
    for i in range(0, len(DF2)):
        Nu_p = ((DF2['biased drops'][i]*(DF2['SIs'][i]/SIref)**ni)).astype(int) # Calculate the
equivalent cycle.
        if Nu_p == 0:
            Nu_p = 1
        delta_n += ((np.log (DF2['biased drops'][i]*(DF2['SIs'][i]/SIref)**n0)-
np.log(Nu_p))*(np.log(DF2['SIs'][i])-np.log(1000)))
    delta_n = delta_n/len(DF2)
    n0 = ni+lam*delta_n # update the n value
    Delta_n.append(delta_n)
    N0.append(n0)
    Fun0.append(results.fun)
plt.rc("figure", figsize=[8, 6])
font = 20
csfont = {'fontname': 'Times New Roman'}
plt.plot(range(1, iterations), N0)
```

```
plt.ylabel('n', fontsize=font, **csfont)
plt.xlabel('Number of iterations', fontsize=font, **csfont)
plt.xticks(fontsize=20, rotation=0, **csfont)
plt.yticks(fontsize=20, rotation=0, **csfont)
plt.title('Gradient Descent', fontsize=font, **csfont)
plt.axhline(y=0.49, color='red',linestyle='--',label="True Value ", linewidth=3)
plt.show()
SI=1000
scale_value=1

#Collect the second portion of the survey data:
DF2=DF2[DF2['biased drops']>1]
DF2.reset_index(drop=True, inplace=True)
ns11 = DF2['biased drops']
SIs11=DF2['SIs']/scale_value

#Build the log-likelihood using the second portion of the survey data after removing its bias.
def logp1(SIs11, ns11, SI):
  summ1 = 0
  for i in range(0,len(DF2)):
    print(i)
    F=DF2['failure'][i]
    nu=ns11[i]*(SIs11[i]/SI)**n
    nu=nu*1.35 # 1.35 is the inverse of the bias value.
    PDF = (B*(nu**(B-1)))/((A*SI**-n)**B)
    R = np.exp(-(nu/(A*SI**-n))**B)
    logLik = np.log ((PDF**F)*R)
    summ1 += logLik
  return(summ1)

#Bayesian analysis using the bias-removed survey data (first step of the sequential Bayesian
analysis):
with pm.Model() as model_ss2:
  MuB =1.6
  SigmaB= pm.HalfNormal("SigmaB",0.16)
  B = pm.Normal('B', mu=MuB, sigma=SigmaB) #prior distribution of beta
  MuA = 192
  SigmaA= pm.HalfNormal("SigmaA", 19.2)
  A = pm.Normal('A', mu=MuA, sigma=SigmaA) #prior distribution of A
  n=0.4 # n=0.4 was obtained using the gradient descent algorithm.
  y = pm.DensityDist ('y', logp1, observed={ 'SI': SI,'SIs11': SIs11.values.astype
(int), 'ns11': ns11.values.astype (int)})
  trace_ss2 = pm.sample (1000, tune=1000, chains = 2, target_accept=0.99, cores=1)
  print ('The code is running')
```

```
    Bi = pm.summary(trace_ss2, var_names=['B'])['mean'][0] #primary posterior distribution of
Beta
    Ai = pm.summary(trace_ss2, var_names=['A'])['mean'][0] #primary posterior distribution of A
    az.plot_trace(trace_ss2, var_names=['B','A'])
    print(Bi, Ai)

# Collect the test data about the similar product:
DF4=DF4[DF4['No. of drops_test_SI1']>0.5]
DF4.reset_index (drop=True, inplace=True)
nts2 = round (DF4['No. of drops_test_SI1'],2)
SIts2=round (DF4['st1_SI1']*DF4['st2_SI1']*DF4['st3_SI1']/scale_value, 2)

#Build the log-likelihood using the second portion of test data about the similar product.
def logp3(SIts2,nts2):
    summ1 = 0
    for i in range(0,len(DF4)):
        print(i)
        SI=1000
        F11=DF4['failure'][i]
        nu=nts2[i]*(SIts2[i]/SI)**n
        PDF = (B*nu**(B-1))/(A*SI**-n)**B
        R = np.exp(-(nu/(A*SI**-n))**B)
        logLik = (np.log ((PDF**F11)*R))
        summ1 += logLik
    return(summ1)
traces = [trace_ts2]
model = Model()

