Na prawach rękopisu

INSTYTUT CYBERNETYKI TECHNICZNEJ
POLITECHNIKI WROCŁAWSKIEJ

Raport serii: PREPRINTY nr 12/97

DATA DRIVEN RELIABILITY PREDICTION BASED ON ACCELERATED LIFE TEST RESULTS

(rozprawa doktorska)

Henryk MACIEJEWSKI

PROMOTOR:
prof. dr hab. inż. Janusz Biernat

Keywords:

- reliability of computer hardware and integrated circuits
- accelerated life test
- accelerated deterioration test
- life distribution function
- life-stress model

WROCŁAW 1997

FACULTY OF ELECTRONIC ENGINEERING
TECHNICAL UNIVERSITY OF WROCŁAW
POLAND

DATA DRIVEN RELIABILITY PREDICTION BASED ON ACCELERATED LIFE TEST RESULTS

(a Ph.D. thesis)

Henryk Maciejewski

Professor conferring the degree:
Prof. D.Sc. Janusz Biernat

Keywords:
reliability of computer hardware and integrated circuits, accelerated life test, accelerated deterioration test, life distribution function, life-stress model

Foreword

Computer hardware and other products of contemporary electronic industry usually exhibit very high reliability. Under normal use conditions, the products reveal so few failures that it is impossible or impractical to gather enough data to estimate the product failure rate or other reliability characteristics. Products are therefore often tested under more severe conditions than normal to shorten product life. Such test methods are known as Accelerated Life Tests (ALT). In many cases, accelerated life testing turns out to be the only practical way to get enough data for reliability estimation under limited testing time and resources.

More severe test conditions (i.e., higher values of some stress factors) yield more failures. Data obtained during accelerated life tests should be properly modelled and analyzed in order to get information about product life under normal conditions. Although various models and approaches to data analysis have been put forward, and despite growing availability of sophisticated tools for statistical analysis of data, ALT data analysis still creates a challenge for analysts.

Data analysis is usually based on some assumptions concerning the data analyzed and the effect of stress on product life.

The thesis proposes a versatile method of ALT data analysis which allows to drop or relax assumptions usually made about the type of distribution function that the life data should follow, and about relationship between life and stress. The method proposed allows to derive life distribution function for the nominal stress level based on a number of accelerated tests which can be complete or censored. The method can be especially useful in a variety of applications where the assumptions on life distribution function and/or life-stress relationship are not well justified by an empirical or theoretical model.

The method is based on the assumption that under each test stress level as well as under the nominal conditions the product lifetime distribution can be represented by a member of a general family of distributions built on Laguerre polynomials. The lifetime cumulative distribution function (c.d.f.) at the nominal conditions is estimated through extrapolation from the c.d.f.s obtained under increased stress levels. The extrapolation procedure does not require that the analyst knows the proper relationship between value of stress and product lifetime. Instead, the method presented makes data driven selection of the most suitable model, out of some specified types of life-stress relationships. An important aspect of the methodology presented is that models used at all stages of analysis are checked for being adequate for data analyzed. In this way, verification of results of ALT data analysis is achieved, which is otherwise difficult due to very scarce field data at use conditions. Performance of the ALT data analysis method presented is illustrated in some examples using sample ALT data.

Contents

Introduction

Evaluation of reliability

Product reliability has drawn attention of manufacturers as one of the major factors contributing to quality. Reliability estimation has become an integral part of the manufacturing process. Products of contemporary electronic, semiconductor and computer industries are now highly reliable with lifespans of many years. In such cases, life testing performed at use (nominal) conditions in order to evaluate product reliability is becoming prohibitively time consuming, expensive and eventually impractical.

To save time and money, many products can be life tested at more severe conditions (higher stress levels) than during actual use to yield more failures during limited testing time. Life testing at higher stress levels imposed on products as a means of speeding-up deterioration processes and shortening time to failure is called *accelerated life testing* (ALT). Typically, data obtained from an accelerated life test consists of *times to failure* observed under increased stress.

Accelerated life testing is not the only methodology available for estimating reliability of highly reliable products. Other approaches include e.g.: *accelerated degradation testing* or using *reliability prediction procedures*.

Similarly to ALT, accelerated degradation tests (ADT) involve submitting products to overstress conditions. However, instead of life, product performance or the amount of degradation is observed as it changes over time. Models required for ADT data analysis differ from those used for accelerated life testing. Nelson (1990) and Meeker and Escobar (1993) provide a survey of models and data analyses for ADT data. Carey and Koenig (1991) describe some real life accelerated degradation test that was performed and analyzed by AT&T. Lu and Meeker (1993) develop statistical methods for estimation of a lifetime distribution from observed degradation measures.

Reliability prediction procedures for electronic devices and equipment are included in well known reliability handbooks, such as MIL-HDBK-217F, "Reliability Prediction of Electronic Equipment". Other widely used prediction procedures are: British Telecom Handbook, NTT Procedures (Nippon Telegraph and Telephone), CNET Procedure or Siemens Procedure. The prediction procedures are used to predict reliability of products in the useful life period, i.e., ignoring infant mortalities and long-term wearout. They are based on knowledge of equipment (system) reliability structure (e.g., a series system) and on some (simplified) assumptions concerning component reliability (e.g., it is commonly assumed that components' hazard rates are constant). Effect of environmental and loading conditions is also modelled using simplified formulas or correction factors. The standard prediction procedures are based on analysis of extensive field failure data and on reliability expert knowledge. Although considered standard and widely used, the prediction procedures cause controversies as different procedures tend to give very different results when used to predict failure rate of the same kind of equipment. Bowles (1992) and Jensen (1995) critically review currently used procedures and discuss incoherence in predicted reliability. Pecht and Nash (1994) give a thorough treatment of the role of reliability prediction in the development of electronic equipment and discuss problems with some methods and opportunities for improvement of prediction procedures.

A new perspective has been emerging that various reliability data sources should be combined and integrated into one reliability database that would include expert opinion, data from accelerated life and degradation tests and actual field failure data. The idea is discussed by Meeker and Hamada (1995). Some ways of combining different reliability data sources are also discussed in Clarotii and Lindley (1988). These ideas show that different approaches to evaluation of product reliability should not be considered competing but rather complementary in the rapidly changing manufacturing environment of computer and electronic industry where new demands are placed on timely and efficient estimation of reliability of newly developed products.

Accelerated life testing - basic concepts

In this thesis, accelerated life testing is considered which has become an important technology for estimation of reliability of highly reliable electronic equipment.

Accelerated life testing consists in observing times to failure under *increased stress*. In principle, acceleration of life tests could be achieved by increasing *use-rate* of items tested (this is known as the method of time compression), or by submitting items to such environmental or loading conditions that increase deterioration and consequently shorten life time of items under test. Although the latter approach, known as *overstress* testing, is considered in the thesis, some results might be applied in the case of time compression ALT.

An accelerated life test is a complex procedure involving several statistical and engineering considerations (Nelson, 1990) such as:

- test plan;
- models for test data analysis.

When an ALT is planed, a number of issues should be determined, including:

- recognition of an accelerating factor (stress). It is important to ensure that failure mechanisms occurring under increased stress are the same as those occurring under nominal stress level. Although *single* accelerating factors are most commonly used, it is possible to organize an ALT with *multiple* accelerating factors. Typical examples of accelerating factors for tests of electronic equipment are temperature voltage level, current density.
- stress loading policy. Stress could be applied in various ways, e.g. constant stress, step stress, continuously increased stress, stress varied cyclically or randomly. Constant stress is most commonly used in practice and consists in submitting different groups of specimens to different levels of stress remaining constant during the test. ALT with step stress or otherwise increased stress have become less important due to growing popularity of tests with censored data.
- other test parameters such as number of stress levels, value of stress factor and number of specimens tested at each stress level. Two opposite objectives should

be met: stress levels should be high enough to yield sufficient number of failures, yet they should not cause failure modes that do not occur under nominal conditions. Nelson (1990) and Meeker and Escobar (1993) survey methods and criteria for development of statistically optimum test plans; other authors develop specific methods for test design (e.g., Bai and Chung, 1992, or Chaloner and Larntz, 1992). The objective in the methods is to minimize statistical uncertainty of estimates of parameters evaluated in an ALT.

In order to obtain reliability of products in normal use, data acquired in an accelerated life test have to be extrapolated from test stress levels to the nominal conditions. Typical accelerated test data analysis employs *models* which usually describe:

- the type of distribution that lifetimes follow at each test stress level and at the nominal conditions;
- effect of stress on lifetime. The life-stress relationship (known also as the time transformation function or ageing model, or acceleration model) usually specifies functional relationship between parameters of the assumed life distribution function and the stress level or relationship between a specified percentile of the distribution function and the stress level.

In many practical cases, there are well established models for data analysis. They are either justified by a physical model or based on experience. For example, Arrhenius-Weibull model is typically used for some electronic devices tested under increased temperature. In this model, lifetimes are assumed to follow Weibull life distribution at each stress level; Arrhenius life-stress relationship is used to relate life under increased stress to life under normal conditions. (Although Arrhenius relationship is considered one of the life-stress models with strong physical foundations, recent research have cast some doubt if application of this model is justified for temperature stressed semiconductor devices. This idea, illustrating difficulty with selection of a suitable life-stress model, will be discussed in Chapter 3).

In some cases there are no underlying physical or empirical models of product deterioration. It is often difficult to create a suitable model due to the fact that electronic devices tested are often highly heterogeneous (i.e., consist of various parts with different failure mechanisms), or due to rapid technology advances and strong pressure for short development time that give no chance to obtain enough data or do enough research to build a model. Nelson states that "Lack of adequate physical models hinders the use of accelerated testing with some products" (Nelson, 1990).

The thesis addresses the issue of models for ALT data analysis. The thesis is motivated by an observation that accelerated tests for which no good models are available would need a more general approach to estimation of reliability through analysis of test data. The purpose of the approach would be to determine the product life distribution at nominal conditions without making too strong assumptions which do not let 'the data speak for themselves' (Müller, 1988).

Aim of the thesis

The aim of the thesis is to show that it is possible to estimate lifetime distribution of equipment under nominal conditions based on results of accelerated life tests without making strong assumptions on distribution that life data should follow and with no a priori knowledge of a life-stress relationship that should be used for the data analysis. The estimation is possible from complete or censored test data (failure or time-censored). It is also possible to check if models used for reliability estimation are adequate for data and thus it is possible to verify results of reliability estimation under use (nominal) conditions.

In the thesis a versatile statistical method of ALT data analysis is proposed. The method allows to estimate lifetime cumulative distribution function of equipment working under nominal conditions based on data from constant stress accelerated life tests. The ALT data to be analyzed can be complete or Type I (time) or Type II (failure) censored. No strong assumption on lifetime distribution is required. The method uses a generic class of distribution functions based on Laguerre polynomials that are fitted to data at each stress level and used to model distribution of lifetime at the test and nominal conditions. A life stress relationship does not have to be known; the estimation procedure makes a data driven selection of the most suitable life-stress model out of a set of specified relationships. An important aspect of the method proposed is methodology to check if models used are adequate for the analyzed data.

Versatility of the method proposed and no need for strong assumptions on life distribution and life-stress relationships make the method especially useful for life testing of electronic equipment such as computer and control device hardware. Accelerated lifetesting of electronic equipment lacks well established models, as technology in electronic manufacturing changes very fast and products or devices are usually highly heterogeneous, consisting of various parts with different deterioration mechanisms.

Survey of existing ALT data analysis methods

Accelerated life test data analyses, that still create a challenge for analysts, have been in recent years extensively covered in literature. A survey of ALT data analysis methods is provided by Meeker and Escobar (1993) and Nelson (1990). Other important overviews are by Jensen (1995), Crowder (1991), Padgett (1984), Lawless (1982), Nelson (1982), Kalbfleish and Prentice (1980). Large number of papers devoted to specific methods of ALT data analysis appeared in recent years in reliability and statistical journals and conference proceedings (e.g., Glaser, 1995, Mackisack and Stillman, 1996, Dietrich and Mazzuchi, 1996, Tseng and Hsu, 1994, Dzerjinski and Tzafestas, 1995, Yang and Kim, 1995).

Standard accelerated life data analysis methods can be broadly classified as *parametric* and *nonparametric*. These methods are slightly different from standard statistical data modelling techniques, because data usually cannot be modelled with a normal distribution; data are often censored and analysis is often focused on the lower tail of the life distribution.

Parametric methods of ALT data analysis employ an assumed parametric model of life time distribution (or a family of distributions). In other words, it is assumed that at all stress levels the times to failure follow the same type of distribution. Unknown parameters of the life distribution model are estimated by fitting the model to data. Various methods are available for data analysis ranging from graphical to purely analytical. Nelson (1990) and Lawless (1982) provide detailed description of standard data analysis methods. Graphical methods are based on data plots. By plotting life times on a probability paper (such as Weibull or lognormal probability paper), validity of an assumed model can be assessed visually and the unknown parameters of life distribution can be estimated graphically. Although less intuitive than graphical methods, analytical methods provide objective estimates of sought parameters as well as uncertainties of estimates. Analytical methods for complete data usually involve least squares calculations. Censored data are typically analyzed using maximum likelihood methods. Many practical parametric data analyses use graphical and analytical methods together.

Another important component of most parametric ALT data analysis models is an assumed ageing model (i.e., a relationship between parameter(s) of life distribution and stress). The ageing model describes how stress effects degradation or life and is used to extrapolate results of accelerated test to obtain information on product life at nominal conditions. Similarly to data analysis at a specified (constant) stress level, extrapolation can be done using graphical or analytical methods. Graphical methods use plots of data (life times versus stress) on life-stress relationship papers (e.g., log-log paper for power ageing model), from which extrapolation to the use stress level is done. Analytical methods are usually based on regression analysis.

Nonparametric or distribution-free methods can be used when it is not possible or desirable to assume a specific parametric model.

Although less efficient (i.e. giving wider confidence intervals) than parametric methods with a fixed distribution, nonparametric methods have an advantage of being free from assumptions about a specific parametric model. Obviously, tighter confidence intervals given by parametric methods do not account for possible model inadequacy which can be magnified through extrapolation resulting in errors highly exceeding statistical uncertainties.

Some nonparametric methods do assume a parametric ageing model with unknown parameters estimated from data; other methods do not assume such models but usually require that at least a few observations are available from the nominal stress level.

Nonparametric methods use various statistics to provide estimates and confidence intervals for distribution parameters such as mean, standard deviation, and percentiles of the distribution. Application of nonparametric methods is not limited to complete data; censored data can also be analyzed.

Some specific nonparametric methods have been developed e.g., by Proschan and Singpurwalla (1980), Shaked and Singpurwalla (1982) or Basu and Ebrahimi (1982). Other methods have been discussed by Lawless (1982), Nelson (1982) and Padgett (1984).

The method of ALT data analysis proposed in the thesis can be regarded as a general parametric method. Contrary to typical parametric methods, the method does not assumed a specific life distribution model or a particular life-stress relationship. A very general life distribution and data driven selection of a life stress relationships are used which place the method somewhere between parametric and nonparametric methods.

Summary of results

In this section, main results obtained in the thesis are summarized:

- A generalized method of ALT data analysis has been proposed that can be used to obtain life distribution under nominal conditions from complete and Type I and Type II censored data. Very weak assumption are required about the type of life distribution at each stress level and the data analysis procedure makes a data driven selection of the best life-stress model. Development of the ALT data analysis method required that some specific optimization (as well as numerical) problems were solved. A key element was the development of a method that allows to fit a distribution from a generalized family (based on Laguerre polynomials) to complete or censored data. Goodness of fit testing is applied in order to verify if the model fitted to data is appropriate.
- The generalized method of inference from ALT data does not assume a specific ageing model for extrapolation of results to the nominal conditions. A statistical method has been developed that makes data driven selection of the most suitable ageing model (out of a set of specified ageing models).
- Methodology has been developed that allows to check if reliability predicted for the nominal conditions (inferred from ALT data) is valid and meaningful. This statistical method addresses a difficult problem of verification of ALT data analysis results.
- A set of computer programs has been developed implementing methods and algorithms considered. The programs allow to analyze existing accelerated life test data in order to infer the lifetime distribution under nominal conditions from ALT data. Development of the computer tools required that some complex optimization and numerical tasks were solved.
- Some sample ALT data sets were analyzed in order to verify performance of the data analysis method. Results were compared with results obtained with other models and approaches. Model verification procedure developed in the thesis indicates that some methods and models used in literature do not seem fully justified for specific data sets considered.

Thesis organization

The structure of the thesis is as follows:

Chapter 1 gives an overview of specific features of digital computer hardware that make reliability assessment through accelerated testing a challenging task that cannot be tackled with standard models and techniques of reliability engineering. Chapter 2 presents analysis of test data at a specified constant stress level. Generalized family of life time distribution functions is introduced and methods are devised for fitting the distributions to complete and censored data. Chapter 3 deals with commonly used ageing models and discusses problems, as revealed be recent research, with using well known models for analysis of ALT data in real life situations. A more general approaches to the problem of extrapolation using an ageing model are discussed. Ideas presented in Chapters 2 and 3 are extended in Chapter 4 which introduces a generalized method of ALT data analysis. Verification of results obtained from ALT data analysis is dealt with in Chapter 5. Material presented in this chapter is also the basis for the data driven selection of the most adequate ageing model. Chapter 6 gives some examples of how the methodology presented performs with real data ALT analysis. Reliability estimated with the method presented is compared with the estimates obtained under other models and approaches. Also, data driven selection of an ageing model is demonstrated. Chapter 7 contains conclusions and discusses directions for future work. Appendix A treats the generalized family of distributions used in the thesis. Appendix B gives an overview of different goodness of fit techniques that model adequacy checking is based on, and Appendix C discusses regression techniques used for extrapolation of results from test to the nominal stress levels.

Acronyms and Notation

ALT	accelerated life test/testing
ADT	accelerated degradation test/testing
EDF	empirical distribution function
ML	maximum likelihood (method)
MLE	maximum likelihood estimator
MTTF	mean time to failure
p.d.f.	probability distribution function
c.d.f.	cumulative distribution function
i.i.d.	independent and identically distributed
d.f.	degree(s) of freedom
r.v.	random variable
S_0	value of the stress in nominal condition
$S_1, S_2, \dots, S_k$	consecutive values of the increased stress
$n_1, n_2, \dots, n_k$	number of items tested at a stress level $S_1, S_2, \dots, S_k$, respectively
$t_{1i}, t_{2i}, \dots, t_{n_i i}$	times to failure observed at condition S_i, $i = 1,2,\dots,k$ (note: n_i times to failure observed only for *complete* data - see Chapter 2)
$t_{(1)i}, t_{(2)i}, \dots, t_{(n_i)i}$	order statistic of the random sample $t_{1i}, t_{2i}, \dots, t_{n_i i}$
$F_0(t)$	population (i.e. true) life distribution function at the nominal stress level
$F_1(t), F_2(t), \dots, F_k(t)$	population (i.e. true) life distribution function at the corresponding stress level
$\hat{F}_1(t), \hat{F}_2(t), \dots, \hat{F}_k(t)$	empirical distribution function (EDF) for the corresponding stress level
$\tilde{F}_0(t)$	estimated life distribution function at the nominal stress level
$\tilde{F}_1(t), \tilde{F}_2(t), \dots, \tilde{F}_k(t)$	estimated life distribution function at the corresponding stress level
$q_1, q_2, \dots, q_m$	quantile levels for which regression analysis is performed
$X \sim N(\mu, \sigma^2)$	the random variable X is normally distributed with the mean μ and variance σ^2
$F_{1-\alpha}(v_1, v_2)$	quantile of the level (1-α) of the F Snedecor distribution with v_1, v_2 degrees of freedom
$R(t)$	reliability function (=1-$F(t)$, where F is the c.d.f.)
$h(t)$	hazard function (of a r.v. T, defined as $h(t) = \lim\limits_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}$, and equal $-\frac{d \log R(t)}{dt}$)

Chapter 1 Reliability of Computer Hardware

This Chapter focuses on specific features of digital computer hardware that make assessment or evaluation of its reliability a challenging task. Stringent reliability requirements for such electronic devices are discussed and various approaches are outlined for assurance of very high reliability levels. Problems with application of standard methods of reliability assessment are discussed. The standard methods of reliability engineering designed for evaluation of materials or relatively simple products are of limited use for products of contemporary computer engineering industry. Current industry practice in evaluation of reliability of increasingly complex computer hardware is discussed, which calls for some improved methods of reliability testing and evaluation.

1.1 Demand for highly reliable computing hardware

Majority of core systems of our modern society are heavily dependent on reliable digital computing power. Digital computers are the mainstay of various world wide or nation wide networks or systems such as e.g., telecommunication, transportation or banking systems. High reliability or availability, i.e., the ability to provide virtually uninterrupted, error-free service, is the primary demand for life-critical or mission critical systems, such as airport and aircraft control, hospital life-sustaining systems, aerospace or military applications, etc. It is also highly desirable that other vital for society services such as telephone networks, industry process control systems or various on-line transaction processing facilities are not interrupted by computer outages, as this may result in high losses or even risk of life.

Today's high reliability and high availability computing systems (such as telephone networks) are designed to provide 99.999% availability (Availability Class 5), which translates to at most five minutes unavailability per year (Grey, Siewiorek, 1991). Demand for even more reliable devices and systems is growing. The automobile industry, for instance, requires that defect levels due to infant mortalities of application specific integrated circuits (ASIC) are less than 5 ppm (parts per million) which is an extremely stringent requirement (Jensen, 1995).

Improved reliability of digital electronic devices over the last years has encouraged many industries, including the safety-conscious areas like railway systems or military, to start turning to commercial, off the shelf (COTS) devices. The COTS policy is intended to reduce cost by giving up expensive, specially designed and manufactured devices. However, this policy requires that manufacturers of massive, non-custom electronic devices keep very high reliability standards.

1.2 Assurance of computer hardware reliability

Meeting so stringent demands is a real challenge that computer engineers have to face. It must by ensured that both hardware and software components of a system or device are designed for reliability and properly manufactured, tested and maintained. This work concentrates on reliability of computer hardware.

Since digital computing or control systems usually consist of a large number of components, high system reliability demands require that very reliable components (modules) are used. Building blocks and elements of early computers offered low reliability with MTTF of about days, which led to a practical limit in system availability of about 60%. Reliability of highly integrated building blocks of modern computers is at the level of several FIT (failures in 10^9 hours of operation). Although ultra high reliability of components and modules is crucial in order to meet system or device reliability requirements, this is often not sufficient because of growing complexity of systems. Several other approaches have emerged to address the problem. Since it is not feasible to avoid component failures altogether, redundancy is built into a device or system so that some failures are tolerated or masked. Devices are featured with off-line or on-line self testing capabilities. Although still not a common practice, built in self tests and on-chip reliability monitors are believed to become a standard tool for assurance of very high reliability or availability of computer and microelectronic devices. In addition to reliability and testability design objectives, proper operational and maintenance policies are also vital for very high reliability. Könemann et al. (1996), Kapur and Miller (1996) and Murray and Hayes (1996) give a comprehensive overview from the industry perspective of these methods and technologies for reliability assurance.

1.3 Reliability assessment of highly reliable computer devices: problems and current practice

Although electronic components, such as integrated circuits, are now very reliable, a number of recent studies, surveyed by Jensen (1995), indicate that as much as about 50% of reported failures of ICs are intrinsic failures. Intrinsic failures are due to product design, materials, processing, assembly, packaging, etc., and are the responsibility of the manufacturer. Remaining (extrinsic) failures are the responsibility of the user who overstressed, overloaded or otherwise mishandled the product. High percentage of manufacturer related failures emphasizes the need for constant monitoring of reliability of products.

Reliability engineering offers a number of tools for reliability assessment of components and multicomponent products. Unfortunately, specific features of modern computer hardware render some standard reliability models and methods obsolete.

Standard techniques for complex product or system reliability assessment involve:

- a reliability model of the product concerned (e.g., a series structure), and
- estimates of reliability of individual components (modules).

Once reliability of building blocks of a product (system) is known, the product reliability can be readily estimated based on its reliability structure. Various standard techniques are available for estimation of reliability of individual components or modules. As summarized in the Introduction, reliability can be either predicted, which is essentially based on expert knowledge, or tested, usually through accelerated life of degradation tests. A number of standard models for prediction or analysis of test results have been developed.

Although some of the standard models and methods can be (and in fact are) used for estimation of reliability of simple devices, a number of problems arise when the standard methods are to be used for high reliability complex computer devices or components:

1. Distinction between components (elements) and devices in modern computer hardware is getting blurred. Jensen (1995) describes an electronic component as a device performing some electronic function, and consisting of one or more elements so joined together that they cannot normally be disassembled without destruction. Typical components of digital computing or control devices are now highly integrated integrated circuits (ICs) or multichip modules (e.g., surface-mounted boards). Such components are now very complex devices. Reliability assessment of such components has become a difficult task. In other words, from the reliability perspective, computer hardware components should themselves be looked at as complex systems.
2. One of the major difficulties for reliability assessment comes from the fact that computer hardware devices or components are very complex and heterogeneous systems. Various failure mechanisms affect different parts of a device. For instance, VLSI ICs involve extreme variability of the dominant failure mechanisms (e.g., hot carrier effect, oxide breakdown, electrostatic diffusion, contact electromigration, corrosion, package-related and other failures). For this reason, it is very difficult to find appropriate models for prediction of reliability or for analysis of accelerated reliability tests. Temperature and other stress factors have different effect on various failure mechanisms of electronic devices. Besides, activation energies for each failure mechanism vary over a wide range which make standard physical models based on activation energies not very accurate. The problem of models for reliability of digital hardware is well illustrated by recent controversies about the Arrhenius life-stress model (refer also to Chapter 3). Recent research, surveyed by e.g., Jensen (1995) and Lall (1996), have indicated that the standard, well known Arrhenius model for constant temperature accelerated life tests is not appropriate for digital hardware components. (If the model *is* used in spite of this conclusion, it seems that it is used merely in a 'curve-fitting' fashion). In has been found that in the useful ranges of temperature, other than steady state applications of temperature are more important failure causing factors. Lall (1996) describes a study of new electronic hardware for military aircraft (F22 and Comanche helicopters) in which steady state temperature reliability analysis led to misleading results. It turned out that major cause of failure was related not to steady state temperature but to temperature cycling (temperature gradients). Generally, due to high complexity (heterogeneity) of computer hardware devices or components, it is difficult to build models for reliability testing of such devices.
3. Standard reliability procedures for complex devices or systems require that first reliability of individual components is estimated. As hinted before, this is not feasible for electronic hardware, because components are too complex or it is not possible to assess reliability of individual parts of a component. Recent studies have shown that there is more to that. It has been found that a large number of hardware failures over the past decade were *not* component failures, but could be attributed to interconnects, connectors, system design, etc. (Lall, 1996) It suggest that reliability testing of individual components is not enough, the whole devices should be tested, too.

These features make reliability assessment (through prediction or accelerated testing) of highly reliable hardware a challenging task. This work is devoted to the problems of accelerated testing. Despite difficulties with complexity of devices and lack of well justified models for analysis of results, the current industry practice shows that accelerated testing is a commonly used technique for estimation of reliability of more and more complex hardware devices. Some examples of recent tests reported in literature are given in Table 1.1.

TABLE 1.1

RECENT ACCELERATED TESTS OF ELECTRONIC HARDWARE

Year	Company / Country	Devices tested
1991	AT&T	integrated logic family (IFL), element of supervisory logic circuit for submarine cables
1995	South Korea	CMOS ICs
1995	NASA	multi-chip modules (MCMs)
1995	USA	HEMT MMIC 44GHz
1995	DEC	SRAM
1995	Motorola	ICs
1995	Electr. & Telecommunication Research Institute, S. Korea	circuit boards for telecommunication systems
1995	AMD	VLSI devices

Various models and approaches are used for analysis of accelerated tests and extrapolation of results to the nominal conditions, as described in Chapters 2 and 3. Many of the models are simplified or otherwise not appropriate for very complex products. The work in the thesis was motivated by the current practice of extensive use of accelerated reliability testing of complex computer hardware, confronted with shortage of flexible models for analysis of results.

Chapter 2 Analysis of Data at a Specified Stress Level

This Chapter is devoted to the problem of modelling failure data of electronic equipment observed during a life test at a specified constant stress level. The purpose of data analysis is to derive an estimate of the life distribution of the population observed based on lifetimes of items put on test. The life data from which the population life distribution is inferred can be either complete (i.e., the test is carried out until all the items tested fail) or censored (i.e., the test is finished after a specified time or after a specified number of items fail).

The Chapter starts with setting the scene for accelerated life test data analysis. Basic concepts are introduced such as types of ALT data and basic assumptions on the ALT experiment organization. Then standard approaches to data analysis are surveyed.

A flexible data analysis method is discussed based on a generic family of cumulative distribution functions comprising Laguerre polynomials. The generic family is fitted to data in order to model the life distribution function. Methods for testing adequacy of models fitted to data are presented. The methods are based on statistical goodness of fit testing.

2.1 Foundations: life test organization and types of data

The method for life data analysis considered in the thesis is based on several assumptions about the organization of an accelerated test in which failure data to be processed is obtained.

- Samples of products are put on test at k *constant* stress levels denoted $S_1, S_2, \ldots, S_k$ (and listed here in the order that if $i<j$ then stress S_j is *more severe* than S_i, which means that average lifetime of items tested under S_j is expected to be shorter than under S_i). During a test, times to failure are observed (ADTs are not considered, although some results could be possibly applied for ADT data). The number of units tested at the stress level S_i is denoted n_i, $i = 1, 2, \ldots, k$. S_0 is the stress level at the nominal working conditions. Product life distribution under S_0 will be inferred from data obtained under $S_1, S_2, \ldots, S_k$. It is assumed that a single accelerating factor is allowed (multiple stresses are not considered in the thesis).
- Stress loading policies other than constant stress (e.g., step stress; stress increased continuously etc.) are not considered in the thesis. Originally, step stress or otherwise increased stress was applied in order to force all units to fail more quickly then under constant stress. Now, this seems no longer necessary since methods for inference from censored data exist. Besides, analysis of data from non constant stress requires much more complicated models and methods than with constant stress levels.
- Life data to be analyzed can be complete or censored. Table 2.1 describes basic policies for an accelerated life test, from which different types of life data to be processed are obtained.

TABLE 2.1

POLICIES FOR AN ACCELERATED LIFE TEST AND TYPES OF DATA OBTAINED

Test organization	Type of data obtained
Under stress level S_i, the test is carried out until all the n_i units tested fail. Life data collected in such a test consist of n_i lifetimes denoted $t_{(1)i} \le t_{(2)i} \le \ldots \le t_{(n_i)i}$.	Complete (uncensored) data
Under stress level S_i, the test is carried out over a fixed time period denoted T_i. Exact lifetimes are observed only for those units that fail by the censoring time T_i; lifetimes of all other units are only known to exceed T_i. The number of failures observed at this test policy is random.	Type I (time) censored data. (More precisely, singly right Type I censored data, see section 2.4.2.1).
Under stress level S_i, the test is carried out until a fixed number r of units fail. The test is terminated at the time of the rth unit failure. Only r smallest lifetimes are known exactly $t_{(1)i} \le t_{(2)i} \le \ldots \le t_{(r)i}$; lifetimes of all other units are only known to exceed $t_{(r)i}$. At this test policy, test duration is random.	Type II (failure) censored data

Life tests involving Type I or Type II censoring are used to reduce test time (and cost) since they usually require shorter test duration than complete tests.

Complete and censored data require that different analytical methods are used in order to fit a model to the data. Inference procedures developed in the thesis allow to analyse complete as well as Type I and Type II censored data which will be discussed in section 2.4.

