Contributions to Accelerated Reliability Testing

HERBERT HOVE

Supervisor: Professor Frank Beichelt

School of Statistics and Actuarial Science

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa, in fulfilment of the requirements for the degree of Doctor of Philosophy

Johannesburg, December 2014

DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

____

(Signature of candidate)

____________________ day of ________________________________ 2014

Abstract

Industrial units cannot operate without failure forever. When the operation of a unit deviates from industrial standards, it is considered to have failed. The time from the moment a unit enters service until it fails is its lifetime. Within reliability and often in life data analysis in general, lifetime is the event of interest. For highly reliable units, accelerated life testing is required to obtain lifetime data quickly. Accelerated tests where failure is not instantaneous, but the end point of an underlying degradation process are considered. Failure during testing occurs when the performance of the unit falls to some specified threshold value such that the unit fails to meet industrial specifications though it has some residual functionality (degraded failure) or decreases to a critical failure level so that the unit cannot perform its function to any degree (critical failure). This problem formulation satisfies the random signs property, a notable competing risks formulation originally developed in maintenance studies but extended to accelerated testing here. Since degraded and critical failures are linked through the degradation process, the open problem of modeling dependent competing risks is discussed. A copula model is assumed and expert opinion is used to estimate the copula. Observed occurrences of degraded and critical failure times are interpreted as times when the degradation process first crosses failure thresholds and are therefore postulated to be distributed as inverse Gaussian. Based on the estimated copula, a use-level unit lifetime distribution is extrapolated from test data. Reliability metrics from the extrapolated use-level unit lifetime distribution are found to differ slightly with respect to different degrees of stochastic dependence between the risks. Consequently, a degree of dependence between the risks that is believed to be realistic to admit is considered an important factor when estimating the use-level unit lifetime distribution from test data.

Keywords: Lifetime; Accelerated testing; Competing risks; Copula; First passage time.

To my wife Judith and our two daughters Rejoice and Praise

Acknowledgements

I am indebted to my doctoral supervisor, Professor Frank Beichelt for suggesting this research area following a false start. Without his guidance, unquestionable support, timely feedback and the constructive meetings we frequently held throughout the four and half years, this journey would not have been possible. I feel privileged and favoured to have worked under the tutelage of such a renowned researcher.

Special mention also goes to my lovely wife Judith for being by my side throughout this journey, enduring the times she spent without me and having to look after our young kids alone while I was busy with research. I cannot thank you enough, but just to say you are a treasured gift from God and I am a lucky man to have you in my life.

Not forgetting the head of School of Statistics and Actuarial Science, Professor Peter Fridjhon for the way he managed my teaching load over the years, Professor Kevin Balkwill for financial assistance towards statistical software, my workmates Charles Chimedza, Honest Chipoyera, Anna Kaduma, Mark Dowdeswell for the informal but useful discussions on research related matters.

I am also deeply appreciative of the examiners, Professor Waltraud Kahle, Professor P. K. Kapur and Doctor Caston Sigauke for carefully going through the thesis. Their comments and suggestions are hugely appreciated.

Above all, I acknowledge God's unfailing love and abounding grace. Wisdom and understanding are His.

Presented papers from the thesis

(1) Hove, H. (2012), Contributions to accelerated reliability testing, *Proceedings of the* 4^{th} *International Conference on Accelerated Life Testing and Degradation Models*, June 4–6, INSA Rennes, France

(2) Hove, H. (2013), Accelerated life testing under dependent competing risks, *Proceedings of the* 8^{th} *International Conference on Mathematical Methods in Reliability: Theory, Methods and Applications*, pp. 135–139, July 1–4, Stellenbosch, South Africa.

(3) Hove, H. (2013), On the use of expert opinion to characterise the joint behaviour of competing risks in industrial accelerated life testing, In: Yeomans, J.S., Montemanni, R. and Norlander, T.E. (eds), *Lecture Notes in Management Science*, Vol. **5**, pp. 33–38, Tadbir, ISSN 2008-0050.

(4) Hove, H. (2014), Stochastic process model for competing risks in accelerated reliability testing, *Proceedings of the* 5^{th} *International Conference on Accelerated Life Testing and Degradation Models*, pp. 28–34, June 11–13, Pau, France.

The first paper contains initial ideas, and hence the title of the thesis. The materials in the second paper are taken from chapter 1 while the materials presented in the third and fourth papers are adopted from chapters 2 and 3 respectively.

Contents

List of Figures

List of Tables

List of symbols

α	Scale parameter
β	Shape parameter
$\beta(t)$	Shape function
ξ	Random variable
κ	Nominal life
θ	Copula dependence parameter
τ	Kendall's tau
ψ	Archimedean generator function
$\psi^{[-1]}$	Pseudo inverse of the Archimedean generator function
ϕ	Shape parameter of the inverse Gaussian distribution
$\Phi(\cdot)$	Standard form of the normal distribution function
$\Gamma(\cdot)$	Gamma function
μ	Distribution mean
σ	Distribution standard deviation
δ_i	Indicator variable
Σ	State space
$\lambda_j(\cdot)$	Cause-specific hazard function for the j^{th} failure mode
ν	Positive drift parameter
$\mathscr{L}$	Log likelihood function
$\hat{\mathscr{L}}_1, ..., \hat{\mathscr{L}}_j$	Maximum log likelihood values
$\{B(t), t \in \mathbb{R}^+\}$	Wiener process
$\{B^+(t), t \in \mathbb{R}^+\}$	Wiener maximum process

$\{B^{-}(t), t \in \mathbb{R}^{+}\}$	Wiener minimum process
$c(\cdot, \cdot)$	Copula density
$C(\cdot, \cdot)$	Copula function
$\{D(t), t \in \mathbb{R}^{+}\}$	Shifted gamma process
E	Activation energy
$f_j^*(\cdot)$	Subdensity function for the j^{th} failure mode
$F_j^*(\cdot)$	Subdistribution function for the j^{th} failure mode
$\{G(t), t \in \mathbb{R}^{+}\}$	gamma process
$h(\cdot, \cdot)$	Joint density function
$H(\cdot, \cdot)$	Joint distribution function
$I_j(\cdot)$	Cumulative incidence function for the j^{th} failure mode
k	Boltzmann's constant
L	Likelihood function
$M_{T_s}(\cdot)$	Moment generating function
$\{M(t), t \in \mathbb{R}^{+}\}$	Latent failure causing process
$\{M(t), R(t), t \in \mathbb{R}^{+}\}$	Bivariate stochastic process
p_c	Concordance probability
q	Probability of detecting a degraded failure in a life test
$\{R(t), t \in \mathbb{R}^{+}\}$	Performance degradation (marker) process
$S(\cdot)$	Survival function
$S(\cdot, \cdot)$	Joint survival function
$S_j^*(\cdot)$	Subsurvival function
T_s	First passage time to s of the degradation process
U_j	Uniformly distributed random variable on $[0, 1]$
$v_1, ..., v_m$	Test stress levels
$\{W(t), t \in \mathbb{R}^{+}\}$	Wiener process with drift
$\{W^{+}(t), t \in \mathbb{R}^{+}\}$	Maximum of a Wiener process with drift
$\{X(t), t \in \mathbb{R}^{+}\}$	Stochastic process
X_j	Life variables (competing risks)
Z	Minimum of the competing risk variables

Nomenclature

AD	Alert delay
ALT	Accelerated life testing
AFT	Accelerated failure time
AGAN	As good as new
AIC_c	Corrected Akaike information criterion
CDF	Cumulative distribution function
DT	Delay time
FON	Failure order number
HMP	Hidden Markov process
IPL	Inverse power law
LED	Light emitting diode
LK-Value	Log likelihood value
ln	Natural logarithm
log	Logarithm to base 10
MAP	Markov additive process
ML	Maximum likelihood
MLE	Maximum likelihood estimation
MON	Mean order number
MPI	Most powerful invariant
MRR	Median rank regression
OREDA	Offshore reliability data
PERT	Project evaluation and review technique

pdf	Probability density function
PIT	Probability integral transformation
RA	Repair alert
T_{LR}	Likelihood ratio test statistic
TS_{MLR}	maximum Likelihood ratio test statistic

Chapter 1

Introduction

Industrial units cannot be in service forever or at least they cannot remain in the same condition while in service forever. When the operation of the unit breaks down or a predefined change occurs in its mode of operation (deviation from industrial standards), the unit is generally considered to have failed. The time from the moment a unit enters service until it fails is its lifetime. To continuously improve the quality and reliability of a unit, an important aspect is that of assessing reliability information such as mean lifetime of the unit. This explains why lifetime is often the subject of interest in reliability and in all life data analysis in general.

Traditional sources of lifetime data in reliability include field tracking studies and warranty databases. The collected lifetime data are utilised to quantify the lifetime distribution of the unit. It is from this distribution that unit lifetime information regarding warranty periods, unit safety and the reliability specification of the unit are derived. Needless to say that bad estimation, particularly of lower percentiles of the unit's lifetime distribution may potentially result in huge losses to industry due to excessive warranty returns.

For industrial units with longer lifetimes, accelerated life testing (ALT) is required to expedite unit failure by stressing these highly reliable units beyond what they would normally experience when in actual use. The goal is to obtain lifetime data in a timely and cost effective manner. However, ALT poses the following problems. Firstly, a decision must be made on how to accelerate failure. For some units, failure modes are known in advance from physical/ chemical theory or experience with similar tests. To be valid however, ALT must only lead to those failure modes which may occur under normal use conditions and should not generate other new failure modes. If a new failure mode occurs, it must be identified and accounted for in the subsequent lifetime data analysis.

Secondly, a decision must also be made on how much to accelerate. Obviously, excessive stress will cause the unit to fail in an extremely short time span, but such failure time data may not provide useful information about the lifetime of the unit. Usually, the norm is to choose test stress levels that fall outside the product's specification limits but within its design limits as shown in Figure 1.1.

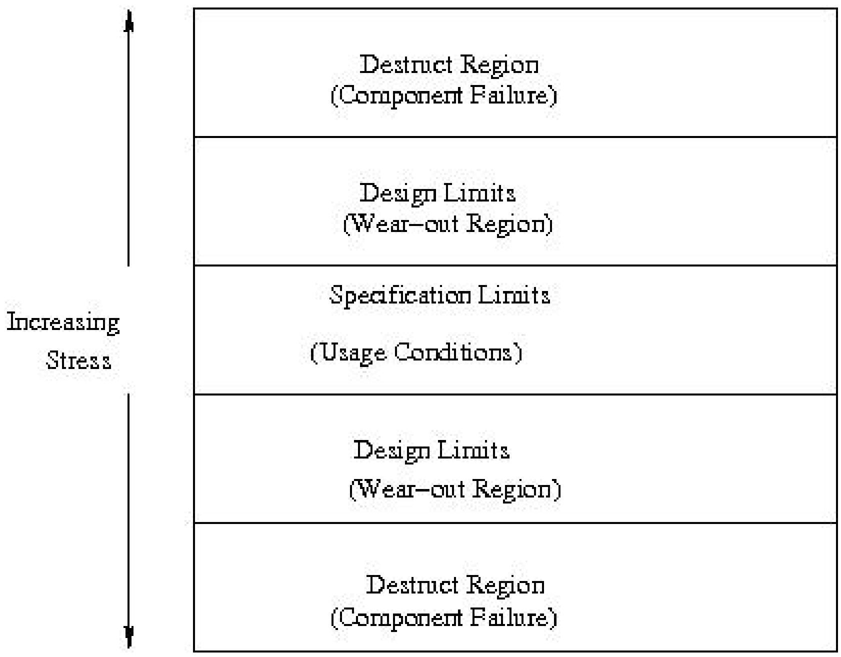


Figure 1.1: *General guide on how to choose test stress levels.*

Test stress levels that fall outside the unit's design limits but within its destruction limits will likely introduce new failure modes. Consider an example where temperature is the acceleration variable. Testing a unit at excessively high temperatures may result in the unit failing by melting.

1.1 Statement of the research problem

By design, ALT forces units to fail more quickly than they normally would. In this sense, it is generally considered to be a destructive practice. However, an important problem in reliability, see for example Padgett and Tomlinson (2004) is that of modeling degradation together with failure data in an accelerated testing setup. Research in this direction include Zhao and Elsayed (2004) who investigate the competing risks problem involving catastrophic and degradation failures in an accelerated life test. In their study, they assume catastrophic and degradation failure modes occur independently of each other. More recently, Pan, Zhou and Zhao (2010) consider the competing risks problem of accelerated failure that combines degradation failure mode with multiple independent traumatic failure modes where the latter depends on the degradation level.

Focus in this thesis is on accelerated tests where failure is not instantaneous, but the end point of an underlying degradation process. As a direct consequence of the general definition of failure, a failed unit does not necessarily imply it has reached the end of its time of potential use. Consider a light emitting diode (LED) for example. If it cannot perform its function to any degree during testing (complete loss of function), this definition of failure implies the end of its useful life. But if a LED is considered to have failed because its luminosity falls below an acceptable industrial standard during testing, then failure does not mean the end of its time of potential use.

Consequently the following situation is considered in this thesis: A unit is assumed to continuously degrade during testing so that the degradation path is the sample path of some stochastic process $\{X(t), t \geq 0\}$. Because ALT deals with new units, then $\{X(t), t \geq 0\}$ has initial condition $X(0) = 0$ and monotonically increases with time. It is also assumed that a dominant measurable performance parameter of the unit exists and that its deterioration over time can be associated with unit reliability. A unit is thus removed from observation during testing if its performance upon inspection:

(1) decreases to a specified threshold value such that the unit fails to meet specifications even though it still has some residual functionality. Such a unit is said to have experienced a degraded failure and consequently its lifetime is right censored. According to the Offshore Reliability Data (OREDA) database, degraded failures prevent the unit from performing

its functions according to the manufacturer's specifications and could develop to critical failures with time.

(2) decreases to a critical failure level so that the unit cannot perform its function to any degree. Such a unit is said to have experienced a critical failure due to critical degradation of its state and has reached the end of its lifetime. According to the OREDA database, critical failures mean immediate and complete loss of a major function such as the capability of a unit to provide its output.

Hence removal from observation in a life test may be from a mode other than the end of the unit's useful life. In particular, life tests where the lifetime of a unit is subject to right censorship are considered in this thesis. It is therefore assumed throughout that two failure modes namely critical failure and degraded failure are distinguished at each stress level. Examples of units that exhibit these failure modes in reliability testing include semiconductors, mechanical systems and microelectronic units where soft or non-catastrophic failures occur in life tests.

A degraded failure is thus a signal that a critical failure is likely to follow soon if the unit is kept on test. The object of interest in this thesis and often in life testing studies is the lifetime of the unit, and hence the occurrence of a critical failure. When detected first, a degraded failure leads to removal from observation during testing since by definition the unit no longer meets specified industrial standards and is therefore considered failed. In this sense, a degraded failure has the interpretation of a censoring variable since it has the effect of censoring a critical failure, the outcome of interest. Consequently the lifetime of a test unit is subject to right censorship with censoring occurring whenever a degraded failure removes the unit from observation in a life test.

1.1.1 Competing risk application to accelerated reliability testing

Denote by $q \in [0, 1]$, the probability of detecting a degraded failure when the performance parameter of the unit decreases to a threshold s_1. But detection or non-detection thereof depends only on the alertness of the crew running the life testing experiment assuming there is no automatic monitoring. It is therefore plausible to assume detection of a degraded failure during testing is independent of the degradation process. If a degraded failure is not detected, the unit is kept on test until the performance parameter decreases to a critical level $s_2 > s_1$ and the unit experiences a critical failure. This happens with probability $1 - q$. Lindqvist and Skogsrud (2009) give a related application but in maintenance studies where a potential unit failure may be avoided by a preventive maintenance.

Most studies in degradation modeling consider unit degradation an observable process and use measured degradation data to assess the lifetime of the unit. This modeling viewpoint appears in the work on degradation modeling by Doksum (1991), Lu and Meeker (1993), Lu, Meeker and Escober (1996) and the numerous citations therein. In this thesis however, it is assumed, and is often the case in practice that degradation paths of test units cannot be monitored continuously during testing. Accordingly, the degradation process leading to unit failure is not fully observable. But by definition, a unit fails the first time the degradation process crosses a failure level. Consequently, unit lifetime is estimated by obtaining the first crossing time of the degradation process over a failure threshold in a life test.

Thus given unit lifetime is censored, the time at which this materialises is the first passage time with regards to level s_1 of the degradation process, denoted by X_1. Otherwise the unit experiences a critical failure at X_2, the first passage time with regards to level $s_2 > s_1$ of the degradation process. Consequently X_1 depends on the degradation process and may also depend on unit lifetime X_2. The random time at which a unit is removed from observation at each test stress level is therefore the minimum of the censoring variable X_1 (the time unit lifetime would be censored if it were not ended first) and the lifetime variable X_2 (the lifetime of the unit if it were not censored). It is denoted by $Z = min(X_1, X_2)$.

This problem formulation satisfies the random signs property due to Cooke (1996) which is captured in the definition that follow:

Definition 1.1.1: *Let X_1 and X_2 be the censoring and lifetime variables respectively with $X_1 = X_2 - \xi$ where $\xi \leq X_2$ is a random variable whose sign does not depend on X_2 and satisfies $P(\xi = 0) = 0$. Then the observed variable $Z = min(X_1, X_2)$ and identification of the variable which achieves the minimum is referred to as the random signs censoring of X_2 by X_1.*

The relationship $X_1 = X_2 - \xi$ implies that if $\xi > 0$, then $X_1 < X_2$ and unit lifetime is censored. But if a degraded failure is not detected and the unit is kept on test until it experiences a critical failure at X_2, then X_1 is not observed. Thus the unobserved X_1 may in theory be assigned any time greater than X_2 which is however never observed. This corresponds to the case where $\xi < 0$ such that $X_2 < X_1$ and hence, the unit reaches the end of its useful life during testing. Thus the random sign of ξ determines whether unit lifetime is censored or not. Hence the name random signs censoring.

Random signs censoring applies to situations where unit lifetime is subject to right censorship, which could either be dependent or independent. It is a well established competing risks model developed originally in maintenance studies (see also Cooke and Bedford, 2002; Bunea and Bedford, 2002; Lindqvist and Skogsrud, 2009) but extended to reliability testing in this thesis. In the sense of the random signs censoring model due to Cooke (1996), degraded and critical failures are competing to remove the unit from observation in a life test. As a result, the observable competing risks data at each stress level are $Z = min(X_1, X_2)$ along with the identity of the mode $J = j \in (1, 2)$ which succeeded in removing the unit from observation in a life test.

Assuming n_k units from the same population are tested at the k^{th} test stress level, then Z_i, $i = 1, 2, ...$ are independent copies of Z. In practice however, only a few and often prototype units are available for testing due to cost constraints. In order to obtain sufficient failure time data, the failed unit is repaired and tested continuously. Since ALT deals with new units, any repair action following a degraded or a critical failure must leave the unit in a state as good as new (AGAN) as depicted in Figure 1.2. That is, the effective age of the unit must be reduced to zero after a repair, called perfect repair. For more details on the scope of repair actions, see for example Barlow and Proschan (1975). Obviously, repairing a degraded failure implies that a censoring occurs with respect to the corresponding lifetime variable.

Perfect repair is a plausible assumption in the case of complex, highly reliable repairable units such as in nuclear industry, space, undersea and in electronics where failure costs are prohibitive. Under this perfect repair assumption, $Z_i, i = 1, 2, ...$ are regeneration points. That is the sequence of nonnegative, independent and identically distributed (iid) random variables $\{Z_i;\ \ i = 1, 2, ...\}$ defines an ordinary renewal process. Thus whether n_k units from the same population are tested

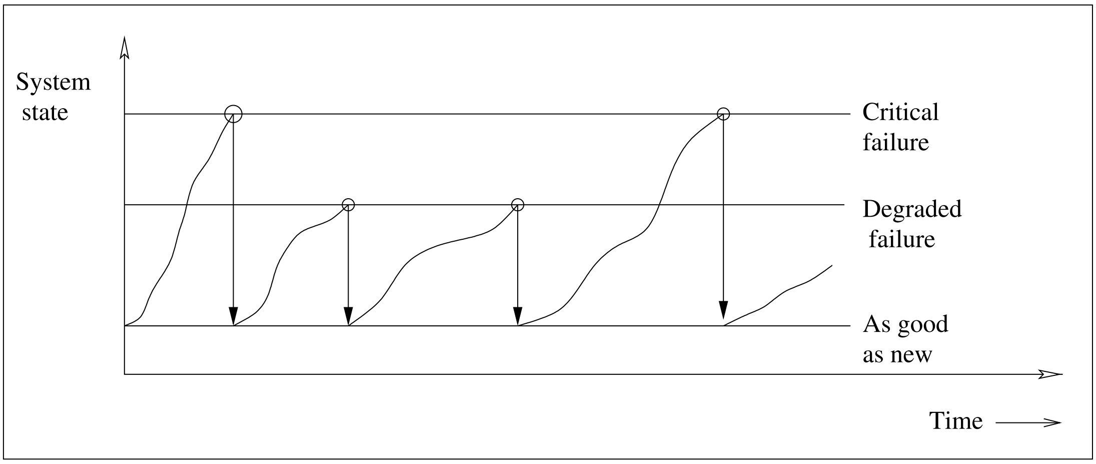


Figure 1.2: *Degraded and critical failures in a life test under perfect repair*

at the k^{th} test stress level or sufficient failure time data are obtained by repairing the failed unit and testing continuously, Z_i, $i = 1, 2, ...$ are iid random variables.