#Bayesian analysis using the second portion of the test data about the similar product (second step
of the sequential Bayesian analysis):
with model:
    B = from_posterior("B", trace_ts2["B"])
    A = from_posterior("A", trace_ts2["A"])
    n=0.4 # n=0.4 was obtained using the gradient descent algorithm.
    y = pm.DensityDist ('y', logp4, observed= {'SItn': SItn.values.astype
(int),  'ntn': ntn.values.astype(int)})
    trace_tn = pm.sample (1000, tune=1000, chains = 2, target_accept=0.99)
    traces.append(trace_tn)
    az.plot_trace(trace_tn, var_names=['B','A'])
    Bii = pm.summary(trace_tn, var_names=['B'])['mean'][0] # Final posterior of beta
    Aii = pm.summary(trace_tn, var_names=['A'])['mean'][0] # Final posterior of A
    print(Bii, Aii)

# Collect the test data about the new product:
scale_value=1
```

```
SI=1000
df2=df2[df2['No. of drops_test_SI1']>0.5]
df2.reset_index(drop=True, inplace=True)
ntn = round(df2['No. of drops_test_SI1'],2)
SItn=round(df2['st1_SI1']*df2['st2_SI1']*df2['st3_SI1']/scale_value,2)


Build the log-likelihood using the test data about the new product:
def logp4(SItn,ntn):
    summ1 = 0
    for i in range(0,len(df2)):
        print(i)
        SI=1000
        F11=df2['failure'][i]
        nu=ntn[i]*(SItn[i]/SI)**n
        PDF = (B*nu**(B-1))/(A*SI**-n)**B
        R = np.exp(-(nu/(A*SI**-n))**B)
        logLik = (np.log ((PDF**F11)*R))
        summ1 += logLik
    return(summ1)
traces = [trace_ts2]
model = Model()