2.2 Standard approaches to data analysis

Several standard life data analysis methods have been described in literature; extensive overviews are given by Nelson (1990 and 1982), Lawless (1982), Kalbfleish (1980). As classified in the Introduction, there are two major approaches: parametric and nonparametric methods. Here, specific most important methods are discussed in more detail.

2.2.1 Parametric methods for electronic equipment

Parametric data analysis involves fitting a parametric model to data. Usually, the methods of ALT data analysis are based on a parametric family of functions used to model product life distribution. Several models are well established and widely used for reliability estimation of electronic equipment. The most commonly used models for life analysis of electronic devices are the Weibull, extreme value and lognormal distributions.

The Weibull distribution, with the c.d.f.

$$F(t) = 1 - \exp\left(-(t/\alpha)^{\beta}\right), \quad t > 0, \alpha > 0, \beta > 0 \tag{2.1}$$

is widely used to describe the life of different kinds of electronic equipment or components, dielectrics, capacitors, etc. A special case of the distribution is the exponential distribution, with β=1 and mean α and constant hazard rate $\lambda=1/\alpha$. The exponential distribution is sometimes used to model life of e.g. certain dielectrics.

The extreme value distribution is closely related to the Weibull distribution in that if a r.v. T has a Weibull distribution (2.1), then $\ln T$ has an extreme value distribution with a c.d.f.:

$$F(y) = 1 - \exp\left(-\exp\left(\frac{y-u}{b}\right)\right), \quad -\infty < y < \infty \tag{2.2}$$

and with $u=\ln\alpha$ and $b=1/\beta$. This distribution may be suitable for "weakest link" products (Nelson 1990).

The lognormal distribution has been widely used for analysis of ALT data. It may be suitable for modelling life distributions of many solid state components, semiconductors, diodes, GaAs FETs, electrical insulation etc. The lognormal c.d.f. is:

$$F(t) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right), \quad t > 0 \tag{2.3}$$

where Φ is the standard normal c.d.f. (2.5). The parameters μ ($-\infty<\mu<\infty$) and σ (positive) are the mean and standard deviation of logarithm of lifetime (and *not* of lifetime). (Mean of the lognormal distribution is $\exp(\mu+\sigma^2/2)$).
The lognormal distribution describes products whose log of lifetime is normally distributed. The distribution is widely used for ALT data despite its (unpleasant) feature that the hazard function decreases to zero for large values of time.

The normal distribution has not been widely used for electronic equipment ALT data modelling. Its applications include some products with wear-out failure, such as electrical insulation or light bulbs. The normal c.d.f. is

$$F(t) = \Phi\left(\frac{t-\mu}{\sigma}\right), \quad -\infty < t < \infty \tag{2.4}$$

where Φ is the standard normal c.d.f. defined as

$$\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} e^{-u^2/2} du \tag{2.5}$$

The normal hazard rate increases without limit, that is why it is used for wear-out failure. Since the distribution is defined over (-∞,+∞), its fraction for times below zero must be small for practical life data analysis.

Although less widely used for ALT data analysis, a number of other models are available for specific cases of ALT. Examples include: the gamma distribution with the c.d.f. provided in statistics textbooks, and the three parameter generalized gamma distribution described by Lawless (1982).

Nelson (1990) presents some other specific models such as distributions with failures at time zero and distributions with eternal survivors.

Mixtures of distributions (presented by Nelson 1990 and Lawless 1982) are used when a population of products tested consists of some number k of subpopulations with different distributions F_i (that may be due to differences in fabrication, raw materials, environment, usage, etc.) If proportions of subpopulations are denoted p_i with $\sum_1^k p_i = 1$ then the population has the c.d.f. (Lawless 1982):

$$F(t) = \sum_{i=1}^{k} p_i F_i(t) \tag{2.6}$$

If F_i's are *not* completely known, mixture models (even for k as low as 2 or 3) are difficult to use due to a large number of unknown parameters.

Drapella (1986) proposed another specific approach to handling nonhomogeneous populations of products comprising subpopulations of faulty and good units. He developed a family of generalized distribution functions, such as the generalized Weibull distribution. It is based on (2.1) with the β parameter not constant but changing over time. The purpose of this modification is to model bimodal distributions that often arise with heterogeneous populations. He also proposed a five parameter Highly Flexible Reliability Distribution (HFR) that is a flexible model whose parameters have physical interpretation.

The inference procedures based on a parametric model involve three steps:

- choice of a model;
- fitting the chosen model to life data;
- checking if the model is adequate for the data.

In some situations there are good theoretical or empirical reasons to justify selection of a particular parametric model. However, there are ALT studies where it is not feasible to choose an appropriate model. This is often the case if products are analysed for which no theoretical (physical) model of deterioration exists, or little past experience is available, or products are highly heterogeneous with unknown proportions and distributions of subpopulations (which makes models like (2.6) very difficult to use).

In most parametric analyses, the choice of a model is a crucial and delicate matter. It is important to remember that confidence areas obtained with parametric methods (that give information about statistical uncertainty of results due to limited sample sizes) do not account for errors due to the choice of an inadequate model.

2.2.2 Nonparametric methods (or distribution-free methods)

Unlike parametric methods, nonparametric methods do not involve an assumption about a parametric family of distributions to be used to model the life data. They can be used when it is not possible or practical to find a specific parametric model.

Nonparametric methods give estimates and confidence intervals for various parameters of life distribution, such as the mean, standard deviation, percentiles or reliability at a specified age (point in time). Lawless (1982) discusses several standard nonparametric methods such as: life tables (giving estimates of survival probabilities); nonparametric estimation of the survivor or hazard functions (by plotting e.g., the empirical survival or empirical distribution functions, or the empirical cumulative hazard function). Examples of some specific nonparametric methods can be found in literature. For instance, Shaked and Singpurwalla (1982) developed a nonparametric method in which inference from ALT data can be made with an assumption that distributions belong to a location-scale family of distributions with the shape parameter remaining constant for all levels of stress.

Nelson (1990) argues that nonparametric methods are little used for practical accelerated test data analysis due to the facts that nonparametric estimates are less accurate than parametric ones (unless the parametric model is inadequate), and percentiles of distributions can be estimated only within the range of data (no extrapolation outside the available data range is possible).

2.3 Generic family of distributions

The first step of the ALT data analysis method proposed in the thesis involves modelling the life data obtained under each stress level. This section presents the approach to model complete as well as Type I and Type II censored data using a generalized parametric family of distributions.

We assume that for each stress level $S_1, S_2, \ldots, S_k$ and the nominal stress level S_0 the corresponding life distribution functions $F_1(t), F_2(t), \ldots, F_k(t)$ and $F_0(t)$ belong to the parametric family of functions $F_{\alpha,\lambda}(t)$ based on Laguerre polynomials and defined by the following formula (Baskin 1988):

$$F_{\alpha,\lambda}(t) = \sum_{j=0}^{2} b_j c_j(t) \tag{2.7}$$

where:

$$c_j(t) = \int_0^{\lambda t} x^{j+\alpha} e^{-x} dx \tag{2.8}$$

$$b_0 = \frac{1}{\Gamma(1+\alpha)}\left(3 + \frac{5}{2}\alpha + \frac{\alpha^2}{2} - \lambda m_1(3+\alpha) + \frac{\lambda^2 m_2}{2}\right) \tag{2.9}$$

$$b_1 = -\frac{1}{\Gamma(1+\alpha)}\left(3 + \alpha - \lambda m_1 \frac{5+2\alpha}{1+\alpha} + \frac{\lambda^2 m_2}{1+\alpha}\right) \tag{2.10}$$

$$b_2 = \frac{1}{2\Gamma(1+\alpha)}\left(1 - \frac{2\lambda m_1}{1+\alpha} + \frac{\lambda^2 m_2}{(1+\alpha)(2+\alpha)}\right) \tag{2.11}$$

$$m_i = \int_0^{\infty} t^i dF(t), \qquad i = 1,2 \tag{2.12}$$

with α>-1, λ>0.

Functions of the family $F_{\alpha,\lambda}(t)$ include four parameters: α (the shape parameter), λ (the location parameter), m_1 and m_2 (the first and the second distribution moment). (For simplicity of notation, the symbol $F_{\alpha,\lambda}$ will be used to denote the four parameter family (2.7), instead of perhaps more consistent $F_{\alpha,\lambda,m_1,m_2}$). The functions (2.7) are not always probability distributions; the values of parameters for which (2.7) have the properties of probability distributions are provided in Appendix A.

Functions $F_{\alpha,\lambda}(t)$ approximate several typical distribution functions such as the normal, lognormal, Weibull and gamma distributions and mixtures of distributions. In Appendix A, properties of the family of functions (2.7) are discussed in more detail.

2.4 Fitting generic life distribution functions to complete and censored data

The life distribution function for each test stress level S_i is estimated by fitting the generic family $F_{\alpha,\lambda}(t)$ to observed data. Life data obtained at each stress level can be either complete or time or failure censored. The next sections present methods of fitting the generic family of parametric functions $F_{\alpha,\lambda}(t)$ to complete and censored data.

2.4.1 Complete data

The most straightforward method of fitting a model to complete data is based on least squares estimation of the model parameters. It involves minimization of the distance between the fitted model and the empirical distribution at the stress level S, defined as:

$$\hat{F}(t,S)=\begin{cases}0 & \text{for} \quad t \le t_1 \\ \frac{k}{n} & \text{for} \quad t_k < t \le t_{k+1}, \; k=1,2,\ldots,n-1 \\ 1 & \text{for} \quad t > t_n\end{cases} \tag{2.13}$$

where, for simplicity, throughout this section $t_1,t_2,\ldots t_n$ denote the order statistic of the data (i.e., times to failure ordered so that $t_1 \le t_2 \le \ldots \le t_n$). More specifically, for each stress level S, for which complete data is available, a function of the class (2.7) is fitted to test data by solving the following optimization problem:

$$\min_{(\alpha,\lambda,m_1,m_2)\in D} \sum_{j=1}^{n} \left(F_{\alpha,\lambda}(t_j)-\hat{F}(t_j,S)\right)^2 \tag{2.14}$$

where $\hat{F}(\cdot,S)$ denotes the empirical distribution at the stress level S (2.13) and $D \subset R^4$ is a domain of $F_{\alpha,\lambda}(t)$ parameters in which the function has the properties of a c.d.f. (see Appendix A).

The vector of parameters obtained from (2.14), denoted $\hat{\Theta}=\left(\hat{\alpha},\hat{\lambda},\hat{m}_1,\hat{m}_2\right)$, defines the least squares estimate $F_{\hat{\alpha},\hat{\lambda}}$ of the population distribution function for the stress S.

Methods of solving the optimization task (2.14) are discussed in Appendix A.

2.4.2 Censored data

In the case of censored data, parameters of the model (2.7) are obtained by the Maximum Likelihood (ML) estimation. Time and failure censored data require slightly different approach which is presented in the following sections.

2.4.2.1 Type I (time) censored data

We assume that for the stress level S, n items are put on test which would give in the case of complete data a random sample $t_1,t_2,\ldots,t_n$ of n individual lifetimes with the p.d.f. $f(t,\Theta)=\frac{dF_{\alpha,\lambda}(t)}{dt}$ and the survivor function $R(t,\Theta)=1-F_{\alpha,\lambda}(t)$. Θ denotes the vector of parameters of the model, $\Theta=(\alpha,\lambda,m_1,m_2)$.

In the cases where, due to time or cost constraints, it is not possible to run the test until all the items fail, only some of the lifetimes are observed. If T_i, $i=1,2,\ldots,n$ denotes the fixed censoring time for the i-th item on test, then we observe:

the failure time t_i for the item i for which $t_i < T_i$, and

the censored time T_i for the item i for which $t_i \geq T_i$.

The sample likelihood function for the time-censored data is equal (Nelson 1982):

$$L(\Theta) = \prod_{i=1}^{n} f(t_i,\Theta)^{\delta_i} R(T_i,\Theta)^{(1-\delta_i)} \tag{2.15}$$

where the indicator δ is defined as:

$$\delta_i = \begin{cases} 1 & t_i < T_i \text{ (i.e. failure observed)} \\ 0 & t_i \geq T_i \text{ (i.e. censored)} \end{cases} \tag{2.16}$$

The most common way of running an accelerated life test is to assume that all units tested are subject to the same censoring time ($T_1 = T_2 = \ldots = T_n = T$, this is known as *singly* Type I censored data), which gives the likelihood function:

$$L(\Theta) = \prod_{i=1}^{n} f(t_i,\Theta)^{\delta_i} R(T,\Theta)^{(1-\delta_i)} \tag{2.17}$$

The model (2.7) is fitted to the time censored data by maximizing the likelihood function (2.17) with respect to the vector of parameters Θ, with the same constraints as in (2.14).

2.4.2.2 Type II (failure) censored data

Using similar notation as in the previous section, the Type II (failure) censored data consists of the r smallest lifetimes out of n values of the random sample $t_1, t_2, \ldots, t_n$ (that would be observed if the data was not censored). In other words, only the values: $t_{(1)}, t_{(2)}, \ldots t_{(r)}$ are observed, whereas the remaining n-r units on test yield times to failure that are only known to exceed $t_{(r)}$.

The joint p.d.f. of $t_{(1)}, t_{(2)}, \ldots t_{(r)}$ is (Lawless 1982):

$$L(\Theta) = \frac{n!}{(n-r)!} \cdot f\left(t_{(1)},\Theta\right) \cdot f\left(t_{(2)},\Theta\right) \cdot \ldots \cdot f\left(t_{(r)},\Theta\right) \cdot R\left(t_{(r)},\Theta\right)^{n-r} \tag{2.18}$$

This gives the sample likelihood in the case of failure censored data. The model (2.7) is fitted to the failure censored data by maximizing the likelihood function (2.18) with respect to the vector of parameters Θ, with the same constraints as in (2.14).

It is noteworthy that the form of the sample likelihood function L is very similar in both cases of Type I (2.17) and Type II (2.18) censoring. The units whose lifetimes were observed during the test contribute a term $f(t_i,\Theta)$ to the likelihood function; all other units contribute a term $R(T,\Theta)$ where T is the censoring time (T is fixed for Type I censoring or random for Type II censoring).

The vector of parameters obtained by maximizing (2.17) or (2.18), denoted $\hat{\Theta}=\left(\hat{\alpha},\hat{\lambda},\hat{m}_1,\hat{m}_2\right)$, defines the ML estimate of the population distribution function $F_{\hat{\alpha},\hat{\lambda}}$ under the stress S for which censored data was obtained.

The maximum likelihood estimate $\hat{\Theta}$ of the model parameter is the value of Θ that maximizes the sample likelihood $L(\Theta)$ or log likelihood $\mathsf{L}(\Theta)=\ln(L(\Theta))$. The standard approach to find $\hat{\Theta}$ is to solve the likelihood equation $\delta\mathsf{L}(\Theta)/\delta\Theta=0$ and to choose solutions that maximize $\mathsf{L}(\Theta)$. However, for the model (2.7) we find it more convenient to find the ML estimates by direct maximization of $L(\Theta)$ rather than solving the likelihood equations. This complex optimization problem is solved by using global optimization methods (such as simulated annealing) which will be dealt with in Appendix A.

2.5 Checking adequacy of the model

Since further inference is based on the model (2.7) fitted to the data by solving (2.14) or maximizing likelihood functions (2.17) or (2.18), it is important to check if the model is adequate. An inadequate model may results in incorrect estimates of distribution quantiles that are computed in the following stages of ALT data analysis, as described in Chapter 4.

Some simple methods of checking the model's adequacy, presented by Lawless (1982) and Nelson (1990), consist in plotting the data on an appropriate distribution paper. For instance, $\log\left[-\log\hat{R}(t)\right]$ (empirical reliability function) plotted versus log t (time) should be approximately linear if the Weibull model assumption is met. Similar checks are available for other standard models. Another approach based on data plotting, used by Barlow et al. (1988), is a nonparametric technique known as total-time-on-test (TTT) plot. TTT plots display the total time on test versus the fraction of units failed and allow to check if some specific models are adequate.

Checks based on data plotting are generally rather subjective. Adequacy of the generalized model (2.7) will be checked using statistical goodness of fit tests. Goodness of fit testing techniques used in the thesis are presented in Appendix B.

In order to check if the model $F_{\hat{\alpha},\hat{\lambda}}$ fitted to complete data (see (2.14)) is adequate, statistics based on empirical distribution functions should be computed using the formulae (B.5), (B.8) and (B.9). The statistics estimate distance between empirical data and the model fitted to data. If the statistics exceed values in Table B.1 then the model should be considered inadequate under a specified test significance level. If the model $F_{\hat{\alpha},\hat{\lambda}}$ was fitted to censored data, then relevant statistics should be obtained using (B.10), (B.15) and (B.16) (in the case of failure censoring), and (B.11), (B.17) and (B.18) (in the case of time censoring). Large values of the statistics indicate that the model fitted is inadequate. Detailed motivation of the procedures outlined here is given in Appendix B.

Chapter 3 Extrapolation in Stress

This Chapter is devoted to extrapolation of results of accelerated life tests in order to make predictions about life distribution of products under the nominal stress level. Extrapolation of results in stress usually involves an assumed relationship between life and stress, also known as the ageing or acceleration model. Several standard life-stress relationships are discussed as well as some other techniques for extrapolation based on specific regression models. This Chapter concludes with presentation of an approach to extrapolation that can be used with the model for constant stress data based on the generalized distribution discussed in Chapter 2.

Extrapolation in stress is inevitable with accelerated life data analysis. However, selection of an appropriate acceleration model is considered one of the most serious difficulties in inference from accelerated tests. Although most authors agree that models for extrapolation should be based on knowledge of physics of the important failure mechanisms, Meeker and Escobar (1993) state that extrapolation is, in most cases, difficult or impossible to justify completely. Indeed, justification for extrapolation should come from the fact that models used have physical or chemical basis, and moreover, the models agree the extensive previous experience in testing similar products. Although desirable, such justification is in real life ALTs of electronic equipment hardly ever possible due to rapid technology changes and high complexity of products. The problem has become even more difficult as some models, so far believed to have a good basis in the physics, turn out to be inappropriate for the failure processes involved in accelerated testing of e.g., modern integrated circuits. This topic will be discussed in section about the Arrhenius relationship.

One the other hand, an alternative approach in which models are based entirely on the test data (i.e., with no attempt to incorporate any physical or chemical considerations) also incurs difficulties, especially in extrapolation.

3.1 Survey of methods for extrapolation in stress

Various life stress relationships appropriate for some specific products and stress factors have been discussed in literature. Extensive surveys of available methods are given by Nelson (1990) and also by Crowder et al. (1991), Jensen (1995), Lawless (1982), Meeker and Escobar (1993), Padgett (1984). MIL-HDBK-217F (1991) provides a number of models suitable for specific kinds of electronic equipment.

Extrapolation in stress is in fact a problem of regression. In many commonly used parametric regression methods, it is assumed that lifetime distributions can be modelled with a specified parametric family of functions (such as e.g., Weibull). The specified distribution model is extended to include the stress as a parameter (parameters included in the model are also referred to as regressor or concomitant variables, or covariates). It is commonly assumed that only one parameter of the specified distribution model depends on stress. For instance, in the widely used Weibull regression model (based on (2.1)), only the α parameter depends on stress, with the β parameter remaining the same for all levels of stress. Similarly, with other commonly used regression models, it is typically assumed that stress effects only the *scale* parameter of the lifetime distribution, and not the *shape* parameter.

The most widely used regression models for extrapolation of ALT results are so called "accelerated failure time" models. Other models (methods) are also available, such as the proportional hazard (PH) models.

3.1.1 Accelerated failure time models

Accelerated failure time regression models (know also as the "accelerated life models"), are based on the assumption that increased stress shrinks the time to failure by a deterministic factor (Meeker, Escobar, 1993):

$$F_S(t) = F_0(\rho(S) \cdot t) \tag{3.1}$$

where F_S is the product lifetime c.d.f., F_0 is some baseline lifetime c.d.f., and $\rho(S)$ is some (positive) function of stress.

The generic model (3.1) includes, as special cases of the function $\rho(S)$, the widely used specific models such as the Arrhenius, inverse power or Eyring relationships. By assuming some specific forms of the life distribution F_0, we get standard models for ALT data analysis, such as the Arrhenius-Weibull, Arrhenius-lognormal, power-Weibull models, etc.

3.1.2 The Arrhenius relationship

The Arrhenius life-stress relationship is commonly used as the regression model for various kinds of electronic equipment lifetested under increased temperature. Applications of the Arrhenius models include (MIL-HDBK-217F, 1991, Nelson 1990): very high speed integrated circuits (VHSIC), VLSI CMOS devices, dielectrics, semiconductors and solid state devices, capacitors, insulating materials, battery cells etc.

The Arrhenius model is defined by the formula:

$$\tau = A \exp\left(\frac{E}{kT}\right) \tag{3.2}$$

where A and E are model parameters that depend on the product tested, k is Boltzmann's constant, T is the value of stress (the absolute temperature in Kelvin) and τ denotes *life* under stress T. If the model is used with e.g., the exponential life distribution, then τ is the mean time to failure; if the Weibull model is used, then τ denotes the α parameter or the model (known as the characteristic life); generally, τ denotes some quantile of the life distribution.

The Arrhenius relationship is based on a physical model of deterioration processes due to chemical reaction or metal diffusion. E has the physical interpretation as the activation energy of the process. It is also believed that when the Arrhenius relationship turns out to hold, it suggests degradation due to a chemical reaction or metal diffusion (Nelson 1990).

The relationship has been extensively used for a vast number of electronic products subjected to temperature stress. However, recent research has shown that application of the model for accelerated testing of integrated circuits (ICs) is doubtful (Jensen 1995, Lall 1996). The major failure mechanisms in ICs are caused by other applications of temperature then steady temperature (which is described by the Arrhenius relationship). For these other failure activating applications of temperature (such as temperature cycling, spatial temperature gradient etc.), even if the Arrhenius model is assumed, it is not based on the physics of failure analysis, but it is used merely as a parametric model fitted to data.

The Arrhenius life-stress relationship is a two parameter model, i.e., parameters A and E are to be fixed then fitting the model to data. When a linearized form of the relationship is used:

$$\log(\tau) = \gamma_0 + \gamma_1 \cdot \frac{1}{T} \tag{3.3}$$

then it is straightforward to fit the model to data by using linear regression (a straight line has to be fitted to a series of points of the form (log life, inverse stress) by using the least squares minimization).

3.1.3 The Inverse Power Law

Contrary to the Arrhenius relationship, the inverse power law (or power law) is not based on theoretical model of deterioration. It has proved to be empirically adequate for many products subjected to various stress factors such as voltage, load etc.

Applications of the inverse power model include: dielectrics and electrical insulation in voltage-endurance tests, incandescent lamps, flash lamps, metal fatigue. The model is also applied to ageing of multicomponent systems (Padgett, 1984). The following formula defines the power relationship (Nelson 1990):

$$\tau = \frac{A}{V^{\gamma_1}} \tag{3.4}$$

where A and γ_1 are parameters characteristic of the products tested, V is a positive stress variable and τ is the product life, usually defined as the mean or characteristic life (if F_0 in (3.1) is assumed to be the exponential or Weibull distribution, respectively), or in general case, it can be a specified quantile (percentile) of the life distribution.

The linearized form of the relationship:

$$\log(\tau) = \gamma_0 - \gamma_1 \cdot \log(V) \tag{3.5}$$

is used in a similar manner as the linearized Arrhenius model. To fit the model to data, a straight line is fitted to the points ($\log(\tau)$, $\log(V)$) by using least squares minimization.

3.1.4 The Eyring relationship

The Eyring relationship, based on quantum mechanics, is an alternative to the Arrhenius relationship for temperature acceleration. The model is of the form (Nelson 1990):

$$\tau = \frac{A}{T}\exp\left(\frac{B}{kT}\right) \tag{3.6}$$

where A and B are parameters characteristic of the product, T is the absolute temperature and k is Boltzmann's constant. For small changes of T, the Eyring model is close to the Arrhenius model.

3.1.5 Endurance limit relationships

Endurance limit relationships may be useful to model life as a function of voltage stress of certain dielectrics and insulations (Nelson 1990). ALT data obtained for such materials suggest that samples tested below a certain stress level (called endurance limit) exhibit virtually infinite life. The model is useful even if there is no physical endurance limit stress as it allows to get a better fit to data than other models with the same number of parameters.

The power-type endurance limit relationship is defined by the formula:

$$\tau = \begin{cases} \gamma_0 (V - V_0)^{-\gamma_1}, & V > V_0, \\ \infty, & V \le V_0. \end{cases} \tag{3.7}$$

where V_0, γ_0 and γ_1 are parameters of the model, with V_0 being a positive endurance limit. V denotes a stress level and τ is life (usually a certain quantile (percentile) of the life distribution).

Another useful form of endurance limit relationship is the following:

$$\log(\tau) = \begin{cases} \gamma_0 + \gamma_1 \log(V), & V > V_0, \\ \infty, & V \le V_0. \end{cases} \tag{3.8}$$

where γ_0 and γ_1 are parameters characteristic of the product, V_0 is a positive endurance limit, V is a stress level and τ is life.

3.1.6 Other relationships

Nelson (1990) and MIL-HDBK-217F (1991) provide a number of other life-stress relationships that have proved useful for specific ALT studies of electronic equipment. For example, the exponential relationship for life τ as a function of stress V:

$$\tau = \exp(\gamma_0 - \gamma_1 V) \tag{3.9}$$

with two parameters γ_0 and γ_1; or the polynomial relationships, e.g., the quadratic relationship:

$$\log(\tau) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 \tag{3.10}$$

Here x represents a stress level or some transformation of stress, and γ_0, γ_1 and γ_2 are model parameters.

MIL-HDBK-217F (1991) also uses the *exponential-power relationship* for a vast number of electronic devices or components:

$$\tau = \exp\left(\gamma_0 + \gamma_1 x^{\gamma_2}\right) \tag{3.11}$$

with three parameters γ_0, γ_1 and γ_2 that depend on the product tested. In the model x denotes voltage stress or inverse of absolute temperature. For example, MIL-HDBK-217F (1991) uses (3.11) with the following parameters for different products:

- γ_0=0 and γ_2 in the range of 1 to 10 (depending on the technology used) for the life of resistors;
- γ_0=0 and γ_2 in the range of 1 to 18 (depending on the technology used) for the life of capacitors;
- γ_0=0 and γ_2 in the range of 10 to 15 (depending on the insulation class) for the life of inductive devices;

The stress variable x used in these cases is inverse of the absolute temperature.

The same relationship is applied for some other types of components used in electronic devices.

3.1.7 Proportional hazards models

The proportional hazards (PH) model is based on the assumption that stress has a multiplicative effect on the hazard functions of units tested under different stress levels, which can be written as:

$$h_S(t) = h_0(t) \cdot g(S) \tag{3.12}$$

where h_S is the hazard function at the stress level S, h_0 is the hazard at some base level of stress (at which $g(S)=1$), and g is some positive function of stress. The models (3.12) can be used either with the assumption that h belongs to a specified parametric family of functions (such as e.g., the Weibull hazard function $h_0(t) = \frac{\beta}{\alpha}(t/\alpha)^{\beta-1}$, refer to (2.1)), or fully nonparametrically, i.e., with no assumption for h_0. The particular model often used is the nonparametric PH Cox model that assumes that g is of the form $g(S)=\exp(S{\cdot}p)$, where p is the model parameter (or a vector of parameters if multiple regressor variables $S=(S_1,S_2,\ldots,S_k)$ are used; then $S{\cdot}p$ stands for $\sum S_i p_i$).

Nelson (1990) argues that the PH models, developed mainly for biomedical data analyses, are little used for ALT data analysis in engineering applications.

Crowder et al. (1991) present a related model know as the proportional odds model, in which stress acts multiplicatively on the term $\Pr\{T<t\}/\Pr\{T\geq t\}$ (which is the 'odds' on the event $T<t$). The model can be written as:

$$\frac{F_S(t)}{R_S(t)} = g(S)\frac{F_0(t)}{R_0(t)} \tag{3.13}$$

where F_S and R_S are the c.d.f. and reliability function at the stress level S, F_0 and R_0 are the corresponding baseline functions and g is some function of stress. This model has not been very popular with ALT studies, either.

3.2 Extrapolation with the generalized life distribution

This section discusses how a life-stress relationship can be used to estimate product life for the nominal conditions. It is assumed that life data at each stress level is modelled using the generalized family of distributions (2.7), as described in Chapter 2.

3.2.1 Motivation

The majority of life-stress relationships presented in previous sections are intended to be used together with standard life distribution models, such as the Weibull or the lognormal distributions. Nonparametric PH models (e.g., the Cox model) are an exception to that; however, they are rarely used for extrapolation of ALT data.

Typically, life-stress models express the relationship between stress and some parameter of the assumed life distribution. To simplify computations, it is usually assumed that only one parameter of the distribution changes with stress while other parameters remain the same for all stress levels. Depending on the distribution assumed, usually the 'shape' or variance parameters are constant, and the distribution 'scale' or mean parameters change according to a specified life-stress model.

Extrapolation in stress with the generalized life distribution (2.7) uses life-stress models in a slightly different manner. All parameters of the life distribution model will be allowed to change, with the life stress function used to model the relationship between distribution *quantiles* and stress, and not between distribution *parameters* and stress. Details of this follow in the next section.

Another problem with the use of life stress models in standard ALT inference procedures is related to selection of a suitable ageing model. As explained before, the models should be based on physical and chemical models of deterioration. However, physical justification of ageing models is rarely possible due to a number of reasons. Often, no physics-of-failure analysis is carried out. If standard relationships are used (such as the Arrhenius model), then the model's parameters (e.g., the activation energy E, refer to (3.2)) are used only as curve fitting parameters. Even if the Arrhenius relationship is claimed correct for some electronic product and the activation energy parameter *is* provided by manufactures, it is often of limited use. Usually, the parameter is specified with relatively big uncertainty (e.g., 0.45 up to 1.2eV) which results in orders of magnitude difference in estimated lifetimes (see Jensen, 1995).

Accelerated temperature testing of modern ICs is yet another example of difficulties with the ageing model selection. As explained in section 3.1.2, the Arrhenius relationship, believed correct and well justified for temperature accelerated tests, now seems to be misused if applied for temperature ALTs of ICs.

These concepts illustrate the current practice in extrapolation in stress with accelerated life tests. If enough research and knowledge is available, it is desirable that physical considerations are incorporated into the model. However, the models often have to be used with no justification other than a reasonably good fit to data.

3.2.2 Generalized extrapolation method

If the generalized life distribution (2.7) is assumed, extrapolation using standard methods, with only the 'scale' distribution parameter changing with stress, is no longer feasible. This section provides an alternative approach to extrapolation in stress.

In order to extrapolate in stress, a life-stress model is fitted to data which is in the form of a series of points (pairs) (t_i, S_i), $i = 1,\ldots,k$, where t_i denotes life, S_i denotes value of stress, and k is the number of stress levels for which accelerated testing has been performed. Similar procedure was performed with the standard life distributions (see e.g., (3.3) or (3.5)), with life being some parameter of the distribution such as e.g., mean or the characteristic life. However, if the generalized life distribution is used the term *life* will be given a different interpretation. More specifically, the life t_i denotes a fixed percentile of the time to failure distribution obtained from test data at the stress level S_i (as suggested by Baskin, 1988).

The generic form of a life-stress model fitted to data can be represented by a parametric family of functions $\tau(S, \mathbf{p})$, where τ denotes life, S - stress and $\mathbf{p}$ - vector of model parameters. Fitting a model to data is realized by solving the following optimization problem:

$$\min_{\mathbf{p}} \sum_{i=1}^{k} (t_i - \tau(S_i, \mathbf{p}))^2 \tag{3.14}$$

where t_i is life corresponding to stress S_i (i.e., a quantile at some fixed level, e.g., 0.1, from the life distribution at stress S_i).