1.2 Statistical approaches to analysing competing risks data

When mode and failure time information are available, competing risks theory provides the appropriate model for analysing failure time data. The following are typical statistical problems that arise when analysing competing risks data (see for example Prentice *et al.*, 1978).

(1) Inference on the effects of covariates on specific failure modes.

(2) Studying interrelations among failure modes and the effect of removing a specific failure modes on remaining failure modes.

Different approaches to analysing competing risk data exist but none addresses all problems. Theoretical approaches assign latent or hypothetical failure times $X_1, ..., X_d$; $0 \leq X_j < \infty$ representing unit failure times from the corresponding d competing failure modes. Hence the name latent failure time approach. The observable data are the pair (Z, J) where $Z = min(X_1, ..., X_d)$ is formally lifetime of a series system. Interrelations among failure modes and the effects of removing a failure mode on the remaining mode(s) are formulated in terms of the $X_j's$ and the absolutely continuous joint survival function

$$S(t_1, ..., t_d) = P\left[\bigcap_{j=1}^{d} (X_j > t_j)\right] \tag{1.1}$$

satisfying $S(0, ..., 0) = 1$ and $S(\infty, ..., \infty) = 0$ where each $t_j \geq 0$. Study questions in reliability are often formulated with regards to the marginal survival (or distribution) functions of the multiple decrement function in Equation 1.1 since they best reflect the underlying failure process. A typical example of a study question is:

What would be the effect of eliminating a failure mode on the reliability specification of the unit?

Other approaches base all statistical analysis on estimable quantities from the observable competing risks data only. These quantities of statistical interest include the cause-specific hazard functions and cumulative incidence functions (Lawless, 2003). The cause-specific hazard function of the j^{th} failure mode, defined as the probability that mode j is responsible for removing a unit from observation in a life test in the small interval $(t, t + \Delta t)$ given that the unit survived all failure modes until time t is

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(t < X_j \leq t + \Delta t, J = j | T > t)}{\Delta t} = \frac{f_j^*(t)}{S(t)} \tag{1.2}$$

if the density exists. Thus $\lambda_j(t)$ represents the hazard of failing from mode j when all other failure modes are acting. The overall survival function, denoted by $S(t) = P\,(T > t)$ is the probability that a unit has not failed from any mode at time t.

The cumulative incidence function of the j^{th} failure mode, denoted $I_j(t)$ is the probability of a unit failing from the j^{th} mode when all other failure modes are present. Assuming independent right censorship, it is easily expressed as

$$I_j(t) = P(T \leq t, J = j) = \int_0^t \lambda_j(t) \text{exp}\left(-\sum_{j=1}^{d} \int_0^u \lambda_j(v) dv\right) du. \tag{1.3}$$

where $\lambda_j(t)$ is the cause-specific hazard function.

Under dependent right censorship however, the expression for the cumulative incidence function of the j^{th} failure mode is less straightforward. Obviously, the function in Equation 1.3 is not a proper distribution function in the sense that $I_j(\infty) = P(J = j)$, and not one as expected. Hence it is often called the *subdistribution* or *crude cumulative incidence* function. Accordingly, the function $f_j^*(t)$ is also called the subdensity function. Cause-specific hazard and cumulative incidence functions are useful when study questions are on the effects of covariates on specific failure modes.

In this thesis and often in accelerated reliability testing, the variable of interest is unit lifetime. Consequently, the object of estimation is the lifetime distribution of the unit. When unit lifetime is subject to right censorship as is the case here, this problem is clearly formulated with regards to the marginal distribution of unit lifetime X_2 from the observable competing risks data (Z, J). The problem is therefore best answered by adopting the latent failure time approach because it suggests a simple way to answer study questions on failure mode removal.

1.2.1 Latent failure time model and right censorship

Consider a unit that fails from mode j in a life test. When all modes are operative, this leads to observing $Z = min(X_1, ..., X_d)$ and $J = \{j | X_j < X_k, \ \ k = 1, ..., d\}$. This latent failure time interpretation is followed in Gail (1975). Obviously the random variable X_j has a clear physical meaning. It is the observed time to failure of the unit from mode j. The $X_k{}'s$ on the one hand are unobserved and generally do not have physical meaning attached to them. Thus the latent failure time approach does seem to present problems of data interpretation and this is at the core of its criticism particularly in the Biostatistical literature for example (see e.g. Prentice *et al.*, 1978).

But for the competing risks problem envisaged in this thesis (unit lifetime subject to right censorship), latent failure times following the first do have physical meaning as follows. A unit is removed from observation during life testing when it experiences either a degraded or a critical failure and hence $d = 2$. Assume the unit experiences a critical failure in a life test when both degraded and critical failure modes are operative. Then the random variable X_2 is the observed lifetime of the unit being tested. The unobserved time X_1 is the time the lifetime of the unit would have been censored had the unit not experienced a critical failure (complete loss of function) first. Because a clear physical meaning is attached to latent failure times following the first, the latent failure time model is applicable to the dependent competing risks formulation considered in this thesis.

Obviously, a unit that is at risk of experiencing a critical failure in a life test is systematically at high (low) risk of experiencing a degraded failure since the two failure modes are linked through the degradation process. Consequently, all study questions associated with interrelations between these two failure modes are identified with interrelations between their corresponding failure times. Specifically, such questions are posed in terms of the observed $X_j's$ and the joint survival function

$$S(t_1, t_2) = P\left(X_1 > t_1, X_2 > t_2\right). \tag{1.4}$$

Recall that unit lifetime is the variable of interest in life testing studies. Hence the main object of estimation in this thesis is the lifetime distribution of the unit with the censoring variable removed. In general, measures that remove a specific failure mode may well alter the failure modes remaining in the study. But random censorship (competing risk) arises from a mechanism external to the underlying failure process since it depends on the alertness of the persons conducting the life test. Accordingly, the competing risks situation considered in this thesis guarantees a failure mechanism that is not influenced by censorship. Hence unit lifetime is unaffected by the removal of the censoring variable and the lifetime distribution is the marginal distribution of X_2 from the joint survival function in Equation 1.4.

1.2.2 Identifiability problems of the latent failure time model

A well documented problem with adopting the latent failure time model (Tsiatis, 1975) is that both the joint and the marginal survival (or distribution) functions of the competing risks variables are in general non-identifiable. This difficulty arises because competing risks cannot be observed directly. Rather, only the pair (Z, J) is observable and such data only allow estimation of subsurvival functions $S_j^*(t) = P\left(X_j > t, X_j < X_k\right)$ for $j \neq k$ and not the survival functions $S_j(t) = P\left(X_j > t\right)$; $j \in (1, 2)$. If however there is no (physical) reason to suggest stochastic dependence the risk variables, then there are often no distribution identifiability issues. This is because subsurvival and the survival functions will coincide and the analysis is just as difficulty as analysing a single failure mode. Consequently the marginal survival (or distribution) functions can be consistently estimated from observations on the pair (Z, J) and are therefore identifiable.

But degraded and critical failures are linked through the underlying failure causing degradation process. Specifically, a degraded failure is reasonably assumed to occur close to a critical failure. For example if a LED is considered failed because the luminosity falls below an acceptable industrial standard in a life test, then its lifetime is likely to end soon if it is kept on test. As a result, there is a physical reason to suggest that unit lifetime is subject to dependent right censorship. In this case, the set of subsurvival functions $S_j^*(t)$ is consistent with a number of joint survival functions $S(t_1, t_2)$ and is therefore generally not identifiable as follows:

A key result in Tsiatis (1975) is suggestive of the following. If the set $S_j^*(t)$ is given for some joint model where unit lifetime is subject to dependent right censorship, then there also exists a joint model where unit lifetime is subject to independent right censorship yielding the same set $S_j^*(t)$. As a direct consequence of this result, the correct model remains unknown from observations of (Z, J) alone since both the independent and dependent risks models may fit the data equally well. Thus over and above the uncertainty that is a result of sampling error, there is also an added problem of model uncertainty, called the *identifiability problem* of competing risks.

Assuming unit lifetime is subject to independent right censorship, the lifetime distribution is i-dentifiable. This explains why this simplifying assumption is often made in practice. In the present application however, unit lifetime is subject to dependent right censorship and the problem is to estimate the lifetime distribution of the unit from test data. Accordingly, the analysis of data of the form (Z, J) inevitably has to rely on unverifiable parametric restrictions about the probability structure of the joint survival function $S(t_1, t_2)$. Otherwise at best, only some bounds on the lifetime distribution of the unit at all stress levels will be obtained (Crowder, 2001).

1.3 Novelty of the thesis

Literature on ALT and competing risks is vast. Needless to say ALT often generates multiple failure modes, but a considerable research gap still remains for problems involving both. The predominant assumption made by the few studies on competing risks in ALT is that the risks act independently to simplify the analysis. But the competing risks problem which is yet to be fully resolved and is therefore an open problem is to resolve issues of distribution identifiability and failure dependence. This thesis seeks to contribute to the debate in the context of accelerated reliability testing as follows:

(i) The question of assessing reliability information of an industrial unit is formulated in terms of the marginal probability distribution of the unit's lifetime. For a highly reliable unit, ALT is required to obtain lifetime information in a timely manner. When unit lifetime is subject to dependent right censorship during testing, the lifetime distribution is in general unidentifiable at all stress levels. The theoretical latent failure time model has been postulated as the natural modeling framework for identifying the lifetime distribution of the unit at all stress levels with the censoring variable (a competing risk) removed.

(ii) The difficult and non trivial problem of modeling dependence between unit lifetime and the censoring variable in a competing risks framework is discussed. The thesis utilises parametrisation of families of copula by rank correlation (Kendall's tau) to estimate the copula dependence parameter using expert opinion by means of a simulation study. Copula model estimation is important because it enables the lifetime distribution of the unit to be uniquely determined from observations of the competing risks data at all stress levels (Zheng and Klein, 1995). Assuming true acceleration (scale transformation only), stochastic dependence between the competing risks at all stress levels is captured by the estimated copula model.

(iii) Functional forms of the observed occurrences of unit lifetime and the censoring variable are derived from a stochastic process point of view. Because this investigation assumes that the degradation path of a unit cannot be monitored continuously during testing, the underlying failure causing process is therefore not fully observable. What is observed during testing is the performance degradation process acting as a marker (or surrogate) process. This motivates the adoption of the framework of hidden Markov processes (HMP) when assessing unit lifetime during testing. Under this HMP modeling framework, the lifetime of a test unit is estimated by the first passage time of the underlying failure causing process over a deterministic threshold. This modeling approach differs from the common degradation modeling viewpoint which considers unit degradation an observable process and uses measured degradation data to assess the lifetime of a test unit.

(iv) For accelerated reliability tests yielding unit lifetimes only (single failure mode) or if unit lifetime is subject to independent right censoring, the derived functional forms are estimates of the lifetime distribution of the unit at all test stress levels. Consequently, the lifetime distribution of the unit at normal operating conditions can be extrapolated from test data. When unit lifetime is subject to dependent right censoring, the derived functional forms are subsurvival (or subdistribution) functions. Their derivatives (subdensity functions) together with the estimated copula model and test data identify the marginal distribution functions at each test stress level. Accordingly, the lifetime distribution of the unit at normal operating conditions can be extrapolated from test data under dependent random censorship.

1.4 Format of the thesis

The remainder of the thesis is organised as follows. Chapter 2 focuses on modeling the stochastic dependence between unit lifetime and the censoring variable during testing. The derivation of functional forms of the observed occurrences of unit lifetime and the censoring variable in a competing risks framework is the subject of Chapter 3. Chapter 4 discusses the statistical methodology for extrapolating the use-level lifetime distribution of the unit from test data while the results of the investigation are presented and discussed in Chapter 5. Finally, Chapter 6 concludes the study and identifies areas for further research.

Chapter 2

Dependence model for unit lifetime and random censorship in life testing

2.1 Introduction

This chapter discusses statistical models that capture the stochastic dependence between unit lifetime and the censoring variable in a life test. Dependence arises because degraded and critical failures are linked through the degradation process of the unit. Thus a basic dependence structure is assumed between the censoring variable X_1 and unit lifetime X_2 at each stress level. In terms of Cooke's random signs censoring model, these two failure modes are competing for the removal of a unit from observation in a life test. Consequently, the problem under consideration here is that of modeling dependent competing risks.

An obvious approach to studying interrelation between competing risks is to place parametric restrictions on the joint survival function $S(x_1, x_2)$ in order to study interrelations more generally. Within such parametric models, parameters that describe possible dependencies between the censoring variable X_1 and unit lifetime X_2 may be estimated. Crucially however, there must be external evidence to justify the assumed parametric model since dependence arises from a model assumption that cannot be tested by the competing risk data alone. Otherwise a non-zero value of the estimated dependence parameter within such parametric models is not necessarily an indication of the stochastic dependence between the competing risk variables (Hove, 2013).

A well known difficulty with placing parametric restrictions on $S(x_1, x_2)$ is that the observable competing risks data (Z, J) do not allow one to distinguish between the assumed model and one with independent risks. Hence in addition to the uncertainty due to sampling error, there is also the extra problem of model uncertainty (Hove, 2013). Besides, utilising classical families of multivariate life distributions to describe stochastic dependencies among competing risks variables is in general restrictive. This is because the same parametric family of univariate distributions must characterise the marginal behavior of the variables. By not allowing for different marginal distributions, well known multivariate life distributions cannot describe different dependence structures.

Other approaches (see e.g. David and Moeschberger, 1978; Meeker, Escobar and Hong, 2009) collapse several related failure modes into fewer groups which are presumably approximately independent and hence identifiable. Alternative approaches include mixtures, specifications of conditional distributions and copulas. Admittedly however, stochastic models are often used for specific purposes and clearly no single concept addresses all problems of stochastic dependence. Hence the choice of a particular approach must be guided by a clear definition of the notion of dependence structure being modeled.

Dependence between the censoring variable X_1 and unit lifetime X_2 is assumed since they are linked through the degradation process of the unit during testing. Recall that a degraded failure occurs when the dominant measurable performance parameter falls below an acceptable level in a life test. When detected, a degraded failure is a signal that the useful life of the unit is likely to end soon if it is kept on test. In this context, stochastic dependence between the censoring variable X_1 and unit lifetime X_2 is the degree to which the occurrence of high (low) values of the one risk variable impacts on the probability of occurrence of values of the other risk variable. This notion of the dependence structure is a matter of relative ranks and is thus completely based on copulas. For other extensively studied notions of multivariate dependence, see Renyi (1959) and Zografos (2000) for example.

2.2 Copulas and stochastic dependence modeling

The copula argument stems from the cumulative distribution function (CDF) transformation or the probability integral transformation (PIT) as follows: Given any arbitrary continuous random variables X_j with invertible CDFs F_j for $j = 1, 2$, the relation

$$X_j = F_j^{-1}(U_j) \Leftrightarrow U_j = F_j(X_j) \tag{2.1}$$

holds. The resulting random variables U_j are uniformly distributed on the interval $[0, 1]$ and correspond to the respective ranks of the distribution. The relation in Equation 2.1 is the basis for the sampling of random variables in Monte Carlo simulation studies.

A remark is however required here. The transformation of the the marginal CDFs of a joint CDF to a standard uniform distribution is not motivated by any mathematical reason. It is often useful in statistical modeling, especially when the resulting distribution has a simpler and easily accessible representation when calculating probabilistic quantities. In multivariate extreme value theory for example, transforming to a standard distribution is standard practice. In general however, several transformations are possible and deciding on which transformation to use often depends on the context. The copula representation (Qu, Zhou and Shen, 2010) standardises to a uniform distribution function on $[0, 1]$.

In the bivariate case (higher order extension is straightforward), the copula representation of a joint distribution function $H(\cdot, \cdot)$ with marginal distribution functions F_j is given by

$$H(x_1, x_2) = P(X_1 \le x_1, X_2 \le x_2) = P(U_1 \le F_1(x_1), U_2 \le F_2(x_2)) = C(F_1(x_1), F_2(x_2)).$$

where the $2-$dimensional df $C(\cdot, \cdot)$ is the copula of the vector (X_1, X_2). Its univariate marginal distributions are uniformly distributed on $[0, 1]$ and hence the distribution function $C(\cdot, \cdot)$ has support $[0, 1]^2$. Consequently, a bivariate copula is a bivariate distribution function $C(\cdot, \cdot)$ defined on the unit square with univariate marginal distributions transformed to uniform. The basis of the copula approach in statistical modeling (Genest and Favre, 2007) is established firmly in Sklar (1973)'s representation theorem. A version of the theorem is:

Theorem 2.1 : Sklar's Theorem. *Given a joint distribution function $H(x_1, x_2)$ for random variables X_1 and X_2 with marginal distribution functions $F_1(x_1)$ and $F_2(x_2)$, then there exists a copula $C(\cdot, \cdot)$ for X_1 and X_2 such that*

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{2.2}$$

If F_1 and F_2 are continuous, then $C(\cdot, \cdot)$ is unique; otherwise $C(\cdot, \cdot)$ is uniquely determined in $RanF_1 \times RanF_2$, where $RanF_j$ is the range of F_j.

Thus by Sklar's theorem, there exists a copula $C(\cdot, \cdot)$ for any joint distribution function $H(\cdot, \cdot)$ that completely captures the stochastic dependence of the random variables. Because one can express any joint distribution in copula form, the theorem is therefore completely general. If the marginal distribution functions F_j are continuous and strictly increasing, then they have unique ordinary inverse functions F_j^{-1} and Theorem 2.1 implies the following corollary.

Corollary 2.1: *Let $H(\cdot, \cdot)$ be a $2-dimensional$ distribution function with continuous univariate distribution functions F_j, $j = 1, 2$ and copula $C(\cdot, \cdot)$ satisfying equation 2.2. Then for any $(u_1, u_2)^T \in [0, 1]^2$, there holds the relationship*

$$C(u_1, u_2) = H(F_1^{-1}(u_1), F_2^{-1}(u_2)).$$

Assume the marginal distribution functions F_j are not strictly increasing and are constant on some interval $[x_{j1}, x_{j2}]$. Then any t such that $x_{j1} \leq t \leq x_{j2}$ satisfies $F_j(t) = u_j$. To ensure that F_j^{-1} is single valued,

$$F_j^{-1}(u_j) = inf\{t : F_j(t) \geq u_j\}; \;\; 0 \leq u_j \leq 1$$

defines the *quasi (or pseudo) inverse* of distribution functions F_j. Thus copulas transform the random vector (X_1, X_2) into another random vector $(U_1, U_2) = (F_1(X_1), F_2(X_2))$ with marginal distribution functions uniform on $[0, 1]$ but maintains the dependence structure of the original variables. It follows from Equation 2.2 that multivariate distributions with different dependence structures are obtained by combining any copula with flexible univariate distribution functions. This is the main advantage of the copula modeling approach over classical families of multivariate distribution when studying dependence properties of random variables.

If the marginal distributions F_j and the copula $C(\cdot, \cdot)$ are continuous and differentiable, then by Sklar's theorem the canonical representation

$$h(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \prod_j^2 f_j(x_j) \tag{2.3}$$

exists where $c(u_1, u_2) = \frac{\partial^2 C(u_1,u_2)}{\partial u_1 \partial u_2}$ is the copula density reflecting the strength of stochastic dependence of the two random variables and $\prod_j^2 f_j(x_j)$ is the joint density under the independence assumption. It follows from Equation 2.3 that it is the copula that captures the stochastic dependence of the random variables X_1 and X_2. This explains why the copula is also called the *dependence function*.

To better comprehend the copula approach to stochastic dependence modeling, consider the fitting of a multivariate distribution in classical statistics. Typically, this entails using maximum likelihood to extract information out of the data about the chosen multivariate parametric family of distributions. But by choosing a multivariate parametric family of distributions, one determines a specific dependence structure. If for example a multivariate Gaussian distribution is chosen, its dependence structure is completely characterised by the covariance matrix.

The copula method transforms the random variables to a common uniform $[0, 1]$ domain and dependence modeling occurs in this common domain. Obviously the dependence structure of two uniformly distributed random variables is clearly unidentifiable since there are infinitely many such dependencies. Just as in classical statistics however, copula based methods ensure a unique dependence structure as follows. The copula, and hence the chosen dependence structure is determined by simply choosing the marginal distributions.

2.3 Families of commonly used multivariate copulas

Different families of copulas exist in the literature and within each are a number of copulas which may be useful when constructing stochastic models with different dependence structures. For families of copulas to be considered useful in statistical applications (see e.g. Durante and Sempi 2010), they need to possess some probabilistic interpretation which suggests situations where the family could be naturally applicable. Within a given family, there must also exist a variety of copulas that describe a wide range of dependence. Above all, members of a parametric family need to have closed form expressions for easy simulation and hence goodness-of-fit tests. This section discusses copula families with extensive applications in the literature.