# Bayesian analysis using the test data of the new product (third step of the sequential Bayesian
analysis):
with model:
    B = from_posterior ("B", trace_ts2["B"])
    A = from_posterior ("A", trace_ts2["A"])
    n=0.4
    y = pm.DensityDist('y', logp4, observed={'SItn': SItn.values.astype(int), 'ntn': ntn.values.astype
(int)})
    trace_tn = pm.sample(1000, tune=1000,chains = 2,target_accept=0.99)
    traces.append(trace_tn)
    az.plot_trace(trace_tn, var_names=['B','A'])
    Bii = pm.summary(trace_tn, var_names=['B'])['mean'][0]
    Aii = pm.summary(trace_tn, var_names=['A'])['mean'][0]
    print(Bii, Aii)
```

Glossary

Term	Definition
Stress	An agent that causes damage to the product.
Stress adjustor	A variable that increases or decreases the stress magnitude or stress absorption.
Stress profile	A combination of stresses and stress adjustors.
Stress block	A given stress profile when repeated multiple times.
Reliability	The probability that a product or system performs its intended function adequately for a specified duration under specified use conditions [68].
Critical Elements	The essential data needed for estimating the reliability model of a consumer product.
New Product	A product that is under development and has not been introduced to the market yet.
Similar Product	A product whose features, dimensions, and materials are close to the new product.
Damage Cycle	A usage cycle that causes partial cumulative damage that when reaches a threshold, the devices will fail.

Bibliography

[1] H. Li, R. Yuan and J. Fu, "A Reliability Modeling for Multi-Component Systems Considering Random Shocks and Multi-State Degradation," *IEEE Access,* vol. 7, pp. 168805-168814, 2019.

[2] C. Yunxia, Z. Wenbo and X. Dan, "Reliability assessment with varying safety threshold for shock resistant systems," *Reliability Engineering and System Safety,* vol. 185, pp. 49-60, 2019.

[3] E. Taylor-Powell and M. G. Marshall, Questionnaire Design: Asking questions with a purpose, Cooperative Extension Service, University of Wisconsin-Extension, 1998.

[4] K. Kelley, B. Clark, V. Brown and J. Sitzia, "Good practice in the conduct and reporting of survey research," *International Journal for Quality in Health Care,* vol. 15, no. 3, pp. 261-266, 2003.

[5] M. S. D. P. Nayak and K. Narayan, "Strengths and weaknesses of online surveys," *IOSR Journal of Humanities and Social Sciences,* vol. 24, no. 5, pp. 31-38, 2019.

[6] F. H. S. Pratama, "Designing an online-based questionnaire application for mobile devices," *Journal of Physics: Conference Series,* vol. 1496, no. 1, p. 012042, 2020.

[7] S. Sreejesh, S. Mohapatra and M. R. Anusree, "Questionnaire Design," in *Business Research Methods*, Springer, 2014, pp. 143-159.

[8] J. Akrosnick, "Questionnaire design," in *The Palgrave handbook of survey research*, Cham, 2018, pp. 439-455.

[9] M. R. Hibberts, "Handbook of survey methodology for the social sciences," in *Handbook of survey methodology for the social sciences*, New York, Springer, 2012, pp. 53-74.

[10] F. Arlene, How to ask Survey Questions, Sage, 2002.

[11] D. F. Alwin, Margins of Error: A Study of Reliability in Survey Measurement, John Wiley & Sons, 2007.

[12] A. C. Scherpenzeel and W. E. Saris, "The Validity and Reliability of Survey Questions: A Meta-Analysis of MTMM Studies," *Sociological Methods & Research,* vol. 25, no. 3, pp. 341-383, 1997.

[13] M. S. Deniz and A. A. Alsaffar, "Assessing the Validity and Reliability of a Questionnaire on Dietary Fibre-related Knowledge in a Turkish Student Population," *Journal of Health, Population and Nutrition,* vol. 4, pp. 497-503, 2013.

[14] R. Heale and A. Twycross, "Validity and reliability in quantitative studies," *Evidence-based nursing,* vol. 18, no. 3, pp. 66-67, 2015.

[15] S. G. Heeringa, B. T. West and P. A. Berglund, Applied Survey Data Analysis, CRC Press, 2017.

[16] A. O'Hagan and J. E. Oakley, "Probability is perfect, but we can't elicit it perfectly," *Reliability Engineering and System Safety,* vol. 85, pp. 239-248, 2004.

[17] C. Chatfield, "Model Uncertainty, Data Mining and Statistical Inference," *J. R. Statist. Soc,* vol. 158, no. 3, pp. 419-466, 1995.

[18] R. Ghanem, D. Higdon and H. Owhadi, Handbook of Uncertainty Quantification, Springer, 2016.

[19] G. L. Mazzi, J. Mitchell and F. Carausu, Data uncertainties: their sources and consequences, Eurostat, 2019.

[20] M. A. Parrish, H. Moradkhani and C. M. DeChant, "Toward reduction of model uncertainty: Integration of Bayesian model averaging and data assimilation," *Water Resources Research,* vol. 48, pp. 1-18, 2012.