For some specific forms of $\tau(S, \mathbf{p})$, the generic optimization task can be reduced to a linear regression problem. The function $\tau(S, \hat{\mathbf{p}})$, where $\hat{\mathbf{p}}$ realizes the minimum distance as defined above, is used to determine the corresponding quantile at the nominal stress level S_0:

$$t_0 = \tau(S_0, \hat{\mathbf{p}}) \tag{3.15}$$

More detailed treatment of the way of estimating quantiles at the nominal stress level as well weights associated with quantiles follow in Chapter 4 where the generalized algorithm for ALT data analysis is presented.

As discussed in previous sections, results of extrapolation are sensitive to a selection of a suitable life stress model which is a difficult and crucial problem. The model assumed should be checked with the data. These ideas need a more careful consideration which follows below.

For some products of electronic industry (especially components or materials) and for some specific stress factors, the choice of the suitable life-stress model could be straightforward since standard models may prove appropriate. In these cases, standard models are justified by either a theoretical models of deterioration or by extensive previous experience. However, in some ALT studies of new technology or non-homogeneous products it may be difficult to find an adequate life-stress model. In such cases, an option would be to use a 'trial and error' approach to selection of a suitable ageing model. An alternative approach would be to use a generalized life-stress relationship proposed in this chapter that should help determine the most appropriate life-stress relationship.

This section proposes a generalized life-stress relationship to be used as a data analysis tool for the ALT studies where there is no a priori justification for any of the standard life-stress models. The generalized relationship is intended to combine in one model the most important life-stress relationships used for ALT data analysis of electronic equipment. By fitting the new flexible model to data, extrapolation in stress is possible without knowledge about the suitable model. Also, the chance of errors of extrapolation due to the incorrect model assumed is less likely.

The generalized life-stress relationship is defined by the formula:

$$\tau(S,\mathbf{p}) = \alpha \exp\left(\gamma_0 + \gamma_1 S^{\gamma_2}\right) + \beta S^{\gamma_3}, \quad \text{for } S > 0 \tag{3.16}$$

where τ is *life*, i.e., a certain parameter of the product lifetime distribution, e.g., the mean life or some fixed quantile of the distribution; S denotes the stress level, and $\mathbf{p} = (\alpha, \gamma_0, \gamma_1, \gamma_2, \beta, \gamma_3)$ is the vector of model parameters.

The parametric family of functions (3.16) comprises the most commonly used relationships for electronic equipment subjected to various types of stress. Specifically, the models included in (3.16) are e.g., the Arrhenius relationship ($\tau = \alpha \cdot e^{\frac{\gamma_1}{S}}$); the inverse power law ($\tau = \beta \cdot S^{\gamma_3}$); and the power-exponential model ($\tau = e^{\gamma_0 + \gamma_1 S^{\gamma_2}}$) (widely used by the standard MIL-HDBK-217F, 1991).

The difficulty incurred in using the flexible model is numerical complexity of fitting the model to data. However, working with flexible models is becoming more feasible thanks to growing availability of powerful number crunching computers and modern optimization techniques (such as e.g., simulated annealing).

Also, when fitting the model (3.16) to data, a sufficient number of data points is required. The model imposes a requirement that accelerated testing should be performed for at least 7-8 stress levels. Alternatively, some of the model parameters could be fixed (such as e.g., γ_0=0, see MIL-HDBK-217F, 1991), which results in a model that is easier to handle.

Chapter 4 Estimation of Life of Electronic Equipment in Nominal Conditions Based on ALT Data

This Chapter presents the generalized inference procedure to estimate a cumulative distribution function (c.d.f.) of electronic equipment operating in nominal conditions based on life data from a series of accelerated tests. Once a c.d.f. of life is available, all other reliability measures can be obtained, such as the hazard function, survivor function, etc.

Estimation of a c.d.f. of life at nominal conditions based on data obtained at several test stress levels is a complex numerical procedure. The major stages of the procedure are: analysis of life data at each test stress level in order to estimate time to failure distributions under increased stress levels; and extrapolation of results obtained for test stress levels to the nominal conditions.

Ways of performing these stages of analysis were considered in Chapters 2 and 3. This Chapter brings together methods presented in the previous Chapters and gives an algorithm that comprises all the stages of ALT data analysis needed to estimate product life distribution at nominal conditions.

4.1 Outline of the inference procedure

Using the notation introduced before, we formulate the purpose of the inference procedure presented in this Chapter as follows:

given results (i.e., observed times to failure) $t_{(1)i}, t_{(2)i}, \ldots, t_{(n_i)i}$, $1 \le i \le k$ of a series of k accelerated life tests performed under stress levels $S_1, S_2, \ldots, S_k$ (specified here in ascending order), estimate the population life distribution $F_0(t)$ under the stress level S_0. Life data obtained at each stress level can be either complete of censored.

An estimate of $F_0(t)$ obtained is denoted $\tilde{F}_0(t)$. Major steps of ALT data analysis required to arrive at $\tilde{F}_0(t)$ are described below. They are summarized in Fig. 4.1.

For each test stress level S_i, the product life c.d.f. $F_i(t)$ is estimated. An estimate of $F_i(t)$ is denoted $\tilde{F}_i(t)$. Once the estimates $\tilde{F}_1(t), \tilde{F}_2(t), \ldots, \tilde{F}_k(t)$ are obtained and tested for adequacy, regression analysis is performed in order to obtain $\tilde{F}_0(t)$. The regression analysis is performed for quantiles of a specified level q_j from the distributions $\tilde{F}_1(t), \tilde{F}_2(t), \ldots, \tilde{F}_k(t)$. This allows to obtain a quantile of the level q_j for the nominal stress level through extrapolation. Regression and extrapolation is done for a number of predetermined quantile levels, thus giving a series of quantiles of the life distribution at the nominal conditions. By fitting a suitable family of models to the quantiles, an estimate $\tilde{F}_0(t)$ of product life distribution is obtained. Once $\tilde{F}_0(t)$ has been obtained, all other reliability parameters can be computed.

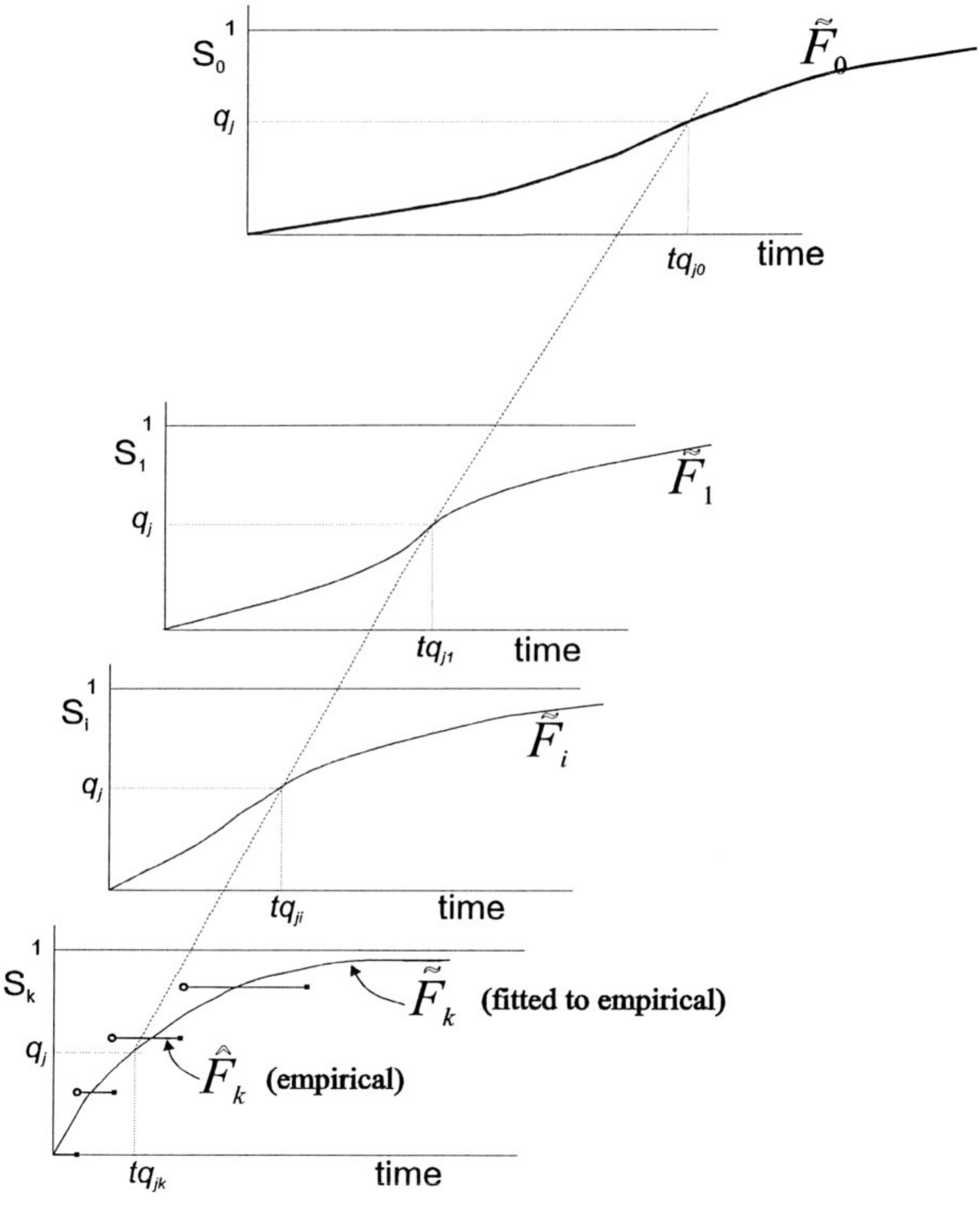

Figure 4.1 Illustration of the inference procedure to estimate the life distribution at the nominal conditions S_0. Each distribution $\tilde{F}_i$ $(i = 1,\dots,k)$ is fitted to the corresponding empirical distribution $\hat{F}_i$ (only $\hat{F}_k$ is presented in this Figure). Regression analysis is performed for quantile levels $q_1,\dots,q_j,\dots,q_m$.

The procedure outlined in this Chapter includes one more important stage. A question should be raised about adequacy of the estimate $\tilde{F}_0(t)$ of the true life distribution

$F_0(t)$. This issue is addressed in Chapter 5 where methodology is developed for verification of results of reliability prediction through ALT data analysis.

4.2 Estimation of life distributions at test stress levels

We assume that for each test stress level S_i, $1 \le i \le k$, the product life distribution F_i belongs to the family of distributions $F_{\alpha,\lambda}$ defined by (2.7). Thus, an estimate $\tilde{F}_i$ of F_i can be found by fitting to data $t_{(1)i}, t_{(2)i}, \dots, t_{(n_i)i}$ a function of the family $F_{\alpha,\lambda}$.

If data obtained under S_i is complete, then the fit is done by minimizing the least squares distance between the empirical distribution $\hat{F}_i$ (defined by (2.13)) and a function of the family $F_{\alpha,\lambda}$, as specified by (2.14): $\min_{(\alpha,\lambda,m_1,m_2)\in D} \sum_{j=1}^{n_i} \left(F_{\alpha,\lambda}\left(t_{(j)i}\right) - \hat{F}_i\left(t_{(j)i}\right)\right)^2$.

If data $t_{(1)i}, t_{(2)i}, \dots, t_{(n_i)i}$ is censored (which means that for some $p \ge 1$, the times to failure $t_{(p+1)i}, t_{(p+2)i}, \dots, t_{(n_i)i}$ are only known to exceed a certain value: a common censoring time L ($t_{(p)i} < L \le t_{(p+1)i}$) in the case of Type I censored data, and the time to failure of the pth item in the case of Type II censored data), then the best fit member of $F_{\alpha,\lambda}$ is determined by maximizing the likelihood function $\prod_{j=1}^{p} f\left(t_{(j)i}, \Theta\right) \cdot \prod_{j=p+1}^{n_i} R(L, \Theta)$, (see formulae (2.17) or (2.18) for Type I and Type II censored data). Here $f = F'_{\alpha,\lambda}$ and $R = 1 - F_{\alpha,\lambda}$.

In either case, i.e., for complete or censored data, a function denoted $\tilde{F}_i$ is obtained. The function is claimed to be the life distribution of products tested under S_i. This needs to be checked by testing goodness of fit as described in detail in section 2.5 and in Appendix B. Only the functions $\tilde{F}_i$, $1 \le i \le k$, that pass the test are taken into consideration in further stages of ALT data analysis.

4.3 Estimation of quantiles

Quantiles $tq_{1i}, tq_{2i}, \dots, tq_{mi}$ from the distribution $\tilde{F}_i$ ($1 \le i \le k$) of the predefined levels $q_1, q_2, \dots, q_m$, where

$$q_j = \frac{j}{m+1}, \quad j = 1, 2, \dots, m \tag{4.1}$$

are estimated from:

$$tq_{ji} = \tilde{F}_i^{-1}\left(q_j\right), \tag{4.2}$$

See Fig. 4.1 for illustration. Dispersions of quantiles, denoted $Dtq_{1i}, Dtq_{2i}, \ldots, Dtq_{mi}$, can be estimated from the formula:

$$Dtq_{ji} = \frac{q_j(1-q_j)}{n_i \tilde{f}_i^2(tq_{ji})} \tag{4.3}$$

where

$$\tilde{f}_i(t) = \frac{d\tilde{F}_i(t)}{dt} \tag{4.4}$$

is the p.d.f. of the distribution under S_i. The formula (4.3) is based on well known asymptotic distribution of a quantile (see e.g., Fisz 1965). For censored data, n_i is taken as the actual number of failures observed. Based on dispersions (4.3), weights $p_{1i}, p_{2i}, \ldots, p_{mi}$ of quantiles obtained from the distribution $\tilde{F}_i$, $1 \le i \le k$, can be defined as:

$$p_{ji} = \frac{\min\limits_{1 \le i \le k} Dtq_{ji}}{Dtq_{ji}} \tag{4.5}$$

This makes it possible to use weighed regression techniques to estimate a nominal conditions quantile at the level q_j, based on quantiles tq_{j1}, $tq_{j2}, \ldots$, tq_{jk}. The purpose of weighed regression is to ensure that quantiles with small dispersions have more significant effect on the regression curve (and consequently on the quantile at nominal conditions) than quantiles with large dispersions.

4.4 Regression analysis

The purpose of this stage is to obtain quantiles of life distribution under nominal conditions S_0. A quantile of the level q_j of the life distribution under S_0 is estimated (extrapolated) from a series of data points $\{(S_i, tq_{ji})\}$, $i = 1,2,\ldots,k$, using regression analysis. The regression procedure presented here is repeated independently for each quantile level ($j = 1,2,\ldots,m$).

4.4.1 Generic case: generalized life stress relationship

In the generic case, regression is based on the generalized life stress relationship $\tau(S,\mathbf{p})$, proposed in Chapter 3, (for definition refer to (3.16)). For each quantile level q_j, a parametric curve $\tau(S,\mathbf{p})$ is fitted to the data points $\{(S_i, tq_{ji})\}$, $i = 1,2,\ldots,k$, by solving the following optimization problem for a fixed j:

$$\min_{\mathbf{p}} \sum_{i=1}^{k} \left(t_{ji} - \tau(S_i, \mathbf{p})\right)^2 p_{ji} \tag{4.6}$$

where $\mathbf{p}$ is a vector of parameters of the model (3.16) and p_{ji} denote weights of quantiles defined by (4.5).

As a results, the vector $\hat{\mathbf{p}}_j$ is obtained for $j = 1, 2, \ldots, m$, for which the function $\tau(S, \hat{\mathbf{p}}_j)$ is the best fit member approximating the series of points $\{(S_i, tq_{ji})\}$, $i = 1, 2, \ldots, k$, for a given j. To solve the difficult task of minimization of the expression (4.6) it is proposed that global optimization techniques are used (such as e.g., simulated annealing).

It should be checked if the regression model $\tau(S, \hat{\mathbf{p}}_j)$ fitted to data is adequate. This can be done using the analysis of variance in regression.

We define the goodness of fit measure D_j by

$$D_j = \sum_{i=1}^{k} \left(tq_{ji} - \tau(S_i, \hat{\mathbf{p}}_j)\right)^2 p_{ji} \tag{4.7}$$

which is the residual sum of squares ("deviations from regression" sum of squares), with $(k\text{-}p)$ degrees of freedom, where p is the dimension of the vector of parameters $\mathbf{p}$.

Checking adequacy of the nonlinear model:

$$t = \tau(S, \hat{\mathbf{p}}_j) \tag{4.8}$$

obtained from (4.6) can be based on a heuristic procedure that involves analysis of variance. The test procedure presented below is discussed in the Appendix C.

First, we estimate the sum of squares of the "pure error" of the observations:

$$S_{pe,j}^2 = \sum_{i=1}^{k} (n_i - 1) Dtq_{ji} \tag{4.9}$$

Since the number of degrees of freedom of $S_{pe,j}^2$ is $\sum_{i=1}^{k} n_i - k$, the mean square of the pure error equals: $S_{pe,j}^2 \Big/ \left(\sum_{i=1}^{k} n_i - k\right)$.

Although the standard F test is *not* appropriate in this situation (see Appendix C), we can obtain some information whether the model fits the data reasonably well by examining the ratio:

$$F_j = \frac{D_j / (k - p)}{S_{pe,j}^2 \Big/ \left(\sum_{i=1}^{k} n_i - k\right)} \tag{4.10}$$

Let $F_{1-\alpha}(\nu_1, \nu_2)$ denote the quantile at the level (1-α) of the F Snedecor distribution with ν_1, ν_2 degrees of freedom, and $N = \sum_{i=1}^{k} n_i$.

If:

$$F_j < F_{1-\alpha}(k-p, N-k) \tag{4.11}$$

then *some indication* is obtained that the model may be appropriate. In this case, the model (4.8) is used to estimate the quantile at level q_j of the life distribution under the stress level S_0:

$$tq_{j0} = \tau(S_0, \hat{\mathbf{p}}_j) \tag{4.12}$$

Goodness of fit terms D_j (4.7), indicating how well regression models fitted data, can be used to define weights associated with quantiles $tq_{10}, tq_{20} \ldots, tq_{j0}, \ldots, tq_{m0}$:

$$p_{j0} = \frac{\min\limits_{1 \le j \le m} D_j}{D_j} \tag{4.13}$$

The weights will be used when fitting, through weighed least squares minimization, the generalized distribution $F_{\alpha,\lambda}$ to quantiles $tq_{10}, tq_{20} \ldots, tq_{m0}$., (section 4.5). In this way relative uncertainties of quantiles due to regression errors are taken into consideration.

4.4.2 Special cases: linearized relationships

Imposing some constraints on $\mathbf{p} = (\alpha, \gamma_0, \gamma_1, \gamma_2, \beta, \gamma_3)$, which yield specific, yet important in practical applications, forms of the generalized life-stress relationship $\tau(S, \mathbf{p})$, the optimization problem (4.6) can be reduced to a simpler linear regression task.

Specific linearizable forms of $\tau(S, \mathbf{p})$ that are important in practice are:

- the inverse power law: $\tau(S, \mathbf{p}) = \beta S^{\gamma_3}$, ($\alpha = 0$, $\beta > 0$);
- the Arrhenius relationship: $\tau(S, \mathbf{p}) = \alpha \exp\left(\frac{\gamma_1}{S}\right)$, ($\beta = 0$, $\gamma_0 = 0$, $\gamma_1 = -1, \alpha > 0$);
- the exponential model: $\tau(S, \mathbf{p}) = \alpha \exp(\gamma_1 S)$, ($\beta = 0$, $\gamma_0 = 0$, $\gamma_1 = -1, \alpha > 0$)
- the exponential-power model: $\tau(S, \mathbf{p}) = \exp(\gamma_1 S^{\gamma_2})$, ($\beta = 0$, $\alpha = 1$, $\gamma_0 = 0$). Actually, this is a special case of the exponential-power model $\tau(S, \mathbf{p}) = \exp(\gamma_0 + \gamma_1 S^{\gamma_2})$ with $\gamma_0 = 0$. According to MIL-HDBK-217F (1991), setting $\gamma_0 = 0$ is justified for a vast variety of electronic devices and components.

These special cases turn out to be of practical importance especially for accelerated life testing of electronic equipment (MIL-HDBK-217F 1991), (Nelson 1990).

These relationships can be presented in a linearized form: $y = a + bx$, where y is some function (transformation) of life τ and x is some function of stress S. Linearization of these relationships is straightforward (recalling that, as assumed in Chapter 3, S is a *positive* value of stress):

- the inverse power law:

$$\ln \tau = \ln \beta + \gamma_3 \ln S \tag{4.14}$$

hence

$$\begin{cases} y = \ln \tau \\ x = \ln S \end{cases} \tag{4.15}$$

- the Arrhenius relationship

$$\ln \tau = \ln \alpha + \frac{\gamma_1}{S} \tag{4.16}$$

and

$$\begin{cases} y = \ln \tau \\ x = \dfrac{1}{S} \end{cases} \tag{4.17}$$

- the exponential model

$$\ln \tau = \ln \alpha + \gamma_1 S \tag{4.17'}$$

and

$$\begin{cases} y = \ln \tau \\ x = S \end{cases} \tag{4.17''}$$

- the exponential-power model

$$\ln \ln \tau = \ln \gamma_1 + \gamma_2 \ln S \tag{4.18}$$

with $\ln \tau > 0$, which can be achieved by choosing a proper scale for τ, (e.g., all quantiles obtained for a specified quantile level q_j could be divided by half of the smallest one). Then it can be also assumed that $\gamma_1 > 0$.

It follows from (4.12) that:

$$\begin{cases} y = \ln \ln \tau \\ x = \ln S \end{cases} \tag{4.19}$$

It can be easily seen from the properties $Var(aX) = a^2 Var(X)$ and $\Delta y \approx \frac{dy(t)}{dt}\Delta t$ that the measure of uncertainty (dispersion) Dy_{ji} of $y_{ji} = f(tq_{ji})$ due to uncertainty of tq_{ji} given by (4.3), can be expressed as:

$$Dy_{ji} = \frac{Dtq_{ji}}{tq_{ji}^2} \tag{4.20}$$

in the cases where y=ln t, and

$$Dy_{ji} = \frac{Dtq_{ji}}{\left(tq_{ji} \ln tq_{ji}\right)^2} \tag{4.21}$$

in the cases where y=ln ln t.

Weights (expressing relative uncertainties of data points used in regression) associated with y_{ji} for a given j and $i = 1, 2, \dots, k$ can be defined as:

$$p_{ji} = \frac{\min_{1 \le i \le k} Dy_{ji}}{Dy_{ji}} \tag{4.22}$$

Adopting one of the transformations (4.9), (4.11) or (4.13) for a specified quantile level q_j, a set of data points $\{(S_i, tq_{ji})\}$, $i = 1, 2, \dots, k$ is transformed into the set $\{(x_i, y_{ji})\}$, $i = 1, 2, \dots, k$, which can be analyzed using weighted linear regression techniques, as explained in Appendix C, with weights given by (4.22).

For the special cases of life stress relationships considered in this section, a linear model is assumed of the relationship between the transformed life y and the transformed stress x:

$$y_{ji} = \beta_{j0} + \beta_{j1} x_i + \varepsilon_{ji} \tag{4.23}$$

with a fixed j and $i = 1, 2, \dots, k$.

The ε's in the above formula are assumed to be independently normally distributed with means equal 0 and unequal variances (that are unknown):

$$\varepsilon_{ji} \sim N\left(0, \frac{\sigma_j^2}{p_{ji}}\right) \tag{4.24}$$

The weights p_{ji} associated with subsequent points (x_i, y_{ji}) are defined by (4.22).

Based on the formulae given in Appendix C, estimates b_{j0} and b_{j1} of the parameters β_{j0} and β_{j1} are obtained as:

$$b_{j0} = \frac{\left(\sum p_{ji}y_{ji}\right)\left(\sum p_{ji}x_i^2\right) - \left(\sum p_{ji}x_i\right)\left(\sum p_{ji}x_i y_{ji}\right)}{\left(\sum p_{ji}\right)\left(\sum p_{ji}x_i^2\right) - \left(\sum p_{ji}x_i\right)^2} \tag{4.25}$$

$$b_{j1} = \frac{\left(\sum p_{ji}\right)\left(\sum p_{ji}x_i y_{ji}\right) - \left(\sum p_{ji}x_i\right)\left(\sum p_{ji}y_{ji}\right)}{\left(\sum p_{ji}\right)\left(\sum p_{ji}x_i^2\right) - \left(\sum p_{ji}x_i\right)^2} \tag{4.26}$$

where all the sums are taken for $i = 1,2,\ldots,k$ $(\sum \cdot \equiv \sum_{i=1}^{k} \cdot)$.

Formulae (4.25) and (4.26) provide estimates of parameters of the *model* (4.23). The resulting form of the model is given by the linear regression line:

$$\hat{y}_j = b_{j0} + b_{j1}x \tag{4.27}$$

Adequacy of the model should be checked using tests described in Appendix C. Based on concepts introduced in Appendix C (see Table C.2), the analysis of variance table for regression can be built as follows (Table 4.1):

TABLE 4.1

ANALYSIS OF VARIANCE TABLE FOR THE LINEAR REGRESSION MODEL ASSOCIATED WITH A QUANTILE q_j.

Source of Variation	Sum of Squares	Degrees of freedom	Mean Square
Due Regression	$SSR'_j = \sum_{i=1}^{k} n_i \left(b_0 + b_1\sqrt{p_i}x_i - N^{-1}\sum_{i=1}^{k} n_i\sqrt{p_i}y_i \right)^2$ where $N = \sum_{i=1}^{k} n_i$	1	$msr'_j = SSR'_j/1$
Deviation (of means) from regression	$SSE_j = \sum_{i=1}^{k} p_{ji}y_{ji}^2 - b_{j0}\sum_{i=1}^{k} p_{ji}y_{ji} - b_{j1}\sum_{i=1}^{k} p_{ji}x_i y_{ji}$	k-2	$mse_j = SSE_j/(k-2)$

"Pure Error"	$S_{e,j}^2 = \sum_{i=1}^{k}(n_i - 1)Dy_{ji}$	$\sum_{i=1}^{k} n_i - k$	$s_{e,j}^2 = \dfrac{S_{e,j}^2}{\sum_{i=1}^{k} n_i - k}$
Residual	$SSE_j' = SSE_j + S_{e,j}^2$	$\sum_{i=1}^{k} n_i - 2$	$mse_j' = \dfrac{SSE_j'}{\sum_{i=1}^{k} n_i - 2}$

The linear regression model (4.23) assumed for the quantile level q_j is considered *inadequate*, with the test significance level α, if:

$$\frac{mse_j}{s_{e,j}^2} > F_{1-\alpha}(k-2, N-k) \tag{4.28}$$

or

$$\frac{msr_j'}{mse_j'} < F_{1-\alpha}(1, N-2) \tag{4.29}$$

where $N = \sum_{i=1}^{k} n_i$ and $F_{1-\alpha}(\nu_1, \nu_2)$ is the quantile at the level (1-α) of the F Snedecor distribution with ν_1, ν_2 degrees of freedom. Interpretation of the tests (4.28) and (4.29) can be found in Appendix C.

If the model proves adequate, the formula (4.27) is used to arrive at the corresponding quantile of the life distribution under S_0:

$$\hat{y}_{j0} = b_{j0} + b_{j1}x_0 \tag{4.30}$$

where x_0 is a proper transformation of S_0 as specified by (4.15), (4.17) or (4.19).

It follows immediately from (4.15), (4.17) or (4.19) that, depending on the transformation adopted, the quantile tq_{j0} (j=1,2,...,m) is:

$$tq_{j0} = \begin{cases} \exp \hat{y}_{j0} & \text{for (4.15) and (4.17)} \\ \exp(\exp \hat{y}_{j0}) & \text{for (4.19)} \end{cases} \tag{4.31}$$

and, considering (4.20) and (4.21), its dispersion Dtq_{j0} is:

$$Dtq_{j0} = \begin{cases} tq_{j0}^2 D\hat{y}_{j0} & \text{for (4.15) and (4.17)} \\ \left(tq_{j0} \ln tq_{j0}\right)^2 D\hat{y}_{j0} & \text{for (4.19)} \end{cases} \tag{4.32}$$

where, following (C.20), the variance of $\hat{y}_{j0}$ obtained from the regression line (4.30) at x_0, is:

$$D\hat{y}_{j0} \approx mse'_j \cdot \frac{\sum p_{ji}x_i^2 - 2x_0 \sum p_{ji}x_i + x_0^2 \sum p_{ji}^2}{\left(\sum p_{ji}\right)\left(\sum p_{ji}x_i^2\right) - \left(\sum p_{ji}x_i\right)^2} \tag{4.33}$$

where the term mse'_j, as defined in Table 4.1, is the estimate of the unknown model variance σ^2 in (4.24).

To complete this stage of analysis for the special case of $\tau(S,\mathbf{p})$, for which linear regression can be applied, we compute weights associated with subsequent quantiles tq_{j0}, j=1,2,...,m:

$$p_{j0} = \frac{\min\limits_{1 \le j \le m} Dtq_{j0}}{Dtq_{j0}} \tag{4.34}$$

Analogously to (4.13), the weights express relative uncertainties of quantiles due to regression errors, and will be used in weighed least squares minimization as described in section 4.5.

4.5 Estimation of the life c.d.f. at nominal conditions

Based on quantiles tq_{j0}, j=1,2,...,m, obtained in the previous stage of analysis, the product life distribution under S_0 can be arrived at by fitting a member of the family of functions $F_{\alpha,\lambda}(t)$ (2.7) to the step function $\hat{F}(t,S_0)$ built upon quantiles obtained:

$$\min_{(\alpha,\lambda,m_1,m_2)\in D} \sum_{j=1}^{m} \left(F_{\alpha,\lambda}\left(tq_{j0}\right) - \hat{F}\left(tq_{j0},S_0\right)\right)^2 p_{j0} \tag{4.35}$$

where D is the domain of $F_{\alpha,\lambda}$ parameters in which the function has the properties of a c.d.f. (see Appendix A). The function $\hat{F}(t,S_0)$ is defined as:

$$\hat{F}(t,S_0) = \begin{cases} 0 & t < tq_{10} \\ q_1 & tq_{10} \le t < tq_{20} \\ \dots & \\ q_m & tq_{m0} < t \end{cases} \tag{4.36}$$

The function obtained by solving the optimization problem (4.35) is denoted $\tilde{F}_0$ and *is claimed* to estimate the product life distribution at the nominal stress level S_0. Methodology for verification of this claim is developed in Chapter 5.

4.6 Estimation of other reliability measures

Once the life distribution has been estimated, any other reliability measure of interest can be readily obtained, such as the *mean time to failure*:

$$MTF = \int_0^{\infty} \left(1 - \tilde{F}_0(t)\right) dt \tag{4.37}$$

the *failure rate* (or *hazard function*):

$$h(t) = \frac{\tilde{F}_0'(t)}{1 - \tilde{F}_0(t)} \tag{4.38}$$

the *mean residual life* (defined as: $m(t) = E\left(lifetime - t \middle| lifetime \geq t\right)$):

$$m(t) = \frac{\int_t^{\infty} \left(1 - \tilde{F}_0(u)\right) du}{1 - \tilde{F}_0(t)} \tag{4.39}$$

etc.