2.3.1 Copulas of elliptical distributions

Elliptical copulas are obtained from elliptical distributions by a direct application of Sklar's theorem. Examples of commonly used elliptical distributions are the multivariate normal and the multivariate Student-t distributions. A $d-$dimensional random vector $\mathbf{X} = (X_1, X_2, ..., X_d)$ possesses an elliptical distribution with a deterministic mean vector $\mu \in \mathbb{R}^d$, a positive definite covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ and a characteristic generator of the distribution $g : [0, \infty) \rightarrow (-\infty, \infty)$, denoted $\mathbf{X} \sim E_d(\mu, \Sigma, g)$ if it can be expressed as

$$\mathbf{X} \stackrel{d}{=} \mu + R\mathbf{A}\mathbf{U} \tag{2.4}$$

assuming the matrix $\mathbf{A}$ exists where $\stackrel{d}{=}$ stands for equality in distribution. The remaining components of Equation 2.4 are defined as follows. If it exists, the matrix $\mathbf{A} \in \mathbb{R}^{d \times k}$ is such that $\mathbf{A^T A} = \Sigma$ is the rank factorisation of Σ, the *d-dimensional* random vector $\mathbf{U}$ is uniformly distributed on the sphere

$$\mathbb{S}^{d-1} = \left\{ \mathbf{u} \in \mathbb{R}^d : u_1^2 + ... + u_d^2 = 1 \right\}$$

and R is a non-negative random variable having density

$$f_g(r) = \frac{2\pi^{d/2}}{\Gamma(d/2)} r^{d-1} g(r^2), \quad r > 0.$$

It is stochastically independent of $\mathbf{U}$. For more details on elliptical distributions, see for example Durante and Sempi (2010) and the numerous references therein. When the characteristic generator of the distribution is

$$g(s) = (2\pi)^{-\frac{d}{2}} exp\left(-\frac{s}{2}\right),$$

then the $d-$dimensional random vector $\mathbf{X} = (X_1, X_2, ..., X_d)$ is distributed as multivariate Gaussian. In the same way, when it is

$$g(s) = m\left(\frac{v+s}{v}\right)^{-\frac{d+v}{2}}$$

where the constant m is carefully chosen, then $\mathbf{X}$ has a multivariate Student's-*t* distribution with v degrees of freedom. Elliptical copulas are obtained from their respective multivariate dfs by a simple application of Sklar's Theorem. Thus the multivariate Gaussian copula is represented as

$$C_\theta(u_1, ..., u_d) = \Phi_\theta\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right)$$

where $\theta \in [-1, 1]$ is the copula dependence parameter and Φ^{-1} is the inverse of the univariate Gaussian distribution function. In the same way, denote by $\theta = \left\{(v, \Sigma) : v \in (1, \infty), \Sigma \in \mathbb{R}^{d\times d}\right\}$ the dependence parameter for the multivariate Student's-*t* copula. Also let t_v denote the univariate *t* distribution with v degrees of freedom. Then by Sklar's Theorem, the multivariate Student's-*t* copula is expressed as

$$C_\theta(u_1, ..., u_d) = \mathbf{t}_{v,\Sigma}\left(t_v^{-1}(u_1), ..., t_v^{-1}(u_d)\right)$$

where $\mathbf{t}_{v,\Sigma}$ is the multivariate Student's *t* distribution with v degrees of freedom and correlation matrix Σ. In general, the multivariate Gaussian distribution arises through the multivariate Central Limit Theorem and is thus natural. The multivariate Student's-*t* distribution is also known (see e.g. Fang, Kotz and Ng, 1990) to fall into the multivariate normal variance mixtures class with representation

$$\mathbf{X} \stackrel{d}{=} \mu + \sqrt{W}\mathbf{Z}$$

where $\mathbf{Z} \sim N_d(\mathbf{0}, \Sigma)$, $\mu \in \mathbb{R}^d$ and the random variable W is independent of $\mathbf{Z}$. Consequently, the multivariate Student's-*t* distribution is also natural. But since elliptical copulas are derived from their respective multivariate dfs by simply applying Sklar's Theorem, then the multivariate *t* and Gaussian copulas do not necessarily suggest natural situations where they could be applicable.

2.3.2 Extreme value and Marshall - Olkin copulas

Denote by $\mathbf{X}_i = (X_{i,1}, ..., X_{i,d})$ independent copies of the random vector $\mathbf{X} = (X_1, ..., X_d)$ with joint df $H(x_1, ..., x_d)$. The case considered in this thesis is when $d = 2$ where the censoring variable X_1 and unit lifetime X_2 are in competition to remove the unit from observation in a life test. Accordingly, $\mathbf{X}_i = (X_{i,1}, X_{i,2})$, $i = 1, ..., n$ is a bivariate sample at each stress level. Assume the pairs $(X_{i,1}, X_{i,2})$, $i = 1, ..., n$ are i.i.d. at each stress level. Then they have a common bivariate df $H(x_1, x_2)$ with univariate margins $F_1(x_1)$ and $F_2(x_2)$. Denote by

$$M_{n,1} = max\,(X_{1,1}, ..., X_{n,1})$$
$$M_{n,2} = max\,(X_{1,2}, ..., X_{n,2})$$

the component-wise maxima for the censoring variable X_1 and unit lifetime X_2 respectively. Also denote by $M_n = (M_{n,1}, M_{n,2})$ the vector of these component-wise block maxima. Therefore $M_1 = (max(X_{11}), max(X_{12})) = (X_{11}, X_{12})$; $M_2 = (max(X_{11}, X_{21}), max(X_{12}, X_{22}))$ and so on. Clearly, the sequence $\{M_n\}$ is non-decreasing in n and interest is in the multivariate limiting distribution for M_n as $n \to \infty$ when appropriately normalised (Embrechts, Kluppelberg and Mikosch 1997). In particular, interest is in weak convergence results for centred and normalised maxima.

Assume sequences of normalising constants $a_{nj} = (a_{n1}, ..., a_{nd})$ and $b_{nj} = (b_{n1}, ..., b_{nd})$ such that $a_{nj} > 0$ and $b_{nj} \in (-\infty, \infty)$ exist. For the bivariate case, $j \in (1, 2)$. If component-wise maxima over n are properly rescaled, then

$$\lim_{n\to\infty} P\left(\frac{M_{n,1} - b_{n1}}{a_{n1}} \leq x_1, \frac{M_{n,2} - b_{n2}}{a_{n2}} \leq x_2\right) = \lim_{n\to\infty} H^n\left(a_{n1}x_1 + b_{n1}, a_{n2}x_2 + b_{n2}\right) \quad (2.5)$$

converges in distribution to a proper bivariate extreme value distribution. In the same way,

$$\lim_{n\to\infty} P\left(M_{n,1} \leq a_{n1}x_1 + b_{n1}\right) = F_1^n(a_{n1}x_1 + b_{n1})$$

and

$$\lim_{n\to\infty} P\left(M_{n,2} \leq a_{n2}x_2 + b_{n2}\right) = F_2^n(a_{n2}x_2 + b_{n2})$$

also converge in distribution to marginal limiting distributions of $M_{n,1}$ and $M_{n,2}$ as $n \to \infty$. Often these limiting marginal distributions are continuous univariate extreme value distributions of the Fréchet, Gumbel or Weibull type (McNeil, Frey and Embrechts 2005). By Sklar's Theorem, the multivariate limiting distribution function $H^n\left(a_{n1}x_1 + b_{n1}, a_{n2}x_2 + b_{n2}\right)$ admits a copula, say C^n

such that

$$H^n(a_{n1}x_1 + b_{n1}, a_{n2}x_2 + b_{n2}) = C^n \left\{ [F_1^n(a_{n1}x_1 + b_{n1})]^{1/n}, [F_2^n(a_{n2}x_2 + b_{n2})]^{1/n} \right\}.$$

Following Equation 2.5, let

$$\lim_{n\to\infty} H^n (a_{n1}x_1 + b_{n1}, a_{n2}x_2 + b_{n2}) \to W(x_1, x_2); \quad n \to \infty$$

where $W(x_1, x_2)$ is a proper bivariate extreme value distribution. Assume $W(x_1, x_2)$ has an associated copula $C_n(u_1, u_2)$. Then

$$C^n \left(u_1^{1/n}, u_2^{1/n} \right) \to C_n(u_1, u_2); \quad n \to \infty \tag{2.6}$$

where C_n is the copula of the limiting distribution $W(x_1, x_2)$, called the extreme value copula. Equation 2.6 justifies (asymptotically) the use of an extreme value copula when modeling component-wise maxima. Hence extreme value copulas are a natural choice when describing multivariate extremes in the data.

Within the reliability context, extremal dependence tends to be induced by random extremal events. Typically, these are random fatal shocks causing common cause failures. An extreme value copula that can capture this extremal dependence is the Marshall-Olkin copula which comes from the fatal shock model of Marshal and Olkin (1967) as follows.

Consider a unit with two components which experiences three independent fatal shocks. A type $k \in (1, 2)$ shock occurs at a random time T_k with probability $P(T_k > t) = e^{-\lambda_k t}$ and destroys component C_k. A shock causing a common cause failure at random time T_{12} with probability $P(T_{12} > t) = e^{-\lambda_{12} t}$ destroys both components simultaneously. Assuming no other failure causes, then the lifetimes of the components are $X_1 = min(T_1, T_{12})$ and $X_2 = min(T_2, T_{12})$ with univariate survival functions $S_1(x_1) = exp(-(\lambda_1 + \lambda_{12}) x_1)$ and $S_2(x_2) = exp(-(\lambda_2 + \lambda_{12}) x_2)$.

But a common cause failure induces statistical dependence between component lifetimes. Hence their joint survival probability is given by

$$\begin{aligned} S(x_1, x_2) &= P(X_1 > x_1, X_2 > x_2) = P(T_1 > x_1) P(T_2 > x_2) P(T_{12} > max(x_1, x_2)) \\ &= exp\{-(\lambda_1 + \lambda_{12}) x_1 - (\lambda_2 + \lambda_{12}) x_2 + \lambda_{12} min(x_1, x_2)\} \\ &= S_1(x_1) S_2(x_2) min[exp(\lambda_{12} x_1), exp(\lambda_{12} x_2)]. \end{aligned}$$

Letting $\alpha_k = \frac{\lambda_{12}}{\lambda_k+\lambda_{12}}$ for $k \in (1,2)$, then $exp\left(\lambda_{12}x_k\right) = S_k(x_k)^{-\alpha_k}$ and consequently the Marshall-Olkin (or survival) copula of $(X_1, X_2)^T$ is given by

$$\bar{C}_{\alpha_1,\alpha_2}(u_1,u_2) = u_1u_2min\left(u_1^{-\alpha_1}, u_2^{-\alpha_2}\right).$$

Thus the Marshall-Olkin copula is motivated by reliability considerations and hence, provides physical justification for natural situations where it can be applied.

2.3.3 The Archimedean family of copulas

Archimedean copulas are easily constructed and they possess nice properties. Following Ling (1965), a copula is called Archimedean if it admits the simple algebraic representation

$$C(u_1,...,u_d) = \psi^{[-1]}(\psi(u_1) + ... + \psi(u_d)), \;\; (u_1,...,u_d) \in [0,1]^d \tag{2.7}$$

for some univariate Archimedean generator function ψ and its pseudo-inverse $\psi^{[-1]}$. The Archimedean generator function ψ satisfies

(i) $\psi : [0,1] \to [0,\infty)$ with $\psi(1) = 0$.

(ii) ψ is a continuous and strictly decreasing function and its pseudo-inverse $\psi^{[-1]}$ with domain $[0,\infty)$ and range $[0,1]$ is given by

$$\psi^{[-1]}(x) = \begin{cases} \psi^{-1}(x) & \text{if} \quad 0 \le x \le \psi(0) \\ 0 & \text{if} \quad \psi(0) \le x \le \infty \end{cases}$$

(iii) $\psi^{[-1]} = \psi^{-1}$ if $\psi(0) = \infty$.

The necessary and sufficient condition for the generator function ψ to generate an Archimedean copula in dimension d is that it must possess an analytical property called d - *monotonicity* in $(a,b) \in [0,1]$. That is

(i) ψ is differentiable in $[0,1]$ up to order $d-2$.

(ii) The continuous derivatives satisfy $(-1)^k\psi^{(k)}(x) \geq 0$ for any $k \in \{1, ..., d-2\}$ and for any $x \in [0, 1]$.

(iii) $(-1)^{d-2}\psi^{(d-2)}$ is non-negative, nonincreasing and convex in $(a, b) \in [0, 1]$.

For more details on the *d - monotonicity* of the Archimedean generator function ψ, see for example McNeil and Neslehova (2009) and the numerous references therein. For the bivariate case however, ψ induces a bivariate Archimedean copula if and only if it is convex, that is, if its second derivative $\psi^{(2)} \geq 0$ (Ling, 1965; Schweizer and Sklar, 1983).

Assuming a bivariate Archimedean copula is a suitable dependence model for the bivariate competing risks test data, then the dependence structure of the vector (X_1, X_2) is completely characterised by the univariate generator function ψ. This is because Archimedean copulas arise through special stochastic modeling and their dependence properties reduce to certain technical conditions (analytical properties) of single-valued Archimedean copula generators being met. Consequently, inference for Archimedean copulas is considerably simple when compared to other copula families.

Different Archimedean generators also provide different dependence structures and a number of different Archimedean copulas already exist in the literature. Accordingly, the Archimedean family of copulas flexibly describes a wide range of dependence structures. The connection of Archimedean copulas with frailty models (see for example Marshall and Olkin, 1988; Oakes, 1989) provides a probabilistic interpretation of Archimedean dependence structures. Commonly used Archimedean copulas are available in closed form and examples of some important Archimedean generators and their respective Archimedean copulas are listed in Table 2.1.

Name	Copula $C_\theta(u_1, u_2)$	Generator $\psi(x)$
AMH	$\frac{u_1 u_2}{(1-\theta(1-u_1)(1-u_2))}$	$ln\frac{1-\theta(1-t)}{t}$
Clayton	$max\left((u_1^{-\theta}+u_2^{-\theta}-1)^{-\frac{1}{\theta}}, 0\right)$	$\frac{1}{\theta}(t^{-\theta}-1)$
Frank	$-\frac{1}{\theta}ln\left(1+\frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{e^{-\theta}-1}\right)$	$-ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$
Gumbel	$exp\left((-ln(u_1))^\theta+(-ln(u_2))^\theta\right)^{\frac{1}{\theta}}$	$(-ln(t))^\theta$
Joe	$1-\left((1-u_1)^\theta+(1-u_2)^\theta+(1-u_1)^\theta(1-u_2)^\theta\right)^{\frac{1}{\theta}}$	$-ln(1-(1-t)^\theta)$

Table 2.1: *Prominent Archimedean families of copulas. AHM denotes Ali - Mikhail - Haq.*

2.4 Copula model for dependent competing risks

Multivariate models aside, which model to use is in general a very difficult problem with no obvious answer even in the univariate case. In the words of McCullagh and Nelder (1983), "all models are wrong but some are more useful than others". Model selection strategies tend to be based on model properties, predictive accuracy, diagnostic tools etc. Other important considerations in statistical modeling (Genest and Remillard 2006) include mathematical simplicity and convenience, interpretability and fitness for purpose.

It is however not always enough to use any statistical technique simply because it is mathematically convenient. Rather, the theoretical ideas on which the technique in question is based and the situation being modeled crucially have to be understood. When unit lifetime X_2 can be right censored by a dependent variable X_1 in a life test, then the random variables $min(X_1, X_2)$ and $max(X_1, X_2)$ are order statistics for X_1 and X_2. Denote by F_1 and F_2 the distributions of X_1 and X_2 respectively. Then the random time at which a unit is removed from observation at each stress level is $Z = min(X_1, X_2)$ and its distribution function is given by

$$P\left(min(X_1, X_2) \leq z\right) = F_1(z) + F_2(z) - C\left(F_1(z), F_2(z)\right)$$

where $P\left(min(X_1, X_2) \leq z\right)$ is the distribution function of Z and $C(\cdot, \cdot)$ is the associated copula of the vector (X_1, X_2). For more details, see Nelsen (1999), p. 25. Since unit lifetime X_2 and the censoring variable X_1 are competing for the removal of a unit from observation in a life test, only the minimum of the competing risks variables is observable. Assuming absolute continuity of the risk variables, a characteristic feature of the observable data is that one variable must be smaller than the other. Consequently, the joint distribution (and hence the copula model) should be reflective of this feature. In this sense, it seems natural to model dependence between unit lifetime X_2 and the censoring variable X_1 by the order statistics copula.

But obtaining a copula representation of the joint distribution of order statistics corresponding to X_1 and X_2 presents problems. For example, the representation by Anjos, Kolev and Tanaka (2005) requires the associated copula of the vector (X_1, X_2) or both of F_1 and F_2 to be completely known. In this study however, only the observable competing risks data are available at each stress level. Accordingly, neither the associated copula of the vector (X_1, X_2) nor the marginal distributions of the risk variables are estimable from competing risks data alone.

The work of Zheng and Klein (1995) suggest that a reasonable estimate of the copula dependence parameter θ, not the functional form of the copula is the important factor when modeling dependence between competing risks variables. This is the modeling approach adopted in for example Bunea and Mazzuchi (2007), Escarela and Carriere (2003), Kaishev, Dimitrova and Haberman (2007) and will also be followed in this thesis. The difficulty with this approach is that the copula dependence parameter cannot be estimated from the competing risks data alone. Consequently, among the important criteria for selecting a family of parametric copulas is that the dependence structure must be summarised by a dependence measure that can be assessed using expert opinion. Within the chosen family, different copulas must exist to model different dependence structures and the copulas must also be available in closed form for easy simulation.

The common dependence measure within the class of elliptical distributions is linear correlation. However, it has no simple direct interpretation in terms of probabilities and is thus not easily assessed by experts in a defensible way. Besides, linear correlation is not invariant under general transformations that are not necessarily linear. Further, copulas of elliptical distributions are typically not available in closed form and are therefore difficult to work with. Regarding extreme value copulas, there is no obvious reason to suggest existence of multivariate extremes in data on unit lifetime and the censoring variable. The Marshall-Olkin type distributions on the other hand have a singularity along $x_1 = x_2$. This is a major drawback in that in a competing risks framework, the risk variables are assumed to be absolutely continuous.

The dependence structure within the class of Archimedean copulas can be summarised by rank correlation. As a dependence measure, rank correlation (Kendall's τ henceforth) has a simple and direct interpretation in terms of probabilities of observing concordance and discordance pairs (Hove, 2013). Accordingly, it is easily assessed by experts in a defensible way. A number of Archimedean generators exist and these provide different dependence structures. As a direct consequence, the Archimedean class of copulas describes a wide range of dependence structures. Above all, copulas within the Archimedean family are generally available in closed form. Hence they provide for mathematical simplicity and convenience. On the basis of the discussed copula classes and their properties, the Archimedean class of copulas is preferred in this thesis.

2.4.1 Which Archimedean copula

Assume the stochastic dependence between unit lifetime X_2 and the censoring variable X_1 at each test stress level is captured by a copula from within the Archimedean family. However, the exact Archimedean copula is never known in advance and an important problem (see e.g. Nelsen 2005) is:

Assuming an Archimedean copula is the appropriate dependence model for test data, what are the statistical procedures for choosing the right family?

But dependence between unit lifetime X_2 and the censoring variable X_1 can potentially vary from extreme negative through independence to extreme positive dependence. Accordingly, the chosen Archimedean copula must attain its Frechet lower and upper bound as well as coinciding with the product copula as $\theta \to 0$. Thus, the important criteria used in this thesis when selecting the appropriate Archimedean copula is that it must capture the full range of dependence. Such copulas are called comprehensive copulas and the only examples in the Archimedean class are the Clayton and the Frank families (Nelsen, 1999).

The Clayton copula is asymmetric, exhibiting lesser dependence in the positive tail and greater dependence in the negative tail. On the other hand, the Frank copula is a symmetric Archimedean copula. There is however no physical justification to suggest asymmetries between unit lifetime X_2 and the censoring variable X_1. As a result, the symmetric Frank copula is preferred ahead of the asymmetric Clayton copula in this thesis.

Assume the dependence between unit lifetime X_2 and the censoring variable X_1 at each test stress level is adequately described by Frank's copula introduced by Genest (1987). It is given by

$$C_\theta^F(u_1, u_2) = -\frac{1}{\theta}\ln\left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{(e^{-\theta} - 1)}\right] \tag{2.8}$$

where $\theta \in (-\infty, +\infty)\backslash\{0\}$ is the copula dependence parameter. The relationship between Kendall's τ and the Frank copula dependence parameter θ, see e.g. Escarela and Carriere (2003), is given by

$$\frac{[1 - D_1(\theta)]}{\theta} = \frac{1 - \tau}{4}$$

where

$$D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{e^t - 1}dt$$

is a Debye function of the the first kind for $\theta > 0$. That is

$$\tau = 1 - \frac{4}{\theta}\left(1 - \frac{1}{\theta}\int_0^\theta \frac{t}{e^t - 1}dt\right). \tag{2.9}$$

But Debye functions, even of the first kind, do not have explicit integral-free expressions. Hence for an estimated τ from expert opinion, the Frank copula dependence parameter θ is obtained by solving Equation 2.9 for θ using numerical methods.