[21] P. P. Biemer, "Total Survey Error: Design, Implementation, and Evaluation," *Public Opinion Quarterly,* vol. 74, no. 5, pp. 817-848, 2010.

[22] D. T. Campbell and D. W. Fiske, "Convergent and discriminant validation by the multitrait-multimethod matrix," *Psychological bulletin,* vol. 56, no. 2, p. 81, 1959.

[23] P. H. Garthwaite, J. B. Kadane and A. O'Hagan, "Statistical Methods for Eliciting Probability Distributions," *Journal of the American Statistical Association,* vol. 100, no. 470, pp. 680-701, 2005.

[24] F. Shen, "Multitrait-Multimethod Matrix," in *The International Encyclopedia of Communication Research Methods*, Wiley-Blackwell, 2017, pp. 1-6.

[25] F. C. Okafor, "Addressing the Problem of Non-response and Response Bias," *Journal of Applied Statistics ,* vol. 1, no. 1, pp. 91-97, 2010.

[26] A. Tversky and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science,* vol. 185, no. 4157, pp. 1124-1131, 1974.

[27] S. Yang, A. Bryan, D. Xiang and P. Tavner, "An Industry-Based Survey of Reliability in Power Electronic Converters," *IEEE Transactions on Industry Applications,* vol. 47, no. 3, pp. 1441-1451, 2011.

[28] J. Cochrane, M. Emerson and J. Donahue, "A Survey of HVDC Operating and Maintenance Practices and their Impact on Reliability and Performance," *IEEE Transactions on Power Delivery,* vol. 11, no. 1, pp. 514-518, 1996.

[29] G. Tollefson, R. Billinton, G. Wacker, E. Chan and J. Aweya, "A Canadian Customer Survey To Assess Power System Reliability Worth," *EEE Transactions on Power Systems,* vol. 9, no. 1, pp. 443-450, 1994.

[30] J. Guo, Z. (. Li and J. (. Jin, "System reliability assessment with multilevel information using the Bayesian melding method," *Reliability Engineering and System Safety,* vol. 170, pp. 146-158, 2018.

[31] R. Pan, "A Bayes Approach to Reliability Prediction Utilizing Data from Accelerated LifeTest sand Field Failure Observations," *Quality and Reliability Engineering International Journal,* vol. 25, pp. 229-240, 2009.

[32] P. Wang, B. D. Youn, Z. Xi and A. Kloess, "Bayesian Reliability Analysis With Evolving, Insufficient, and Subjective Data Sets," *Journal of Mechanical Design,* vol. 131, no. 11, pp. 1-11, 2009.

[33] R. Yuan, M. Tang, H. Wang and H. Li, "A Reliability Analysis Method of Accelerated Performance Degradation Based on Bayesian Strategy," *IEEE Access,* vol. 7, pp. 169047 - 169054, 2019.

[34] A. Barbiero, "Parameter estimation for type III discrete Weibull distribution: a comparative study," *Journal of Probability and Statistics,* vol. 2013, 2013.

[35] Z. Tondpour and G. Yari, "Some New Discretization Methods with Application in Reliability," *Applications and Applied Mathematics: An International Journal (AAM),* vol. 13, no. 2, pp. 1-13, 2018.

[36] A. M. Sarhan and M. Zaindin, "Modified Weibull distribution," *APPS. Applied Sciences,* vol. 11, pp. 123-136, 2009.

[37] Q. Dan, D. Ye, Y. Teng and L. An, "Imprecise reliability analysis for lifetime modeling based on four-parameter Weibull distribution," in *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)*, 2021.

[38] J. Wang and C. Zhou, "Statistical Modeling and Analysis of Satellite Failure Based on 2-Weibull Segmented Model," *IEEE Access,* vol. 9, pp. 128747-128754, 2021.

[39] C. Bracquemond and O. Gaudoin, "A survey on discrete lifetime distributions," *International Journal of Reliability, Quality and Safety Engineering,* vol. 10, no. 01, pp. 69-98, 2003.

[40] J.-M. Jia, Z.-Z. Yan and X.-Y. Peng, "A new discrete extended weibull distribution," *IEEE Access,* vol. 7, pp. 175474-175486, 2019.

[41] X. Dong, Y. Tsong and M. Shen, "Statistical Considerations in Setting Product Specifications," *Journal of Biopharmaceutical Statistics,* vol. 25, pp. 280-294, 2015.

[42] Y.-L. Lee, S. Makam, S. McKelvey and M.-W. Lu, "Methods, Durability Reliability Demonstration Test," *Procedia Engineering,* vol. 133, pp. 31-59, 2015.

[43] G. Yang, "Heuristic Degradation Test Plans for Reliability Demonstration," *IEEE Transactions on Reliability,* vol. 62, no. 1, pp. 305-311, 2013.

[44] Gerokostopoulos, Athanasios, H. Guo and E. Pohl, "Determining the Right Sample Size for Your Test: Theory and Application," in *Annual Reliability and Maintainability Symposium Tutorial Notes*, 2015.

[45] Z. Yang, Y. Chen, Y. Li, E. Zio and R. Kang, "Smart electricity meter reliability prediction based on accelerated degradation testing and modeling," *Electrical Power and Energy Systems,* vol. 56, pp. 209-219, 2014.