4.7 Discussion

The function $\tilde{F}_0$ obtained in section 4.5 is estimated through a complex numerical procedure that entails several models, such as models for the life distribution at each test stress level, as well as regression models for extrapolation of quantiles. Although adequacy of individual models utilized for inference is always tested, the question remains whether the claim can be justified that $\tilde{F}_0$ estimates the product life distribution under nominal conditions. Obviously, no direct justification is possible unless some reliability data is available for the nominal conditions that could be used for comparison with $\tilde{F}_0$.

The next Chapter proposes the methodology in order to indirectly verify results obtained from the inference procedure discussed in this Chapter. The methodology is intended to give evidence for or against $\tilde{F}_0$ regarded as the estimate of product field reliability. Based on the evidence provided by the verification procedure proposed, the reliability analyst can decide what degree of trust he can put in results of analysis of accelerated life test data.

Chapter 5 Verification of Results of Reliability Prediction

This Chapter presents a method for verification of validity of the product life distribution obtained for the nominal working conditions. Obviously, verification of results of reliability prediction through comparison with the field life data is usually not feasible. However, some method for results verification can be devised. The method developed in this Chapter is based on the idea that the ALT data analysis procedure described in Chapter 4 can be used to extrapolate ALT results not only to the nominal stress level, but to any other stress level, in particular to the stress levels for which empirical data is available. Then it is possible to compare predicted reliability with empirical data.

This procedure can be regarded as a way of examining if the set of models (specially the ageing model), upon which the ALT data analysis is based, is adequate for the study concerned. If adequacy of models is proven for stress levels for which empirical data is available (i.e., stress levels other than nominal), then it may be believed that the models work equally well when used to obtain results at the nominal conditions. On the other hand, if the models used turn out to be inadequate for some stress levels, then it gives indication that results obtained for the nominal conditions may not be valid.

Although individual models (i.e., the life distribution model and the ageing model) used with the generalized ALT data analysis procedure are always checked for adequacy, this is not always sufficient. It turns out for instance that for some data different ageing models that yield considerably different results may not be rejected by the adequacy of regression tests (see (4.11), (4.28), (4.29), also refer to Appendix C). For such cases, methodology developed in this Chapter gives a more definite verification if models used for estimation of reliability at nominal conditions can be considered adequate.

5.1 Problem formulation

Due to very high reliability of electronic equipment typically observed, no life testing is usually performed at the nominal working conditions. Field failure data is also very scarce. For these reasons, reliability analysts usually face the problem that there is no or hardly any empirical life data available for the nominal stress level that could be used in order to verify results of reliability prediction based on ALT data. Empirical life data is usually available only for stress levels higher than nominal under which accelerated life testing is carried out.

The question arises how in these circumstances it could be checked whether the life cumulative distribution function (that has been obtained through a complex mathematical procedure of ALT data analysis) really represents the true life distribution of products working at nominal conditions.

Obviously, the most convincing and ultimate way of doing such a check would be to compare the estimated life distribution (i.e., the distribution function obtained from the ALT data analysis based on some models) with the empirical reliability data measured under the nominal conditions. However, since that is usually not doable, (as no empirical data is available), some ways would be needed of checking if results of reliability prediction conform to observed lifetimes.

The check in order to verify validity of reliability prediction can be done by assuming that the target stress level (i.e., the stress level for which reliability is predicted) is equal to one of the test stress levels for which empirical life data is available. The life distribution can be estimated for the target stress level by carrying out ALT data analysis based on data obtained at stress levels other than the target stress level. The estimated life distribution can be then compared against empirical data thus giving indication whether the ALT data analysis tools and models are adequate for the particular ALT study. This procedure could be repeated for all the stress levels for which data is available. If the estimated life distributions (especially those at the stress levels closest to the nominal level) prove to fit well to the corresponding empirical data, one can *be confident* that the life distribution estimated for the nominal stress level would fit to the data as well. Otherwise, evidence is gathered against the results of the ALT data analysis method.

Detailed description of the method follows in the next section.

5.2 Method for verification of results of ALT data analysis

Verification of results of ALT data analysis, as introduced in section 5.1, will be done by repeating the ALT data analysis (as presented in Chapter 4) in order to estimate product reliability at the target stress levels equal to the test stress levels. This makes it possible to compare the predicted reliability with empirical data. Based on this procedure, results obtained for the nominal stress level are either justified or rejected.

The purpose of the ALT data analysis presented in Chapter 4 is to estimate the cumulative distribution function of product life at the nominal stress level S_0. The estimation is based on life data obtained from a series of k accelerated life tests carried out at stress levels $S_1, S_2, \ldots, S_k$. (Notation introduced in Chapter 2 is used throughout this Chapter).

Figure 5.1 illustrates the two main stages of ALT data analysis that are performed in order to estimate the life distribution $\tilde{F}_0(t)$ at the nominal conditions. The stages are:

- analysis of life data at each test stress level $S_1, S_2, \ldots, S_k$ (as considered in Chapter 2). The purpose of this stage is to estimate the cumulative distribution functions for stress levels $S_1, S_2, \ldots, S_k$ based on data observed at the corresponding stress level.
- extrapolation of results obtained for test stress levels $S_1, S_2, \ldots, S_k$ to the nominal stress level S_0 at which no experimental data has been collected (extrapolation in stress was discussed in Chapter 3). This gives an estimate of the product life distributions $\tilde{F}_0(t)$.

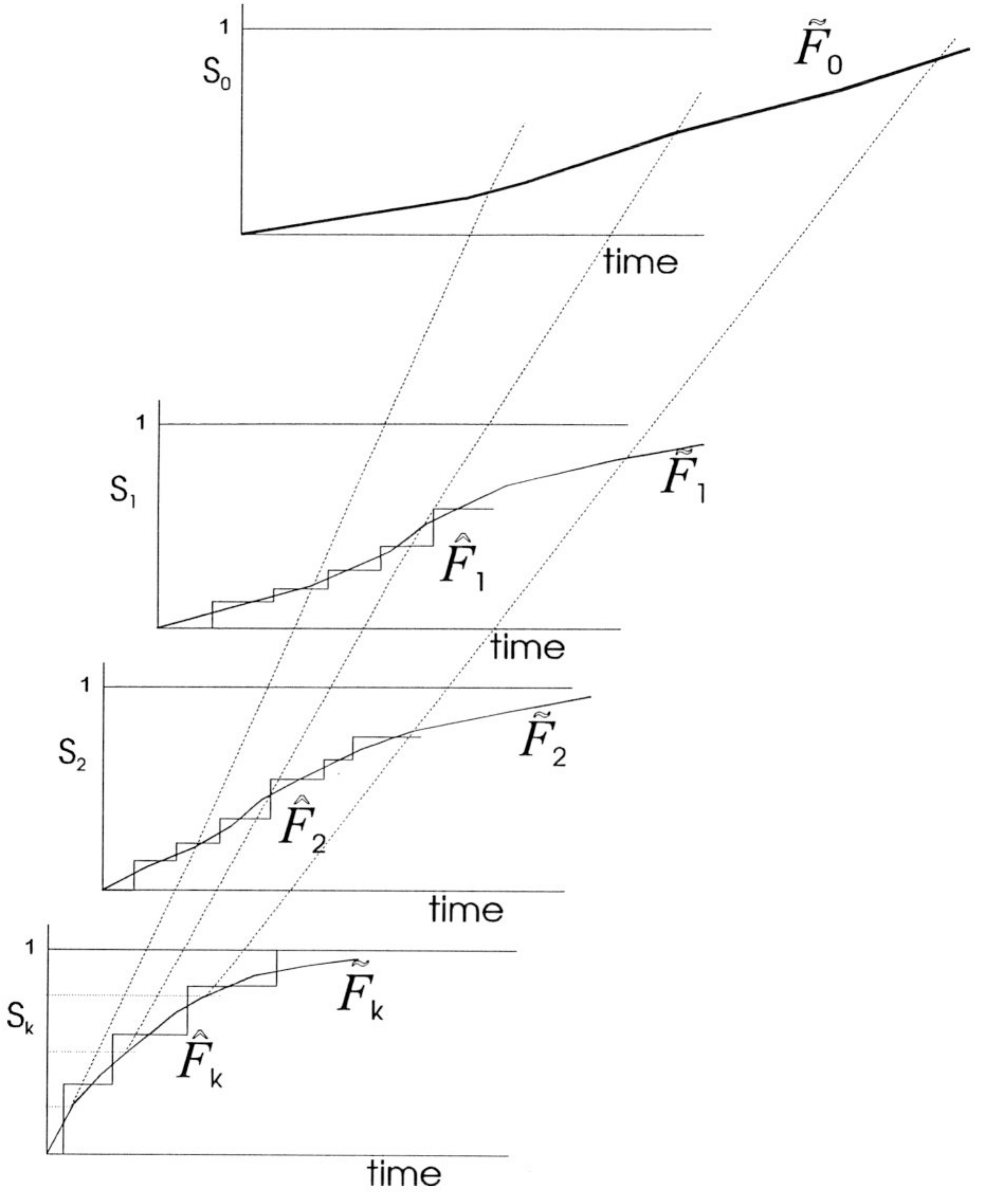

Figure 5.1 Main stages of ALT data analysis: analysis of data at test stress levels $S_1, S_2, \ldots, S_k$ and extrapolation in stress to the nominal stress level S_0.

The step functions denoted $\hat{F}_i$ (i=1,2,...,k) represent empirical distributions obtained from test data at stress levels $S_1, S_2, \ldots, S_k$, and the functions denoted $\tilde{F}_i$ (i=1,2,...,k) represent the estimated life distributions for the corresponding stress levels. The function $\tilde{F}_0$ represents the distribution function claimed to be the product life distribution at the nominal stress level.

The purpose of the verification procedure considered here is to collect evidence for or against the claim that $\tilde{F}_0$ represents the product life distribution at the nominal stress level.

The way of achieving the goal is to verify against experimental data the results obtained by applying the ALT data analysis method considered for target stress levels equal $S_1, S_2, \ldots, S_k$.

More specifically, for a specified test stress level S_j $(1 \leq j \leq k)$ the cumulative distribution at S_j is estimated based on data obtained at the test stress levels other than S_j, i.e., empirical data at S_j is not used. $\tilde{F}_j^{\text{est}}$ denotes the estimated life c.d.f. The procedure to estimate $\tilde{F}_j^{\text{est}}$ is performed in the same way and under the same models as the procedure used to arrive at the function $\tilde{F}_0$, see Chapter 4. The only (obvious) difference concerns the target stress level (which is the nominal stress S_0 in Chapter 4, and the test stress S_j here).

The distribution function $\tilde{F}_j^{\text{est}}$ is claimed to represent the product life c.d.f. corresponding to the stress level S_j, i.e., the distribution of the population from which data at the stress level S_j have been obtained.

In order to verify the claim, statistical goodness of fit tests can be used, such as the tests based on empirical distribution functions, as described in Appendix B. Tests are available both for complete and for censored data at the stress level S_j. The c.d.f.s tested can be considered fully specified (this was discussed in Appendix B), as their parameters have not been obtained from the data against which the functions are tested. Goodness of fit testing is considered in detail in section 5.3.

The method considered here is illustrated in Figure 5.2. As before, $\hat{F}_j$ represents the empirical distribution obtained from test data at stress levels S_j, and $\tilde{F}_i$ $(i \neq j)$ represent the life distributions fitted to empirical data at corresponding stress levels S_i. $\tilde{F}_j^{\text{est}}$ denotes the life distribution obtained for the target stress level S_j, based on $\tilde{F}_i$ $(i \neq j)$. The functions $\hat{F}_j$ and $\tilde{F}_j^{\text{est}}$ are compared using statistical goodness of fit tests, as outlined in section 5.3.

If predicted reliability distribution functions do not match empirical data (especially at stress levels close to S_0, e.g., S_1) then this indicates that reliability prediction at S_0 may be invalid.

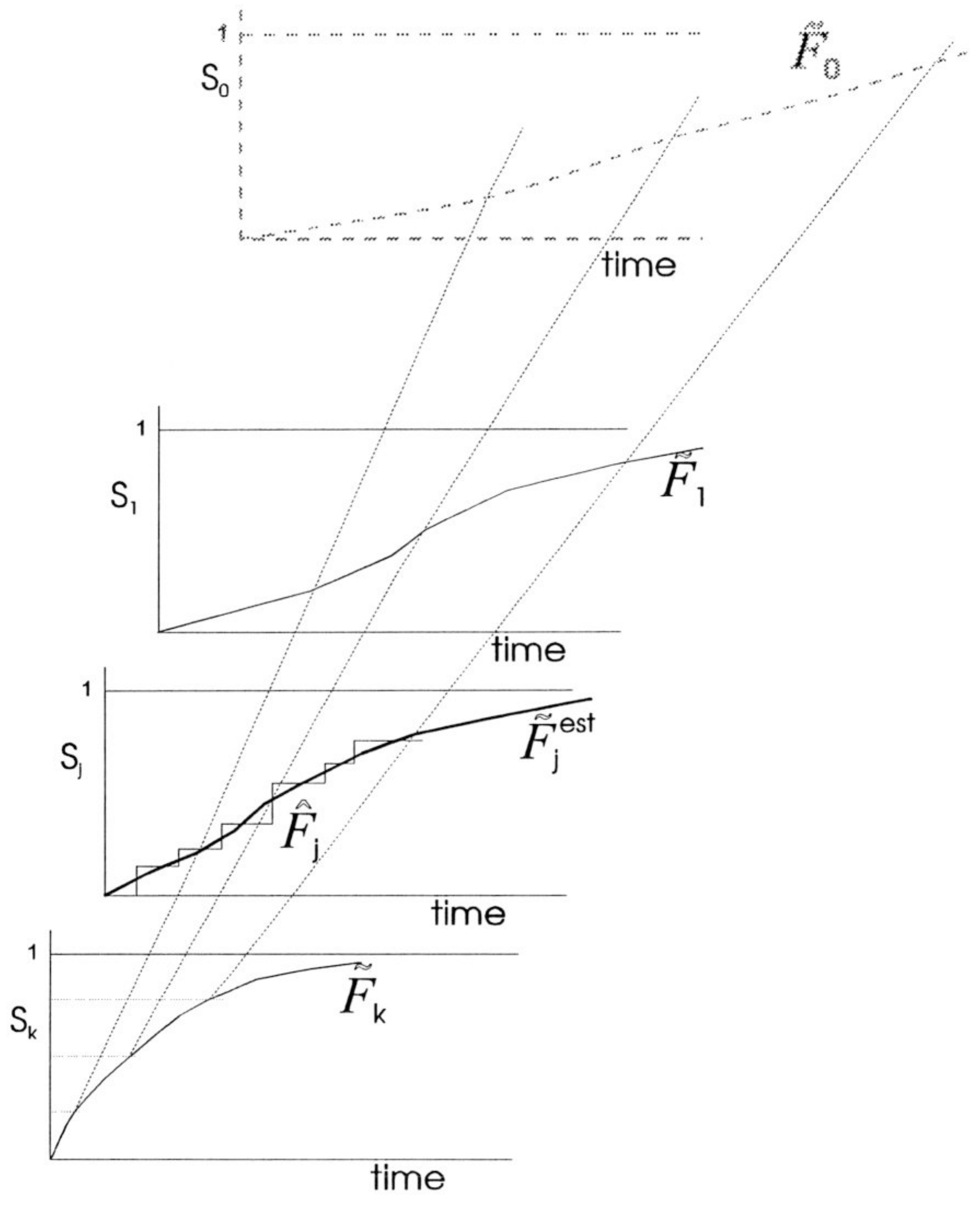

Figure 5.2 Verifying adequacy of ALT data analysis models: the estimated life distribution $\tilde{F}_j^{est}$ is compared with the empirical data ($\hat{F}_j$) at the target stress level S_j.

5.3 Testing goodness of fit

This section considers methods for testing whether the function $\tilde{F}_j^{\text{est}}$ fits well the empirical data at the stress S_j. The goodness of fit methods used here are explained in detail in Appendix B.

The task is to test the (null) hypothesis that the estimated cumulative distribution function $\tilde{F}_j^{\text{est}}$ is the c.d.f. of the time to failure of the population (of products) at the stress level S_j:

$$H_0: F_j = F_j^{\text{est}} \tag{5.1}$$

where F_j is the (true) distribution function of the population at S_j.

The goodness of fit tests used here are based on empirical distribution functions (EDF) for continuous ungrouped data.

It should be noted that once obtained from the ALT data analysis procedure, the distribution function $\tilde{F}_j^{\text{est}}$ can be regarded as fully specified, i.e., it does not involve any unknown parameters that have been estimated from data at S_j. Recall that $\tilde{F}_j^{\text{est}}$ has been estimated from data at the stress levels other than S_j, so parameters of the function do not depend on data at this stress level. This means that the EDF tests based on statistics described in Appendix B can be used. Tests for uncensored and censored data at the stress level S_j will be considered separately.

5.3.1 Uncensored data

For uncensored data, the classical goodness of fit test based on the Kolmogorov-Smirnov statistics (B.5) can be used:

$$d_n = \sup_t \left| \tilde{F}_j^{\text{est}}(t) - \hat{F}_n(t) \right| \tag{5.2}$$

where $\hat{F}_n$ is the empirical distribution defined as:

$$\hat{F}_n(t) = \begin{cases} 0 & \text{for} \quad t \le t_1 \\ \dfrac{k}{n} & \text{for} \quad t_k < t \le t_{k+1}, \ k = 1, 2, \ldots, n-1 \\ 1 & \text{for} \quad t > t_n \end{cases} \tag{5.3}$$

where n is the number of samples and $t_1 \le t_2 \le \ldots \le t_n$ is the life data at S_j.

Due to the fact that, for uncensored data, with $\tilde{F}_j^{\text{est}}$ fully specified, the test based on d_n is distribution-free, the critical value λ_α is readily determined, as described in Appendix B, where α is the significance level of the test (e.g., α=0.05).
Large values of the d_n statistic indicate evidence against H_0. If

$$d_n > \frac{\lambda_\alpha}{\sqrt{n}} \tag{5.4}$$

then H_0 is rejected. This means that the hypothesised distribution function $\tilde{F}_j^{\text{est}}$ cannot be regarded as the distribution of the population at the target stress level S_j.

5.3.2 Censored data

Statistics to be used in the case of Type II and singly Type I censored data are given by (B.10) and (B.11), respectively.
For Type II (failure) censored data, the modified Kolmogorov-Smirnov statistic (B.10) has the form:

$$d_{n,r} = \sup_{t \le t_{(r)}} \left| \tilde{F}_j^{\text{est}}(t) - \hat{F}_n(t) \right|, \tag{5.5}$$

where r is the number of observations available in the random sample at S_j, i.e., $t_{(r)}$ is the rth smallest lifetime observed and all remaining n-r lifetimes exceed $t_{(r)}$.

For singly Type I (time) censored data, the modified Kolmogorov-Smirnov statistic (B.11) has the form:

$$d_{n,p} = \sup_{t \le L} \left| \tilde{F}_j^{\text{est}}(t) - \hat{F}_n(t) \right|, \tag{5.6}$$

where L is the (common) censoring time and

$$p = \tilde{F}_j^{\text{est}}(L). \tag{5.7}$$

A critical value λ_α should be determined (please refer to Appendix B) for a chosen test significance level α (e.g., α=0.05).
Since large values of the statistics $d_{n,p}$ and $d_{n,r}$ indicate evidence against H_0, we check if:

$$d_{n,p} > \frac{\lambda_\alpha}{\sqrt{n}} \tag{5.8}$$

or

$$d_{n,r} > \frac{\lambda_\alpha}{\sqrt{n}} \tag{5.9}$$

for Type I and Type II censored data, respectively. If the inequalities hold then H_0 should be rejected, i.e., $\tilde{F}_j^{\text{est}}$ cannot be regarded as the distribution at the stress level S_j.

5.3.3 Tests with more power for uncensored and censored data

Alternative EDF tests to be used for testing goodness of fit in the case of uncensored as well as censored data are the tests based on the Cramer-von Mises statistic and the Anderson-Darling statistic (formulae (B.8) and (B.9) for complete data and (B.15-B.18) for censored data). The advantage of these tests over the tests based on the Kolmogorov-Smirnov statistic is more power for a broad range of alternatives.

For complete data, value of the Cramer-von Mises statistic W_n^2 can be computed using (B.8a) and value of the Anderson-Darling statistic using (B.9a), where $F_0 = \tilde{F}_j^{\text{est}}$ is the hypothesized distribution estimated for the test level S_j as considered in section 5.2.

If (modified) values of the statistics exceed critical values (quantiles) obtained from the Table B.1 then H_0 should be rejected and $\tilde{F}_j^{\text{est}}$ should not be considered a distribution at S_j.

Generalizations of these statistics to be used for Type II and singly Type I censored samples are also available (see (B.15)-(B.18)).

Chapter 6 Results of Analysis of Sample ALT Studies

This Chapter presents results of reliability prediction obtained through analysis of sample accelerated life tests data using methodology developed in the thesis. The data sets analysed here have been either computer-generated on the basis of assumed life distribution and life-stress models or come from real life ALT studies considered in literature. In addition to analysis of complete ALT data, some examples of inference from time and failure censored data are also provided.

The purpose of numerical examples included in this chapter is to give a chance to compare results obtained using the methodology presented in the thesis with 'true' life characteristics (where available) or with results obtained with other data analysis approaches reported in literature.

It is also intended that numerical examples set the scene for discussion how well some stages of analysis perform for real life data. This may reveal points in the methodology proposed that require e.g., too much experimental data and thus call for some amendments through more investigation and research. Alternatively, experience with real life data may suggest modifications to accelerated life test plans that yield data sets most suitable for accurate prediction using the method presented.

Last but not least, the purpose of this chapter is to show that the ALT data analysis method discussed can be realized in practice using available computational facilities thus giving foundation to possible development of a flexible ALT data analysis tool for reliability analysts.

6.1 Tools, methods and materials

Real life experimental data analysed in this chapter comes from a well know paper by Nelson (1975). The data provided by Nelson have been extensively analysed by many authors using various approaches to ALT data analysis, see e.g., Nelson (1990).

Some numerical experiments involve analysis of simulated life data. Data was generated using a computer program based on a random number generator by Park and Miller, with the period of about 2.1×10^9 (Press et al 1992).

The numerical results presented in this chapter have been obtained using a computer program written by author implementing subsequent stages of ALT data analysis. The program comprises a set of numerical tools used for curve fitting, testing goodness of fit, regression analysis, checking adequacy of regression, maximum likelihood analysis.

Computer tools used for data and results visualization were well known Matlab and Mathematica software packages and the program ALT - accelerated life test data analysis and visualization tool, written by author.

Most of results presented here have been obtained on Sun 10 or Pentium PC computers.

6.2 Study 1: inverse power - exponential model data

This section presents a sample accelerated life test data analysis example that involves inference from simulated (computer generated) test data. First, models are introduced from which the test data was generated. Then results of reliability prediction are presented based on different models (both correct and incorrect ones). Results obtained are compared with the true life distribution at the nominal conditions.

Note that using *computer generated* input data implies that all underlying models as well as life distribution at the nominal conditions are known, thus allowing for verification of results by comparing them with the "true" life information.

6.2.1 ALT data

Test data analysed in this example is based on the following assumptions:

- at any stress level (including the nominal stress level) product lifetimes come from an exponential distribution, that is, a life p.d.f. at any stress level S_i is given by: $f_i(t) = \lambda_i \exp(-\lambda_i t)$;
- relationship between life and stress is given by the inverse power law: $\tau_i = \beta S_i^{\gamma}$, where τ_i is the mean life at stress S_i, and β, γ are model parameters;
- data at any stress level is complete (no censoring was imposed).

The following settings were used:

the nominal stress level: S_0=1.5 [$\times 10^4$ V];

test stress levels: S_1=2.4, S_2=2.6,..., S_9=4.0 [$\times 10^4$ V];

model parameters: $\beta = 10^8$, $\gamma = -12.82$. The stress is in [10^4 V=10 kV], then (mean) life is expressed in minutes, e.g., for $S_1 = 2.4\times 10^4$ V, mean life $\tau_1 = 10^8 \times 2.4^{-12.82}$ [minutes] $\approx$ 1336 [minutes].

Although physical interpretation of numbers processed throughout this example is irrelevant, it may be of interest why these particular values of model parameters and stress levels were selected. The motivation was to have an ALT study that resembles the Nelson data (Nelson 1990, p. 86) in terms of ranges of lifetimes and values of stress levels. So, if we postulate that mean life at the arbitrarily chosen stress level of 3.0 [$\times 10^4$ V] be 76 [min], and at the stress level of 3.4 [$\times 10^4$ V] be 15 [min], we find that the inverse power law model parameters are $\beta \approx 10^8$, $\gamma \approx -12.82$.

For each test stress level S_i 200 data points were generated from an exponential distribution with $\lambda_i = 1/\tau_i$, where τ_i is given by the inverse power law. The exponential data was generated using a random number generator of Park and Miller (Press 1992) returning a uniform random deviate between 0 and 1. Considering the well known property that a random variable $F(Y)$ has a uniform distribution on [0,1], where Y is a r.v. with a continuous cumulative distribution F, see e.g., Fisz (1965, p.158), we get immediately that:

$$F^{-1}(X) = -\frac{\ln(1-X)}{\lambda} \tag{6.1}$$

is exponentially distributed with the mean $1/\lambda$, where X is uniformly distributed on $[0,1]$ and $F(t) = 1 - \exp(-\lambda t)$ is for $t \geq 0$ and $\lambda > 0$ the exponential cumulative distribution function.

The following table summarizes population means for subsequent stress levels and some characteristics of the generated samples: sample mean, variance coefficient (defined as $\sqrt{\text{variance}}/\text{mean}$) and smallest and largest sample values (for sample size = 200 and random number generator seed = 123456).

TABLE 6.1

POPULATION MEANS AND CHARACTERISTICS OF GENERATED DATA FOR SUBSEQUENT STRESS LEVELS

stress level	population mean	generated sample mean	variance coefficient	sample min value	sample max value
2.4	1335.7	1508.2	0.89	1.41	7728.2
2.6	478.7	502.3	0.97	3.08	2601.0
2.8	185.1	192.0	0.97	0.70	1163.6
3.0	76.44	83.55	1.00	0.93	372.97
3.2	33.42	35.20	0.89	0.04	168.97
3.4	15.36	16.96	1.00	0.0001	88.43
3.6	7.38	6.66	0.89	0.003	26.67
3.8	3.69	3.67	1.04	0.003	17.71
4.0	1.91	1.95	0.98	2.6e-7	11.22
1.5 (nominal)	552726	-	-	-	-

6.2.2 Results of reliability prediction

6.2.2.1 Results based on correct models

First, results of reliability prediction based on correct models are presented, i.e., the inverse power law is used as an ageing model for the data analysis.

Obviously, in many real life ALT studies no prior information is available on what models are the 'true' ones - it is the data analysis process itself that reveals adequacy or inadequacy of adopted models. The purpose of this examples is to demonstrate how the data analysis method performs when correct models are applied.

Following figures (Fig. 6.1a - 6.1i) show continuous distribution functions of the family $F_{\alpha,\lambda}$ (2.1) fitted to test data at subsequent stress levels, according to (2.7). Step functions represent empirical distributions, as defined by (2.8), and the dashed lines represent the $F_{\alpha,\lambda}$ distributions. The table 6.2 provides α, λ, m_1 and m_2 parameters defining the $F_{\alpha,\lambda}$ functions at subsequent stress levels.

TABLE 6.2

PARAMETERS OF $F_{\alpha,\lambda}$ FUNCTIONS FITTED TO DATA AT SUBSEQUENT STRESS LEVELS (SEE ALSO FIGS. 6.1A - 6.1I)

stress level	α	λ	m_1	m_2
2.4	0.299	0.0013846	1508.2	4.08e+6
2.6	-0.112	0.002873	502.3	490694
2.8	0.188	0.007492	192.0	71557
3.0	0.497	0.02341	83.6	13991
3.2	-0.300	0.0411	35.2	2218.8
3.4	-0.189	0.0709	17.0	574.3
3.6	0.559	0.3195	6.67	79.6
3.8	0.073	0.277	3.67	28.2
4.0	-0.347	0.6384	1.95	7.43

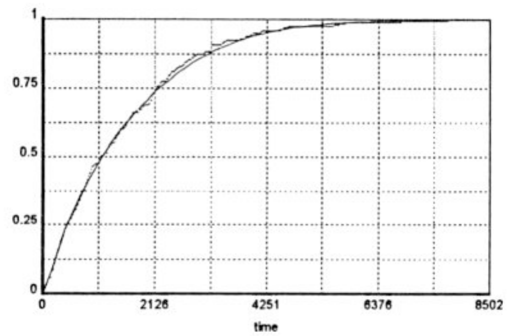

Figure 6.1a Life distribution at the stress level 2.4

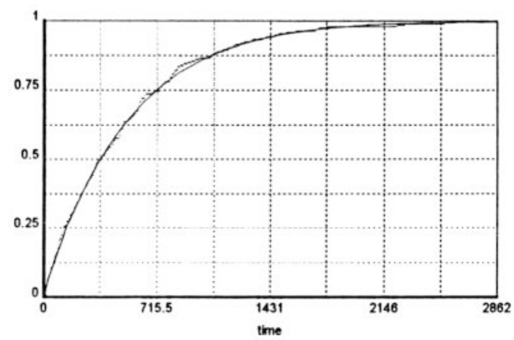

Figure 6.1b Life distribution at the stress level 2.6

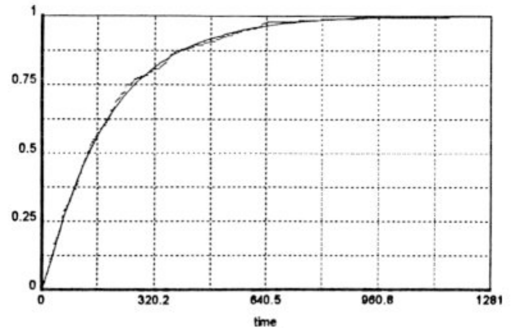

Figure 6.1c Life distribution at the stress level 2.8

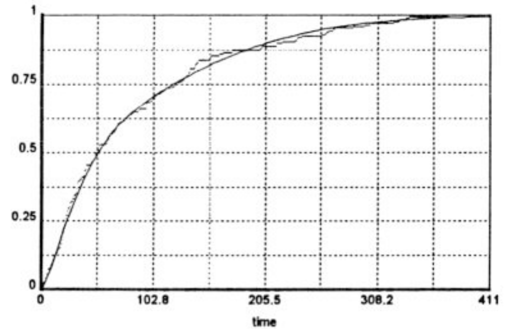

Figure 6.1d Life distribution at the stress level 3.0

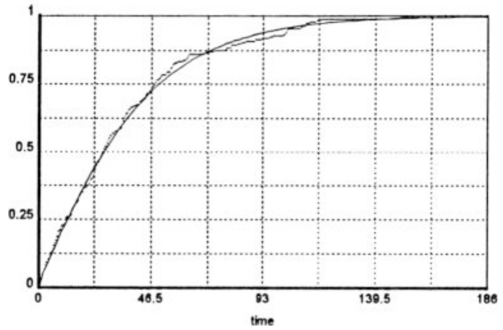

Figure 6.1e Life distribution at the stress level 3.2

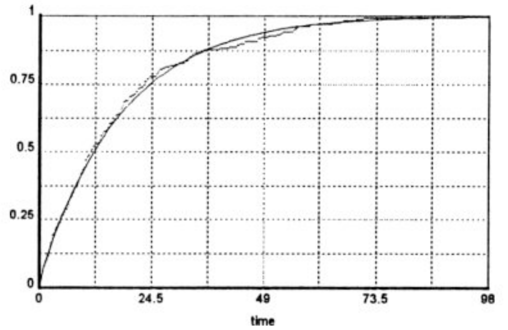

Figure 6.1f Life distribution at the stress level 3.4

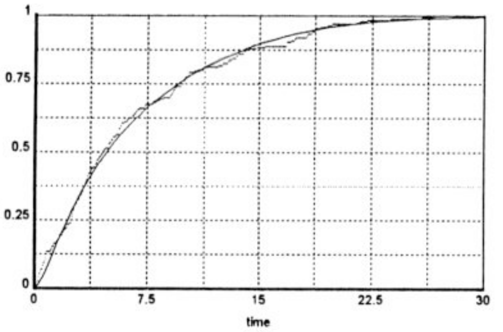

Figure 6.1g Life distribution at the stress level 3.6

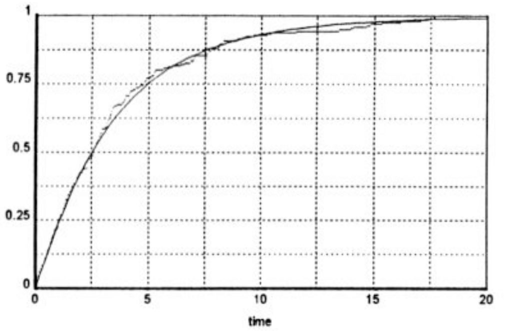

Figure 6.1h Life distribution at the stress level 3.8

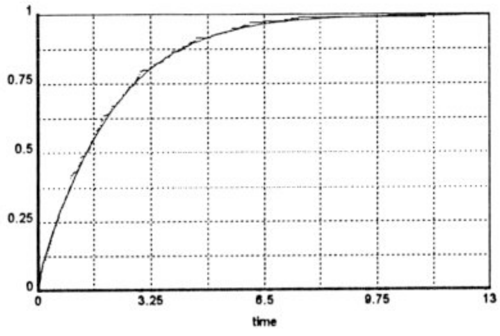

Figure 6.1i Life distribution at the stress level 4.0

The smooth distribution functions depicted in Figures 6.1a-6.1i are claimed to be population distribution functions at the corresponding stress levels. To verify the claim, statistical goodness of fit tests are performed. The Kolmogorov, Cramer-von Mises and Anderson-Darling tests give no reason to reject the H_0 hypotheses that the $F_{\alpha,\lambda}$ functions are population distribution functions at test levels 2.4 up to 4.0. We observe that computed values of all the three statistics are smaller than relevant critical values corresponding to the test significance level α=0.15 (and obviously, since smaller values of α yield bigger critical values, this guarantees that the statistics do not exceed critical values for all α<0.15). Computed values of the Kolmogorov, Cramer-von Mises and Anderson-Darling statistics are provided in the Table 6.3; for critical values of the statistics refer to Table B.1.