2.5 Estimation of the Frank copula dependence parameter

The Frank copula parameter θ measures the strength of functional dependence between unit lifetime X_2 and the censoring variable X_1 at each stress level. To compute its estimate $\hat{\theta}$, a random sample $(x_{11}, x_{21}), (x_{12}, x_{22}), ..., (x_{1n}, x_{2n})$ from (X_1, X_2) is in general required. But since unit lifetime is subject to dependent random censorship in a life test, it is the random sample $(min(x_{11}, x_{21})), (min(x_{12}, x_{22})), ..., (min(x_{1n}, x_{2n}))$ from $min\,(X_1, X_2)$ that is observable. Obviously, such data alone contain not sufficient information to estimate θ. Hence, the analysis inevitably has to rely on subjective judgements by subject matter experts (Hove, 2013). This is common practice in engineering reliability and science in general where uncertainties in knowledge often exist and individual expertise are increasingly recognised as just another type of data (see e.g. van Noortwijk, Dekker, Cooke and Mazzuchi, 1992; Kurowicka and Cooke, 2006) and the numerous references therein.

2.5.1 Aspects of the problem to elicit

A unique copula associated with the pair (X_1, X_2) is invariant under strictly increasing transformations of the marginal distribution functions $F_1(x_1)$ and $F_2(x_2)$. Since the dependence between X_1 and X_2 is characterised by this copula, it follows that an appropriate measure of dependence must also exhibit the same scale-invariance property. One such dependence measure is the rank correlation which measures the degree of monotone relationships between variables.

The best known rank based distribution-free measures of dependence are Spearman's ρ and Kendall's τ. They are both calculated on the ranks of the data and not on the actual data values themselves. It is therefore not surprising that a relationship exists between these rank based dependence measures and copula functions. In terms of the copula function (see e.g. Carriere, 1994; Nelsen, 1999), Spearman's ρ and Kendall's τ are correspondingly given by

$$\rho_{X_1X_2} = 12 \int\int_{I^2} C^F(u_1, u_2) du_1 du_2 - 3$$

and

$$\tau_{X_1X_2} = 4 \int\int_{I^2} C^F(u_1, u_2) dC^F(u_1, u_2) - 1.$$

Parameterisations of families of copulas by the rank correlation implies that knowledge of these rank based dependence measures identifies the copula. Hence, the rank correlation can be taken as the uncertain quantity to be elicited when fitting the Frank copula model to dependent competing risk data. When planning an elicitation for a single uncertain quantity (rank correlation in this case), good guidance exists in the literature on the aspects of the problem to elicit. Cooke (1991), Cooke and Goossens (2000) and Bedford and Cooke (2001) all stress the need to elicit on observable quantities only.

The practical reason for eliciting on observable quantities only is that experts are more comfortable with answering questions on such quantities. In particular, experts must be asked to give subjective judgements on quantities that in principle can be measured by a physical though not necessarily practical procedure they are familiar with. Whenever possible, assessment questions must be asked in frequency terms instead of directly to minimize cognitive biases (Gigerenzer 1991, Gigerenzer, Hoffrage and Kleinbolting, 1991).

Admittedly, the rank correlation is clearly not an observable quantity. To indirectly infer the appropriate rank correlation number, other derived elicitation or query variables which can be regarded as observable and are related to rank correlation have to be identified. Because experts will assess these query variables, it is important that the query variables are also familiar to experts. Spearman's rho is a widely used measure of rank correlation largely because its computation is very simple. That is, it is Pearson's product moment correlation computed on the ranks rather than the actual data values themselves.

The major disadvantage of Spearman's ρ is that it has no simple direct interpretation in terms of probabilities and is thus difficult to quantify. Though usually considered more difficult to compute than Spearman's rho, Kendall's τ, an alternative rank correlation does have a simple and direct interpretation in terms of probabilities of observing concordance and discordance pairs (Conover, 1999). Its distribution also rapidly converges to the normal distribution and thus, has a better normal approximation. Hence Kendall's τ (rank correlation henceforth) is the uncertain quantity to be elicited and concordance probability is the query variable to be assessed by experts in this thesis.

2.5.2 The elicitation process

Elicitation is a part of the general process of statistical modeling (Garthwaite, Kadane and O'Hagan, 2005) in that the usual principles of statistical modeling apply. It involves the following:

(1) The expert makes a finite (usually small) number of specific judgements about the uncertain quantity, rank correlation in this case. These judgements are summaries of his or her distribution and examples include the mean, minimum, maximum or some specified percentiles. The basic assumption for the elicited distribution to be considered useful is that judgements about these distributional summaries capture the important features of the expert's beliefs.

(2) The analyst constructs a fully specified (joint) probability distribution from the elicited summaries. The goal is to express in probabilistic form, the expert's current knowledge since an elicited distribution rather than a single point estimate for the unknown quantity best describes the uncertainty about the quantity of interest. This however remains an important but ill-posed problem because many other possible distributions may fit the expert's judgements equally well.

(3) The next stage is to check that the fitted distribution matches the expert's beliefs. Hence there is always the option of returning to check if the fitted distribution accurately represents the expert's knowledge. This makes elicitation almost invariably an iterative process.

Elicitation is thus defined as the process of formulating the beliefs of an expert about an uncertain quantity into a probability distribution for that quantity. The term *expert* is in general often associated with special knowledge regarding the subject matter under consideration. As used here however, an expert refers to one whose knowledge is to be elicited. As a result, elicitation is considered a success if the fitted distribution represents the expert's knowledge accurately, with no regard to how good the expert's knowledge is.

Various approaches for eliciting expert knowledge regarding dependence are suggested in Clemen and Reilly (1999). A thorough investigation is reported in Clemen, Fischer and Winkler (2000) where a discussion of some desirable characteristics for dependence assessment methods is given. To be practically useful, a dependence assessment method should:

(1) Have a sound and defensible probabilistic foundation for modeling.

(2) Be general enough to allow dependence modeling in a number of situations.

(3) Be directly linked to the modeling procedure and have a clear and natural interpretation for easy of assessment..

Eliciting probability distributions from experts is a complex process. Typically, it entails selecting and training experts regarding the identified summaries to be elicited for the problem at hand. Instead of choosing real experts and obtaining their distributional summaries, the elicitation process is demonstrated by means of a simulation study in this thesis. Crucially however, all stages of the elicitation process are followed in the simulation study.

2.5.3 Expert elicitation: A simulation study

Denote by $\left(X_1^{(1)}, X_2^{(1)}\right)$ and $\left(X_1^{(2)}, X_2^{(2)}\right)$ two independent random draws from a population of test units (X_1, X_2). Label them units 1 and 2 respectively where $X_1^{(1)}$ and $X_1^{(2)}$ are the random censoring times and $X_2^{(1)}$ and $X_2^{(2)}$ are the lifetimes for the two test units. The expert could be asked the assessment question:

Suppose it turns out in a life test that unit 2 has a longer lifetime than unit 1, that is $X_2^{(1)} < X_2^{(2)}$, what is your probability that a degraded failure for unit 1 would also occur before the degraded failure for unit 2 in an ALT experiment?

This assessment question is asking directly for a concordance probability

$$P\left[\left(X_1^{(1)} - X_1^{(2)}\right)\left(X_2^{(1)} - X_2^{(2)}\right) > 0\right]$$

and denote this probability be p_c. Conversely, the probability that a degraded failure for unit 2 would occur before the degraded failure for unit 1 given that $X_2^{(1)} < X_2^{(2)}$ is a discordance probability. In terms p_c, it is given by

$$P\left[\left(X_1^{(1)} - X_1^{(2)}\right)\left(X_2^{(1)} - X_2^{(2)}\right) < 0\right] = 1 - p_c.$$

If experts can assess concordance probability, then Kendall's τ (uncertain quantity) is obtained from the assessed concordance probability by

$$\tau = 2p_c - 1. \tag{2.10}$$

The concordance probability p_c can be considered an observable quantity because it is a physically realisable quantity and its assessment has a natural interpretation in frequency terms. Specifically, p_c is the relative frequency for $\left\{X_1^{(1)} < X_1^{(2)} | X_2^{(1)} < X_2^{(2)}\right\}$ when a large sample of pairs of independent draws from a population of test units is observed (Bunea & Bedford 2002). Clearly,

$$p_c = P\left(X_1^{(1)} < X_1^{(2)} | X_2^{(1)} < X_2^{(2)}\right)$$

is a probability statement on a single variable X_1 and is therefore easier to communicate to experts than say a joint probability. Since any two independent draws from the population of test units would resolve into one of two states (either concordant or discordant), then a single subjective probability from the expert is the required response.

Recall that often in industrial ALT, very few units (usually prototype) are available for life testing due to cost constraints. To obtain sufficient failure data quickly and in a more cost effective manner, the few test units are repaired after failure and tested continuously. In this sense, the two independent draws $\left(X_1^{(1)}, X_2^{(1)}\right)$ and $\left(X_1^{(2)}, X_2^{(2)}\right)$ will have the interpretation of repairable test units from the population (X_1, X_2) of test units. When a degraded or critical failure is detected during testing, the failed unit will have to be repaired to a state as good as new since ALT deals with new units.

To yield the right data structure for the present setup, failure times are simulated from models tailored for situations where the variable of interest (unit lifetime) is subject to random censorship. Few such models have been proposed in the reliability literature. The one that is well established is the random signs censoring model due to Cooke (1996). Different refinements of this model have also been introduced in the reliability literature. They include the delay-time (DT) model (Christer, 2002) and the repair-alert (RA) model (Lindqvist, Stove, and Langseth, 2006). The later has since been modified further by Bedford and Alkali (2009). In this thesis however, unit failure times are simulated from the alert-delay (AD) model of Dijoux and Gaudoin (2009), a dependent competing risk model that is midway between the DT and RA models.

Though originally developed in maintenance studies, the AD model is extended to an ALT setting in this thesis as follows: Consider a unit which should experience a critical failure in a life test at some random time X_2. Assume the system exhibits inferior performance during testing at some random time pX_2 before X_2 where $p \in [0, 1]$. Inferior performance acts as a signal (an alert) to the life testing team that a critical failure is approaching. But a critical failure is not expected to occur immediately after the alert. Hence an additional time ξ after pX_2 is introduced and corresponds to the delay allowed before the performance of the unit falls below industrial standards (degraded failure) assuming the signal is detected. Otherwise the system is kept on test until it experiences a complete loss of function (critical failure).

This yields the AD model

$$X_1 = pX_2 + \xi \tag{2.11}$$

where X_2 and ξ are two independent lifetime variables since there is no particular reason to link the two life variables. Particular cases of the AD model for special choices of the model parameters are:

(1) If $p = 0$, then $X_1 = \xi$. This implies that X_1 and X_2 are statistically independent. But since degraded and critical failure are assumed to be linked through the degradation process of the unit, $p \neq 0$.

(2) If $p = 1$, then $X_1 = X_2 + \xi$ and hence $Z = min(X_1, X_2) = X_2$ always. This implies that a unit is removed from observation in a life test only if it reaches the end of its useful life. In the present application however, unit lifetime X_2 is subject to random censorship. Accordingly, a unit can also be removed from observation in a life test whenever a degraded failure is detected during testing even if it has not reached the end of its useful life. This happens with probability $q \neq 0$ and consequently $p \neq 1$.

(3) $X_1 = pX_2$ implies that a degraded failure is always responsible for removing the unit from observation during testing. In the present formulation however, the signal may not be detected with probability $1 - q \neq 0$. In this case, the unit is kept on test until it experiences a complete loss of function (critical failure).

Observed sequences of the failure times $Z = min(X_1, X_2)$ along with the identity of which mode removed the unit from observation in a life test are generated as follows:

(1) Any life distribution is in theory possible for unit lifetime X_2 and the life variable ξ. To account for system degradation, lifetimes for the two units $\left(X_1^{(1)}, X_2^{(1)}\right)$ and $\left(X_1^{(2)}, X_2^{(2)}\right)$, that is $X_2^{(1)}$ and $X_2^{(2)}$ respectively are simulated from the Weibull distribution. For simplicity, the life variable ξ is simulated from the exponential distribution.

(2) For a specified value of p, the corresponding times to a degraded failure for the two units namely $X_1^{(1)}$ and $X_1^{(2)}$ are obtained from the AD model. To fully exploit the residual life of the unit, the signal must not be delivered too early. This corresponds to choosing a value of p close enough to one.

(3) Only consider cases when it turns out that $X_2^{(1)} < X_2^{(2)}$. In practice, this can be easily determined in ALT by testing the two systems at higher test stresses and retaining cases when unit 1 has a shorter lifetime than unit 2. Assume the count of such cases from the simulation is m_2 for example. Out of these m_2 cases count how many are such that $\{X_1^{(1)} < X_1^{(2)}\}$, say m_1. Estimate the concordance probability p_c by $\frac{m_1}{m_2}$ and obtain the rank correlation from $\tau = 2p_c - 1$ which is the target of estimation.

(4) Repeat the simulations several times, say k to obtain $\tau_1, ..., \tau_k$. Use these k simulated rank correlation values to obtain an estimate of the expert's distribution using nonparametric methods such as the histogram plot or a kernel density estimate.

A remark is however necessary here. The expert's distribution is not pre-formed and waiting to be extracted. Rather, the expert only responds when prompted. This implies that the way the elicitation question is phrased is an important aspect of the elicitation process. Accordingly, the given simulation design will likely yield a reasonable and hence practically useful estimate of the expert's distribution for the assessment question under consideration. In a typical elicitation however, the expert is asked to only give summaries (usually few) of his or her distribution. Accordingly, few summaries of the estimated expert's distribution are obtained from the k simulated rank correlation values. The choice of what summaries to elicit largely depends on the choice of the distributional form one intends to fit to those summaries.

A subjective distribution that uses the same assumption about the mean (some functional form for the unknown distribution) as Project Evaluation and Review Technique (PERT) networks used for project planning is the PERT distribution. It is frequently used in applications to model expert opinion and requires as input the minimum, maximum and most likely values for the uncertain quantity. The PERT function then finds a distributional shape that fits these restrictions. The syntax in ModelRisk software is `VosePERT(min,mode,max)`. A version of the PERT distribution which offers some degree of control of peakedness and hence uncertainty of the elicited distribution is the modified PERT distribution. In ModelRisk software, the syntax is `VosePERT(min,mode,max,$\gamma$)` where an increase (decrease) in the value of γ assigns more (less) peakedness to the the elicited distribution. When $\gamma = 4$, the modified PERT becomes the standard PERT distribution.

Thus the few summaries that would be obtained from the estimated expert's distribution are the minimum, maximum and most likely values. Uncertainty about Kendall's τ is then modeled in ModelRisk software by fitting a fully specified distribution to these summaries. If the fitted distribution adequately matches the nonparametric expert's distribution estimated from the k simulated rank correlation values, then elicitation is considered to have been a success. A specified percentile or summary of the fitted distribution is then used to obtain an estimated value of Kendall's τ, the uncertain quantity. Given the estimated Kendall's τ value, Equation 2.9 is then solved for the Frank copula dependence parameter θ. Consequently, the copula model that captures the stochastic dependence between the censoring variable X_1 and unit lifetime X_2 in a life test is estimated.

2.5.4 Numerical results

The concordance probability is assessed by means of a simulation study using the `R` code in `Appendix A`. The AD model is preferred for simulating failure times because it is tailored for cases where unit lifetime is subject to random censorship in a competing risk context as is the case in this study. Equation 2.10 is used to generate an estimated value of Kendall's τ from the assessed concordance probability. The simulation is repeated for $k = 1000$ times yielding the same number of estimated values for Kendall's τ.

The expert's distribution of Kendall's τ (the uncertain quantity) is estimated nonparametrically from the k generated rank correlation values. Specifically, the `R` command `plot(density(x))` where **x** is the vector of the generated rank correlation values is used to get a kernel estimate of the expert's density. Figure 2.1 shows the estimated expert's distribution.

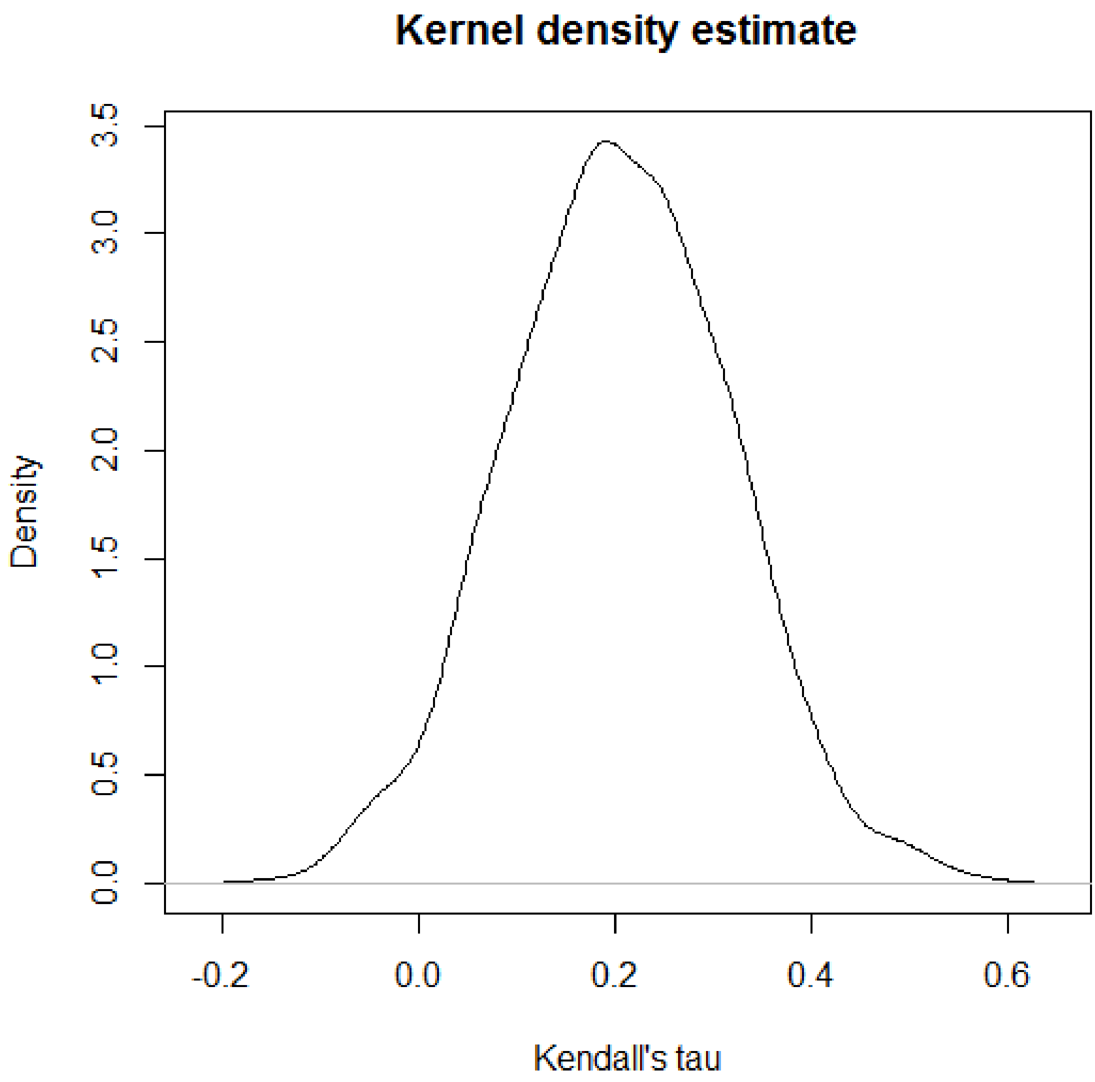


Figure 2.1: *Kernel density estimate of the expert's distribution from simulated data.*

Instead of giving his or her distribution for the uncertain quantity, the expert is asked to give a few summaries of the distribution. Accordingly, only the minimum, mode and maximum values of K-endall's τ are obtained from the simulated data. These values are summaries of the expert's distribution in Figure 2.1. From the vector **x** of generated values of Kendall's τ, the sample minimum, mode and maximum are obtained in `R` using `min(x)`, `names(sort(-table(x)))[1]` and `max(x)` respectively. In this simulation study, the resulting summaries of the expert's distribution are `minimum=-0.1466667`; `mode=0.1733333` and `maximum=0.5733333`.

The negative values of Kendall's τ corresponds to cases where there is a disagreement between the rankings of X_1 and X_2. That is, the ranking of one risk variable is mostly in the reverse of the other risk variable and corresponds to cases where $X_1^{(1)} > X_1^{(2)}$ is observed in a life test even though it has been observed that $X_2^{(1)} < X_2^{(2)}$. However, most of the simulated cases (see Figure 2.1) yield rankings that are mostly in agreement. This makes sense because in practice, random censorship is expected to occur close to the end of the unit's useful life.

Uncertainty about Kendall's τ is then modeled by fitting a fully specified distribution to the sample minimum, mode and maximum values of the uncertain quantity. The motivation is that a real expert would have been asked to give only these summaries in a typical elicitation process. Elicitation will be a success if the fitted distribution closely approximates the expert's distribution in Figure 2.1. There are two interfaces for fitting a fully specified distribution to distribution summaries of the uncertain quantity in ModelRisk. In one, ModelRisk selects the distribution while in the other interface, the analyst (or expert) draws own distribution.