[46] W.-L. Chang, K.-Y. Lin, C.-D. Hsueh and J.-M. Chang, "Vibration Test Specification Design and Reliability Analysis," *International Journal of Materials and Manufacturing,* vol. 4, no. 1, pp. 675-685, 2011.

[47] G. Yang, Life cycle reliability engineering, John Wiley & Sons, 2007.

[48] L. M. Howden and J. A. Meyer, "Age and sex composition, 2010," US Department of Commerce, Economics and Statistics Administration, US Census Bureau, 2011.

[49] C. D. Fryar, D. Kruszon-Moran, Q. Gu and C. L. Ogden, "Mean body weight, weight, waist circumference, and body mass index among adults: United States 1999–2000 through 2015–2016.," National Health Statistics Reports, Hyattsville, MD, USA, 2018.

[50] S. C. Savva, Y. Kourides, M. Tornaritis, M. Epiphaniou-Savva, P. Tafouna and A. Kafatos, "Reference growth curves for Cypriot children 6 to 17 years of age," *Obesity research,* vol. 9, no. 12, pp. 745-762, 2001.

[51] C. C. Gordon, T. Churchill, C. E. Clauser, B. Bradtmiller, J. T. McConville and I. Tebbets, "Anthropometric Survey of US Personnel: Summary Statistics Interim Report," US Army Natick Research Development and Engineering Center, 1989.

[52] P. Strzelecki and T. Tomaszewski, "Application of Weibull distribution to describe SN curve with using small number specimens," in *AIP Conference Proceedings*, 2016.

[53] Y. Chen, W. Zhang and D. Xu, "Reliability assessment with varying safety threshold for shock resistant systems," *Reliability Engineering & System Safety,* vol. 185, pp. 49-60, 2019.

[54] N. Shafiei, J. W. Herrmann, A. Krive, M. Nikiforov, G. Sethi and M. Modarres, "Analyzing Product Reliability Using User Survey Data and Reliability Test Data," *IEEE Transactions on Reliability (Under Review),* 2022.

[55] N. Shafiei, J. W. Herrmann, A. Krive and M. Modarres, "Estimating Reliability Model of Consumer Electronics Using User Survey Data and Reliability Test Data," in *Reliability & Maintainability Symposium (RAMS)*, Tuscon, Arizona, 2022.

[56] J. M. Lewis, S. Lakshmivarahan and S. Dhall, Dynamic Data Assimilation : A Least Squares Approach, Cambridge University Press, 2006.

[57] M. Modarres, M. P. Kaminskiy and V. Krivtsov, Reliability engineering and risk analysis: a practical guide, CRC press, 2009.

[58] Y. Xu and R. Goodacre, "On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning," *Journal of Analysis and Testing,* vol. 2, no. 3, pp. 249-262, 2018.

[59] B. Qin and F. Xiao, "A Non-Parametric Method to Determine Basic," *IEEE Access,* vol. 6, pp. 73509-73519, 2018.

[60] A. Gomez-Alanis, J. A. Gonzalez-Lopez and A. M. Peinado, "A Kernel Density Estimation Based Loss Function and its Application to ASV-Spoofing Detection," *IEEE Access,* vol. 8, pp. 108530-108543, 2020.

[61] S. Woo, Reliability Design of Mechanical Systems: A Guide for Mechanical and Civil Engineers, Singapore: Springer , 2020.

[62] N. Shafiei, J. W. Herrmann, A. Krive and M. Modarres, "Estimating the Reliability of Consumer Electronics from User Survey Data and Test Data," in *Reliability and Maintainability Symposium (RAMS)*, Tuscon, USA, 2022.

[63] S. Raschca, Python Machine Learning, Packt Publishing ltd, 2015.

[64] E. Patel and D. S. Kushwaha, "Clustering cloud workloads: k-means vs gaussian mixture model," in *Procedia Computer Science*, 2020.

[65] M. Modarres, M. Kaminskiy and V. Krivtsov, Reliability engineering and risk analysis : a practical guide, CRC press, 2016.

[66] V. Satopa, J. Albrecht, D. Irwin and B. Raghavan, "Finding a “Kneedle” in a Haystack: Detecting Knee Points in System Behavior," in *International Conference on Distributed Computing Systems Workshops (ICDCSW)*, 2011.

[67] H. T. Nguyen, O. Kosheleva, V. Kreinovich and S. Ferson, "Trade-off between sample size and accuracy: Case of measurements under interval uncertainty," *International Journal of Approximate Reasoning,* vol. 50, pp. 1164-1176, 2009.

[68] S. Lazarova-Molnar and N. Mohamed, "Reliability Assessment in the Context of Industry 4.0: Data as a Game Changer," in *The 2nd International Conference on Emerging Data and Industry 4.0 (EDI40)*, Leuven, Belgium, 2019.

[69] N. Shafiei, J. W. Herrmann, A. Krive, G. Sethi and M. Modarres, "Designing Reliability-Informed Customer Surveys," in *European Safety and Reliability (ESREL)*, Angres, 2021.

[70] N. Shafiei, J. W. Herrmann and M. Modarres, "Estimating the Reliability Model of a New Consumer Product Assuming a Discrete Life Distribution," in *RAMS*, Florida, 2022.

[71] N. Shafiei, J. W. Herrmann and M. Modarres, "Applying User Surveys and Accelerated Tests Data to Estimate Reliability of New Consumer Products Using a Discrete Life Distribution Model," *IEEE Access,* vol. 10, 2022.