TABLE 6.3

GOODNESS OF FIT TESTING FOR THE TEST STRESS LEVELS (SEE ALSO FIGS. 6.1A - 6.1I)

stress level	Kolmogorov statistic d_n (B.5)	modified Kolmogorov statistic d_n^* (Table B.1)	modified Cramer-von Mises statistic W_n^* (Table B.1)	Anderson-Darling statistic A_n^2 (B.9)	decision
2.4	0.028	0.405	0.029	0.254	do not reject H_0
2.6	0.036	0.509	0.027	0.337	do not reject H_0
2.8	0.035	0.496	0.025	0.184	do not reject H_0
3.0	0.039	0.562	0.050	0.427	do not reject H_0
3.2	0.033	0.466	0.041	0.320	do not reject H_0
3.4	0.031	0.441	0.037	0.314	do not reject H_0
3.6	0.056	0.798	0.078	0.969	do not reject H_0
3.8	0.049	0.698	0.047	0.429	do not reject H_0
4.0	0.034	0.484	0.034	0.329	do not reject H_0

Once the cumulative distribution functions at test stress levels are established, extrapolation can be made to arrive at the life distribution at the nominal conditions, as described in sections 4.3-4.5. For the purpose of this example, inverse power law is used as the ageing model (which by definition is the true assumption). Results of the analysis are shown in Figure 6.2. The Figure shows the step function which is built from estimated quantiles (see (4.36)), and the continuous function that is claimed to be the product life distribution under nominal conditions (which is a $F_{\alpha,\lambda}$ function defined by the following parameters: α=-0.9, λ=2.61859×10^{-6}, m_1=652319, m_2=7.1109×10^{11}).

Following characteristics are also available for nominal stress of 1.5:

mean life	652319
variance	7.1109×10^{11}
variance coefficient	1.29

We observe that the mean life obtained from the ALT data analysis is close to the true population mean 552726 (Table 6.1).

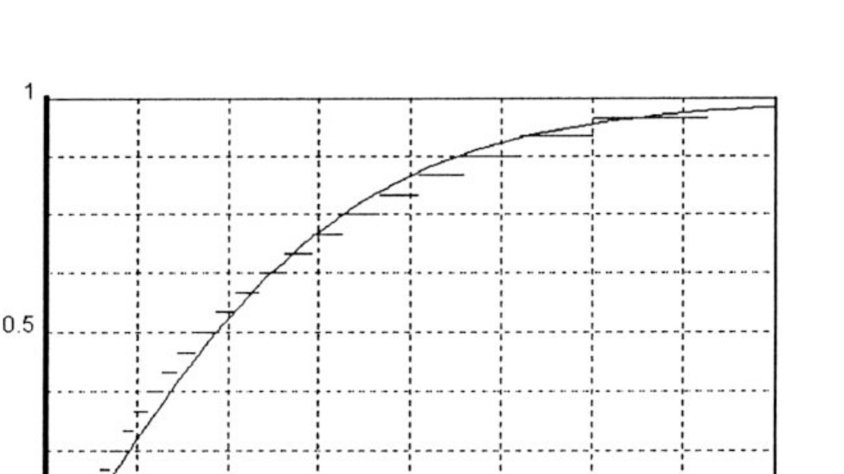

Figure 6.2 Life distribution at the nominal conditions (stress level = 1.5); inverse power law used

We can also verify the estimated distribution function against empirical data we have available for the stress of 1.5.

Results of goodness of fit tests are given in Table 6.4. The null hypothesis H_0 says that the continuous distribution in Fig. 6.2 is the population life distribution under a stress level of 1.5. Since computed values of the test statistics do not exceed critical values corresponding to quantile level 0.85 (Table B.1), the tests give no reason to reject the H_0 hypothesis (under the test significance level of 0.15 or smaller).

TABLE 6.4

GOODNESS OF FIT TESTING FOR THE NOMINAL STRESS LEVEL 1.5. THE C.D.F. ESTIMATED FROM TEST DATA AT STRESS LEVELS 2.4-4.0 IS COMPARED AGAINST DATA GENERATED FROM THE EXPONENTIAL DISTRIBUTION WITH MEAN OF 552726.

nominal stress level	Kolmogorov statistic d_n (B.5)	modified Kolmogorov statistic d_n^* (Table B.1)	modified Cramer-von Mises statistic W_n^* (Table B.1)	Anderson-Darling statistic A_n^2 (B.9)	decision
1.5	1.074	1.096	0.262	1.412	do not reject H_0

This numerical example illustrates the major steps of the inference from ALT data described in previous chapters. It also shows that the complex procedure performs well in the case when the data analyst has *a priori* knowledge about the *correct ageing model* to be used for data analysis.

6.2.2.2 Results obtained when correct models not known

In this numerical example ALT data presented in section 6.2.1 will be analyzed in order to obtain life distribution of product operating at nominal conditions. However, this time no assumption will be made in advance on what ageing model should be used for data analysis. Instead, the most suitable ageing model will be selected on the basis of results verification procedure presented in Chapter 5. This will illustrate the method of automatic (i.e., based solely on data) selection of an ageing model.

In this example we will consider the following three ageing models: power (and inverse power), exponential and Arrhenius[1].

Obviously, any other models might be considered as well, providing

a) the model is linearizable, so that it is possible to use linear regression apparatus presented in section 4.4.2, or
b) an efficient tool is available to fit to data a nonlinearizable model (that is to solve the optimization problem (4.6)).

As before, the analysis starts with estimation of life c.d.f.s for test stress levels (Figures 6.1a-6.1i) and verification of estimated functions by the means of goodness of fit testing (Table 6.3). Then quantiles are computed from the c.d.f.s obtained and linearized versions of the models concerned (as given by (4.15), (4.17) and (4.17''))

[1] by which I mean here merely the exponential model with inverse of stress taken. No other similarities with the well known physics-based Arrhenius model should be sought.

are used in regression analysis to obtain extrapolation of quantiles under nominal conditions. These quantiles are then used to estimate the nominal conditions life distribution. For each of the ageing models concerned results verification procedure (section 5.2) is carried out. The model that performs best in the results verification procedure is considered the most suitable ageing model.

Results of analysis follow below.

Exponential ageing model

Application of the exponential ageing model gives the following estimates of life parameters at the nominal stress level of 1.5:

mean life	47262
variance	3.89163×10^{9}
variance coefficient	1.32

Figure 6.3 depicts the life cumulative distribution function obtained (the $F_{\alpha,\lambda}$ function is defined by the following parameters: α=-0.7, λ=3.30667×10^{-5}, m_1=47262, m_2=3.89163×10^{9}).

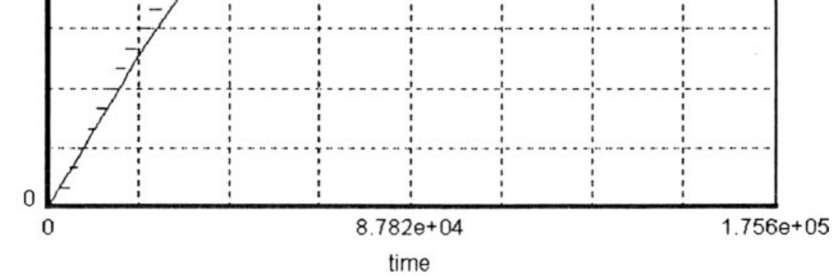

Figure 6.3 Life distribution at the nominal conditions (stress level = 1.5); exponential ageing model used.

Arrhenius ageing model

Following results are obtained under the Arrhenius ageing model assumption:

mean life	36931900
variance	2.10669×10^{15}
variance coefficient	1.24

The life c.d.f. at nominal conditions is shown in Figure 6.4 (the $F_{\alpha,\lambda}$ function parameters: are: α=1.9, λ=1.11522×10^{-7}, m_1=3.69319×10^{7}, m_2=2.10669×10^{15}).

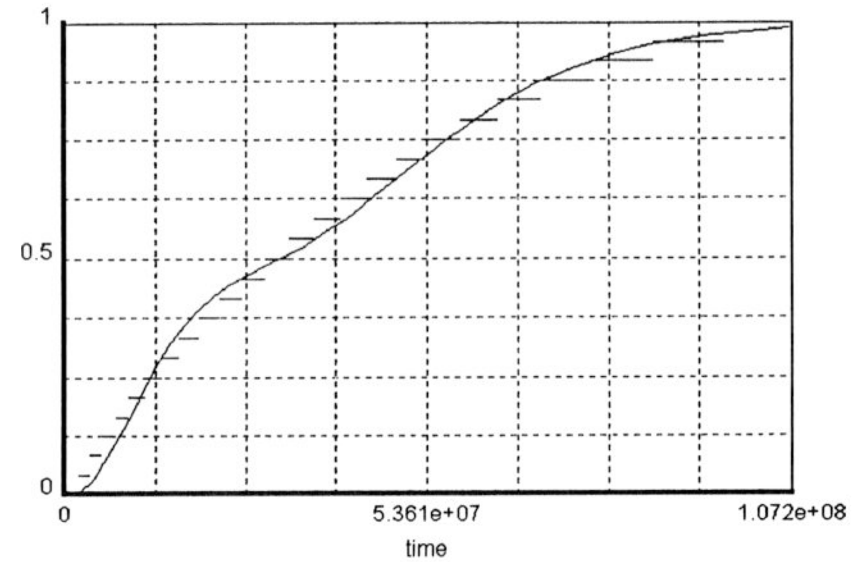

Figure 6.4 Life distribution at the nominal conditions (stress level = 1.5); Arrhenius ageing model used.

Power ageing model

The life c.d.f. at stress level of 1.5 is depicted in Figure 6.2 and the mean and variance are shown above the table 6.2. The estimated mean life of about 652319 is a range of magnitude bigger than the estimate obtained under the exponential model and about two ranges of magnitude smaller than the estimate under the Arrhenius model.

Data driven selection of the most appropriate ageing model

As we see, considerably different results are obtained under different ageing model assumptions. In order to determine which of the ageing models considered is the most appropriate, we run results verification procedure described in Chapter 5.

We predict product life distributions for test stress levels of 2.4, 2.6,...,4.0 based on data obtained for stress levels other than the stress for which life is being predicted. In other words, we repeat the ALT data analysis procedure in turn considering as nominal the stress levels 2.4, 2.6, etc. Results of reliability prediction obtained for a specified stress level are then compared with experimental data available for the stress level. In this way adequacy of models assumed is verified.

Results verification for the power ageing model

The following table contains mean lives obtained for subsequent stress levels (treated as nominal in ALT data analysis), compared with experimental data.

TABLE 6.5

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS FOR POWER AGEING MODEL. COMPARISON OF MEANS.

stress level	experimental data (sample) mean	mean life obtained through ALT data analysis
2.4	1508.2	1382.7
2.6	502.3	500.2
2.8	192.0	189.2
3.0	83.55	75.9
3.2	35.20	32.69
3.4	16.96	14.71
3.6	6.66	7.28
3.8	3.67	3.50
4.0	1.95	1.74

Life cumulative distribution functions obtained for the (nominal) stress levels of 2.4 up to 4.0 are compared against empirical distributions. The comparison is performed by the means of statistical test (Kolmogorov, Cramer-von Mises and Anderson-Darling tests are used). For details refer to Chapter 5.

Results of statistical goodness of fit tests follow in Table 6.6. Entries in the "decision" column of this Table and of the following Tables 6.8 and 6.10 should be interpreted as follows:

"no" - test (under significance level of 0.15 or less) gives no reason to reject the H_0 hypothesis (that the c.d.f. obtained from ALT data analysis is the population distribution at a corresponding stress level; refer also to Eqn. 5.1);

number (e.g., 0.1) means that the H_0 hypothesis should be rejected with the test significance level given by the number concerned. However, tests with smaller significance levels (out of those in the set: 0.15, 0.1, 0.05, 0.025, 0.01) give no reason to reject H_0. For instance, an entry 0.1 implies that H_0 should be rejected only when tested with significance level 0.1, but tests with significance levels of 0.01, 0.025 and 0.05 give no reason to reject the hypothesis.

TABLE 6.6

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS FOR POWER AGEING MODEL. CUMULATIVE DISTRIBUTIONS COMPARED WITH EMPIRICAL DISTRIBUTIONS.

stress	Kolmogorov test		Cramer-von Mises test		Anderson-Darling test	
	d_n^* (Tab. B.1)	decision	W_n^* (Tab. B.1)	decision	A_n^2 (B.9)	decision
2.4	0.905	no	0.147	no	0.952	no
2.6	1.258	0.1	0.261	no	2.609	0.05
2.8	0.824	no	0.141	no	1.293	no
3.0	1.112	no	0.341	0.15	3.406	0.025
3.2	1.067	no	0.205	no	1.803	0.15
3.4	1.089	no	0.232	no	4.087	0.01
3.6	1.478	0.05	0.564	0.05	2.794	0.05
3.8	1.003	no	0.210	no	2.716	0.05
4.0	0.866	no	0.198	no	1.434	no

(Legend to the "decision" column is explained in the paragraph just before the Table):

Comparison of distribution functions with empirical data reveals that in most cases results of ALT data analysis conform well to empirical data. For instance, for test significance level of 0.01, the statistical tests show very good consistency between ALT results and empirical data for all stress levels.

Results verification for exponential ageing model

Table 6.7 presents mean lives obtained from ALT data analysis when exponential ageing model is applied, compared with experimental data.

Table 6.8 provides results of goodness of fit tests comparing life cumulative distribution functions obtained for the stress levels of 2.4 up to 4.0 with empirical distributions.

TABLE 6.7

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS FOR EXPONENTIAL AGEING MODEL. COMPARISON OF MEANS.

stress level	experimental data (sample) mean	mean life obtained through ALT data analysis
2.4	1508.2	906.6
2.6	502.3	479.0
2.8	192.0	213.7
3.0	83.55	91.71
3.2	35.20	39.84
3.4	16.96	17.08
3.6	6.66	7.68
3.8	3.67	3.13
4.0	1.95	1.18

The table 6.8 summarizes results of comparison of empirical distributions with the c.d.f.s obtained from ALT data analysis.

TABLE 6.8

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS FOR EXPONENTIAL AGEING MODEL. CUMULATIVE DISTRIBUTIONS COMPARED WITH EMPIRICAL DISTRIBUTIONS. FOR LEGEND SEE THE PARAGRAPH BEFORE THE TABLE 6.6.

stress	Kolmogorov test		Cramer-von Mises test		Anderson-Darling test	
	d_n^* (Tab. B.1)	decision	W_n^* (Tab. B.1)	decision	A_n^2 (B.9)	decision
2.4	3.170	0.01	4.345	0.01	30.652	0.01
2.6	1.119	no	0.189	no	2.016	0.1
2.8	1.466	0.05	0.565	0.05	3.316	0.025
3.0	1.938	0.01	1.082	0.01	5.670	0.01
3.2	1.363	0.05	0.449	0.1	3.284	0.025
3.4	1.242	0.1	0.360	0.1	3.791	0.025
3.6	1.852	0.01	1.002	0.01	4.926	0.01
3.8	1.093	no	0.152	no	3.631	0.025
4.0	2.723	0.01	2.656	0.01	19.123	0.01

Results verification for Arrhenius ageing model

Table 6.9 provides comparison of mean lives obtained from test with those obtained from ALT data analysis under the Arrhenius ageing model assumption.

Table 6.10 provides results of goodness of fit tests comparing life c.d.f.s obtained for the stress levels of 2.4 -4.0 with empirical distributions.

TABLE 6.9

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS FOR ARRHENIUS AGEING MODEL. COMPARISON OF MEANS

stress level	experimental data (sample) mean	mean life obtained through ALT data analysis
2.4	1508.2	2190.6
2.6	502.3	496.4
2.8	192.0	162.0
3.0	83.55	62.20
3.2	35.20	27.07
3.4	16.96	13.07
3.6	6.66	7.23
3.8	3.67	4.08
4.0	1.95	2.51

In the Table 6.10, empirical distributions are compared with the corresponding c.d.f.s obtained from ALT data analysis, as described in Chapter 5.

TABLE 6.10

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS FOR ARRHENIUS AGEING MODEL. CUMULATIVE DISTRIBUTIONS COMPARED WITH EMPIRICAL DISTRIBUTIONS. FOR LEGEND SEE THE PARAGRAPH BEFORE THE TABLE 6.6.

stress	Kolmogorov test		Cramer-von Mises test		Anderson-Darling test	
	d_n^* (Tab. B.1)	decision	W_n^* (Tab. B.1)	decision	A_n^2 (B.9)	decision
2.4	2.580	0.01	3.099	0.01	16.330	0.01
2.6	1.289	0.1	0.306	0.015	2.813	0.05
2.8	1.161	0.15	0.234	no	2.863	0.05
3.0	1.719	0.01	0.632	0.025	5.197	0.01

3.2	1.668	0.01	0.940	0.01	5.400	0.01
3.4	1.345	0.1	0.492	0.05	4.514	0.01
3.6	1.289	0.1	0.404	0.1	2.060	0.1
3.8	2.072	0.01	1.311	0.01	6.539	0.01
4.0	2.071	0.01	1.813	0.01	11.773	0.01

Large values of the test statistics, in many cases exceeding the 0.99 quantile of the distributions involved, show poor consistency of the empirical distributions and the c.d.f.s obtained under the Arrhenius model assumption.

6.2.3 Discussion

The purpose of calculations summarized in Tables 6.5 through 6.10 is to determine, by the means of results verification procedure (Chapter 5), which of the predefined ageing models is the most suitable for the study concerned.

Results verification procedure clearly gives reasons to reject the exponential and Arrhenius models, as values of the test statistics obtained for these two models are larger for most stress levels than values obtained under the power model. For majority of stress levels, the c.d.f.s obtained under the exponential or Arrhenius models do not fit well to experimental data, which suggests that the models should not be used for extrapolation of ALT results to the nominal conditions. On the other hand, good conformity of c.d.f.s obtained under the power ageing model with experimental data justifies applying the model for extrapolation of ALT data.

It is also interesting that by comparing the mean lives only (empirical means against means obtained under a particular ageing model assumption), we find some evidence in favour of the power model versus the two other models. If we observe that relative discrepancy between means at the stress of 2.4 (which is the nearest stress level to the nominal stress and which involves real extrapolation in regression analysis, just like quantiles under the nominal stress level do), we find that extrapolation of results from test levels of 2.6 through 4.0 to the "nominal" level of 2.4 performs best for the power model. (Note that relative discrepancy is about 8% ($\approx|1383-1508|/1508$) in case of the power model, but as much as 40% ($\approx|907-1508|/1508$) in case of the exponential model and about 45% ($\approx|2191-1508|/1508$) in case of the Arrhenius model).

For the purpose of illustration, in Figures 6.5, 6.6 and 6.7 empirical distribution at test stress level of 2.4 are compared graphically with the c.d.f. obtained under the three ageing models concerned (power, exponential and Arrhenius, respectively). Note that the function have been compared analytically (by the means of goodness of fit tests) in the first rows of Tables 6.6, 6.8 and 6.10.

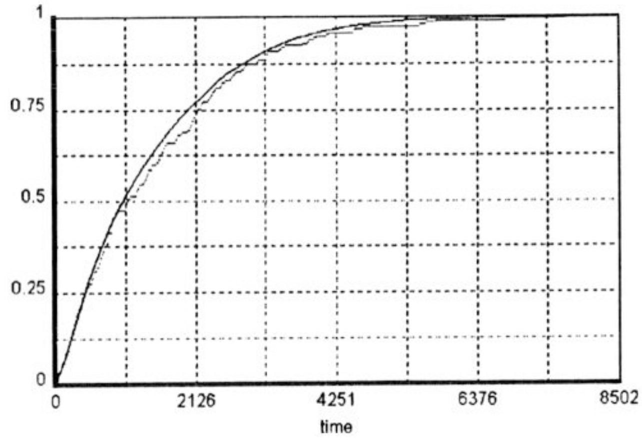

Figure 6.5 Empirical distribution at test stress level of 2.4 compared with the c.d.f. obtained under the *power* ageing model (c.d.f. parameters: α=0.599, λ=1.662×10^{-3}, m_1=1382.7, m_2=3.3374×10^{6})

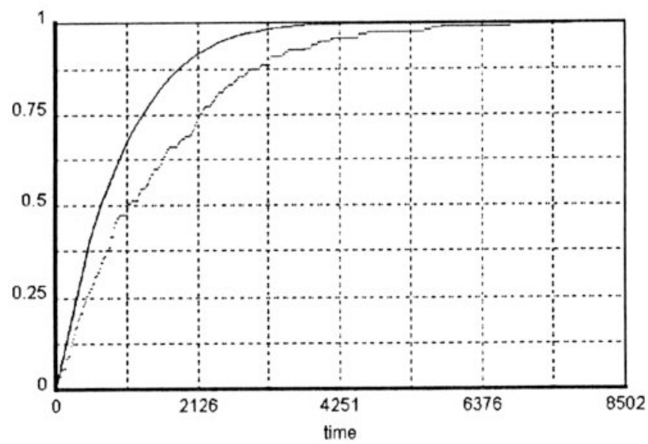

Figure 6.6 Empirical distribution at test stress level of 2.4 compared with the c.d.f. obtained under the *exponential* ageing model (c.d.f. parameters: α=0.499, λ=2.368×10^{-3}, m_1=906.6, m_2=1.4325×10^{6})

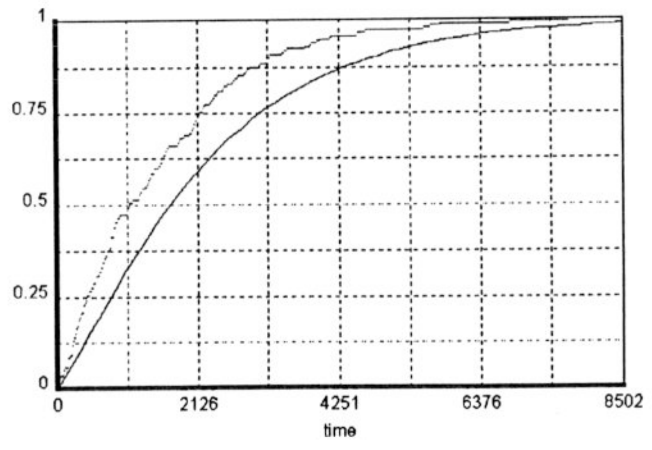

Figure 6.7 Empirical distribution at test stress level of 2.4 compared with the c.d.f. obtained under the *Arrhenius* ageing model (c.d.f. parameters: α=-0.7, λ=7.226$\times10^{-4}$, m_1=2190.6, m_2=8.3092$\times10^{6}$)

The final conclusion is that out of the three models concerned, the power model might be considered the most suitable ageing model. The ageing model selection procedure performed well since the power model is the true ageing model in this study.

6.3 Study 2: Nelson data

In this chapter, data from a real life accelerated life test is analyzed that comes from voltage-endurance tests of liquid electrical insulation. The data was originally published and analyzed by Nelson (1975), and it was later considered by many authors (Proschan, Singpurwalla 1980; Basu, Ebrahimi 1982; Shaked, Singpurwalla 1982; Baskin 1988; Nelson 1990, Biernat et al. 1992) who used different life distribution and ageing models and different approaches to ALT data analysis.

Although this is not justified by theory, the inverse power law is commonly used as the ageing model when electrical insulation is tested under increased voltages. In this study, the Nelson ALT data will be analysed with different ageing model assumptions. Data driven selection of the most suitable ageing model will be demonstrated and results will be compared with the ones obtained by other authors.

Table 6.11 cites the original Nelson ALT data: times to breakdown (in minutes) of insulating fluid specimens tested at voltages from 26 to 38 kV.

TABLE 6.11

TIMES (IN MINUTES) TO BREAKDOWN OF SPECIMENS OF INSULATING FLUID TESTED AT INCREASED VOLTAGES.

26 kV	28 kV	30 kV	32 kV	34 kV	36 kV	38 kV
5.6	68.8	7.70	0.30	0.20	0.30	0.10
1579.0	108.3	17.70	0.40	0.80	0.60	0.40
2324.0	110.3	20.50	0.70	1.00	1.00	0.50
	426.0	21.00	0.80	1.30	1.00	0.70
	1068.0	22.70	2.90	2.80	1.70	0.70
		43.40	3.90	3.20	2.00	1.10
		47.30	9.90	4.20	2.10	1.40
		139.10	13.90	4.70	2.60	2.30
		144.10	15.90	4.80	2.70	
		175.10	27.80	6.50	2.90	
		194.10	53.20	7.30	3.70	
			82.80	8.00	4.00	
			89.30	8.30	5.30	
			100.60	12.10	13.80	
			215.10	31.80	25.50	
				32.50		
				33.90		
				36.70		
				72.90		

6.3.1 Summary of results

The purpose of analysis of ALT data shown in Table 6.11 is to estimate product life distribution under the nominal conditions of 20 kV. In this sample study, data analysis was based on three ageing model assumptions: inverse power law, exponential and exponential with inverse of stress taken as the exponent. (Throughout this example, the last model will be referred to as the Arrhenius model, although there is obviously no correspondence of this model to the well known Arrhenius relationship between product life and temperature).

Application of different models yields considerably different results of reliability prediction at the nominal conditions, which is summarized in Table 6.12. Roughly, mean lives obtained under different ageing models differ by a range of magnitude.

TABLE 6.12

RESULTS OF RELIABILITY PREDICTION AT 20KV OBTAINED FROM NELSON ALT DATA USING DIFFERENT AGEING MODELS

	inverse power	exponential	Arrhenius
mean life [in mins.]	164473	31369	1.418809×10^6
variance of life	2.0879×10^{10}	7.7205×10^8	1.5008×10^{10}
variance coefficient	0.879	0.886	0.863

6.3.2 Selection of the suitable ageing model

Data driven selection of the most appropriate ageing model is based on comparison of results of prediction against empirical data. As before, ALT data analysis is performed in order to estimate life at the stress of 26 kV (based on test data at 28 through 38 kV). Estimated life at 26 kV is compared with test data at 26 kV. The procedure is repeated for other test stress levels, as well. Good conformity of results of prediction with the test data works in favour of a particular ageing model used.

In Table 6.13 mean lives obtained under the three ageing models are compared with the test data obtained at corresponding stress levels. Tables 6.14, 6.15 and 6.16 summarize results of comparison of life c.d.f.s obtained at test stress levels with empirical distributions.

TABLE 6.13

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS UNDER DIFFERENT AGEING MODELS. COMPARISON OF MEANS

stress [kV]	mean life from test	mean life (inverse power model used)	mean life (exponential model used)	mean life (Arrhenius model used)
26	1303	1316	873.3	2065
28	356.3	323.1	310.3	325.1
30	75.7	102.3	118.3	86.3
32	41.2	26.4	30.6	22.6
34	14.3	8.1	9.3	7.2
36	4.6	2.9	3.0	2.9
38	0.9	1.8	1.4	2.4

TABLE 6.14

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS UNDER INVERSE POWER MODEL. CUMULATIVE DISTRIBUTIONS COMPARED WITH EMPIRICAL DISTRIBUTIONS. (FOR LEGEND SEE PARAGRAPH BEFORE TABLE 6.6).

stress	Kolmogorov test		Cramer-von Mises test		Anderson-Darling test	
	d_n^* (Tab. B.1)	decision	W_n^* (Tab. B.1)	decision	A_n^2 (B.9)	decision
26	0.61	no	0 ($W_n \approx 0.07$)	no	0.47	no
28	0.80	no	0.03	no	0.56	no
30	1.06	no	0.22	no	1.28	no
32	1.24	0.1	0.41	0.1	4.24	0.1
34	1.13	no	0.15	no	2.78	0.05
36	0.50	no	0.02	no	0.70	no
38	0.80	no	0.13	no	0.79	no

TABLE 6.15

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS UNDER EXPONENTIAL AGEING MODEL. CUMULATIVE DISTRIBUTIONS COMPARED WITH EMPIRICAL DISTRIBUTIONS. (FOR LEGEND SEE PARAGRAPH BEFORE TABLE 6.6).

stress	Kolmogorov test		Cramer-von Mises test		Anderson-Darling test	
	d_n^* (Tab. B.1)	decision	W_n^* (Tab. B.1)	decision	A_n^2 (B.9)	decision
26	0.90	no	0.12	no	1.65	0.15
28	0.77	no	0.03	no	0.59	no
30	1.15	0.15	0.27	no	1.50	no
32	1.31	0.1	0.48	0.05	4.77	0.01
34	1.06	no	0.23	no	2.26	0.1
36	0.51	no	0.02	no	0.74	no
38	0.56	no	0.03	no	0.39	no

TABLE 6.16

VERIFICATION OF RESULTS OF ALT DATA ANALYSIS UNDER ARRHENIUS AGEING MODEL. CUMULATIVE DISTRIBUTIONS COMPARED WITH EMPIRICAL DISTRIBUTIONS. (FOR LEGEND SEE PARAGRAPH BEFORE TABLE 6.6).

stress	Kolmogorov test		Cramer-von Mises test		Anderson-Darling test	
	d_n^* (Tab. B.1)	decision	W_n^* (Tab. B.1)	decision	A_n^2 (B.9)	decision
26	0.71	no	0.02	no	0.58	no
28	0.82	no	0.04	no	0.56	no
30	0.92	no	0.14	no	0.86	no
32	1.15	0.15	0.35	0.1	3.95	0.01
34	1.12	no	0.12	no	2.32	0.1
36	0.51	no	0.02	no	0.65	no
38	1.05	no	0.30	0.15	1.47	no

A few observations and comments can be made when analyzing mean lives in Table 6.13 and results of goodness of fit tests (Tables 6.14-6.16):

The best estimate of the mean life at the stress level of 26 kV is accomplished when the inverse power model is used.