Using the given minimum, mode and maximum values of Kendall's τ, the parameters of the elicited PERT (modified PERT with $\gamma = 4$) distribution in Figure 2.2

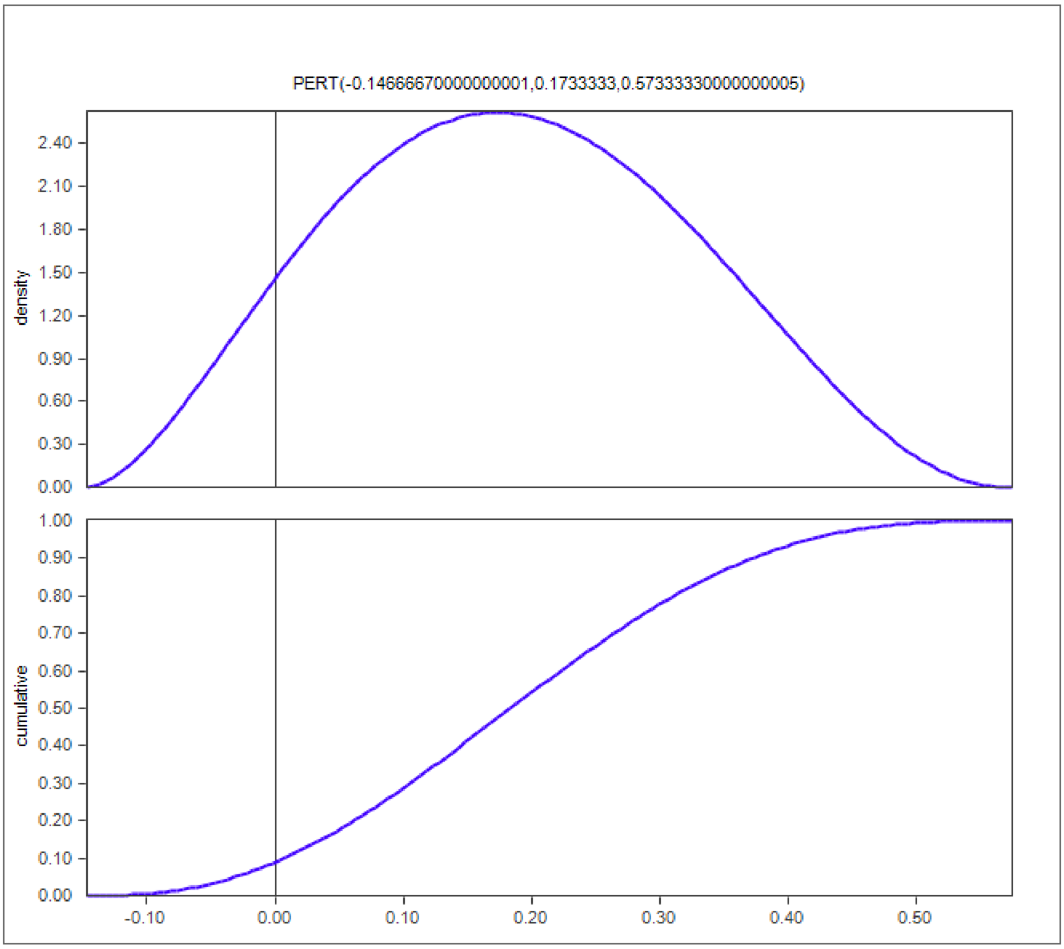


Figure 2.2: *Elicited PERT distribution.*

are given in Table 2.2.

Parameter	PERT
Mean	0.18667
Standard deviation	0.13569
50^{th} percentile	0.18352

Table 2.2: *Parameters of the elicited PERT distribution.*

In terms of the general shape, the elicited distribution in Figure 2.2 to some extent resembles the expert's distribution in Figure 2.1. However, it is flatter and hence more uncertain than the expert's distribution. Recall that the standard PERT distribution is the modified PERT with $\gamma = 4$. Thus to better match the expert's distribution, γ must be increased to assign more peakedness to the elicited distribution while retaining the same distributional summaries. The resulting modified PERT distribution with $\gamma = 6$ is given in Figure 2.3

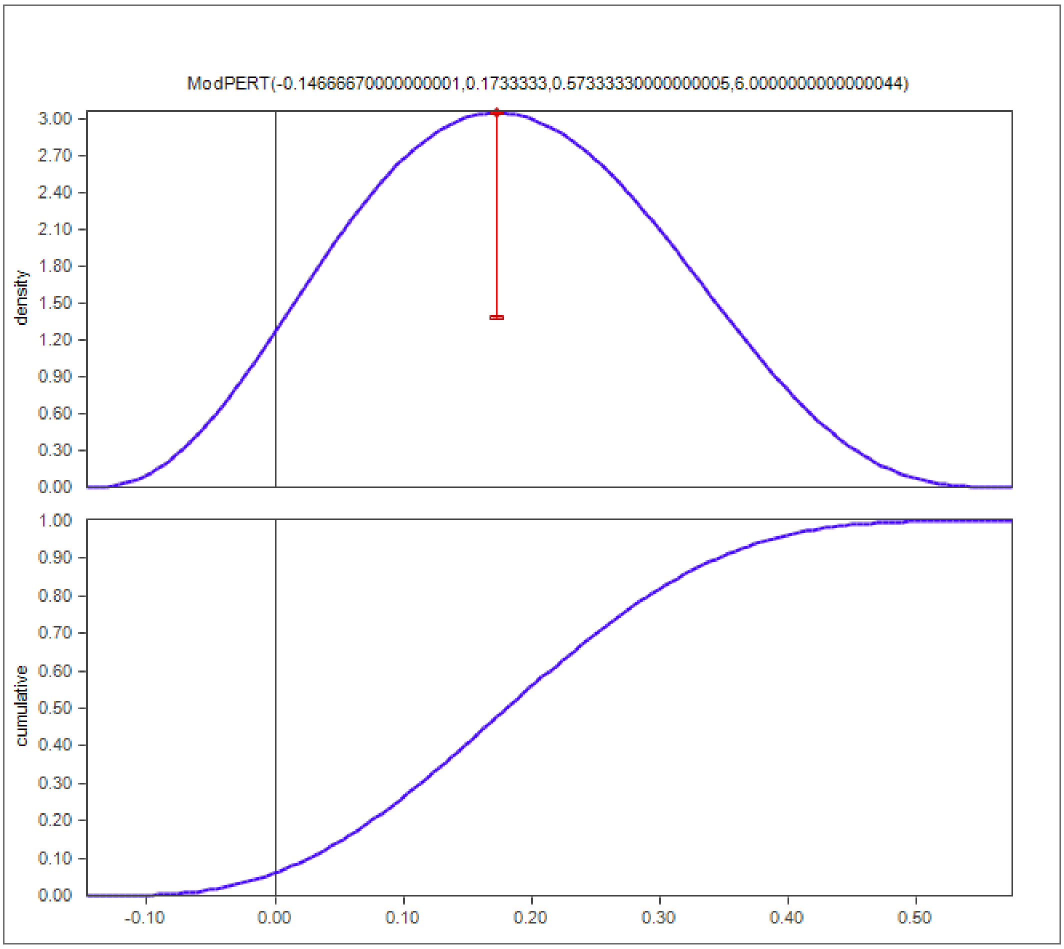


Figure 2.3: *Elicited modified PERT distribution with* $\gamma = 6$.

and its parameters are given in Table 2.3.

Parameter	Modified PERT
Mean	0.18333
Standard deviation	0.11958
50^{th} percentile	0.18071

Table 2.3: *Parameters of the elicited modified PERT distribution.*

Still, the elicited distribution in Figure 2.3 is flatter than the expert's distribution though minor improvement is apparent in the distributional shape. Model risk also allows the analyst (or expert) to draw own distribution within the defined range of values of the uncertain quantity. This offers greater flexibility in terms of distributional shapes to match expert opinion. Figure 2.4

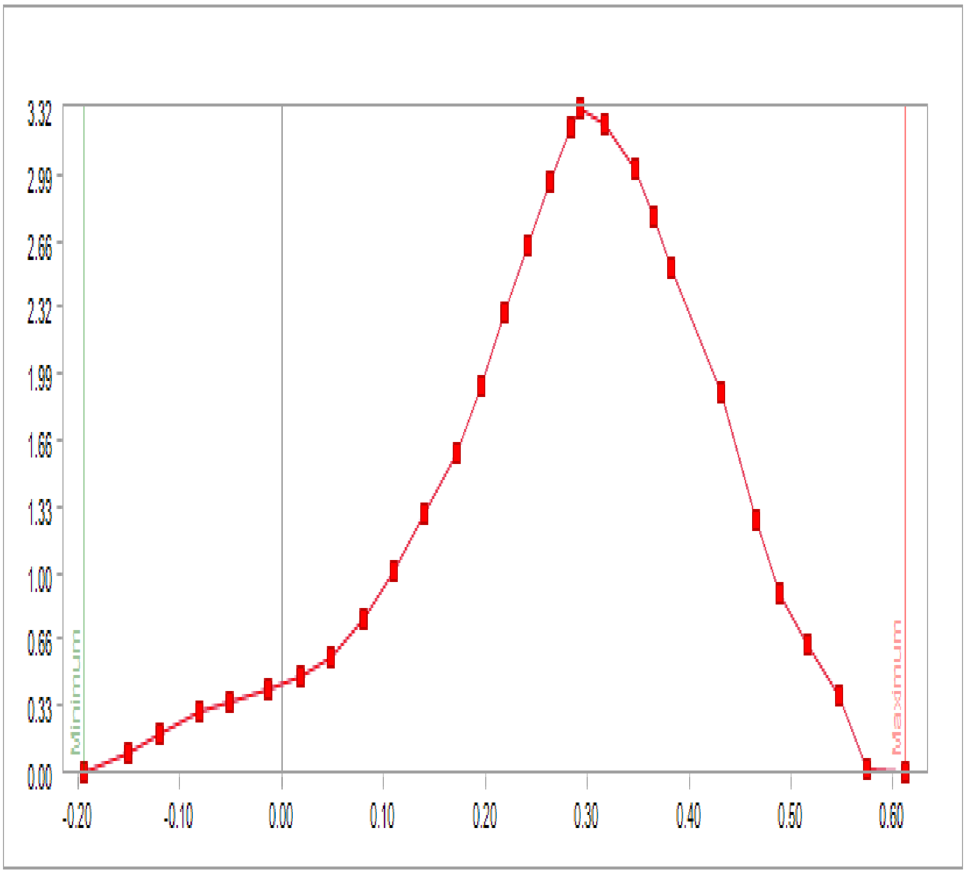


Figure 2.4: *Elicited distribution by plotting.*

gives the plotted distribution defined by the parameters in Table 2.4.

Parameter	Plotted distribution
Mean	0.27816
Standard deviation	0.13823
50^{th} percentile	0.29317

Table 2.4: *Parameters of the elicited distribution by plotting.*

By nature subjective distribution estimates are never precise. For estimates of an uncertain quantity to be useful in a model, the elicited distribution ought to be realistic. The plotted distribution closely matches the expert's distribution and is thus preferred to the PERT and modified PERT distributions. Consequently, the 50^{th} percentile of the plotted distribution is taken as the estimate of Kendall's τ, the uncertain quantity. Thus the assessed rank correlation is $\hat{\tau} = 0.29317$. It represents the difference between the probability that times to degraded and critical failures of test units are in the same order and the probability that they are not in the same order.

Though positive, the assessed rank correlation is less than one which implies that the agreement between the rankings of X_1 and X_2 is not perfect. This makes sense in practice because the fact that $X_2^{(1)} < X_2^{(2)}$ does not guarantee that $X_1^{(1)} < X_1^{(2)}$ will always hold in a life test. Substituting $\hat{\tau} = 0.29317$ into Equation 2.9 yields

$$\frac{4}{\theta}\left(1 - \frac{1}{\theta}\int_0^{\theta} \frac{t}{e^t - 1}dt\right) = 0.70683. \tag{2.12}$$

Numerical methods are required to solve Equation 2.12 for θ. Figure 2.5 is a plot of the left hand side of Equation 2.12 in Mathematica.

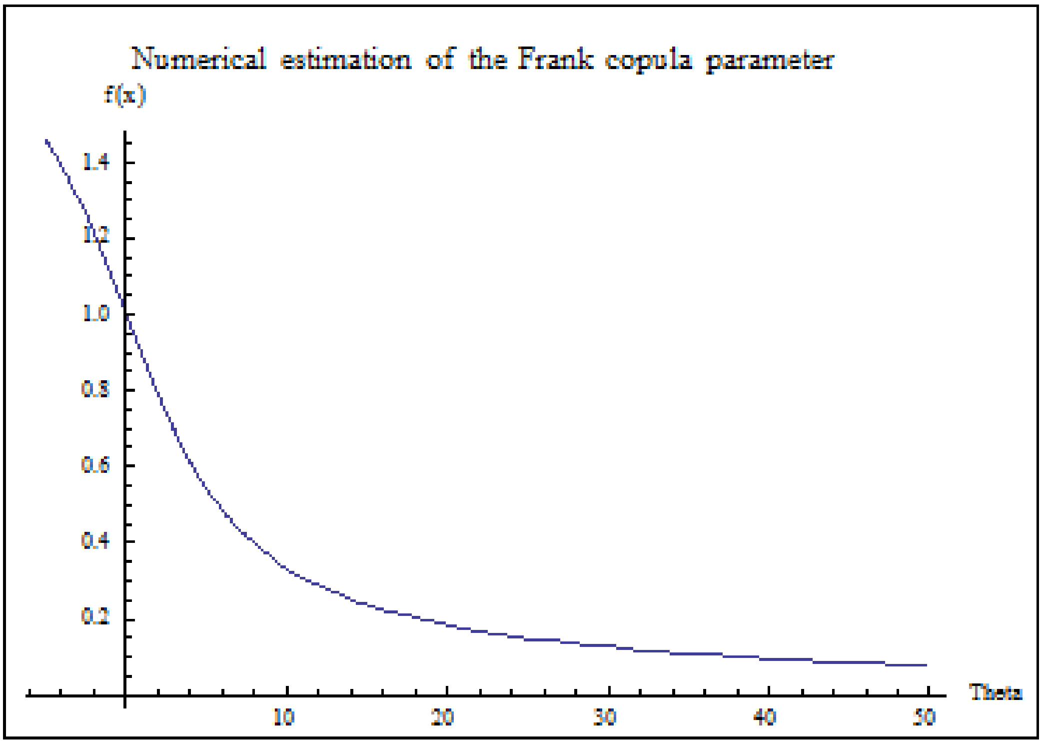


Figure 2.5: *Estimation of the Frank copula parameter.*

Using the `Table` function in Mathematica yields the estimate $\hat{\theta} = 2.8405$ of the Frank copula parameter. If the estimate is required to a large number of decimal places, small step sizes will be used. Thus based on the simulation study, there is positive dependence between the risk variables. Recall that degraded and critical failures are linked through the degradation process and that detection of the former is a signal that the latter will likely follow if the unit is kept on test. Accordingly, degraded failures will likely occur close to critical failures thereby justifying positive dependence between the risk variables. Hence the Frank copula model that captures the stochastic dependence between the censoring variable X_1 and unit lifetime X_2 at all stress levels is estimated from expert opinion. If there are reasons to suggest increased test stresses not only alter the scale but also the dependence structure of the competing risk variables, the expert can easily factor this extra information when assessing concordance probability.

Chapter 3

Lifetime models based on degradation phenomenon

3.1 Introduction

This chapter discusses the derivation of functional forms (or models) for the observed occurrences of degraded and critical failures in a life test. These risk variables are presumed to be competing with each other for the removal of a unit from observation in a life test. Thus the first N observations of the censoring variable X_1 and unit lifetime X_2 after rearranging if necessary are $(z_1, ..., z_N) = (x_{11}, ..., x_{1n}; x_{21}, ..., x_{2m})$; $n + m = N$. These observable competing risks data only allow one to estimate subdistribution functions

$$\begin{cases} F_1^*(x_1) = P\left(X_1 \leq x_1, X_1 < X_2\right) = S_1^*(0) - S_1^*(x_1) \\ F_2^*(x_2) = P\left(X_2 \leq x_2, X_2 < X_1\right) = S_2^*(0) - S_2^*(x_2) \end{cases} \tag{3.1}$$

such that the non-negative and non-increasing real functions S_1^* and S_2^* with support $\mathbb{R}^+ = [0, \infty)$ are continuous at zero and satisfy $S_1^*(0) + S_2^*(0) = 1$. Functional forms of the subdistribution functions in Equation 3.1 are the main object of estimation in this chapter. But failure (degraded or critical) in a life test is defined as the end point of some underlying degradation process. Hence suitable functional forms for $F_j^*(x_j)$ at all stress levels may as well be derived based on the failure mechanism and an understanding of the underlying degradation process. As used in this thesis, unit degradation is the irreversible accumulation of damage in a life test that leads to unit failure.

Degradation processes have the characteristic feature that they are governed by some random mechanism that is conveniently represented by a stochastic process $\{X(t), t \in T\}$. The index t is a time parameter and the index set T contains all possible time points. Hence unit degradation in a life test is assumed to be adequately described by the stochastic process $\{X(t), t \in T\}$. For degradation models that are best described by point processes and their respective counting processes, see Kahle and Wendt (2004) for example.

The modeling approach followed in this thesis assumes that the degradation path of a unit cannot be monitored continuously during testing. Consequently, the underlying failure causing process is not fully observable. Functional forms of observed occurrences of degraded and critical failures are therefore obtained by investigating the first passage time distributions of the underlying failure causing process with regard to failure thresholds, called the first passage time problem. Accordingly, the variable of interest in this investigation is unit lifetime and the target of estimation is its lifetime distribution.

As a result, stochastic processes are discussed from sample paths and other related properties as well as lifetime estimation points of view. If the values of $\{X(t), t \in T\}$ are observed over the entire index set T, then the function $x = x(t)$ over the domain T is called a sample path (equivalently, a trajectory or a realisation) of the stochastic process. Candidate stochastic processes for $\{X(t), t \in T\}$ are the simple Wiener process (Brownian motion) and the Wiener process with drift.

But since unit degradation is defined as the irreversible accumulation of damage leading to unit failure, candidate stochastic processes for $\{X(t), t \in T\}$ are extended to strictly monotone stochastic processes. These include among others the maximum of the Wiener process and the gamma process. The problem is now to choose from among these stochastic processes the one that best describes unit degradation in a life test and the first passage time distributions.

3.2 The Wiener process model for unit degradation

The basis of the Wiener process (or Brownian motion) $\{B(t), t \in \mathbb{R}^+\}$ as a degradation model is that the degradation increment in an immeasurably small time interval can be regarded as the sum of a large number of small, independent random stress effects (additive superposition). Denote this sum by B_n such that $B_n = R_1 + R_2 + ... + R_n$ where the random variables R_i are independent but not necessarily identically distributed, having finite means $\mu_i = E(R_i)$ and finite variances $\sigma_i^2 = Var(R_i)$. Assume also that none of these n independent random variables dominates the rest. Then by the simplest variation of the central limit theorem, the standardization of the sum B_n, denoted by

$$Zn = \frac{B_n - E(B_n)}{\sqrt{Var(B_n)}} = \frac{B_n - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

converges under the Lindeberg condition (see Beichelt, 2006) to a normal distribution. That is,

$$\lim_{n\to\infty} P\left(Z_n \leq x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2}\right) du \tag{3.2}$$

where $\Phi(x)$ is the standard form of the normal distribution function. It is therefore reasonable to assert that the degradation increment $B(t + h) - B(t)$ over the time interval $(t, t + h)$ of a Wiener process is a random variable that is normally distributed. Under the assumption of additive accumulation of degradation, the increment $B(t + h) - B(t)$ is dependent on the length h of the time interval only and not on the time one begins observation in a life test. This assertion implies that the Wiener process $\{B(t), t \in \mathbb{R}^+\}$ is a homogeneous increment stochastic process with the following properties (Beichelt, 2006):

(1) For all $0 \leq s < t$, the degradation increment $B(s, t) = B(t) - B(s)$ is a normally distributed random variable with mean 0 and variance $\sigma^2(t - s)$, that is, $B \sim N(0, \sigma^2(t - s))$. Hence the process has homogeneous increments

(2) Let $[t_1, t_2]$, $[t_3, t_4]$, ... ,$[t_{n-1}, t_n]$ be disjoint time intervals for arbitrary $n \in \mathbb{Z}^+$. Then the increments $B(t_n) - B(t_{n-1})$, ... ,$B(t_4) - B(t_3)$ and $B(t_2) - B(t_1)$ are independent random variables distributed as described in property 1.

(3) $B(0) = 0$ with probability 1.

For being real-valued and having independent, homogeneous increments, the Wiener process is a Levy process.

3.2.1 Markov and sample paths properties

Given the present state of the Wiener process $B(t_0) = b_0$, the probability law governing the degradation increment $B(t + t_0) - B(t_0)$ is independent of any additional knowledge of values of past states $B(s)$ for $s < t_0$, called the Markov property of the process. Mathematically, the Markov property states that if $t_0 < t_1 < ... < t_n < t$, then

$$P\left(B(t) \leq b | B(t_n) = b_n, ..., B(t_1) = b_1, B(t_0) = b_0\right) = P\left(B(t) \leq b | B(t_n) = b_n\right)$$

In many applications, the Markov property is a reasonable assumption. As used here, unit degradation in a life test is an accumulation of damage over time that leads to failure. In this sense, it is reasonable to assume that unit degradation is continuous in time. Consequently, sample paths of the stochastic process describing unit degradation ought to be restricted to continuous functions. Continuity is a convergence property and different kinds of convergence for random variables exist. It therefore follows that a stochastic process $\{X(t), t \in \mathbb{R}^+\}$ can be considered continuous in various ways.