Relative discrepancies (distance) between mean life at 26 kV obtained through extrapolation and the true value of mean life (which is assumed to be the value obtained form test) are:

power model used: relative discrepancy (r.d.) is about 1% ($\approx$|1316-1303|/1303);

exponential model used: r.d. is about 33% ($\approx$|873-1303|/1303);

Arrhenius model used: r.d. is about 58% ($\approx$|2065-1303|/1303);

In this example, performance of regression at 26 kV will be used to draw conclusions about suitability of a particular ageing model. This seems justified by the fact that estimation of quantiles at the nominal conditions of 20 kV involves the same type of inference. Quantiles at 20 and 26 kV are obtained in a similar way by extrapolation in regression (extrapolation is based on results obtained at stress levels of 26, 28 up to 38, and 28, 30, up to 38 kV, respectively). It can be argued that since extrapolation from stress levels 28, 30, up to 38 kV performs quite well under inverse power model, and rather poorly under exponential and Arrhenius models, the inverse power model is more appropriate then the two other models for extrapolation of results from the stress levels 26, 28 up to 36 kV to the nominal conditions of 20 kV.

Although goodness of fit tests performed with the significance level α of 0.15 or less for the three ageing models do not give reasons to reject the H_0 hypothesis at 26 kV (refer to Eqn. 5.1), and in consequence to reject a particular ageing model, it can be observed that values of test statistics are slightly smaller for the inverse power model than for the other models (see the first rows of Tables 6.14, 6.15 and 6.16). This seems to support the conclusion just drawn, based on comparison of means, that extrapolation performed at 26 kV under the inverse power model produces results that give best fit to experimental data.

It can be observed that goodness of fit tests, as summarized in Tables 6.14-6.16, do not lead to as clear conclusions as in Section 6.2 as far as adequacy of ageing models is concerned. This can be due to very small amount of experimental data available at virtually all test levels (e.g., only three data points at the test stress level of 26 kV).

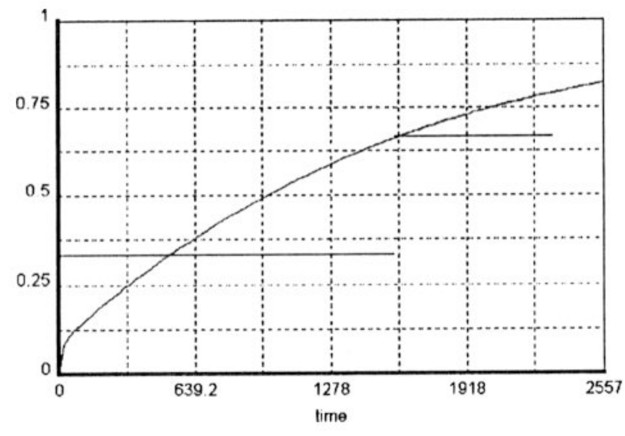

Figure 6.8 Empirical distribution at test stress level of 26 kV compared with the c.d.f. obtained under the inverse power model (c.d.f. parameters: α=-0.75, λ=9.722×10^{-4}, m_1=1315.5, m_2=3.5391×10^{6})

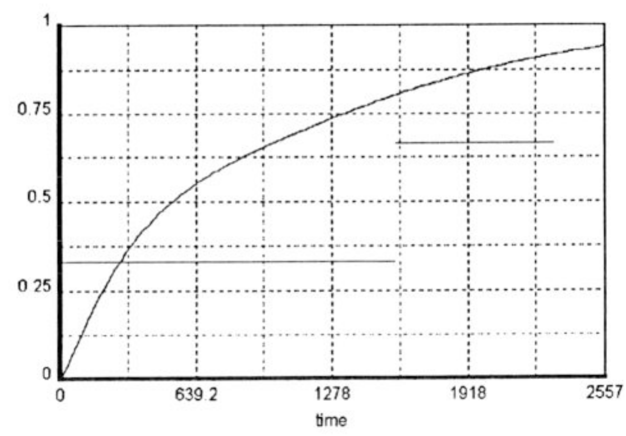

Figure 6.9 Empirical distribution at test stress level of 26 kV compared with the c.d.f. obtained under exponential ageing model (c.d.f. parameters: α=0.2486, λ=2.194×10^{-3}, m_1=873.3, m_2=1.5559×10^{6})

The problem is illustrated by Figures 6.8 and 6.9 depicting c.d.f.s obtained at 26 kV under the inverse power (Fig. 6.8) and exponential (Fig.6.9) model (based on test data from test levels 28 through 38 kV) compared with empirical distribution. Although these two c.d.f.s differ considerably in shape and in expected mean life, goodness of fit tests (for α=0.15) give no reason to reject either of them. (As mentioned before, goodness of fit testing shows a slight advantage (smaller values of test statistics) of the power model over the exponential and Arrhenius models).

The final conclusion from the data driven selection of the best ageing model is that the inverse power model seems more suitable than two other competing models. Results of reliability prediction obtained under the inverse power model are:

expected mean life	164473 minutes
lifetime variance	2.0879×10^{10}

and the lifetime cumulative distribution function is shown in Figure 6.10.

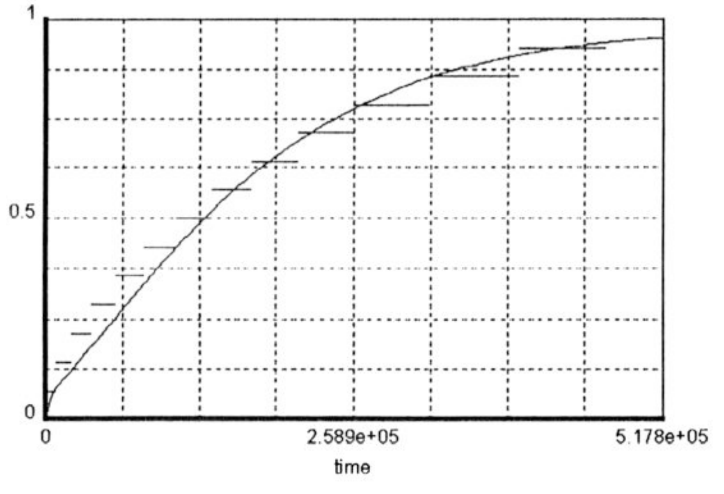

Figure 6.10 Life distribution at the nominal stress level of 20 kV obtained under the inverse power ageing model (c.d.f. parameters: α=-0.75, $\lambda=1.01\times10^{-5}$, m_1=164473, $m_2=4.793\times10^{10}$

It would be interesting to compare results of reliability prediction obtained using the approach presented in the thesis with some results obtained by other authors, e.g.,

author(s)	estimated mean life (in minutes)
Nelson (1990)	31067
Baskin (1988)	194342
Shaked, Singpurwalla (1982)	106000
Basu, Ebrahimi (1982)	13686

Results presented by Nelson (1990) are based on the assumption that life time at each stress level follows Weibull distribution with the c.d.f. $F(t) = 1 - \exp\left(-(t/\alpha)^{\beta}\right)$ and that the shape parameter β of the distribution remains constant throughout all stress levels. Nelson applies the inverse power ageing model for results extrapolation. His models give following estimates of mean lives (see Nelson 1990, p.195) for the test stress levels:

voltage [kV]	estimate of mean life [min.]	mean life observed at test [min]
26	421.4	1302.9
28	125.1	356.3
30	40.4	75.7

which suggests that his assumptions are not quite adequate for the data and tend to underestimate expected lifetimes.

Mean life reported by Baskin is based on the assumption that the life stress relationship is the inverse power law and lifetime distributions belong to the generalized distribution family $F_{\alpha,\lambda}$.

Shaked and Singpurwalla's (1982) approach is based on the inverse power law assumption, but they do not assume that the lifetimes follow a specified distribution. Their only assumption regarding lifetime distribution is that for all stress levels the distributions belong to some (parametric) family of distributions such that the distribution at the stress level V_i is $F_i(t) = F\left(RV_i^{\gamma}t\right)$, where F is a specified parametric distribution and R and γ are unknown constants. This assumption states that shape of the distributions does not change with stress. Results presented by the authors are based on data obtained at test stress levels 30, 32, 34 and 36 kV.

Basu, Ebrahimi (1982) use data from the same test stress levels as Shaked and Singpurwalla, but they treat the data as time-censored at T=120, 90, 30 and 10 minutes (censoring times for stress 30, 32, 34, 36 kV, respectively). This accounts for short mean life obtained.

6.4 Study 3: inference from censored data

In this section, inference from time and failure censored data is presented. The ALT data analyzed here come from Table 6.11, with the following censoring imposed:

stress [kV]	28	30	32	26, 34, 36, 38
censoring	failure; r=4	failure; r=9	time; T=100	no censoring

where T denotes the test duration (i.e., the common censoring time, refer to section 2.4.2.1), and r is the number of items that have to fail for the test to stop (refer to section 2.4.2.2).

Figures 6.11, 6.12 and 6.13 depict the c.d.f.s fitted to the censored data at stress levels 28, 30 and kV, respectively. The c.d.f.s have been obtained by direct maximization of the likelihood functions (2.18) or (2.19). For comparison, the figures also show the c.d.f.s fitted to complete data at the corresponding stress levels (using least squares minimization (2.14)); these c.d.f. are plotted with dashed line.

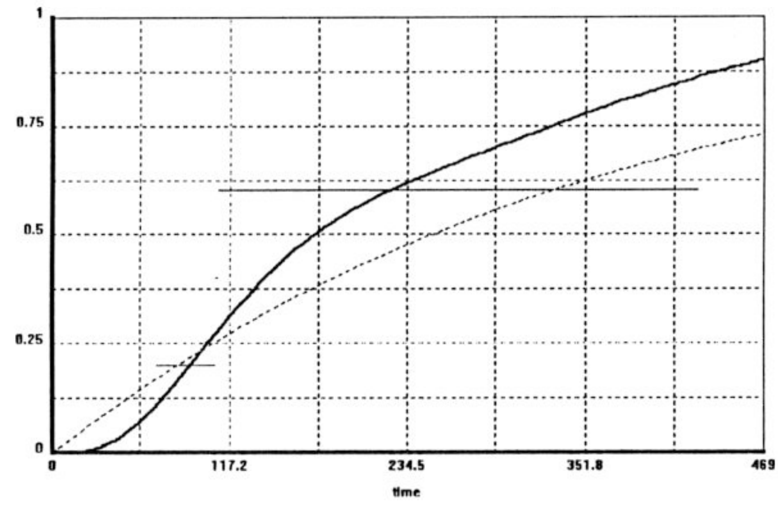

Figure 6.11 The c.d.f. at 28kV fitted to censored data (thick line), with parameters: α=2.69, λ=1.84×10^{-2}, m_1=228, m_2=78317. The dashed lined represents the c.d.f. fitted to complete data.

The c.d.f. estimated for the nominal conditions of 20 kV based on censored data at 28, 30, 32 kV (and complete data at other stress levels) is shown in Figure 6.14.

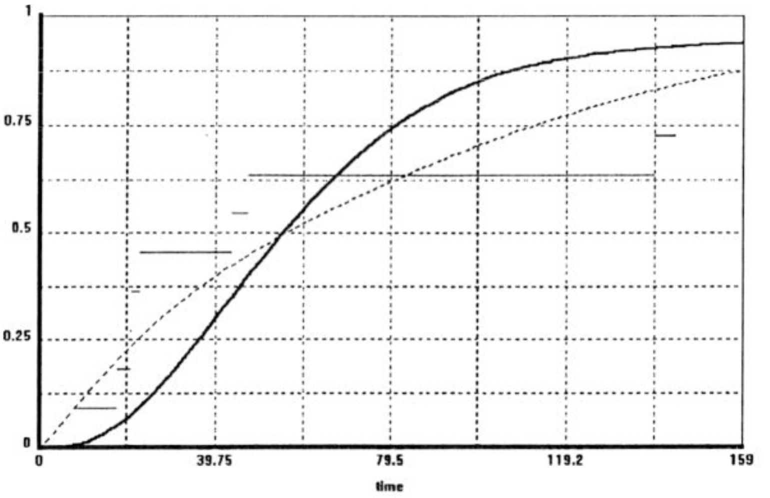

Figure 6.12 The c.d.f. at 30kV fitted to censored data (thick line), with parameters: α=2.08, λ=3.36×10^{-2}, m_1=69.0, m_2=7957. The dashed line represents the c.d.f. fitted to complete data.

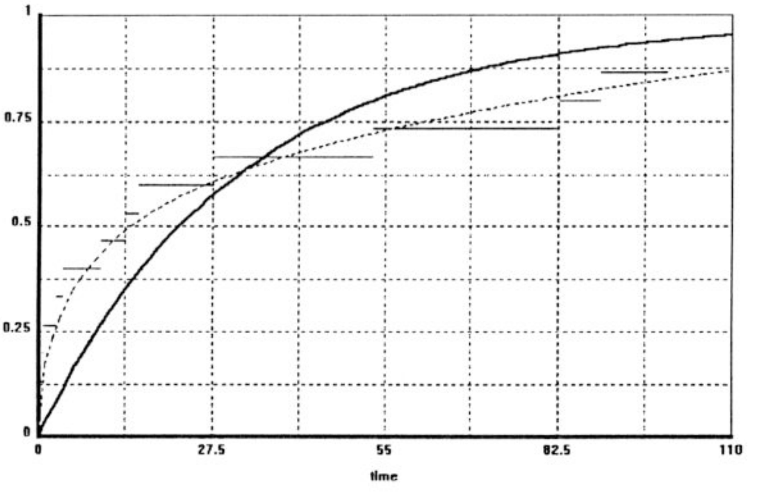

Figure 6.13 The c.d.f. at 32kV fitted to censored data (thick line), with parameters: α=-1.2×10^{-3}, λ=2.59×10^{-2}, m_1=34.0, m_2=2601. The dashed line represents the c.d.f. fitted to complete data.

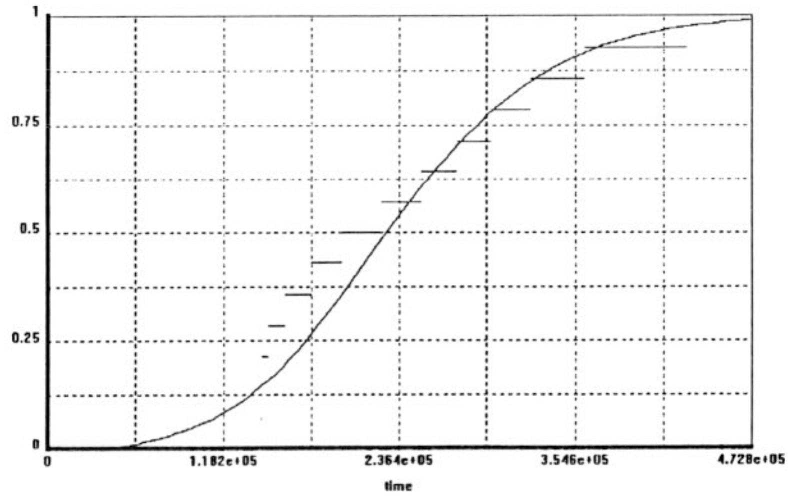

Figure 6.14 The c.d.f. at 20kV fitted to censored data (solid line), with parameters: α=4.5, λ=3.38×10^{-5}, m_1=234398, m_2=6.285×10^{10}.

Surprisingly, mean time to failure ($\approx 2.3\times10^5$) predicted in this study significantly exceeds the mean time of about 1.6×10^5 obtained in section 6.3 (see Fig. 6.10). This is probably not what would be expected when censored data is analyzed. Normally, censored samples should yield shorter predicted lifetimes as the inference is based on the lower tail of the empirical distribution. This slightly misleading behaviour of the prediction procedure can be explained if we observe that the c.d.f.s ML-fitted to censored samples (thick lines in Figs. 6.11-6.13) tend to overestimate the lower tails of the distributions. Combined with the effect that dispersions (4.20) of quantiles are small (and weights (4.22) are large) at lower tails of censored data, this leads to perhaps overestimated mean life at the nominal conditions. This problem can be alleviated if slightly modified version of the algorithm for extrapolation in stress is used in which all weights (4.5) and (4.22) are assumed to be equal (=1). Results (c.d.f., p.d.f. and hazard functions) obtained with this modification are presented in Figs. 6.15 and 6.16 (for censored and complete data, respectively). The y-axes are scaled according to values of the c.d.f.s; values of the p.d.f. and hazard functions corresponding to the '1' on the plots are as follows:

	p.d.f.	hazard
Fig. 6.15	4.86×10^{-6}	1.63×10^{-5}
Fig. 6.16	1.08×10^{-5}	1.08×10^{-5}

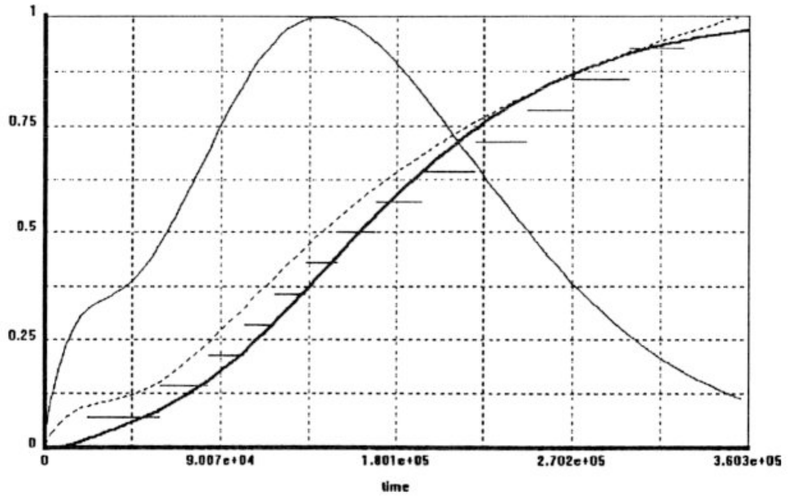

Figure 6.15 Censored data: nominal conditions c.d.f. (thick line, α=1.25, λ=2.52×10^{-5}, m_1=171219, m_2=3.726×10^{10}); hazard function (dashed line) and p.d.f. (solid thin line)

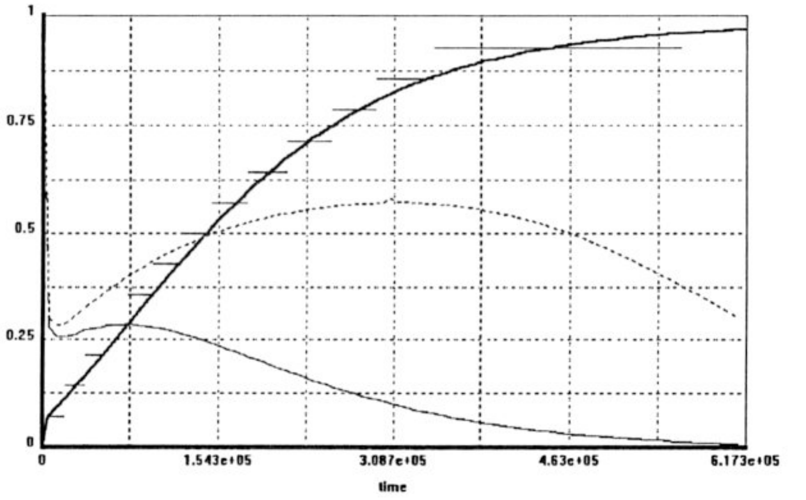

Figure 6.16 Complete data: nominal conditions c.d.f. (thick line, α=-0.8, λ=1×10^{-5}, m_1=171926, m_2=5.137×10^{10}); hazard function (dashed line) and p.d.f. (solid thin line)

Predicted mean lives in both cases are very similar (about 1.71×10^5 for censored and 1.72×10^5 for complete data). However, shapes of the hazard function and the p.d.f.s differ: the censored data prediction does not reveal the early failure evident in Fig. 6.16. Also, the failure rate for complete data is decreasing for distribution quantiles over 0.8, which is not predicted from censored data.

Chapter 7 Conclusions

In the thesis, a versatile method have been developed for analysis of data from accelerated life tests of electronic equipment. The method makes a flexible analysis of complete or censored life data observed at a specified level of stress. Extrapolation in stress does not require that a life-stress relationship is known in advance, but it is based on data driven selection of the most appropriate model. Verification of predicted reliability is an important feature of the methodology developed. Numerical examples in which sample test data was analysed demonstrate feasibility of practical application of the method to ALT data analysis. The examples also show how the results verification procedure performs and how it can be used to verify if models used for ALT studies analysed in literature are appropriate (for instance, it has been found in Chapter 6 that models used by Nelson (1990) are not suitable for his study).

A number of possible directions for further development of the method can be thought of. The method for modelling test data at a specified level of stress is entirely data driven, i.e., no physical considerations are used when fitting models to data. In most practical situations, this is the strength of the method since usually no physically founded models exist for new, high reliability and complex computer hardware devices. However, even though models of deterioration are not known, sometimes some specific behaviour of e.g., the failure rate (hazard) function might be expected, e.g., the increasing failure rate (IFR). In such cases, it would be desirable that a condition of this kind is taken into consideration when the flexible model as described in the thesis is used. It would be necessary that the method for modelling of complete or censored data (Chapter 2) was extended so that functions fitted to data satisfy certain constraints such as the IFR.

Another possible extension of the method might be the ways of obtaining confidence areas for parameters of the estimated reliability distribution functions (or confidence bounds for the distribution functions). Several approaches may be considered here, that are based on e.g., asymptotic behaviour of maximum likelihood estimates (they are asymptotically normally distributed), or on some statistics based on scores vectors and the Fisher information matrix (that exhibit more rapid convergence to their asymptotic distributions). An alternative approach that might be considered is the one demonstrated by Shaked and Singpurwalla (1982) to directly obtain confidence bounds for the life c.d.f. at the nominal stress.

For the data analysis method to be used in practice, it would be desired that a set of directions for the organization of an accelerated test plan is specified. They should be based on the observations that the method performs poorly for the studies with a small number of stress levels involved (such as three); on the other hand, it does not require that all units tested at increased stress actually fail.

Versatility of the method proposed and no need for a priori assumptions on life distribution and life-stress relationships make the method especially useful for life testing of electronic equipment. Although accelerated life testing has recently been extensively used for reliability estimation of electronic devices (e.g., the NASA study of multi chip modules in 1995, the VLSI ICs study at Motorola and AMD, 1995 or the

South Korean study of circuit boards for telecommunication systems, 1995, to name just the few), there are still areas where accelerated life testing is hindered by the lack of adequate models for analysis of results. For such areas, the method presented could inspire more interest in accelerated testing as the means of reliability estimation.

The methods presented in the thesis could be also applied to analysis of results of accelerated deterioration tests (ADT). Instead of times to failure analyzed with ALT, one might analyze times the units on test take to exceed some deterioration (or performance) threshold.

Based on experience with the method of ALT data analysis developed in the thesis, hopefully a versatile set of software tools can be developed to be used by reliability analysts working on real life accelerated test studies.

Appendix A Generalized Life Distributions

This Appendix gives an overview of properties of the generalized parametric family of life distributions used throughout the thesis for modelling of life data. Techniques used for fitting the functions to data are also discussed.

Definition and properties

The generalized family of distributions $F_{\alpha,\lambda}$ is defined as (Baskin 1988):

$$F_{\alpha,\lambda}(t) = \sum_{j=0}^{2} b_j c_j(t), \qquad \text{for } t \geq 0 \tag{A.1}$$

with

$$c_j(t) = \int_0^{\lambda t} x^{j+\alpha} e^{-x} dx \tag{A.2}$$

$$b_0 = \frac{1}{\Gamma(1+\alpha)}\left(3 + \frac{5}{2}\alpha + \frac{\alpha^2}{2} - \lambda m_1(3+\alpha) + \frac{\lambda^2 m_2}{2}\right) \tag{A.3}$$

$$b_1 = -\frac{1}{\Gamma(1+\alpha)}\left(3 + \alpha - \lambda m_1 \frac{5+2\alpha}{1+\alpha} + \frac{\lambda^2 m_2}{1+\alpha}\right) \tag{A.4}$$

$$b_2 = \frac{1}{2\Gamma(1+\alpha)}\left(1 - \frac{2\lambda m_1}{1+\alpha} + \frac{\lambda^2 m_2}{(1+\alpha)(2+\alpha)}\right) \tag{A.4}$$

$$m_i = \int_0^{\infty} t^i dF(t), \qquad i = 1,2 \tag{A.5}$$

for α>-1 and λ>0.

The functions (A.1) are used to approximate lifetime cumulative distribution functions, i.e., monotone functions F defined over $[0,\infty)$, with the properties $F(0)=0$, $F(\infty)=1$. The motivation for such representation comes from the theory of the orthogonal Laguerre polynomials L_j^α (j=0,1,2,...) defined over a weight function $W(x)=x^\alpha e^{-x}$, for $0<x<\infty$ (see e.g., Wawrzyńczyk 1978, Dziubiński 1978). The polynomials are given by the recurrence formula $(j+1)L_{j+1}^\alpha = (-x+2j+\alpha+1)L_j^\alpha - (j+\alpha)L_{j-1}^\alpha$, with $L_0^\alpha = 1$, $L_{-1}^\alpha = 0$, (this yields $L_1^\alpha(x) = 1+\alpha-x$, and $L_2^\alpha(x) = \left[x^2 - 2x(2+\alpha) + (1+\alpha)(2+\alpha)\right]/2$). Any continuous probability density function $f(x)$ defined over $(0,\infty)$ can be approximated with the sum $\sum_{j=0}^{\infty} c_j L_j^\alpha(x)$, where $c_j = \frac{j!}{\Gamma(j+\alpha+1)}\int_0^\infty f(x)x^\alpha e^{-x} L_j^\alpha(x)dx$.

The approximation suggested by Baskin uses three polynomials. This proves sufficient to approximate several typical life distributions dealt with in reliability theory, such as the Weibull, lognormal, gamma and mixtures of distributions.

The following examples illustrate performance of the function family (A.1) used to approximate some sample bimodal distributions.

Mixture of Weibull distributions

In this example, the mixture of two Weibull distributions:

$$F(t) = 1 - \varepsilon \exp\left(-(t/a_1)^{b_1}\right) - (1-\varepsilon)\exp\left(-(t/a_2)^{b_2}\right), \quad t > 0 \tag{A.6}$$

is being approximated using the $F_{\alpha,\lambda}$ family. Specific values are assumed: ε=0.1, a_1=0.5, b_1=1.5, a_2=7, b_2=3. The distribution (c.d.f. and p.d.f.) is shown in Figure A.1.

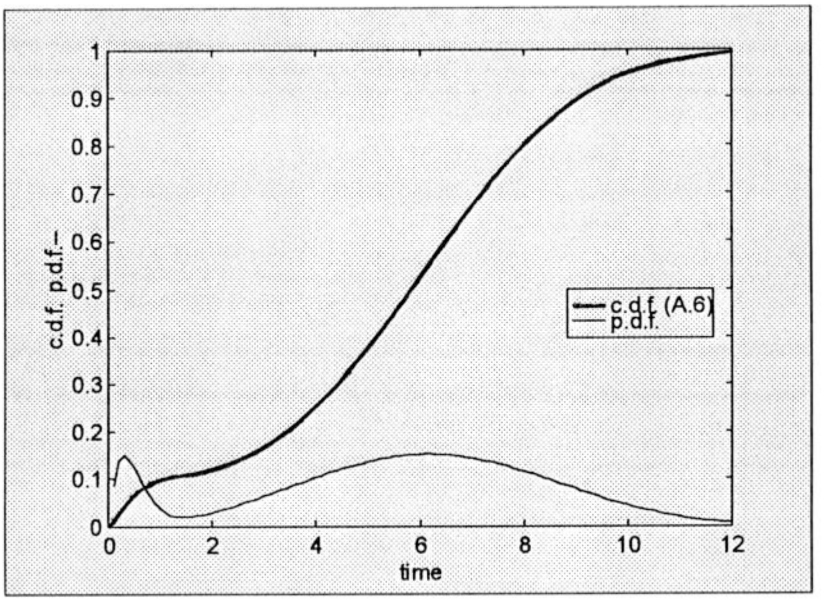

Figure A.1 C.d.f. and p.d.f. of mixture of Weibull distributions (A.6)

The best fit member of the $F_{\alpha,\lambda}$ family approximating the distribution is found by least squares minimization. This yields the following parameters for the $F_{\alpha,\lambda}$ function: α=3.28, λ=1.26, m_1=5.65, m_2=38.29. The original distribution and the approximating function $F_{\alpha,\lambda}$ are compared in Figure A.2.

One can compare the two distribution by the means of goodness of fit tests (as described in Appendix B). E.g., for a 19 element sample of the distribution (A.6) {0.4, 1.2, 2.88, 3.2, 3.8, 4.32, 4.845, 5.25, 5.4, 5.7, 6.3, 6.57858, 6.84897, 7.26284, 7.6, 8.04493, 8.56337, 9.08003, 10.0334}, the EDF statistics (B.5), (B.6) and (B.7) are: d_n=0.079, W_n^2=0.020, A_n^2=0.23. Since this is less than the quantiles (critical values) of the test distributions given in Table B.1, we conclude that the approximating distribution fits the original distribution well.

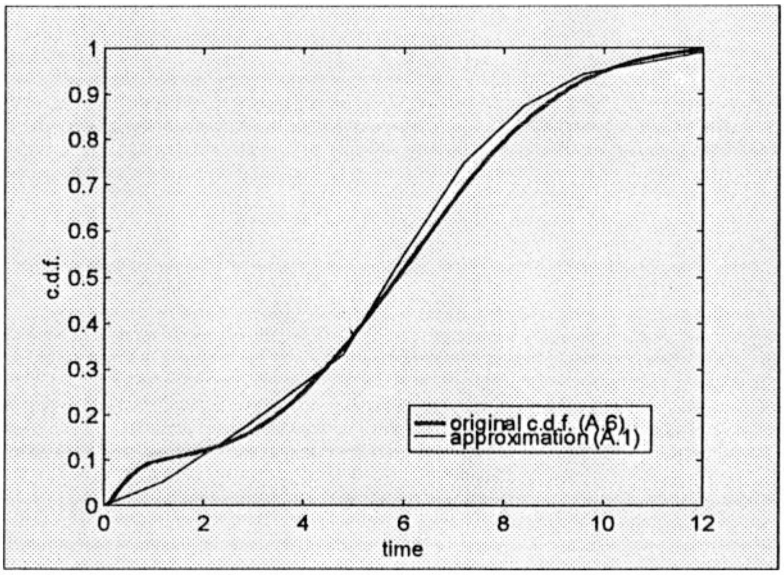

Figure A.2 C.d.f. of the mixture of Weibull distributions (A.6) and the approximation using the $F_{\alpha,\lambda}$ function (A.1).

Generalized Weibull distribution

The distribution used in this example is a specific form of the generalized Weibull distribution developed by Drapella (1986), mentioned in section 2.2.1 (and known as the URW), whose c.d.f. is defined as

$$F(t) = 1 - \exp\left(-(t/a_0)^{b(t)}\right), \quad t > 0, a_0 > 0, b(t) > 0 \tag{A.7}$$

with the linear form of the *shape function* $b(t)=b_0(t/a_1+1)$. In this example, $a_0=1$, $b_0=1$, $a_1=1$. The c.d.f. and p.d.f. of this distribution are shown in Figure A.3 (the p.d.f. is plotted in dashed line).

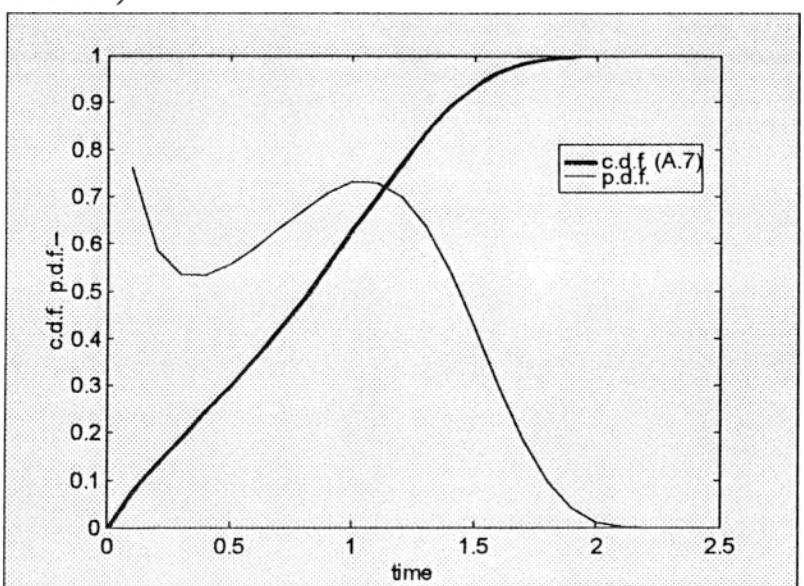

Figure A.3 C.d.f. and p.d.f. of the generalized Weibull distribution (A.7) by Drapella

The $F_{\alpha,\lambda}$ function (A.1) fitted to this distribution, with the parameters: α=3.48, λ=8.03, m_1=0.796, m_2=0.819, is presented in Figure A.4, together with the original distribution (A.7).

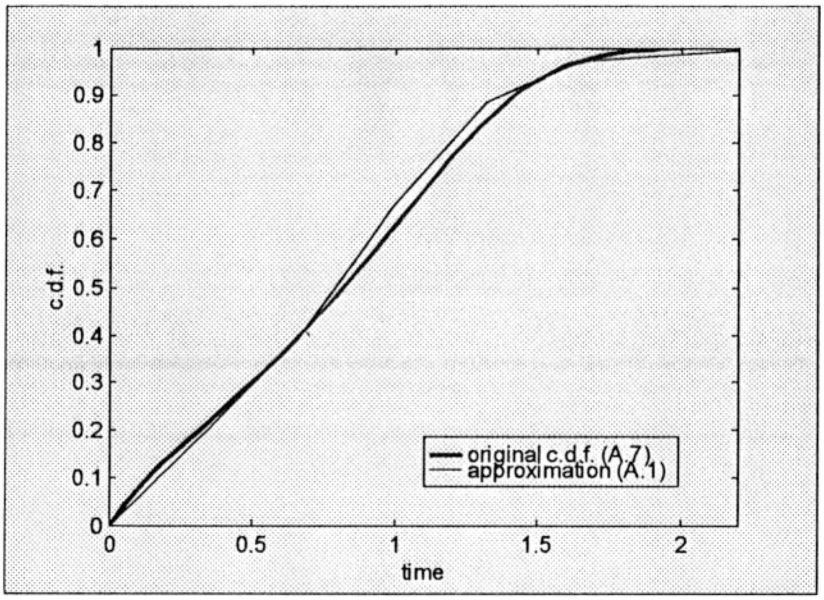

Figure A.4 C.d.f. of the generalized Weibull distribution (A.7) and the approximation using the $F_{\alpha,\lambda}$ function (A.1).

If a sample is selected from the distribution (A.7), then the EDF goodness of fit tests (see Appendix B) can be performed to check how well $F_{\alpha,\lambda}$ approximates the original distribution. For the sample of 19 values from the distribution: {0.060815, 0.138551, 0.227623, 0.32138, 0.414431, 0.503817, 0.588517, 0.6686, 0.744629, 0.817362, 0.887628, 0.956296, 1.02431, 1.09275, 1.16301, 1.23705, 1.31815, 1.41291, 1.54022}, the EDF statistics (B.5), (B.6) and (B.7) are: d_n=0.082, W_n^2=0.017, A_n^2=0.31, which is well below critical values of the test statistics (Table B.1). This means that the fit obtained is statistically significant.

Parameters of $F_{\alpha,\lambda}$

The functions from the family (A.1) do not always (i.e., for all values of the parameters α,λ,m_1,m_2) have the properties of a c.d.f. More specifically, although $F_{\alpha,\lambda}(0)=0$ and $F_{\alpha,\lambda}(\infty)=1$ is always true, the condition $dF_{\alpha,\lambda}(t)/dt\geq 0$ holds only for parameters that satisfy (Biernat, 1989, Maciejewski, 1990):

$$\begin{cases} \delta > (3+\alpha)(2\mu-2-\alpha) \\ \delta \geq (2+\alpha)(2\mu-1-\alpha) \\ \delta < \mu(5+2\alpha)-(3+\alpha)(1+\alpha) \end{cases} \tag{A.8}$$

or

$$A^2+(1+\alpha)B^2<(1+\alpha)(2+\alpha) \tag{A.8'}$$

where $\delta=\lambda^2 m_2$, $\mu=\lambda m_1$, $A=\delta-\mu(5+2\alpha)+(1+\alpha)(2+\alpha)$, $B=\mu-2-\alpha$.

Methods for fitting $F_{\alpha,\lambda}$ to data

Different algorithms are used for fitting members of the family $F_{\alpha,\lambda}$ to complete and censored data. For complete data, least squares distance is minimized between the $F_{\alpha,\lambda}$ function and the empirical distribution based on data, according to the formula (2.14). The algorithm for global minimization of the least squares distance is based on the steepest descent search launched from various starting points. Detailed description of the algorithm as well as heurestics employed for selection of the starting points for minimization are provided in Maciejewski (1990).

In the case of censored data, fitting the $F_{\alpha,\lambda}$ function involves maximization of the likelihood function given by (2.17) or (2.18). The most effective way of maximization found was by direct maximization of (2.17) or (2.18) by using global optimization techniques. The specific algorithm used for global minimization was based on the algorithm of the downhill simplex method of Nelder and Mead combined with genetic simulated annealing algorithm (see Press, 1992).

Computer programs to perform the least squares minimization and the maximum likelihood approach have been written be author.

Appendix B Goodness of Fit Testing

This Appendix presents some ways of testing goodness of fit used in the thesis. Methods considered here are used for checking adequacy of models applied for life data analysis at a specified stress level (as considered in Chapter 2 and Chapter 4), and in the procedure to verify results of reliability prediction which was considered in Chapter 5.

The statistical goodness of fit tests presented here are based mainly on empirical distribution functions (EDF), although applicability in certain cases of tests based on grouped data is also discussed.

Goodness of fit testing performed during the inference procedures presented in the thesis involves several cases that need separate treatment:

- models fitted to complete data or censored data;
- the hypothesized distribution function is fully specified or the distribution function involves unknown parameters.

Ways of testing goodness of fit in each of the cases are discussed in this Appendix.

Preliminaries

Given a random sample $t_1, t_2, \ldots, t_n$ from a population with the c.d.f. F, a goodness of fit test is performed to check if a given distribution function F_0 can be considered the population c.d.f. Testing goodness of fit involves testing the statistical hypothesis of the form:

$$H_0 : F = F_0 \tag{B.1}$$

Tests considered here are referred to as *omnibus tests*, as they are effective against wide classes of alternatives to F_0.

Procedures to test the hypothesis are based on a certain statistic U calculated from the random sample $t_1, t_2, \ldots, t_n$. For a chosen *significance level* α, an area $Q \subset R$ should be identified that comprises values of the statistic U indicating evidence against the hypothesis H_0, i.e., if $U(t_1, t_2, \ldots, t_n) \in Q$, then H_0 should be rejected. The significance level α is defined as the probability of rejecting the hypothesis H_0 although it is true:

$$P(U(t_1, t_2, \ldots, t_n) \in Q | H_0) = \alpha \tag{B.2}$$

Typically, small values of α are selected (e.g., 0.05 or 0.01).

For a given α, Q should be specified in such that the power of test is maximized. The power is defined as:

$$\text{power}=P(U(t_1,t_2,\ldots,t_n)\in Q|H_1)=1-P(U(t_1,t_2,\ldots,t_n)\notin Q|H_1) \tag{B.3}$$

which means that by maximizing power of the test, we minimize probability given by the term $P(U(t_1,t_2,\ldots,t_n)\notin Q|H_1)$ of accepting H_0, despite the fact that it is not true (i.e., the alternative hypothesis H_1 is true). Many goodness of fit tests, including tests presented in subsequent sections, realize best power for a chosen α if

$$Q=\{t:\ t>\lambda_\alpha\} \tag{B.4}$$

where λ_α is the quantile at level (1-α) of the distribution of the statistic U. This is illustrated on Fig. B.1. λ_α is referred to as a *critical value* of the test.

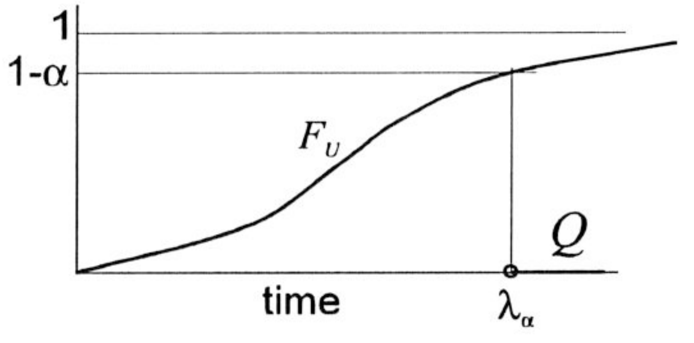

Figure B.1 Critical area Q that yields best power for tests in which large values of the test statistic U give evidence against H_0. The curve labelled F_U represents distribution of the test statistic U; α is a significance level of the test and λ_α is a quantile of F_U at the level (1-α).

Tests for uncensored data

Tests used for the inference procedures presented in the thesis are mainly EDF tests, i.e., they involve measuring distance between the hypothesized distribution function and the empirical distribution function obtained from a random sample.

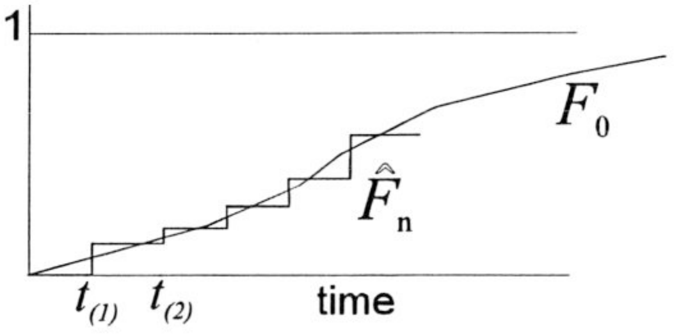

Figure B.2 EDF goodness of fit tests involve measuring distance between empirical distribution function $\hat{F}_n$ and the hypothesized distribution F_0.

For uncensored data $t_1, t_2, \ldots, t_n$, under the assumption that F_0 is fully specified (i.e., it does not involve any unknown parameters), the well known goodness of fit test based on the Kolmogorov-Smirnov statistic can be used (Fisz 1965):

$$d_n = \sup_t \left| F_0(t) - \hat{F}_n(t) \right| \tag{B.5}$$

where $\hat{F}_n$ is the empirical distribution defined as:

$$\hat{F}_n(t) = \begin{cases} 0 & \text{for} \quad t \le t_{(1)} \\ \dfrac{k}{n} & \text{for} \quad t_{(k)} < t \le t_{(k+1)},\ k = 1,2,\ldots,n-1 \\ 1 & \text{for} \quad t > t_{(n)} \end{cases} \tag{B.6}$$

and n is the number of samples observed and $t_{(1)} \le t_{(2)} \le \ldots \le t_{(n)}$ is the order statistic from the random sample.

Large values of the d_n statistic indicate evidence against H_0 which means that the critical area Q should be selected according to (B.4).

For uncensored data, with F_0 fully specified, the test based on d_n is *distribution-free*, i.e., the distribution of the statistic d_n under H_0 does not depend on F_0. The asymptotic distribution of $d_n\sqrt{n}$ is, under H_0, the Kolmogorov distribution (Fisz 1965). Exact distribution of d_n under H_0 is known for all n. Quantiles of small sample distributions (small n) are tabulated (Lawless 1982).

To test the hypothesis, a quantile λ_α of the Kolmogorov distribution (or of a small sample distribution, for a small n) at a specified level (1-α) should be determined, where α is the significance level of the test (e.g., α=0.05).

If

$$d_n > \frac{\lambda_\alpha}{\sqrt{n}} \tag{B.7}$$

then H_0 should be rejected, i.e., F_0 is not a distribution of the population from which the random sample $t_1, t_2, \ldots, t_n$ was selected.

Other goodness of fit tests with good power against broad ranges of alternatives, recommended by Lawless (1982), are the tests based on the Cramer-von Mises statistic:

$$W_n^2 = n \int_{-\infty}^{+\infty} \left[F_0(t) - \hat{F}_n(t)\right]^2 dF_0(t) \tag{B.8}$$

and the Anderson-Darling statistic:

$$A_n^2 = n \int_{-\infty}^{+\infty} \frac{\left[F_0(t) - \hat{F}_n(t)\right]^2}{F_0(t)\left[1 - F_0(t)\right]} dF_0(t) \tag{B.9}$$

According to Stephens (1974), the W_n^2 and A_n^2 statistics tend to be more powerful than d_n for a wide class of alternatives.

For F_0 fully specified, the both statistics are distribution free. Asymptotic distributions of the statistics are known and small sample distributions (quantiles) were obtained by simulation. Based on the results, Table B.1 (Stephens 1974) allows easy determination of approximate values of quantiles of the statistics d_n, W_n^2 and A_n^2 for all n.

TABLE B.1

QUANTILES FOR ALL n OF THE STATISTICS d_n, W_n^2 AND A_n^2 TO BE USED IF F_0 IS FULLY SPECIFIED. IF VALUE OF A MODIFIED STATISTIC EXCEEDS A QUANTILE OBTAINED FROM THE TABLE THEN H_0 SHOULD BE REJECTED WITH THE TEST SIGNIFICANCE LEVEL α=(1-QUANTILE LEVEL).

statistic	modified statistic	Quantile				
		0.85	0.90	0.95	0.975	0.99
d_n	$d_n^* = d_n\left(\sqrt{n} + 0.12 + 0.11/\sqrt{n}\right)$	1.138	1.224	1.358	1.480	1.628
W_n^2	$W_n^* = \left(W_n^2 - 0.4/n + 0.6/n^2\right)\left(1 + 1/n\right)$	0.284	0.347	0.461	0.581	0.743
A_n^2	for all $n \geq 5$	1.610	1.933	2.492	3.070	3.857

Computationally more convenient forms of d_n, W_n^2 and A_n^2 are given by the formulae (Lawless 1982):

$$d_n = \max\left(d_n^+, d_n^-\right) \tag{B.5a}$$

where:

$$\begin{aligned} d_n^+ &= \max_{1\le i\le n}\left(\frac{i}{n} - F_0\left(t_{(i)}\right)\right) \\ d_n^- &= \max_{1\le i\le n}\left(F_0\left(t_{(i)}\right) - \frac{i-1}{n}\right) \end{aligned} \tag{B.5b}$$

$$W_n^2 = \sum_{i=1}^{n}\left(F_0\left(t_{(i)}\right) - \frac{i-0.5}{n}\right)^2 + \frac{1}{12n} \tag{B.8a}$$

$$A_n^2 = -\sum_{i=1}^{n}\frac{2i-1}{n}\left\{\log\left[F_0\left(t_{(i)}\right)\right] + \log\left[1 - F_0\left(t_{(n+1-i)}\right)\right]\right\} - n \tag{B.9a}$$

where $t_{(1)}, t_{(2)}, \ldots, t_{(n)}$ denotes the order statistic from the random sample.

Tests for censored data

In the case of Type II and singly Type I censored data, modified EDF statistics based on (B.5), (B.8) and (B.9) can be used. The Kolmogorov-Smirnov test can be modified as suggested by Barr and Davidson (1973).

For Type II (failure) censored data, the modified Kolmogorov-Smirnov statistic has the form:

$$d_{n,r} = \sup_{t\le t_{(r)}}\left|F_0(t) - \hat{F}_n(t)\right| \tag{B.10}$$

where r is the number of observations available in the failure censored random sample, i.e., $t_{(r)}$ is the rth biggest lifetime observed and all remaining n-r lifetimes are known to exceed $t_{(r)}$.

For singly Type I (time) censored data, the modified Kolmogorov-Smirnov statistic has the form:

$$d_{n,p} = \sup_{t\le L}\left|F_0(t) - \hat{F}_n(t)\right| \tag{B.11}$$

where L is the (common) censoring time and

$$p = F_0(L) \tag{B.12}$$

Using the common asymptotic distribution of $d_{n,p}$ in the cases of time censoring and failure censoring, with a fixed $p=r/n$ as $r \to \infty$ and $n \to \infty$ in the case of failure censoring (Koziol and Byar 1975), or using tabularized small sample distributions of $d_{n,p}$ and $d_{n,r}$ (Dufour and Maag 1978), a quantile λ_α of the predetermined level $(1-\alpha)$ should be determined, where α is the significance level of the test (e.g., α=0.05).
Large values of the statistics $d_{n,p}$ and $d_{n,r}$ indicate evidence against H_0. If

$$d_{n,p} > \frac{\lambda_\alpha}{\sqrt{n}} \tag{B.13}$$

or

$$d_{n,r} > \frac{\lambda_\alpha}{\sqrt{n}} \tag{B.14}$$

for Type I and Type II censored data, respectively, then H_0 should be rejected, i.e., F_0 cannot be regarded as the distribution of the population form which the (censored) random sample was obtained.

Generalizations of the Cramer-von Mises and Anderson-Darling statistics for Type I and Type II censored samples are given by the formulae (see Lawless (1982)):

$$W_{n,r}^2 = n \int_{-\infty}^{t_{(r)}} \left[F_0(t) - \hat{F}_n(t)\right]^2 dF_0(t) \tag{B.15}$$

$$A_{n,r}^2 = n \int_{-\infty}^{t_{(r)}} \frac{\left[F_0(t) - \hat{F}_n(t)\right]^2}{F_0(t)\left[1 - F_0(t)\right]} dF_0(t) \tag{B.16}$$

where $t_{(r)}$ is the r-th largest observation available, and all other observations are known to be exceed $t_{(r)}$.
Similarly, the version to be used for singly Type I censored data is:

$$W_{n,p}^2 = n \int_{-\infty}^{L} \left[F_0(t) - \hat{F}_n(t)\right]^2 dF_0(t) \tag{B.17}$$

$$A_{n,p}^2 = n \int_{-\infty}^{L} \frac{\left[F_0(t) - \hat{F}_n(t)\right]^2}{F_0(t)\left[1 - F_0(t)\right]} dF_0(t) \tag{B.18}$$

where L is the (common) censoring time and p is defined by (B.12).
Critical areas can be obtained from asymptotic distributions given, together with tables of percentage points, by (Petitt and Stephens 1976).

Computationally more convenient forms of (B.15)-(B.18) are provided by Lawless (1982).

Distributions with unknown parameters

Tests presented in the previous sections are intended to be used with the hypothesized distribution function *fully specified*, i.e., a distribution that does not involve any unknown parameters. The tests are of use for instance for results verification procedures described in Chapter 5. In the case considered in Chapter 5, a distribution can be regarded as fully specified function since it does not involve any unknown parameters that must be estimated from data to which the function is fitted.

However, distribution functions dealt with in Chapter 2 cannot be considered fully specified because they contain parameters that are estimated from data at the corresponding stress level. In that case, tests presented earlier are no longer distribution free (i.e., distributions of statistics presented do depend on F_0) and the distribution theory in not known for the generalized model $F_{\alpha,\lambda}$ as defined by (2.7), with parameters estimated from data.

In this important though difficult case, alternative approaches to testing goodness of fit are proposed:

1) Use EDF tests designed for functions that are fully specified despite the fact that the hypothesized distribution contains parameters estimated from data. Although not recommended, this procedure is actually used by some authors (see e.g., Fisz 1965). One should be aware, however, that in this case, the significance levels may be far off the correct ones.

   Lawless (1982) argues that if unknown parameters are present, tests based on W_n^2 and A_n^2 prove to be substantially more powerful than tests based on d_n, thus they are preferred in this case.

2) Use tests based on grouped data, e.g., the Pearson χ^2 test. Although less powerful that EDF tests (Lawless 1982), the Pearson χ^2 test accommodates unknown parameters of the hypothesized distribution function (Fisz 1965). Improvements of the classical Pearson test, based on recent results (Bogdan 1995) concerning selection of classes, should be taken into consideration when developing the test for the inference procedures presented in Chapter 2.

3) Application of smooth goodness of fit tests for the purposes of this thesis could be also investigated. The idea of smooth goodness of fit tests was originally proposed by Neyman (in 1937) and in recent years considerable interest in the tests have arisen (see e.g., Lawless 1982, Kallenberg&Ledwina 1995).

   The tests consist in embedding the hypothesized c.d.f. $F_0(t,\theta)$ with the p.d.f. $f_0(t,\theta)$ in a more general model with the p.d.f.

$$g(t;\theta;\beta) = f_0(t;\theta)\exp\left(\sum_{j=1}^{k}\beta_j F_0^j(t;\theta) - k(\beta)\right) \qquad \text{(B.19)}$$

where θ is the unknown vector of parameters; $\beta=(\beta_1,\beta_2,\ldots,\beta_k)$ is a vector of parameters and $k(\beta)$ is a constant that does not depend on θ. Testing H_0 is reduced to testing that $\boldsymbol{\beta}=\mathbf{0}$. Kallenberg and Ledwina (1995) propose modified (data driven) Neyman's tests and quote a conclusion resulting from a review of several methods of goodness of fit testing which says "don't use those other methods - use a smooth test!".

In the thesis, the first approach is adopted if data is censored and the distribution function fitted to data is not fully specified.

Appendix C Analysis of Variance for Regression

This Appendix presents details of the weighted linear regression model and is intended to give mathematical foundation for the inference procedure dealt with in Chapter 4. Analysis of variance for regression is also considered as it provides the tool for examining the model's suitability. Although the major part of this appendix concentrates on the weighted linear regression, some ideas are also presented about testing model's adequacy in case of a nonlinear model. The formulae presented here are based mainly on (Dunn 87) and (Draper 73).

Linear regression

We consider two random variables denoted X and Y. The two-dimensional probability distribution function of the pair of variables gives the full probabilistic information about their behaviour. Dependence of one of the variables on the other can be described, considering mean values, by the regression line (function) defined as $E(Y|X=x)=\mu_{yx}$ and treated as a function of x. However, in many practical cases no information of this kind is available about the random variables observed. In such cases a model of the relationship has to be assumed and ways have to be devised to check if the model is adequate.

The model considered here is based on the following assumptions:

- for each (fixed) value x of X (of a certain set), there is a population of Y's, that is normally distributed, with the mean $E(Y|X=x)$ and variance σ_x^2;
- the means lie on the straight line:

$$E(Y|X)=\beta_0+\beta_1 X \tag{C.1}$$

where β_0 and β_1 are the unknown model parameters.

The model parameters β_0 and β_1 can be estimated based on a random sample consisting of k observations $\{(x_i,y_i)\}$, i=1,2,...,k, where y_i is a value of Y observed for a fixed value x_i of X. It is assumed that some of the observations are weighted more heavily than the others, which is expressed by an arbitrary set of fixed positive weights $\{p_i\}$, i=1,2,...,k, such that:

$$Var(Y|X=x_i)=\frac{\sigma^2}{p_i} \tag{C.2}$$

The model assumptions can be summarized by an alternative expression:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1,2,\ldots,k \tag{C.3}$$

where ε_i are independently normally distributed with a mean 0 and variance σ^2/p_i (σ^2 is unknown).

Using the following notation: $\boldsymbol{\beta}=[\beta_0,\beta_1]^T$, $\mathbf{Y}=[y_1,y_2,\ldots,y_k]^T$, $\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \ldots & x_k \\ 1 & 1 & \ldots & 1 \end{bmatrix}^T$, $\mathbf{P} = \begin{bmatrix} 1/\sqrt{p_1} & & 0 \\ & \ldots & \\ 0 & & 1/\sqrt{p_k} \end{bmatrix}$ and $\mathbf{V}=\mathbf{P}^2$, the point estimates $\mathbf{b}=[b_0,b_1]^T$ of the model parameters $\boldsymbol{\beta}$ are given by (Draper 73):

$$\mathbf{b} = \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y} \tag{C.4}$$

under the assumption that $\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}$ is not singular.
(C.4) is equivalent to:

$$\begin{aligned} b_0 &= \frac{\left(\sum p_i y_i\right)\left(\sum p_i x_i^2\right) - \left(\sum p_i x_i\right)\left(\sum p_i x_i y_i\right)}{\left(\sum p_i\right)\left(\sum p_i x_i^2\right) - \left(\sum p_i x_i\right)^2} \\ b_1 &= \frac{\left(\sum p_i\right)\left(\sum p_i x_i y_i\right) - \left(\sum p_i x_i\right)\left(\sum p_i y_i\right)}{\left(\sum p_i\right)\left(\sum p_i x_i^2\right) - \left(\sum p_i x_i\right)^2} \end{aligned} \tag{C.5}$$

where all the sums are for i=1,2,...,k.
The linear regression model becomes:

$$\hat{y} = b_0 + b_1 x \tag{C.6}$$

As considered in Chapter 4, this model is used for extrapolation of quantiles to the nominal stress level, providing a linearized life-stress relationship is assumed. This procedure is justified only if the model (C.3) is *adequate*. The method of checking adequacy of the model is based on the analysis of variance which is outlined below.

It follows from the identity (see Fig. C.1):

$$y_i - \bar{y} = \left(y_i - \hat{y}_i\right) + \left(\hat{y}_i - \bar{y}\right) \tag{C.7}$$

where $\hat{y}_i = b_0 + b_1 x_i$ and $\bar{y} = \sum_{i=1}^{k} y_i$, that (Draper 73):

$$\sum_{i=1}^{k}(y_i-\bar{y})^2=\sum_{i=1}^{k}(y_i-\hat{y}_i)^2+\sum_{i=1}^{k}(\hat{y}_i-\bar{y})^2 \tag{C.8}$$

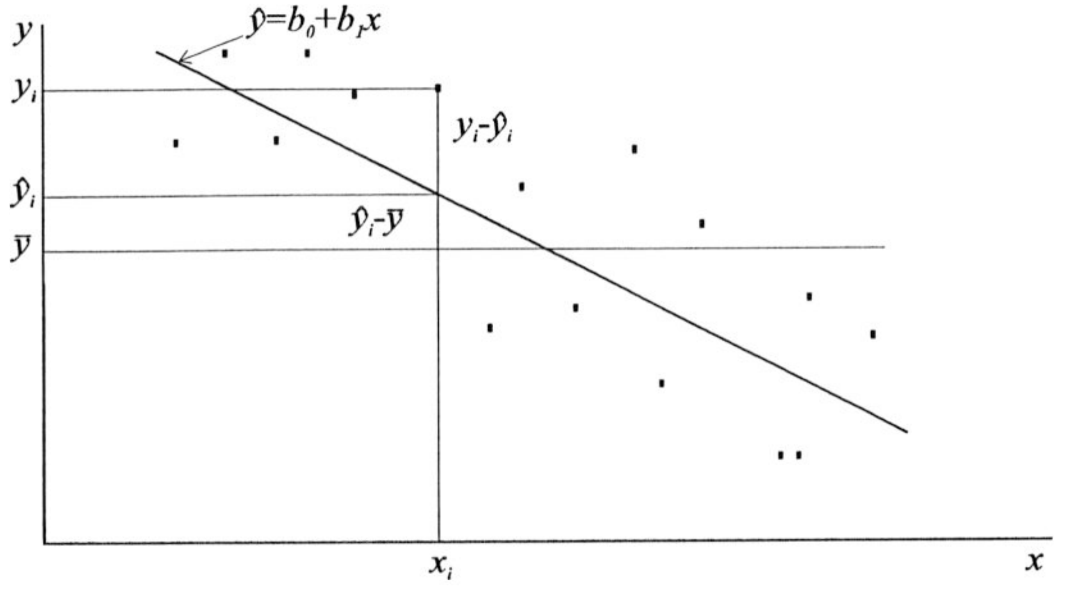

Figure C.1 Linear regression model

The sums of squares in (C.8) form the classical analysis of variance table for linear regression and have the following interpretation:

TABLE C.1

ANALYSIS OF VARIANCE FOR REGRESSION

$SSR=\sum_{i=1}^{k}(\hat{y}_i-\bar{y})^2$	is the *due regression* sum of squares, with 1 degree of freedom;
$SSE=\sum_{i=1}^{k}(y_i-\hat{y}_i)^2$	is the *residual* sum of squares, with k-2 degrees of freedom;
$SST=\sum_{i=1}^{k}(y_i-\bar{y})^2$	is the *total* sum of squares, with k-1 degrees of freedom.

Under the assumptions that the terms ε_i in (C.3) are independent and $\varepsilon_i \sim N(0,\sigma^2)$ for $i=1,2,\ldots,k$, (i.e., their variances are *equal*), and $\beta_1=0$, the statistic $F=(SSR/1)/(SSE/(k-2))$ has the F-Snedecor distribution with (1,k-2) degrees of freedom (Draper 73). Thus, to test the hypothesis H_0: $\beta_1=0$ (versus H_1: $\beta_1\neq0$), we compare F with $F_{1-\alpha}(1,k-2)$. If $F> F_{1-\alpha}(1,k-2)$, then H_0 is rejected (with the test significance level equal α), which implies that the regression model is *adequate* (i.e., it fits the data better than a horizontal line).

In the case of weighted regression, which is considered here, the sums of squares are given by (Draper 73):

$$SSR = \mathbf{b}^T\mathbf{Q}^T\mathbf{Z} - \frac{\left(\sum_{i=1}^{k} z_i\right)^2}{k} \quad \text{(with 1 d.f.)} \tag{C.9}$$

$$SSE = \mathbf{Z}^T\mathbf{Z} - \mathbf{b}^T\mathbf{Q}^T\mathbf{Z} \quad \text{(with } k\text{-2 d.f.)} \tag{C.10}$$

and $SST=SSR+SSE$, with k-1 degrees of freedom. The following notation has been used: $\mathbf{Q}=\mathbf{P}^{-1}\mathbf{X}$, $\mathbf{Z}=\mathbf{P}^{-1}\mathbf{Y}$, $z_i = \sqrt{p_i}y_i$, with $\mathbf{X}$,$\mathbf{Y}$ and $\mathbf{P}$ defined before the formula (C.4). Equivalent versions of (C.9) and (C.10), convenient for computation, follow immediately:

$$\begin{aligned} SSR &= b_1\left(\sum p_i x_i y_i - k^{-1}\left(\sum \sqrt{p_i} x_i\right)\left(\sum \sqrt{p_i} y_i\right)\right) = \\ &= b_0 \sum p_i y_i + b_1 \sum p_i x_i y_i - k^{-1}\left(\sum \sqrt{p_i} y_i\right)^2 \end{aligned} \tag{C.11}$$

$$SSE = \sum p_i y_i^2 - b_0 \sum p_i y_i - b_1 \sum p_i x_i y_i \tag{C.12}$$

with all the sums running for i=1,2,...,k. The F test is performed as described before.