Definition 5.3.2: *A stochastic process* $\{X(t), t \in \mathbb{R}^+\}$ *is continuous*

(1) in mean square at t_0 *if*

$$\lim_{t \to t_0} \boldsymbol{E}\left(X(t) - X(t_0)\right)^2 \to 0.$$

(2) in probability at t_0 *if*

$$\lim_{t \to t_0} \boldsymbol{P}\left[|X(t) - X(t_0)| > \epsilon\right] \to 0.$$

(3) in the almost sure sense at t_0 *if*

$$\boldsymbol{P}\left\{\lim_{t \to t_0} X(t) = X(t_0)\right\} = 1.$$

Any one of these definitions of continuity of a stochastic process can be applied when describing sample path properties of a Wiener process.

Theorem 3.1: *A real-valued Wiener process* $\{B(t), t \in \mathbb{R}^+\}$ *has continuous sample paths almost surely.*

Proof: Let $t \in \mathbb{R}^+$ and $h > 0$. Without loss of generality, assume the degradation increment $B(t+h)-B(t)$ to be distributed as $N(0,h)$. It follows from Equation 3.2 that the random variable $Z = \frac{B(t+h)-B(t)}{\sqrt{h}} \sim N(0,1)$. Symmetry of the probability density function of the standard normal implies that all its odd moments are zero. By definition,

$$E(Z^r) = \int_{-\infty}^{\infty} z^r \left(\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right) dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{r-1} \left(ze^{-z^2/2} \right) dz$$

Letting $u = z^{r-1}$, $dv = ze^{-x^2/2}$ and using integration by parts, the expression for $E(Z^r)$ becomes

$$\begin{aligned} E(Z^r) &= \frac{1}{\sqrt{2\pi}} \left(-z^{r-1} e^{-z^2/2} \mid_{-\infty}^{+\infty} + (r-1) \int_{-\infty}^{\infty} z^{r-2} e^{-z^2/2} dz \right) \\ &= (r-1) \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{r-2} e^{-z^2/2} dz \right\} \\ &= (r-1) E(Z^{r-2}) \end{aligned}$$

Using the fact that $E(Z^0) = 1$, the recursive expression for the r^{th} moment in terms of r is given by

$$E(Z^r) = (r-1)(r-3)...(r-(r-3)(r-(r-1)) = \frac{r!}{\prod_{i=1}^{r/2} 2i} > 0.$$

It follows therefore that there must exist an $r > 2$ such that $E(|Z|^r) > 0$ and consequently $E(|B(t+h) - B(t)|^r) = h^{r/2} E(|Z|^r)$. Let $r = 2(1+\epsilon)$ where ϵ is a positive constant. Then $E(|B(t+h) - B(t)|^r) = Kh^{1+\epsilon}$ with $K = E(|Z|^r)$. By Kolmogorov's continuity theorem, there exists a modification and hence a version of $\{B(t), t \in \mathbb{R}^+\}$, say $\{\widetilde{B}(t), t \in \mathbb{R}^+\}$ whose paths are continuous. That is, for every $t \in \mathbb{R}^+$, $B(t) = \widetilde{B}(t)$ and consequently $P\{\lim_{t \to t_0} X(t) = X(t_0)\} = 1$. This completes the proof.

The general assumption that is often valid in many applications is that the physical degradation process is a continuous process. Because trajectories of a Wiener process are continuous functions, it is therefore not surprising that the Wiener process is the basic model for a degradation process. In this thesis however, the accumulation of damage throughout life testing is assumed to be irreversible. In particular, unit degradation in a life test is considered to be gradual and monotonically accumulating over time in a sequence of tiny positive increments. The development of failure is therefore described as follows. The degradation process gradually increases during testing until it reaches a failure threshold at which point the test unit fails. Hence only stochastic processes with strictly monotone sample paths can adequately describe such unit degradation in a life test.

3.3 Wiener maximum process for monotone degradation

One way of partially mitigating the tooths in the sample paths of a Wiener process is to describe the irreversible accumulation of damage in a life test by its supremum process. Observe that continuity of sample paths of the Wiener process implies that the maximum and minimum random variables are well defined on compact intervals. Let

$$B^+(t) = \sup_{0 \leq u \leq t} \{B(u), u \geq 0\} \tag{3.3}$$

and

$$B^-(t) = \inf_{0 \leq u \leq t} \{B(u), u \geq 0\} \tag{3.4}$$

denote the Wiener maximum process and the Wiener minimum process respectively. Replace the path $B(u)$ by its reflection $-B(u)$ which is also a Wiener process. Then the maximum and minimum are interchanged and consequently

$$B^+(t) \stackrel{d}{=} B^-(t)$$

where $\stackrel{d}{=}$ denotes equality in distribution. Hence it suffices to only determine the distribution of the random variable $B^+(t) = \sup_{0 \leq u \leq t} \{B(u), u \geq 0\}$. In order to determine the probability of the event $\left\{\sup_{0 \leq u \leq t} B(u) \geq s\right\}$ on the closed interval $[0, t]$, the approach adopted here is to use the first passage time.

3.3.1 First passage time distributions

Denote by T_s the time when the event $\{B(t) \geq s\}$ occurs. Put differently, it is the time when the stochastic process $\{B(t), t \in \mathbb{R}^+\}$ hits the failure threshold s for the first time. Continuity of sample paths of the process $\{B(t), t \in \mathbb{R}^+\}$ and the assumption that $B(0) = 0$ guarantees that the occurrence of the event $\{B(t) > s\}$ at time t implies that the event $\{T_s \leq t\}$ has already occurred. Written more formally,

$$\{T_s \leq t \text{ and } B(t) > s\} = \{B(t) > s\}.$$

But the occurrence of the event $\{T_s \leq t\}$ or its non-occurrence is known simply by observing the evolution of the process $\{B(t), t \in \mathbb{R}^+\}$ prior to time t. Hence T_s has the interpretation of a stopping time with respect to the process $\{B(t), t \in \mathbb{R}^+\}$.

Now, for the event of interest $\left\{\sup_{0\leq u\leq t} B(u) \geq s\right\}$ to occur, the process $\{B(t), t \in \mathbb{R}^+\}$ must have crossed the failure threshold s once or more in the closed interval $[0, t]$ given that $B(0) = 0$. It follows therefore that

$$\left\{\sup_{0\leq u\leq t} B(u) \geq s | B(0) = 0\right\} = \{T_s \leq t\} \tag{3.5}$$

since the process $\left\{\sup_{0\leq u\leq t} B(u), u \geq 0\right\}$ is non-decreasing. Using these observations and noting that the event $\{T_s \leq t\}$ can be written as a sum of disjoint events $\{T_s \leq t \text{ and } B(t) > s\}$ and $\{T_s \leq t \text{ and } B(t) < s\}$, then

$$\begin{aligned} P\{T_s \leq t\} &= P\{T_s \leq t, B(t) < s\} + P\{T_s \leq t, B(t) > s\} \\ &= P\{B(t) < s | T_s \leq t\} P\{T_s \leq t\} + P\{B(t) > s\}. \end{aligned} \tag{3.6}$$

Under the condition $B(0) = 0$, the increment $B(t) - B(0)$ is a normally distributed random variable and continuity of sample paths ensures that $B(T_s) = s$. Given that $T_s \leq t$, the Wiener process $\{B(t), t \in \mathbb{R}^+\}$ is equally likely to remain above or to fall below s at time t.

This is a consequence of the Wiener process being symmetric about the *x-axis* and that its future is independent of its past prior to T_s. Accordingly, $P\{B(t) < s|T_s \leq t\} = \frac{1}{2}$ and hence $P\{T_s \leq t\} = 2P\{B(t) > s\}$ from Equation 3.6 and the first passage time distribution is required.

Thus for the Wiener process $\{B(t), t \in \mathbb{R}^+\}$ and any $s \neq 0$,

$$\begin{aligned} P(T_s \leq t) &= 2P(B(t) \geq s) = \frac{2}{\sqrt{2\pi}} \int_{s/\sigma\sqrt{t}}^{\infty} e^{-\frac{u^2}{2}} du \\ &= 2\left[1 - \Phi\left(\frac{s}{\sigma\sqrt{t}}\right)\right]. \end{aligned} \tag{3.7}$$

where Φ is the standard normal distribution function. Now for the first passage time of the maximum of a Wiener process, observe that

$$\left\{\sup_{0 \leq u \leq t} B(u) \geq s|B(0) = 0\right\}$$

if and only if $T_s \leq t$. Consequently,

$$P\left\{\sup_{0 \leq u \leq t} B(u) \geq s|B(0) = 0\right\} = P(T_s \leq t)$$

and the formula for the distribution of the maximum of a Wiener process is as given in Equation 3.7.

3.4 Wiener process with drift model for unit degradation

But unit degradation generally has a non-zero mean. Hence an obvious improvement of the Wiener process as a degradation model is to include a mean or drift measure $\nu > 0$. This yields a one dimensional Wiener process with a fixed positive drift parameter ν and fixed variance parameter σ^2, denoted by $\{W(t), t \in \mathbb{R}^+\}$. It can be represented as

$$W(t) = \nu t + \sigma B(t) \tag{3.8}$$

where $B(t)$ is the standard Wiener process on $[0, \infty)$ capturing the stochastic movements of the degradation process. It therefore follows that $E[W(t)] = \nu t$ and $Var[W(t)] = \sigma^2 t$ such that $W(t) \sim N(\nu t, \sigma^2 t)$. For any n, the realisations of the random process $\{B(t), t \in \mathbb{R}^+\}$ at $t_1, t_2, ..., t_n$ are jointly Gaussian. Hence $\{B(t), t \in \mathbb{R}^+\}$ is a Gaussian process. Following Equation 3.8, $\{W(t), t \in \mathbb{R}^+\}$ is also a Gaussian process. When the drift parameter ν is zero, there is no degradation. Unless indicated otherwise, all test units having the same design are assumed to have common drift and variance parameters which admittedly is a strong assumption.

Degradation phenomena such as unit wear in a life test have the interpretation of accumulation of additive deterioration caused by higher than usual test stresses. Based on this additive accumulation of degradation assumption, a large number of authors use the Wiener process with drift $\{W(t), t \in \mathbb{R}^+\}$ to describe unit degradation. Statistical methods of estimating the parameters of the Wiener process with drift when analysing reliability of technical units are described in Kahle (1994), Kahle and Lehmann (1998) and Kahle and Lehmann (2010). Other examples include Aalen and Gjessing (2001) and the numerous references therein. Application of the Wiener process as a degradation model in accelerated testing include Whitmore and Schenkelberg (1997).

By the CLT, the degradation increment $W(t+h) - W(t)$ can also be reasonably assumed to have the same distribution as $W(h) - W(0)$ for any $h > 0$ if the stress applied in a life test is time independent. Written formally,

$$\{W(t+h) - W(t)\} \stackrel{d}{=} \{W(h) - W(0)\}$$

for all $h \in \mathbb{R}^+$ where $\stackrel{d}{=}$ stands for equality in distribution. But for time-varying stress tests such as step-stress tests, the increased stresses will likely produce time inhomogeneity in the degradation process. Accordingly, the Wiener process with drift may not adequately describe unit degradation for such tests.

3.4.1 First passage time distributions

The Wiener process with drift enjoys wide applications as a model for degradation phenomena mainly because of its mathematical advantages. The one of interest in this study is that it gives rise to mathematically tractable first passage time distributions. Observed occurrences of times to a critical failure X_2 in a life test have the interpretation of the first passage time to s_2 for the degradation process $\{W(t), t \geq 0\}$ regardless of whether or not unit lifetime is censored. That is, $X_2 = T_{s_2} = min\{t \in \mathbb{R}^+ : W(t) = s_2\}$. On the other hand observed occurrences of times to a degraded failure X_1 only have the interpretation of first passage times conditionally given unit lifetime is censored, that is, $X_1 < X_2$. This is because in theory, X_1 may potentially assume a value greater that X_2 when observation during testing is not stopped after the degradation process reached the failure threshold s_1.

But degraded and critical failures are competing as it were to terminate observation during life testing. It is therefore the first occurring failure mode that is observed while the other is only known to occur latter. That is, X_1 is observed provided $X_1 < X_2$ and similarly X_2 is observed if it is the smaller of the two. In this competing risks manner, observed occurrences of both X_1 and X_2 have the interpretation of first passage times of the degradation process $\{W(t), t \geq 0\}$ from zero to failure thresholds. Hence if the Wiener process with drift $\{W(t), t \in \mathbb{R}^+\}$ adequately describes system degradation leading to failure in a life test, then the observed occurrences of the competing risks variables are first passage times to failure thresholds for $\{W(t), t \geq 0\}$.

Assume the degradation process for a test unit satisfies Equation 3.8. Then the lifetime of the test unit is defined as the first time the process $\{W(t), t \in \mathbb{R}^+\}$ exceeds the failure threshold $s > 0$. Let T_s denote the random time the process $\{W(t), t \in \mathbb{R}^+\}$ exceeds the failure threshold s. Then $T_s = inf\, \{t \in \mathbb{R}^+ : W(t) > s\}$. It is well established (see e.g. Chhikara and Folks, 1989) that the first passage time of a Wiener process with drift from zero to a deterministic failure threshold is distributed as inverse Gaussian with density function

$$f(t) = \frac{s}{\sigma\sqrt{2\pi t^3}} \exp\left\{-\frac{1}{2\sigma^2}\frac{(s-\nu t)^2}{t}\right\}, \quad t > 0, \quad \nu > 0. \tag{3.9}$$

Because T_s has the interpretation of a first passage time, it follows therefore that it has the potential of being useful in lifetime studies. For fixed s, a useful parameterisation of the density in Equation 3.9 in terms of the development of statistical properties analogous to those of the normal distribution (Tweedie, 1957a) is obtained by setting $\mu = \frac{s}{\nu}$ and $\lambda = \frac{s^2}{\sigma^2}$. Under this parameterisation, the density function of the inverse Gaussian random variable T_s is

$$f(t,\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi t^3}}\text{exp}\left\{-\frac{\lambda}{2\mu^2}\frac{(t-\mu)^2}{t}\right\}, \quad t > 0 \tag{3.10}$$

where it is assumed that $\mu \in \mathbb{R}^+$ and $\lambda \in \mathbb{R}^+$. The corresponding cumulative distribution function (CDF) is given by

$$F(t) = \Phi\left\{\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu}-1\right)\right\} + \text{exp}\left(\frac{2\lambda}{\mu}\right)\Phi\left\{-\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu}+1\right)\right\} \tag{3.11}$$

where $\Phi(t)$ is the CDF of the standard normal distribution.

The moment generating function of the inverse Gaussian distributed random variable $T_s \sim IG(\mu,\lambda)$, denoted as $M_{T_s}(\omega) = E\left[e^{\omega T_s}\right]$ is given by

$$M(\omega) = \text{exp}\left[\frac{\lambda}{\mu}\left(1-\sqrt{1-\frac{2\omega\mu^2}{\lambda}}\right)\right], \quad \omega < \frac{\lambda}{2\mu^2}. \tag{3.12}$$

It follows therefore that the r^{th} moment of the positive-valued random variable T_s is the r^{th} derivative of the moment generating function of $T_s \sim IG(\mu,\lambda)$ evaluated at $\omega = 0$. By letting $\alpha(\omega) = \sqrt{1-\frac{2\omega\mu^2}{\lambda}}$, the first moment is

$$\begin{aligned} M^{'}(\omega) &= \frac{d}{d\omega}\text{exp}\left[\frac{\lambda}{\mu}(1-\alpha(\omega))\right] = e^{\frac{\lambda}{\mu}}\left[\frac{d}{d\alpha(\omega)}e^{-\frac{\lambda}{\mu}\alpha(\omega)} \times \frac{d\alpha(\omega)}{d\omega}\right] \\ &= e^{\frac{\lambda}{\mu}}\left[-\frac{\lambda}{\mu}e^{-\frac{\lambda}{\mu}\left(1-\frac{2\omega\mu^2}{\lambda}\right)^{\frac{1}{2}}} \times \frac{1}{2}\left(1-\frac{2\omega\mu^2}{\lambda}\right)^{-\frac{1}{2}} \times \left(-\frac{2\mu^2}{\lambda}\right)\right]\Bigg|_{\omega=0} = \mu \end{aligned}$$

Hence the parameter μ is the distribution mean while λ is the scaling parameter. Similarly, the second moment

$$M^{''}(\omega) = \left[\frac{d}{d\omega}\left(e^{\frac{\lambda}{\mu}\left[1-\left(1-\frac{2\omega\mu^2}{\lambda}\right)^{\frac{1}{2}}\right]} \times \mu\left[1-\frac{2\omega\mu^2}{\lambda}\right]^{-\frac{1}{2}}\right)\right]\Bigg|_{\omega=0} = \mu^2 + \frac{\mu^3}{\lambda}$$

implies that $\frac{\mu^3}{\lambda}$ is the variance for the inverse Gaussian distribution. However, there are several other forms of the inverse Gaussian distribution in the literature. Tweedie (1957a) for example gave three forms of 3.10 which he generated by replacing the parameters (μ, λ) by any of (η, λ), (μ, ϕ) or (ϕ, λ) where $\frac{1}{2}\eta^2 = \mu = \frac{\lambda}{\phi}$. Each of these four forms was demonstrated by Tweedie to be useful in some applications.

The parameters μ and λ have the same physical dimensions as the first passage time T_s since a change of scale of T_s will result in both μ and λ being multiplied by the same factor as T_s. Consequently, this results in a new member of the family. On the other hand, the parameter $\phi = \lambda/\mu$ which determines the shape of the distribution is invariant to any scale transformation of T_s. This scale invariant property of the shape parameter ϕ may be useful when analysing accelerated testing data. This is particularly so for cases where the failure causing mechanism as represented by the distribution's shape parameter is assumed to remain the same at all stress levels.

3.4.2 Approximation for monotone degradation

The Wiener process with drift has been widely applied in accelerated life and degradation testing of technical units since it adequately describes most physical phenomena. Often, it may however be the case that the degradation process for a test unit be regarded as gradual and monotonically accumulating over time. In this case, the test unit can be returned to its original state or at least improved by external repair actions only, otherwise deterioration proceeds only in one direction. Since the Wiener process with drift is not a monotone stochastic process, its application to degradation processes that exhibit monotone behaviour presents practical problems (see e.g. Si, Wang, Hu & Zhou, 2011).

In particular, application of the Wiener process with drift to describe the degradation process for test units would imply that the quality characteristic, defined as measured unit performance can increase or decrease during testing. Clearly such behaviour has no physical justification. Thus the matter of why the Wiener process with drift remains popular with practitioners as a model for degradation despite this bi-directional characteristic feature requires a little elaboration.

In a number of applications, the degradation process of interest is only required to be a continuous process and can therefore be described by any stochastic process with continuous sample paths.

When it is important to assume that the degradation process of a test unit is monotonic (see e.g. Whitmore and Schenkelberg, 1997 and the references therein), the Wiener process with drift is proposed only as an approximation. The approximation is good especially when the diffusion parameter is small relative to the drift (mean) parameter. In this case, the tooths in the evolving paths of the Wiener process are significantly smoothed out and the sample paths become approximately monotone. Besides, the wide application of the Wiener process with drift as a degradation model is also due to mathematical convenience. Given its close relation to the normal distribution, the Wiener process facilitates easy computations.

Alternatively, the irreversible accumulation of damage in a life test may be described by the maximum of the Wiener process with drift. The Wiener maximum process

$$W^{+}(u) = \left\{ \sup_{0 \le u \le t} W(u), u \ge 0 \right\}$$

is non-decreasing in its argument and has initial condition $W^{+}(0) = 0$. A test unit experiencing the Wiener maximum process $\left\{ \sup_{0 \le u \le t} W(u), u \ge 0 \right\}$ in a life test fails the first time sample paths of the degradation process hits a failure threshold s. The first passage time T_s is thus defined as

$$T_s = min \left\{ u \in \mathbb{R}^{+} : \sup_{0 \le u \le t} W(u) = s \right\} \tag{3.13}$$

and this coincide with the first passage time of the process $\{W(t), t \in \mathbb{R}^{+}\}$ to the same failure threshold s. Hence the observed occurrences of degraded and critical failure in a life test assuming the maximum of a Wiener degradation process are also distributed as inverse Gaussian with density as in Equation 3.10.

3.5 Gamma process model for monotone degradation

A natural way of describing a stochastic degradation process that proceeds in one direction is often to consider it as a gamma process (see e.g. Abdel-Hameed, 1975). Mathematically, the gamma process is defined from a gamma distributed random variable G as follows: Assume the random variable G has a gamma distribution with density

$$Ga(g|\beta, \alpha) = \frac{\alpha^{\beta}}{\Gamma(\beta)} g^{\beta-1} e^{-\alpha g}, \;\; g \geq 0,$$

where $\Gamma(\beta) = \int_{y=0}^{\infty} y^{\beta-1} e^{-y} dy$ is the gamma function, $\alpha > 0$ is the scale parameter and $\beta > 0$ is the shape parameter. Henceforth, the notation $G \sim Ga(\beta, \alpha)$ implies that the random variable G is distributed as gamma with shape and scale parameters β and α respectively. There is however no loss in generality if only gamma processes with scale parameter 1 are considered.

Let $\beta(t)$, $t \geq 0$ be a strictly increasing, right continuous real-valued shape function with initial condition $\beta(0) \equiv 0$. Then the gamma process with scale parameter $\alpha > 0$ and shape function $\beta(t) > 0$ is a continuous time stochastic process $\{G(t), t \geq 0\}$ satisfying the following properties:

(1) $G(0) = 0$ with probability one,

(2) For all $0 \leq s < t < \infty$, the degradation increment $G(t) - G(s)$ is a gamma distributed random variable with shape parameter $(\beta(t) - \beta(s))$ and scale parameter α. Hence the quantity $G(t) - G(s)$ is non-negative.

(3) For any choices $0 \leq s < t < u < \infty$, the random variables $G(t) - G(s)$, $G(u) - G(t)$ are independent.

In summary, a gamma process is a continuous time stochastic process whose increments are non-negative, independent and distributed as gamma. Property 2 is a direct consequence of the infinite divisibility of the gamma distribution. That is if G is distributed as gamma, then for every $n \in \mathbb{N}$, there exists i.i.d. random variables $Y_1, Y_2, ..., Y_n$ such that $G \stackrel{d}{=} Y_1 + ... + Y_n$ where $\stackrel{d}{=}$ stands for equality in distribution (Sato, 1999). Consequently, the degradation increments of a gamma process and their cumulative sum are distributed as gamma. Property 3 implies that the gamma process is Markovian: given the current state $G(s)$, the process proceeds to a future state $G(t)$ where $t > t$ independently of all states before s.

3.5.1 Sample paths properties of a gamma process

If the shape function $\beta(t) > 0$ is linear, then the gamma process $\{G(t), t \geq 0\}$ has homogeneous increments. Regarding sample paths properties, recall from the properties of the gamma process that for all $0 \leq s < t < \infty$, the degradation increment $G(t) - G(s)$ is a random variable distributed as gamma. Because the increment is gamma distributed, it is therefore never negative with probability 1 and consequently $G(t) > G(s)$ almost surely. Hence the trend function is increasing and sample paths of a gamma process are continuous. It is therefore an appropriate stochastic process model for gradual damage that accumulate over time in a sequence of tiny positive increments. Simulation methods for a gamma process are presented in van Noortwijk (2009).

3.5.2 First passage time distributions

Assume the gamma process is the appropriate model for the stochastic deterioration of units in a life test. Then the degradation of a test unit at time t can be modeled by a gamma process $G(t)$ with positive shape parameter β and scaling parameter α. Assume also that the the gamma process has a starting value $G(0) = g(0) > 0$. Then the level of degradation (cumulative damage) of the test unit at time t is given by

$$G(t) - G(0) = D(t)$$

such that $D(0) = 0$ with probability one. Consequently, the stochastic process $\{D(t), t \geq 0\}$ is a shifted gamma process with shape parameter $\beta > 0$ and scale α on account of $G(t)$ being a stationary gamma process. The test unit fails when its degradation process $G(t)$ reaches a certain known failure threshold s and its failure time T is defined as the first passage time of $G(t)$ to s. Because $G(t)$ has non-decreasing sample paths, the events $\{T > t\}$ and $\{G(t) < s\}$ are equivalent and consequently

$$P(T > t) = P\left[G(t) < s\right] = P\left[\{G(t) - g(0)\} < \{s - g(0)\}\right]$$

where the increment $G(t)-g(0)$ is distributed as gamma with shape coefficient βt if $\beta(t)$ is linear and scale α. Accordingly,

$$\begin{aligned} P(T>t) &= \int_0^{s-g(0)} \frac{\alpha^{\beta t}}{\Gamma(\beta t)} g^{\beta t-1} \exp(-\alpha g) dg \\ &= \frac{1}{\Gamma(\beta t)} \int_0^{s_\alpha} \xi^{\beta t-1} e^{-\xi} d\xi \end{aligned}$$

where $\xi = \alpha g$ and $s_\alpha = \alpha(s-g(0))$. The cumulative distribution function of the first passage time T (unit lifetime) is thus given by

$$F_T(t; g(0), s) = \frac{\gamma(\beta t, s_\alpha)}{\Gamma(\beta t)} \tag{3.14}$$

where $\gamma(b,w)$ denotes the upper incomplete gamma function $\gamma(b,w) = \int_w^\infty u^{b-1} e^{-u} du$. The exact pdf of the first passage time T of the gamma process $G(t)$ to some deterministic failure threshold s has already been derived from Equation 3.14 (see e.g. Park and Padgett, 2005). However, the derived distribution is not feasible for applications.

Observe that continuous time processes with homogeneous increments are often regarded as continuous time versions of partial sum processes. In this sense, the first passage time to s of the gamma process can be regarded as a discrete time version of T. Let N denote this discrete first passage time. Then by the central limit theorem, the exact distribution of T may be approximated by a continuous version of N. This idea can be traced back to the work of Birnbaum and Saunders (1969) and was also adopted by Park and Padgett (2005). It is summarised here as follows:

Let $Y_i = G(i+1) - G(i)$ denote the increments of a gamma process. Then $Y_1, Y_2, ...$ are independent gamma random variables and the partial sum process $\{G_n;\ n \in \mathbb{N}\}$ such that

$$G_n = \sum_{i=1}^{n} Y_i, \ \ n \in \mathbb{N}; \ \ G_0 \equiv 0 \tag{3.15}$$

has the interpretation of the magnitude of accumulated degradation up to n. As a direct consequence of Equation 3.15, the process $\{G_n;\ n \in \mathbb{N}\}$ has the following properties.

(1) For $0 < n_0 < n_1 < ...;\ \ G_{n_0}, G_{n_1} - G_{n_0}, ...$ are independent.

(3) For any choices $m, n \in \mathbb{N}$; $G_{m+n} - G_m \stackrel{d}{=} G_n$.

Thus $\{G_n;\ \ n \in \mathbb{N}\}$ is a discrete time process with homogeneous increments. On account of Y_i being increments of a pure positive jump process, it follows that the partial sum process $\{G_n;\ n \in \mathbb{N}\}$ is non-decreasing. As a result the events $\{N > n\}$ and $\{G_n < s\}$ are equivalent and consequently

$$P(N > n) = P(G_n < s) = P\left[\sum_{i=1}^{n} Y_i < s\right]. \tag{3.16}$$

The degradation increments $Y_i = G(i+1) - G(i)$ may be made sufficiently small (presumably microscopic) and further assumed to be independent and identically distributed random variables having the same finite mean μ and finite variance σ^2. Then the magnitude of accumulated degradation $G_n = \sum_{i=1}^{n} Y_i$ has mean value

$$E\left(G_n\right) = E\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n}\left[E(Y_i)\right] = n\mu$$

and variance

$$Var\left(G_n\right) = Var\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n}\left[Var(Y_i)\right] = n\sigma^2.$$

Denote by Z_n the standardisation of G_n such that

$$Z_n = \frac{G_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$\lim_{n\to\infty} P(Z_n \le y) = \Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{w^2}{2}}\,du.$$

These conditions imply that by the central limit theorem, the magnitude of accumulated degradation $G_n = \sum_{i=1}^{n} Y_i$ is approximately normally distributed with mean $n\mu$ and variance $n\sigma^2$ as $n \to \infty$. Thus the distribution of the discrete first passage time N of G_n to s from Equation 3.16 assuming the partial sum process $\{G_n;\ \ n \in \mathbb{N}\}$ has starting value g_0 is

$$
\begin{aligned}
P(N \leq n) &\cong 1 - P\left(\{G_n - g_0\} < \{s - g_0\}\right) = 1 - \Phi\left(\frac{s - g_0 - n\mu}{\sigma\sqrt{n}}\right) \qquad (3.17)\\
&= \Phi\left(\frac{\mu\sqrt{n}}{\sigma} - \frac{s - g_0}{\sigma\sqrt{n}}\right).
\end{aligned}
$$

Denote by T the continuous approximation of the discrete first passage time N in Equation 3.17. Birnbaum and Saunders (1969) prove that the continuous random variable T has the Birnbaum-Saunders type distribution

$$
F_T(t; g(0), s) = \Phi\left[\frac{1}{\beta^*}\left(\sqrt{\frac{t}{\alpha^*}} - \sqrt{\frac{\alpha^*}{t}}\right)\right] \qquad (3.18)
$$

with support $(0, \infty)$ where $\alpha^* = s_\alpha/\beta$, $\beta^* = 1/\sqrt{s_\alpha}$ and $s_\alpha = \alpha(s - g_0)$. But this derivation of the continuous version of the distribution of N suggests that the accumulated degradation $G_n \sim N(n\mu, n\sigma^2)$ for large n. Consequently, it must therefore assume negative values with non-zero probability. A remark is however in order here. Degradation of test units in a life test is ideally a non-negative random variable and so are the degradation increments. In this sense, it is therefore reasonable to assert that the accumulated degradation G_n assumes negative values with zero probability even though it is approximately normally distributed. In line with this observation, Birnbaum and Saunders (1969) assumed that T is a continuous non-negative random variable in their original derivation.

The relation between the Birnbaum-Saunders distribution and the inverse Gaussian distribution is discussed in Bhattacharyya and Fries (1982). The derivation of the latter involves approximations while the former is an exact first passage time distribution to a failure threshold of a Wiener process with drift. If the normally distributed random variable G_n is assumed to take on negative values with zero probability (non-negative degradation increments), then the distribution functions of the Birnbaum-Saunders and inverse Gaussian distributions are identical. For more details on why the Birnbaum-Saunders distribution is an approximation to the inverse Gaussian distribution with mean $\mu = \alpha^* > 0$ and positive scale parameter $\lambda = s_\alpha^2/\beta$, see Bhattacharyya and Fries (1982). In particular, the approximation is good for large values of the drift (mean) parameter relative to the diffusion parameter of the Wiener process with drift. Consequently, both lifetime distributions may fit test data equally well and are both flexible since they admit different shapes.

3.6 Marker processes and degradation phenomena

To this end, reliability assessments of industrial units when failure in a life test is defined in terms of the observed level of unit degradation has been discussed. Specifically, removal of a unit from observation in a life test is considered to occur the first time the level of unit degradation exceed a failure threshold during testing. Stochastic process models namely the Wiener process with drift, the Wiener maximum process and the gamma process have been discussed as models for describing the accumulation of damage in a life test that leads to unit failure. The lifetime of test units assuming these stochastic process models is of first passage type and accordingly, the inverse Gaussian and the Birnbaum-Saunders distributions are motivated as lifetime models.

Recall that the discussed Wiener and gamma process models for degradation both satisfy the Markov property and are therefore Markov processes. Since the failure causing process is assumed to be not directly observable, a more general statistical and structural approach is to adopt the framework of hidden Markov process (HMP) for modeling the irreversible accumulation of damage in a life test. Being a class of Markov processes, the definition of the HMP requires the Markov process to be introduced first.

A stochastic process $\{X(t), t \in \mathbf{T}\}$ taking values in $\Sigma \subset \mathbb{R} = (-\infty, +\infty)$ is a Markov process with state space Σ if for all ordered (n+1)-tuples $t_1 < t_2 < ... < t_{n+1}$, with $t_i \in \mathbf{T}$ and for any $A_i \leq \Sigma; i = 1, ..., n+1$;

$$P\left[X(t_{n+1}) \in A_{n+1} | X(t_n) \in A_n, ..., X(t_1) \in A_1\right] = P\left[X(t_{n+1}) \in A_{n+1} | X(t_n) \in A_n\right]. \quad (3.19)$$

This implies that the future development of the Markov process depends only on its present value.

A HMP is a doubly stochastic process having an underlying stochastic process whose states are not observable (hence hidden) but can only be observed through another stochastic process called an observation process. Both the underlying and the observation processes are assumed to be Markov processes. This thesis could give a thorough treatment of HMP and the associated problems but to do that would however obscure the study's main focus. Rather the consequences of adopting the HMP modeling framework and hence a bivariate stochastic process as a degradation model are discussed within the context of accelerated life testing. Specifically, emphasis is placed on selecting a suitable probability structure for the doubly stochastic process model.

The implication of adopting this modeling framework is that the irreversible accumulation of damage in a life test is regarded as a latent variable (an unobservable construct) describing the process leading to unit failure. What is observed during testing are manifestations of damage (surrogates) caused by the underlying failure causing process. General observable surrogates when assessing lifetimes of technical units include measured wear, crack growth, corrosion etc. In the present application, the observable surrogate is the performance degradation process which is assumed to decrease with time during testing.

More formally, the idea is to consider the degradation process as an interplay between the latent failure causing process and the performance degradation process acting as a marker. The latent failure causing process is governed by some random mechanism and is thus described by a stochastic process $\{M(t), t \in \mathbb{R}^+\}$. The marker process on the other hand can be reasonably assumed to be a function of time since measured performance decreases with time during testing. Accordingly, it is also governed by some random mechanism and is thus best described by a stochastic process $\{R(t), t \in \mathbb{R}^+\}$, called the performance degradation process. Under this modeling viewpoint, the degradation process in a life test is specified as a bivariate stochastic process $\{M(t), R(t), t \in \mathbb{R}^+\}$ and a test unit fails when $\{M(t), t \in \mathbb{R}^+\}$ first reaches a failure threshold s.

The specification of the degradation process $\{M(t), R(t), t \in \mathbb{R}^+\}$ implies that stochastic processes $\{M(t), t \in \mathbb{R}^+\}$ and $\{R(t), t \in \mathbb{R}^+\}$ must be related in some way. In particular, the marker process $\{R(t), t \in \mathbb{R}^+\}$ must be a useful predictor of the latent failure causing process $\{M(t), t \in \mathbb{R}^+\}$. Why $\{M(t), t \in \mathbb{R}^+\}$ is assumed not to be a fully observable process follows from the observation that $\{R(t), t \in \mathbb{R}^+\}$ may not be monitored continuously during testing. Rather, observations on $\{R(t), t \in \mathbb{R}^+\}$ may only be at discrete times so that data on the process $\{M(t), t \in \mathbb{R}^+\}$ are impossible to collect. Both $\{M(t), t \in \mathbb{R}^+\}$ and $\{R(t), t \in \mathbb{R}^+\}$ are however restricted to stochastic processes with continuous sample paths. But because $\{M(t), t \in \mathbb{R}^+\}$

is the phenomenon of degradation, it is further required to be non-decreasing in t while no such requirement is necessarily imposed on the observation (or marker) process $\{R(t), t \in \mathbb{R}^+\}$.

3.6.1 Probabilistic structure of the bivariate process model

The question of how observations on the observable process $\{R(t), t \in \mathbb{R}^+\}$ may be used to make inferences about the unobservable process of interest $\{M(t), t \in \mathbb{R}^+\}$ has attracted attention in different application areas. In particular, the problem of selecting a probabilistic structure for the bivariate process $\{M(t), R(t), t \in \mathbb{R}^+\}$ from among the many possible choices is at the core of most studies. Intuitively, considering the degradation process as a bivariate structure $\{M(t), R(t), t \in \mathbb{R}^+\}$ makes the Markov additive process (MAP) a natural choice. A bivariate stochastic process $\{M(t), R(t), t \in \mathbb{R}^+\}$ is a MAP with continuous time parameter t if

(1) $\{M(t), t \in \mathbb{R}^+\}$ and $\{R(t), t \in \mathbb{R}^+\}$ are mean-square continuous at t_0. That is,

$$\lim_{t \to t_0} \mathbf{E} \left(M(t) - M(t_0)\right)^2 \to 0$$

and

$$\lim_{t \to t_0} \mathbf{E} \left(R(t) - R(t_0)\right)^2 \to 0.$$

(2) $\{M(t), t \in \mathbb{R}^+\}$ is non-negative and is non-decreasing in t,

(3) $\{R(t), t \in \mathbb{R}^+\}$ takes values in the state space Σ which is ether countable or $\Sigma \in \mathbb{R}$.

The importance of a MAP is captured in *Theorem 2.22* of Cinlar (1972) which states that if the bivariate stochastic process $\{M(t), R(t), t \in \mathbb{R}^+\}$ is a MAP, then

(1) $\{R(t), t \in \mathbb{R}^+\}$ is a Markov process with state space Σ, and

(2) The probability law of $\{M(t), t \in \mathbb{R}^+\}$ given $\{R(t), t \in \mathbb{R}^+\}$ is that of a process that can be represented as a sum of non-negative independent increments.

Property 1 of a MAP implies that $\{R(t), t \in \mathbb{R}^+\}$ may be any Markov process while Property 2 restricts $\{M(t), t \in \mathbb{R}^+\}$ to increasing Levy processes only. The implication in the present application is that the underlying failure causing process $\{M(t), t \in \mathbb{R}^+\}$ in a life test may be considered an increasing Levy process while measured performance (observable surrogate) can be represented by a Markov process. But the consequences of regarding $\{M(t), R(t), t \in \mathbb{R}^+\}$ as a MAP and using observations on the Markov process $\{R(t), t \in \mathbb{R}^+\}$ to make inferences about the failure causing process $\{M(t), t \in \mathbb{R}^+\}$ have no developed statistical theory and is thus an open problem.

Whitmore, Crowder and Lawless (1998) represented $\{M(t), R(t), t \in \mathbb{R}^+\}$ by a bivariate Wiener process. One Wiener process is taken to represent the marker process while the other represents the unobservable failure causing process. Assuming the bivariate Wiener process model implies that a test unit fails when the hidden process crosses a failure threshold. Its main advantage is that statistical inference can be done using data on both the marker process and the times to failure. However, the drawback of adopting this probability structure is that the latent failure causing process $\{M(t), t \in \mathbb{R}^+\}$ is no longer non-decreasing in t when it is assumed to be a Wiener process. As a result, this construction is not adequate when describing monotone degradation as is the case here. In addition, the link between the marker and the failure causing processes is not obvious. Hence the usefulness of the marker process in terms of tracking progress of the underlying failure causing process $\{M(t), t \in \mathbb{R}^+\}$ is subject to debate.

A minor modification of the bivariate process model proposed by Whitmore *et al.* (1998), see for example Singpurwalla (2006b) is to describe the marker process $\{R(t), t \in \mathbb{R}^+\}$ by a Wiener process with drift $\{W(t), t \in \mathbb{R}^+\}$ and the latent failure causing process $\{M(t), t \in \mathbb{R}^+\}$ by its maximum process $W^+(t) = \left\{\sup_{0 \le u \le t} W(u), u \ge 0\right\}$. This probability structure has the following advantages. First, the link between the marker and the underlying failure causing processes is obvious from

$$W^+(t) = \left\{\sup_{0 \le u \le t} W(u), u \ge 0\right\}.$$

Second, both $\{W(t), t \in \mathbb{R}^+\}$ and $\left\{\sup_{0 \le u \le t} W(u), u \ge 0\right\}$ have continuous sample paths and in addition, the latter is non-decreasing in t as required.

3.7 Degradation model and first passage time distributions

The more general degradation modeling viewpoint which treats the degradation process as an interplay between the unobservable failure causing process and the marker process is adopted in this thesis. In particular, system degradation in a life test is considered to be described by a bivariate process $\{W^+(t), W(t), t \in \mathbb{R}^+\}$ where $\{W^+(t), t \geq 0\}$ is the Wiener maximum failure causing process and $\{W(t), t \geq 0\}$ is the Wiener process with drift acting as a marker. Singpurwalla (2006, 2006b) also proposed this probability structure. In his construction however, the latent failure causing process is the cumulative hazard process, a continuously increasing process. Unit failure occurs when the cumulative hazard process first hits a failure threshold, assumed to be random. The uncertainty about this random threshold is further assumed to be described by an exponential distribution.

While assuming a random threshold may be reasonable in theory, it is usually not the case in practice and at least it is assumed so in the present application as follows. Adequate performance of a unit is often specified by industrial standards and a unit fails during testing when performance no longer conforms to the set standard. Hence it is reasonable to assume that the failure threshold is deterministic, otherwise failure during testing will not be well defined. Thus test units experience the Wiener maximum process

$$W^+(t) = \left\{ \sup_{0 \leq u \leq t} W(u), u \geq 0 \right\} \tag{3.20}$$

during testing and will fail if this underlying failure causing process first crosses a fixed failure threshold s. The lifetime of the unit is estimated by obtaining the first passage time of the Wiener maximum process $\{W^+(t), t \geq 0\}$ over the deterministic threshold. Since unit lifetime is of the first passage, no additional degradation data may necessarily be required in order to assess the reliability of the industrial unit. That is, the lifetime of the unit in a life test denoted by T_s is defined as

$$T_s = inf \left\{ u \in \mathbb{R}^+ : \sup_{0 \leq u \leq t} W(u) \geq s \right\}.$$

As a direct consequence of Equation 3.20, the first passage time of $W^+(t)$ to s coincides with the first crossing time of $W(t)$ to the same failure threshold s. Hence observed occurrences of degraded and critical failures in a life test assuming unit degradation is adequately described by the Wiener maximum process $W^+(t)$ are distributed as inverse Gaussian with density

$$f_j(x_j;\mu,\lambda) = \begin{cases} \sqrt{\frac{\lambda}{2\pi x_j^3}}\text{exp}\left(-\frac{\lambda(x_j-\mu)^2}{2\mu^2 x_j}\right); & x_j > 0, \;\; j \in (1,2) \\ 0 & \text{otherwise.} \end{cases}$$

since they both have the interpretation of stopping times. Consequently, the subdistribution functions of the competing risks variables are inverse Gaussian.