The test for H_0: β_1=0 allows to check if the least squares line $\hat{y} = b_0 + b_1 x$ fits the data better than the horizontal line. However, it does not give any indication whether the relationship between means of Y and X is really linear (as assumed in (C.1)). It is possible to check the linearity assumption providing some independent (a priori) information is available about the variance of $Y|X$.

The information can be obtained e.g., from repeated observations y_{ij} made for some x_i (see Fig.C.2). If repeated observations are available, the residual sum of squares (as defined in Table C.1) can be split into two terms: the *means about regression* term (which gives a measure of deviation from the regression line of the means for each x_i), and the residual (or *pure error*) term (which represents deviation of observations obtained for an x_i around their mean). The tests of linearity involves comparing the two terms. If the means about regression term is significantly larger than the pure error term, then it can be suspected that the relationship between $E(Y|X)$ may not be linear. Computational details follow below.

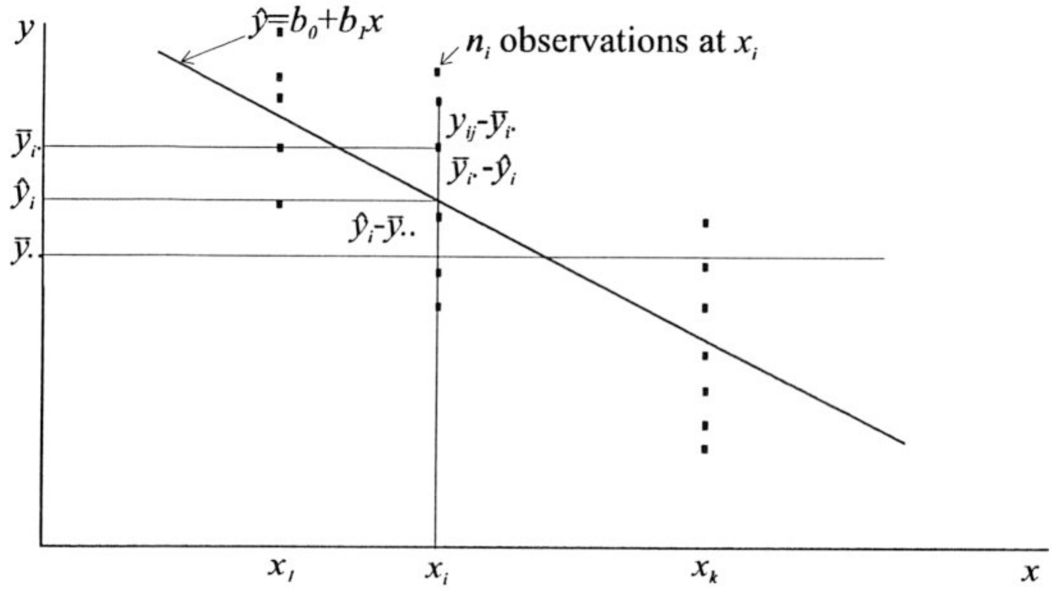

Figure C.2 Testing linearity based on repeated observations

Let y_{ij}, j=1,2,...,n_i, denote individual observations available for x_i, i=1,2,...,k, as illustrated in Fig. C.2. Let $\bar{y}_{i\bullet} = n_i^{-1}\sum_{j=1}^{n_i} y_{ij}$ denote the mean for x_i, and $\bar{y}_{\bullet\bullet}$ the mean of all observations. Obviously, $\bar{y}_{\bullet\bullet} = \left(\sum_{i=1}^{k} n_i\bar{y}_{i\bullet}\right)\Big/\sum_{i=1}^{k} n_i$. Let $N = \sum_{i=1}^{k} n_i$.

Using the notation, the analysis of variance table can be summarized as follows (Dunn 87):

TABLE C.2

ANALYSIS OF VARIANCE FOR REGRESSION (WITH REPEATED OBSERVATIONS)

$SSR^r = \sum_{i=1}^{k} n_i(\hat{y}_i - \bar{y}_{\bullet\bullet})^2$	is the *due regression* sum of squares, with 1 degree of freedom;
$SSE^r = \sum_{i=1}^{k} n_i(\bar{y}_{i\bullet} - \hat{y}_i)^2$	is the *means about regression* sum of squares, with (k-2) degrees of freedom;
$S_e^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{i\bullet}\right)^2$	is the *residual* (or *pure error*) sum of squares, with N-k degrees of freedom;
$SST^r = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{\bullet\bullet}\right)^2$	is the *total* sum of squares, with N-1 degrees of freedom.

The superscript r in Table C.2 was introduced to distinguish the "repeated observations" case for the "single observation" case considered in Table C.1.

To test the linearity hypothesis $E(Y|X) = \beta_0 + \beta_1 X$, we compare the statistic $F = \left(SSE^r/(k-2)\right)/\left(S_e^2/(N-k)\right)$ with the quantile $F_{1-\alpha}(k\text{-}2,N\text{-}k)$ of the F-Snedecor distribution with $(k\text{-}2,N\text{-}k)$ degrees of freedom (Dunn 87). If $F > F_{1-\alpha}(k\text{-}2,N\text{-}k)$, then the linearity hypothesis is rejected, with the test significance level α.

Obviously, in the case considered in Chapter 4, no repeated observations (quantiles) are available for each value of x_i (which corresponds to a specified value of the stress level). However, some a priori information about dispersion of the observations made at each x_i is available (see: (4.20) and (4.21)). As suggested by (Baskin 88), this can be used as an estimate of the pure error, which implies that the test for linearity can be performed.

If we regard the value y_i (obtained by transforming the quantile tq_{ji} (4.2) for a fixed j, as described in Chapter 4 using (4.15), (4.17) or (4.19)), as the *mean* of n_i observations available for x_i (which corresponds the stress level S_i), then the sum of squares of the *pure error* for x_i can be rewritten as $(n_i - 1)Dy_i$, where Dy_i denotes the estimate of variance of the observations for x_i. Hence the term S_e^2 in Table C.2 is equal:

$$S_e^2 = \sum_{i=1}^{k} (n_i - 1) Dy_i \tag{C.13}$$

This gives the pure error sum of squares.

For the *weighted* regression case, the means about regression term SSE^r can be rewritten as (see (C.12)):

$$SSE^r = \sum_{i=1}^{k} p_i y_i^2 - b_0 \sum_{i=1}^{k} p_i y_i - b_1 \sum_{i=1}^{k} p_i x_i y_i \tag{C.14}$$

where y_i now denotes the *mean* of n_i observations for x_i.

As mentioned before, the linearity assumption ($E(Y|X) = \beta_0 + \beta_1 X$) is rejected if:

$$\frac{SSE^r/(k-2)}{S_e^2/(N-k)} > F_{1-\alpha}(k-2, N-k) \tag{C.15}$$

with the test significance level α.

Also observe that the terms (C.13) and (C.14) sum up to the *residual* sum of squares which represents dispersion of observations around the regression line (in the case when repeated observations are available). This is analogous to the residual sum of squares SSE for the case of regression with single observations (Table C.1).

Thus, to test H_0: $\beta_1=0$ with repeated observations, we compare the *due regression mean square* ($SSR^r/1$) with the *residual mean square* ($\left(SSR^r + S_e^2\right)/(N-2)$), (Dunn 87). If

$$\frac{SSR^r/1}{\left(SSE^r + S_e^2\right)/(N-2)} > F_{1-\alpha}(1, N-2) \tag{C.16}$$

then the model is considered adequate (i.e., it fits the data better than the horizontal line). In the case of weighted regression, the *due regression* sum of squares SSR^r has the form:

$$SSR^r = \sum_{i=1}^{k} n_i \left(b_0 + b_1\sqrt{p_i}\,x_i - N^{-1}\sum_{i=1}^{k} n_i\sqrt{p_i}\,y_i \right)^2 \tag{C.17}$$

This follows immediately from Table C.2. The definition of SSR in Table C.2 is appropriate for the case of non weighted regression. By using the transformations: $z_i = \sqrt{p_i}\,y_i$ and $q_i = \sqrt{p_i}\,x_i$, the weighted regression is converted to the nonweighted case (Draper 73). Thus the term SSR can be defined in terms of the variable z. Observe that $\hat{z}_i = b_0 + b_1 q_i = b_0 + b_1\sqrt{p_i}\,x_i$ and $\bar{z}_{\bullet\bullet} = \left(\sum n_i z_i\right)/\sum n_i = \left(\sum n_i\sqrt{p_i}\,y_i\right)/\sum n_i$, which yields (C.17).

If the model (C.1) is adequate, then the equation (C.6) can be used to arrive at the value of a specified quantile in the nominal conditions (corresponding to x_0). The variance $D\hat{y}(x_0)$ of the prediction $\hat{y}(x_0) = b_0 + b_1 x_0$ obtained for a fixed value of x is given by (Draper 73):

$$D\hat{y}(x_0) = \sigma^2 \mathbf{x}_0^T \left(\mathbf{Q}^T\mathbf{Q}\right)^{-1} \mathbf{x}_0 \tag{C.18}$$

under the assumption that that $\mathbf{Q}^T\mathbf{Q}$ is not singular. Following notation has been used: $\mathbf{x}_0=[1,x_0]^T$, $\mathbf{Q}$ has been defined above (after the formula (C.10)), and σ^2 is the unknown model variance. The estimate S^2 of σ^2 is given by (Draper 73):

$$S^2 = \frac{SSE^r + S_e^2}{N-2} \tag{C.19}$$

Hence (C.18) can be rewritten as (with all sums for $i=1,2,\ldots,k$):

$$D\hat{y}(x_0) \approx S^2 \frac{\sum p_i x_i^2 - 2x_0\sum p_i x_i + x_0^2\sum p_i}{\left(\sum p_i\right)\left(\sum p_i x_i^2\right) - \left(\sum p_i x_i\right)^2} \tag{C.20}$$

Nonlinear regression

This section deals with testing adequacy of a nonlinear regression model. A heuristic approach is presented, based on (Draper 73), to test adequacy of a nonlinear model fitted to data. The formulae given here are used in Section 4.4.1. A method is also presented to obtain approximate confidence areas for parameters of the model.

As discussed in the previous section, we consider the random variables X and Y. We drop the assumption (C.1) and assume instead that $E(Y|X)=f(X,\Theta)$ where Θ is the vector of p model parameters and a continuous function f is a specified (nonlinear) model of the relationship between $E(Y|X)$ and X.

Let $\hat{\Theta}$ denote the least squares estimate of the (true) parameter value, i.e., $\hat{\Theta}$ is the value of Θ that globally minimizes the distance:

$$S(\Theta)=\sum_{i=1}^{k}\left(y_i-f\left(x_i,\Theta\right)\right)^2 \tag{C.21}$$

where $\{(x_i,y_i)\}$, i=1,2,...,k, is the data the model is fitted to.

The problem considered here is to test if a nonlinear model $f\left(x,\hat{\Theta}\right)$ fitted to data $\{(x_i,y_i)\}$, i=1,2,...,k, is *adequate* (by which we mean that it represents the *true* relationship between $E(Y|X)$ and X).

If *a priori* information is available about variance of y's for a fixed x, then a heuristic method can be proposed for testing adequacy of a nonlinear model, as suggested by (Draper 73).

Let Dy_i denote an estimate of variance of $Y|X$ for a fixed value $X=x_i$. Then the sum of squares of the pure error is:

$$S_{pe}^2=\sum_{i=1}^{k}(n_i-1)Dy_i \tag{C.22}$$

where it is assumed that n_i independent observations are available for a fixed x_i. These n_i independent observations have been used to arrive at the estimate Dy_i.

The error of approximation of the model $f\left(x,\hat{\Theta}\right)$ is $S\left(\hat{\Theta}\right)$ (see C.21).

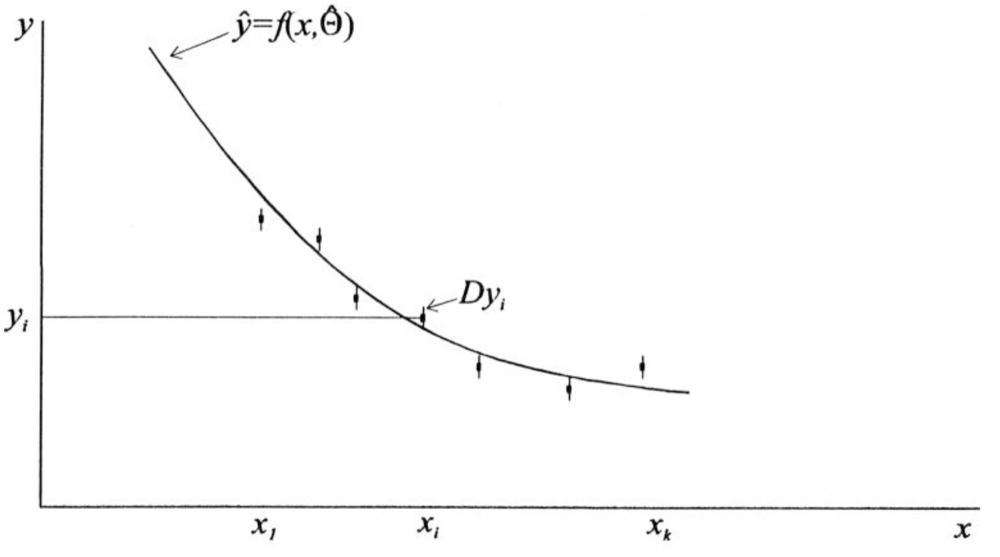

Figure C.3 Nonlinear regression

Testing adequacy of the model $f(x,\hat{\Theta})$ involves comparing the terms $S(\hat{\Theta})$ and S_{pe}^2 in order to check whether or not the error of approximation term $S(\hat{\Theta})$ can be accounted for by the pure error term S_{pe}^2. If $S(\hat{\Theta})$ is significantly larger than S_{pe}^2, then conclusion can be drawn that error of the model exists. Specifically, we compare the ratio of the mean squares:

$$F = \frac{S(\hat{\Theta})/(k-p)}{S_{pe}^2/\left(\sum_{i=1}^{k} n_i - k\right)} \tag{C.23}$$

with a quantile $F_{1-\alpha}(k\text{-}p,N\text{-}k)$ (where $N = \sum_{i=1}^{k} n_i$). If $F<F_{1-\alpha}(k\text{-}p,N\text{-}k)$, then some indication is obtained that the model $f(x,\hat{\Theta})$ may be adequate (Draper 73).

Obviously, assumptions of the F-test do not hold in the case of a nonlinear model. The test presented above should be considered a heuristic method to be used if no other, more accurate indications are available.

Similarly, $(1\text{-}\alpha)$th confidence areas for the model parameters can be estimated as follows (Draper 73):

$$\left\{\Theta\text{: } S(\Theta) \le S(\hat{\Theta})\left[1+\frac{p}{k-p}F_{1-\alpha}(p,k-p)\right]\right\} \tag{C.24}$$

In the nonlinear case, the formula (C.24) gives approximate shapes of the confidence areas (as opposed to the linear case for which (C.24) gives exact confidence areas).

References

1. Bai D.S., Chung S.W., "Optimal Design of Partially Accelerated Life Testing for the Exponential Distribution under Type-I Censoring", IEEE Transactions on Reliability, Vol. 41, No. 3, September 1992.

2. Barlow R.E., Toland R.H., Freeman T., "A Bayesian Analysis of the Stress-Rupture Life of Kevlar/Epoxy Spherical Pressure Vessels", in: "Accelerated Life Testing and Experts' Opinions in Reliability", Proceedings of the International School of Physics 'Enrico Fermi', Villa Marigola 1988, ed. by C.A. Clarotti and D.V. Lindley.

3. Barr D.R., Davidson T., "A Kolmogorov-Smirnov Test for Censored Samples", Technometrics, 15, p.739-757, 1973.

4. Baskin E.M., "Obrabotka rezul'tatov uskorennyh ispytanij pri nieizwestnoj funkcji racpredelenija otkazov izdelij", Izv. AN SSSR Techn. Kib. no. 3, 1988.

5. Baskin E.M., "O sglaživanii charakteristik nadežnosti polynomami Lagera", Izv. AN SSSR Techn. Kib. no. 5, 1971.

6. Baskin E.M., "Približenia zakonov nadežnosti obobščennymi polynomami Lagera", Izv. AN SSSR Techn. Kib. no. 5, 1973.

7. Basu A.P., Ebrahimi N., "Nonparametric Accelerated Life Testing", IEEE Transactions on Reliability, Vol. 31, No. 5, Dec. 1982.

8. Biernat J., Jarnicki J., Kapłon K., Kuraś A., Anders G., "Accelerated Life Testing of Electrical Insulation Material with Generalised Distribution Function", 3rd Int. Conference on Application of Probability Methods in Electric Power Systems, London, UK., July 1991.

9. Biernat J., Jarnicki J., Kapłon K., Kuraś A., Anders G., "Reliability Considerations in Accelerated Life Testing of Electrical Insulation with Generalised Life Distribution Function", IEEE Transactions Power Systems, Vol. 7, No.2, pp.656-664, May 1992.

10. **Biernat J., Jarnicki J., Maciejewski H., "Wyznaczanie charakterystyk niezawodności metodą transfomacji wyników prób forsownych", Zagadnienia Ekspoatacji Maszyn, Zeszyt 1 (101), 1995.**

11. **Biernat J., Jarnicki J., Maciejewski H., "Wyznaczanie niezawodności obiektów niejednorodnych w warunkach forsownych", Raport ICT PWr. Ser. SPR, 1995**

12. Biernat J., Kapłon K., Kuraś A., "Accelerated Life Testing of Electronic Devices", Proceedings of RELCOMEX 1989.

13. Bogdan Małgorzata, "Data Driven Version of Pearson's Chi-square Test for Uniformity", Institute of Mathematics, TU of Wrocław, 1995, to be published.

14. Bowless J.B., "A Survey of Reliability-Prediction Procedures For Microelectronic Devices", IEEE Transactions on Reliability, Vol. 41, No. 1, March1992.

15. Carey M.B., Koenig R.H., "Reliability Assesment Based on Accelerated Degradation: A Case Study", IEEE Transactions on Reliability, Vol. 40, No. 5, Dec 1991.

16. Chaloner K., Larntz K., "Bayesian design for accelerated life testing", Journal of Statistical Planning and Inference, 33, North-Holland, 1992.

17. Clark R.M., "Non-parametric Estimation of a Smooth Regression Function", J.Roy.Statist.Soc., B39, 107-113, 1977.

18. Clarotti C.A., Lindley D.V. (editors), "Accelerated Life Testing and Experts' Opinions in Reliability", Proceedings of the International School of Physics 'Enrico Fermi', Villa Marigola 1988, Elsevier, 1988.

19. Collett D., "Modelling Survival Data in Medical Research", Chapman & Hall, 1995.

20. Crowder M.J., Kimber A.C., Smith R.L., Sweeting T.J., "Statistical Analysis of Reliability Data", Chapman & Hall, 1991.

21. Dietrich, D.L.; Mazzuchi, T.A., "An alternative method of analyzing multi-stress multi-level life and accelerated life tests", Annual Reliability and Maintainability Symposium. 1996 Proceedings. The International Symposium on Product Quality and Integrity, 1996.

22. Drapella A., "Metoda statystycznej predykcji niezawodności elementów elektronicznych na podstawie wyników termicznych prób forsownych", Zeszyty Naukowe Politechniki Gdańskiej Nr 401, Gdańsk ,1986.

23. Draper N.R., Smith H., "Applied Regression Analysis", Wiley, 1973.

24. Dufour R., Maag U., "Distribution Results for Modified Kolmogorov-Smirnov Statistics for Truncated of Censored Samples", Technometrics, 20, p.29-32, 1978.

25. Dunn O.J., Clark V.A., "Applied Statistics: Analysis of Variance and Regression, Second Edition", Wiley 1987.

26. Dziubiński I., Świątkowski T., Poradnik matematyczny, PWN, 1978.

27. Dzerjinski S.M., Tzafestas S.G., "Accelerated Life Testing of Industrial Products y Step Stress", Systems Science Vol. 21, No.3, 1995.

28. Escobar, L.A.; Meeker, W.Q., "Planning accelerated life tests with two or more experimental factors", Technometrics Vol: 37 Iss: 4, 1995.

29. Feller W., "Wstęp do rachunku prawdopodobieństwa", PWN, 1978.

30. Fisz M., "Rachunek prawdopodobieństwa i statystyka matematyczna", PWN, Warszawa, 1965.

31. Fleming T.R., Harrington D.P., "Counting Processes and Survival Analysis", Wiley, 1991.

32. Fortuna Z., Macukow B., Wąsowski J., "Metody numeryczne", WNT, 1982.

33. Geist R., Trivedi K., "Reliability Estimation of Fault-Tolerant Systems: Tools and Techniques", IEEE Computer, July 1990.

34. Glaser R.E., "Weibull Accelerated Life Testing With Unreported Early Failures", IEEE Transactions on Reliability, Vol. 44, No. 1, March 1995.

35. Grey J., Siewiorek D.P., "High-Availability Computer Systems", IEEE Computer, September 1991.

36. Jankowscy J. i M., "Przegląd metod i algorytmów numerycznych", WNT, 1988.

37. **Jarnicki J., Maciejewski H., "Opracowanie komputerowej metody analizy wyników badań przyspieszonych", Raport ICT PWr. Ser. SPR, 1993.**

38. Jensen F., "Electronic Component Reliability, Fundamentals, Modelling, Evaluation, and Assurance", Wiley, 1995.

39. Kalbfleish J.D., Prentice R.L., "The Statistical Analysis of Failure Time Data", Wiley, 1980.

40. Kallenberg W.C.M., Ledwina T., "On Data Driven Neyman's Tests", Institute of Mathematics, TU of Wrocław, 1995, to be published.

41. Kapur R., Miller E.F., "System Test and Reliability Techniques for Avoiding Failiure", IEEE Computer, November 1996.

42. Koziol J.A., Byar D.P., "Percentage Points of the Asymptotic Distributions of One and Two Sample K-S Statistics for Truncated or Censored Data", Technometrics, 17, p.507-510, 1975.

43. Könemann B., Bennetts B., Jarwala N., Nadeau-Dostie B., "Built-In Self-Test: Assuring System Integrity", IEEE Computer, November 1996.

44. Lall P., "Tutorial: Temperature As an Input to Microelectronics-Reliability Models", IEEE Transactions on Reliability, Vol. 45, No. 1, March 1996.

45. Lawless J.F., "Statistical Models and Methods for Lifetime Data", Wiley, 1982.

46. Lu C.J., Meeker W.Q., "Using Degradation Measures to Estimate a Time-to-Failure Distribution", Technometrics, May 1993, Vol. 35, No.2.

47. **Maciejewski H., "Accelerated Life Test Data Analysis with Generalised Life Distribution Function and with no Aging Model Assumption", Microelectronics and Reliability, Vol. 35, No. 7, 1995.**

48. **Maciejewski H., "Inference from Accelerated Life Tests with Generalised Life Distribution Function and with Data Driven Selection of an Ageing Model.", to appear in the Proceedings of the ESREL '97 European Safety and Reliability Conference, Lisboa, 17-20 June 1997.**

49. **Maciejewski H., "Metoda analizy wyników badań przyspieszonych z uogólnioną funkcją rozkładu czasu życia i bez zakładania modelu starzenia", Materiały XXII Zimowej Szkoły Niezawodności, Szczyrk 1994.**

50. **Maciejewski H., "Wybrana metoda predykcji niezawodności w badaniach przyspieszonych", [praca dyplomowa], Politechnika Wrocławska, Inst. Cybernetyki Technicznej, Wrocław 1990.**

51. Mackisack M.S., Stillman R.H., "A Cautionary Tale About Weibull Analysis", IEEE Transactions on Reliability, Vol. 45, No. 2, June 1996.

52. Marsy E., "Probability Density Estimation from Sampled Data", IEEE Transactions on Information Theory, Vol. IT-29, No.5, September 1983.

53. Meeker W.Q., Escobar L.A., "A Review of Recent Research and Current Issues in Accelerated Testing", International Statistical Review, 61, 1, pp.147-168, 1993.

54. Meeker W.Q., Hamada M., "Statistical Tools for the Rapid Development & Evaluation of High-Reliability Products", IEEE Transactions on Reliability, Vol. 44, No. 2, June 1995.

55. Michael J.R., Schucany W.R., "A New Approach to Testing Goodness of fit for Censored Samples", Technometrics, 21, p.435-441, 1979.

56. MIL-HDBK-217F, "Reliability Prediction of Electronic Equipment", Department of Defence, Washington DC 20301, 1991.

57. Mitsubishi Semiconductors Reliability Handbook, Mitsubishi Electric Corp. 1992.

58. Murray B.T., Hayes J.P., "Testing ICs: Getting to the Core of the Problem", IEEE Computer, November 1996.

59. Müller H.G., "Nonparametric Regression Analysis of Longitudinal Data", Springer Verlag, 1988.

60. Nelson W., "Accelerated Testing - Statistical Models, Test Plans, and Data Analyses", Wiley, 1990.

61. Nelson W., "Applied Life Data Analysis", Wiley, 1982.

62. Nelson W., "Analysis of Accelerated Life Test Data - Least Squares Methods for the Inverse Power Law", IEEE Transactions on Reliability, Vol. R-24, No. 2, June 1975.

63. Padgett W.J., "Inference form Accelerated Life Tests", in "Reliability Theory and Models", ed. M.S. Abdel-Hameed et al., Academic Press 1984.

64. Pecht M.G., Nash F.R., "Predicting the Reliability of Electronic Equipment", Proceedings of the IEEE, Vol. 82, No. 7, July 1994.

65. Pettitt A.N., Stephens M.A., "Modified Cramer-con Mises statistic for censored data", Biometrica, 63, p.291-298, 1976.

66. Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., "Numerical Recipies in C", Cambridge University Press, 1992.

67. Proschan F., Singpurwalla N.D., "A New Approach To Inference From Accelerated Life Tests", IEEE Transactions on Reliability, Vol. R-29, No. 2, June 1980.

68. Shaked M., Singpurwalla N.D., "Nonparametric Estimation and Goodness-of-Fit Testing of Hypotheses for Distributions in Accelerated Life Testing", IEEE Transactions on Reliability, Vol. R-31, No. 1, April 1982.

69. Sherif Y.S., Kheir N.A., "Reliability and Failure Analyses of Computing Systems", Comput. & Elect. Engng., Vol. 11, No. 2/3, 1984.

70. Smith R.M., Bain L.J., "Correlation type goodness of fit tests with censored samples", Communications in Statistics, A5, p.119-132, 1976.

71. Stephens M.A., "EDF statistics for goodness of fit and some comparisons", Journal Am. Stat. Assoc., 69, p.730-737, 1974.

72. Tsang A., Jardine A., "Estimators of 2-Parameter Weibull Distributions from Incomplete Data with Residual Lifetimes", IEEE Transactions on Reliability, Vol. 42, No. 2, June 1993.

73. Tseng S.T., Hsu C.H., "Comparison of Type-I @ Type-II Accelerated Life Tests for Selecting the Most Reliable Product", IEEE Transactions on Reliability, Vol. 43, No. 3, September 1994.

74. Wawrzyńczyk A., Współczesna teoria funkcji specjalnych", PWN, 1978.

75. Yang G.B., "Optimum Constant-Stress Accelerated Life Test Plans", IEEE Transactions on Reliability, Vol. 43, No. 4, December 1994.

76. Yang C.R., Kim J.T., "Temperature Accelerated Life Test (ALT) at the Circiut Board Level", 17th IEEE/CPMT Int. Electronics Manufacturing Technologies Symposium, 1995.

Streszczenie

Praca dotyczy metod analizy wyników przyspieszonych badań niezawodności urządzeń komputerowych, charakteryzujących się bardzo wysoką niezawodnością i złożonością. Wyznaczenie niezawodności takich urządzeń na podstawie badań prowadzonych w nominalnych warunkach pracy jest w praktyce niemożliwe, gdyż wymagałoby długotrwałych testów oraz bardzo licznych próbek testowych. W celu skrócenia i zmniejszenia kosztów badań niezawodnościowych urządzenia testowane są w warunkach podwyższonego stresu, który skraca czas życia urządzeń. Na podstawie danych zebranych w wyniku badań przyspieszonych wyznaczana jest niezawodność urządzeń (np. rozkład czasu życia) w warunkach nominalnych. Analiza wyników badań przyspieszonych w celu wyznaczenia niezawodności w warunkach nominalnych jest trudnym zadaniem, realizowanym na ogół w oparciu o szereg modeli. Modele te opisują rozkład czasu życia badanych urządzeń na poszczególnych poziomach stresu oraz zależność pomiędzy poziomem stresu a czasem życia lub parametrami przyjętych rozkładów. Standardowe, znane z teorii niezawodności modele niezbyt dobrze sprawdzają się w przypadku bardzo złożonych (niejednorodnych) urządzeń komputerowych.

Celem pracy jest pokazanie, że możliwe jest wyznaczenie czasu życia takich urządzeń na podstawie wyników badań przyspieszonych, bez przyjmowania silnych założeń dotyczących modeli. Możliwa jest również weryfikacja niezawodności wyznaczonej dla warunków nominalnych jedynie w oparciu o dane otrzymane w testach przyspieszonych, bez konieczności posiadania danych eksperymentalnych dla warunków nominalnych.

Zaproponowana w pracy metoda analizy wyników badań przyspieszonych opiera się na założeniu, że rozkład czasu życia na każdym poziomie stresu może być opisany uogólnioną funkcją rozkładu, zbudowaną w oparciu o wielomiany Laguerre'a, przybliżającą znane z teorii niezawodności rozkłady czasu życia oraz złożenia tych rozkładów. Opracowane zostały metody dopasowania uogólnionego rozkładu (modelu) do danych pochodzących z badań pełnych lub obciętych (I i II rodzaju). Ekstrapolacja wyników badań z warunków testowych na warunki nominalne nie wymaga przyjmowania a priori określonego modelu starzenia (czyli zależności pomiędzy stresem a czasem życia lub parametrami rozkładu czasu życia). Opracowana została metoda doboru, w oparciu o dane eksperymentalne, najbardziej właściwego dla danego badania modelu starzenia, spośród zadanych modeli. Automatyczny dobór modelu starzenia, w oparciu o dane, zmniejsza ryzyko błędów ekstrapolacji wyników spowodowanych przyjęciem nieadekwatnego modelu. Opracowana również została metoda weryfikacji wyników predykcji niezawodności dla warunków nominalnych. Polega ona na wykonaniu ekstrapolacji dla warunków testowych w celu porównaniu wyników predykcji z danymi eksperymentalnymi. Brak dobrej zgodności wyników predykcji z danymi eksperymentalnymi (na tych poziomach stresu, dla których dostępne są takie dane), może świadczyć o niewiarygodności wyników dla warunków nominalnych. W pracy zawarto również wyniki predykcji niezawodności na podstawie przykładowych danych z badań przyspieszonych, mające służyć ilustracji zaproponowanych w pracy metod. Ilustrują one również praktyczną realizowalność tych metod, w oparciu o które może w przyszłości powstać zestaw narzędzi (programów komputerowych) do analizy rzeczywistych danych niezawodnościowych.

mgr inż. Henryk Maciejewski

Instytut Cybernetyki Technicznej
Politechniki Wrocławskiej
ul Janiszewskiego 11/17
50-372 Wrocław

Niniejszy raport otrzymują:

1. OINT	1 egz.
2. Biblioteka Główna PWr.	1 egz.
3. Z-ca Dyrektora Instytutu	1 egz.
4. Promotor	1 egz.
5. Recenzenci	2 egz.
6. Autor	2 egz.
Razem:	8 egz.

Raport wpłynął do Redakcji I-6
w czerwcu 1997 r.