It must be highlighted however that the failure rate of the inverse Gaussian distribution is not monotonic. It initially increases to a maximum, and then decreases to a nonzero asymptotic value as the testing time goes to infinity. Early unit failures tend to dominate the lifetime distribution in ALT. As a result, the failure rate is expected to initially increase and latter decrease with testing time thereby exhibiting a non-monotonic behaviour. In applications where there is apparent skewness in the data, the inverse Gaussian distribution is a possible choice as a lifetime model.

The failure rate of the log normal distribution qualitatively exhibits the same behaviour. But unlike the log normal distribution, the inverse Gaussian distribution has physical justification as first passage time distribution. It also represents a wider class of lifetime distributions ranging from highly skewed to almost increasing failure rate (symmetrical) distributions as the shape parameter ϕ varies from 0 to ∞. The observations explain why as a lifetime model, the inverse Gaussian distribution is generally preferred in practice to the log normal distribution.

3.8 Statistical inference when barrier is assumed known

Based on the first passage time to a deterministic barrier of the failure causing process, observed occurrences of degraded and critical failures at all stress levels in a life test have been postulated to be distributed as inverse Gaussian with the bivariate parameter $\theta = (\mu, \lambda)^T$. Test data collected at each test stress level are utelised to estimate θ. Because of its well-known asymptotic distributional optimality properties, maximum likelihood (ML) remains the standard approach to statistical inference. The idea behind ML is to choose estimators $\hat{\theta} = (\hat{\mu}, \hat{\lambda})^T$ from among all possible values for θ that most likely produced the collected test data.

In general, if $t_1, t_2, ..., t_n$ is a random sample from an inverse Gaussian population with mean μ and scaling parameter λ, the loglikelihood function is given by

$$\mathscr{L}_n(\theta|t_1, ..., t_n) = \frac{n}{2}ln\lambda - \frac{n}{2}ln(2\lambda) - \frac{3}{2}\sum_{i=1}^{n} ln(t_i) - \frac{\lambda}{2\mu^2}\sum_{i=1}^{n}.\frac{(t_i-\mu)^2}{t_i} \tag{3.21}$$

Maximum likelihood estimates of μ and λ are obtained by maximising the likelihood function in Equation 3.21. Traced back to Schrodinger (1915), these estimators are well-known to be given by

$$\hat{\mu} = \overline{T} = \frac{1}{n}\sum_{i=1}^{n} T_i$$

and

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n}\left(\frac{1}{T_i} - \frac{1}{\overline{T}}\right)}. \tag{3.22}$$

In life testing however, the testing period often ends while some units are still to fail. As a result, test data are often right censored and the loglikelihood takes the form

$$\mathscr{L}(\mu, \lambda) = \sum_{i=1}^{n} \delta_i ln(f(t_i; \mu, \lambda)) + (1-\delta_i)ln\left[1 - F(t_i; \mu, \lambda)\right]$$

where

$$\delta_i = \begin{cases} 1 & \text{if unit is uncensored} \\ 0 & \text{if unit is censored.} \end{cases}$$

Consequently, the loglikelihood function assuming a sample from the inverse Gaussian population becomes

$$\mathscr{L}_n(\mu,\lambda) = \sum_{i=1}^{n}(1-\delta_i)ln\left\{\Phi(A_i) + \exp\left(\frac{2\lambda}{\mu}\right)\Phi(B_i)\right\} + \sum_{i=1}^{n}\delta_i\left\{\frac{1}{2}ln\lambda - \frac{1}{2}ln(2\pi) - \frac{3}{2}ln(t_i) - \lambda\frac{(t_i-\mu)^2}{s\mu t_i}\right\} \tag{3.23}$$

where $A_i = \sqrt{\frac{\lambda}{t_s}}\left(\frac{t_s}{\mu} - 1\right)$ and $B_i = -\sqrt{\frac{\lambda}{t_s}}\left(\frac{t_s}{\mu} + 1\right)$ The derivatives of $\ell_n(\mu,\lambda)$ with respect to μ and λ are messy but can be evaluated in a straightforward way (see e.g. Lemeshko, Lemeshko, Akushkina, Nikulin and Saaidia, 2010). But in the present competing risks situation, the contribution to the likelihood function when unit lifetime is censored during testing is the subdensity function of X_1. It is given by

$$f_1^*(x_1;\mu,\lambda) = q\sqrt{\frac{\lambda}{2\pi x_1^3}}\exp\left(-\frac{\lambda(x_1-\mu)^2}{2\mu^2 x_1}\right).$$

Similarly, the contribution to the likelihood function when unit lifetime is observed during testing is the subdensity function of X_2 given by

$$f_2^*(x_2;\mu,\lambda) = (1-q)\sqrt{\frac{\lambda}{2\pi x_2^3}}\exp\left(-\frac{\lambda(x_2-\mu)^2}{2\mu^2 x_2}\right).$$

Contributions to the likelihood function for a different parameterisation of the inverse Gaussian distribution is found in Lindqvist and Skogsrud (2009) for example. Assuming the observed competing risks data at each test stress level are $(z_1,...,z_N) = (x_{11},...,x_{1n};x_{21},...,x_{2m})$, then the likelihood function is given by

$$\begin{aligned} L &= \prod_{i=1}^{n} f_1^*(x_{1i})\prod_{j=1}^{m} f_2^*(x_{2j}) \\ &= q^n(1-q)^m\left(\frac{\lambda}{2\pi}\right)^{\frac{n+m}{2}}\left(\prod_{i=1}^{n}x_{1i}\right)^{-3/2}\left(\prod_{j=1}^{m}x_{2j}\right)^{-3/2} \\ &\quad\times\ \exp\left(-\sum_{i=1}^{n}\frac{\lambda(x_{1i}-\mu)^2}{2\mu^2 x_{1i}} - \sum_{i=1}^{n}\frac{\lambda(x_{2j}-\mu)^2}{2\mu^2 x_{2j}}\right). \end{aligned}$$

Thus ML estimates of model parameters are obtained by calculating the loglikelihood function, taking partial derivatives with respect to the parameter and solving the resulting likelihood equations. For example, the likelihood equation for the probability q of observing a degraded failure in a life test before the unit reaches the end of its useful life is

$$n(1-q) - mq = 0.$$

Hence the ML estimate $\hat{q}$ of q is given by

$$\hat{q} = \frac{n}{n+m}$$

while in practice, readily available optimisation software are used to obtain maximum likelihood estimates of the remaining inverse Gaussian distribution parameters. Consequently, this yields functional forms of the observed occurrences of the competing risks variables at each test stress level.

Chapter 4

Statistical modeling of life data from accelerated tests

4.1 Introduction

This chapter discusses the statistical modeling of life data from accelerated life tests. By design, accelerated life testing always requires extrapolation since test data are utilised to draw inferences about the lifetime distribution of the unit under normal use conditions. The basic idea behind accelerated life testing is the hypothesis that the mechanisms of failure under the right test stress levels remain the same as at normal operating conditions to justify extrapolation. Otherwise model errors will potentially dominate other sources of uncertainty. Under this assumption, accelerated life testing is thus a transformation of the time scale such that the lifetime distribution under a range of test stress levels is the same as under use conditions but evaluated at a compressed time scale.

At the core of the accuracy of the extrapolation is a physically motivated model for life data from accelerated life tests. Typically, the model for life data collected from accelerated life tests is a combination of a lifetime distribution and the life-stress relationship, called the accelerated life test model (ALT model). A diagrammatic representation of the ALT model is given in Figure 4.1.

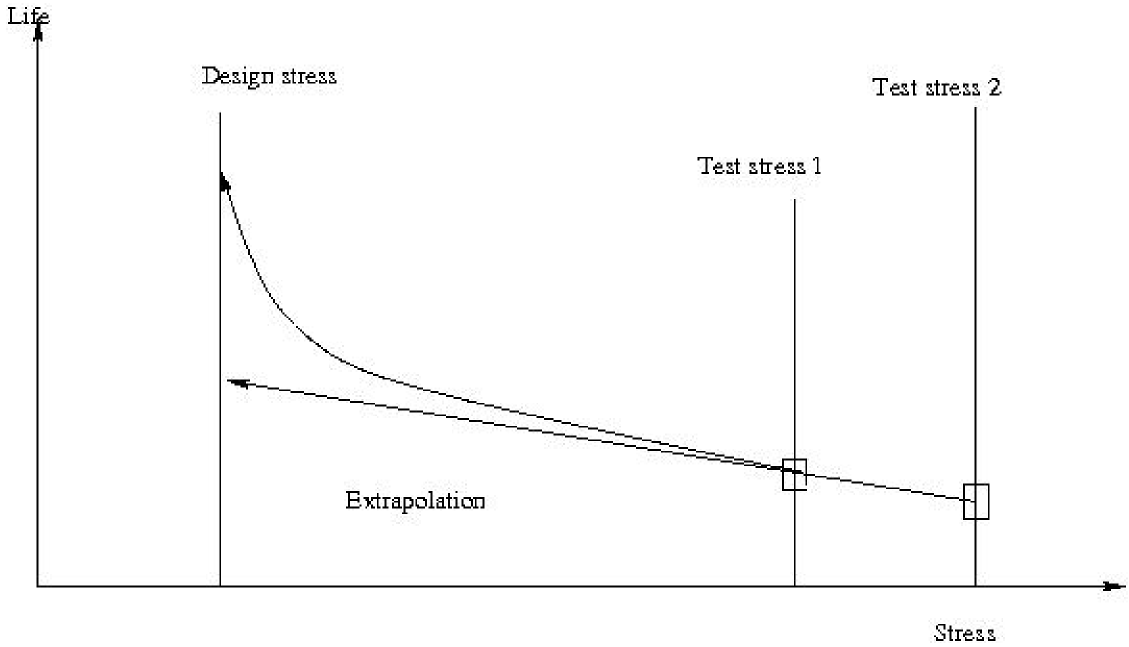


Figure 4.1: *Accelerated life test model.*

For a review of ALT models, see Escober and Meeker (2006). Accordingly, two assumptions are made when analysing test data on unit lifetime from accelerated life testing experiments. First, an appropriate lifetime distribution is chosen to describe the scatter in unit life at each test stress level. Often, the underlying lifetime distribution is assumed to come from a specified parametric family. The next step is to choose a model that describes how a quantifiable life measure of the assumed lifetime distribution varies with stress. In practice, model choice is guided by an understanding of the physics of failure or experience with similar tests. Lifetime distribution and model choice are discussed in turn.

4.2 Probability models for life data from accelerated tests

Typically, life data from accelerated tests are positively skewed because of more early unit failures. Consequently lifetime distributions which can have positively skewed frequency curves may provide a good fit to such test data. Two such distributions which are very popular in reliability and life testing studies are the Weibull and the lognormal distributions. The former can also have negatively skewed frequency curves and reduces to the exponential distribution when the shape parameter $\beta = 1$. Where early failures dominate unit lifetime distribution, the Weibull and the lognormal distributions may provide a similar fit to test data. In reliability analysis however, which model to use also depends on an understanding of the physics of failure. The PhD thesis by Liu (1997) gives a thorough discussion of the use of these two lifetime distributions when analysing reliability data.

4.2.1 Justification for the Weibull and lognormal distributions

A physical motivation for using the Weibull distribution to describe life data stems form its interpretation as a limiting distribution of minima. Specifically, the Weibull distribution has been shown (see for example Gumbel, 1958) to be identical to the type III smallest extreme value distribution for minimum values. Accordingly, it is an acceptable model of the first occurring failure mode for a unit where different failure modes are competing to remove the unit from observation in a life test. Thus whenever test data satisfy the chain model, they can be adequately described by the Weibull distribution.

The use of the lognormal distribution as a time-to-failure distribution can be justified as follows. Assume unit degradation during testing is directly observable such that $Y_1 < ... < Y_n$ is a sequence of random variables describing the state of unit degradation at stages $i = 1, ..., n$. By the proportional growth model (see e.g. Mann, Schafer and Singpurwalla, 1974), the change in the state of unit degradation at stage i, denoted $\Delta Y_i = Y_i - Y_{i-1}$ is randomly proportional to the state of degradation at stage $i - 1$. That is

$$I_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}, \quad i = 1, ..., n$$

where I_i are independent random variables, interpreted as small proportional degradation increments. The unit fails during testing when its degradation state reaches Y_n. As the change in the state of unit degradation at stage i becomes small, that is $\Delta Y_i \to 0$ and $n \to \infty$, then the sum of a large number of small proportional degradation increments

$$\sum_{i=1}^{n} I_i \approx \int_{Y_0}^{Y_n} \frac{1}{Y} dY = \ln Y_n - \ln Y_0 \tag{4.1}$$

where Y_0 is the initial degradation state of the unit. Accordingly, $Y_0 = 0$ if the unit is new. Rearranging Equation 4.1 yields

$$\ln Y_n = \sum_{i=1}^{n} I_i + \ln Y_0$$

and by the central limit theorem $\sum_{i=1}^{n} I_i \xrightarrow{d} N(\mu, \sigma^2)$ where $\xrightarrow{d}$ stands for converges in distribution. Consequently, lnY_n follows a normal distribution and hence Y_n is distributed as lognormal.

But how well the assumed lifetime distribution fits test data depends on the behaviour of its failure rate or hazard function. It is given by

$$\lambda(t) = \frac{f(t)}{S(t)}$$

and it is a measure of how prone units are to failure as a function of testing period (unit age). The bathtub curve in Figure 4.2 shows typical failure patterns over testing time. The exponential

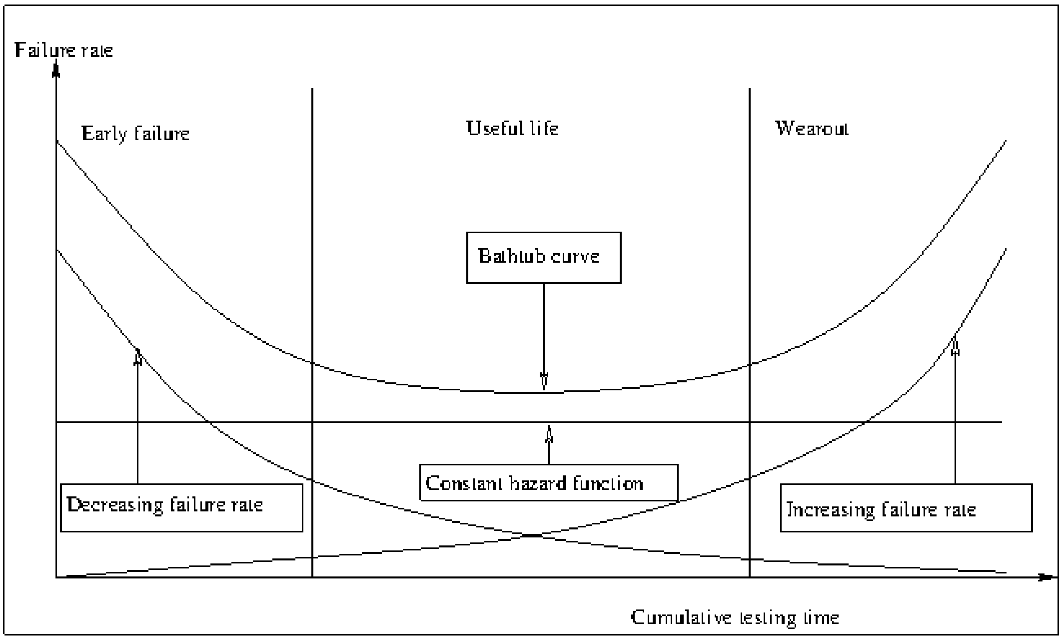


Figure 4.2: *Bathtub curve.*

distribution purely has a constant failure rate. As a result, it is an appropriate model for units that experience random failures during testing, possibly due to external shocks. Accordingly, the exponential distribution adequately describes failure patterns over testing time in the flat portion of the bathtub curve.

The Weibull hazard function is a power function of testing time given by

$$\lambda(t) = \left(\frac{\beta}{\alpha}\right)\left(\frac{t}{\alpha}\right)^{\beta-1}; \; t > 0, \; \alpha > 0, \; \beta > 0$$

where the shape parameter β is such that when $\beta > 1$ ($\beta < 1$), $\lambda(t)$ increases (decreases) with testing time. When $\beta = 1$, the Weibull distribution reduces to the exponential distribution and $\lambda(t)$ is constant over time. Hence the Weibull distribution adequately describes failure patterns over testing time in both early failure and wearout regions of the bathtub curve. This flexibility of the Weibull distribution in describing both increasing and decreasing failure rates makes it a preferred choice as the underlying lifetime distribution.

The lognormal failure rate has this property: it is zero at time zero, increases to a maximum and decreases to zero with increasing testing time. But interest in accelerated life testing is in obtaining failure data quickly. Hence good estimation of the lower percentiles of the assumed lifetime distribution is very important. Accordingly, the lognormal distribution flexibly fits test data particularly over its lower tail.

4.3 Life-stress relationship

The ultimate goal in accelerated life testing is to extrapolate a use-level lifetime distribution of the unit from test data. This can only be accomplished if there is a way to relate life at elevated test stresses to life at normal operating conditions. Hence it is also assumed that a time transformation or acceleration function exists that describes how a quantifiable life measure of the assumed lifetime distribution changes with stress. The quantifiable life measure can be any life measure such as the mean, median or some specified percentile. Typical quantifiable life measures for the exponential, Weibull and lognormal distributions are given in Table 4.1.

Distribution	Parameters	Quantifiable life measure
Exponential	$\frac{1}{\alpha}$	Mean life, α
Weibull	β^*, α	Scale parameter, α
Lognormal	μ', σ'^*	Median life, $\kappa_{0.5} = e^{\mu'}$

Table 4.1: *Quantifiable life measures for the exponential, Weibull and lognormal distributions. Parameters with* * *are assumed to be constant.*

Lognormal distribution parameters have $'$ to differentiate them from those of the normal distribution. The scale parameter μ' and the shape parameter σ' are the mean and standard deviation of the natural logarithm of the times to unit failure respectively.

Commonly used time transformation functions are the Arrhenius, the Eyring and the inverse power law (IPL) relationships or their generalisations. The physics based Arrhenius relationship applies when temperature is the accelerating variable. Its basis is the Arrhenius law which states that the reaction rate r depends on temperature through $r = Ae^{-E/kV}$ where the constant A is characteristic of unit failure mechanism and test conditions, E is the activation energy in electron-volts, k is the Boltzmann's constant (8.6171×10^{-5} electron-volts per $°C$) and V is the accelerating stress. In particular, V is the absolute Kelvin temperature for the Arrhenius relationship. The Arrhenius life-stress relationship expresses nominal life κ as inversely proportional to the rate constant. That is

$$\kappa = \frac{1}{A} e^{\frac{E}{kV}} \tag{4.2}$$

and linearising by taking base `10` logarithm yields

$$\log(\kappa) = \gamma_0 + \frac{\gamma_1}{V}$$

where $\gamma_0 = \log\left(A^{-1}\right)$ and $\gamma_1 = (E/k)\log(e)$. Accordingly, the Arrhenius life-stress relationship combines the lifetime distribution with the Arrhenius dependence of life on temperature.

The Eyring relationship is an alternative to the Arrhenius relationship in that it is also used when temperature is the accelerating variable. It expresses nominal life κ as a function of stress (temperature) through

$$\kappa = \frac{A}{V} e^{\frac{B}{kV}}$$

where the constants A and B are characteristic of unit failure mechanism and test conditions. On the other hand, the IPL is used when the accelerating variable is non-thermal. The IPL relationship between nominal life κ and the accelerating stress V is given by $\kappa(V) = AV^{-\omega_1}$ where A and ω_1 are model parameters to be determined. Taking natural logarithms yields the linearised relationship

$$\ln\left[\kappa(V)\right] = \omega_0 + \omega_1 \left[-\ln(V)\right]$$

which expresses the log of nominal life as a linear function of transformed stress. Consequently, if the lifetime distribution at higher test stresses and an appropriate time transformation function can be reasonably hypothesised, then the lifetime distribution at actual use conditions can be calculated mathematically. Clearly, at least two test stress levels are required to extrapolate a use-level lifetime distribution and the more the test stress levels, the better the fit. Estimates of the parameters for the lifetime distribution and life-stress relationship are obtained from test data.

4.4 ALT model and stress loading schemes

The ALT model is to a large extent determined by the type of stress loads applied in an ALT experiment. Stress loads are classified according to how the applied stress relates to time. In very broad terms, commonly applied stress loading methods can be classified into two schemes namely constant (in time) stress and time-varying stresses. Only a brief summary of these stress loading schemes is given here assuming units are tested at m different levels of stress, denoted by $v_1, ..., v_m$. Their merits and demerits are also briefly discussed. For a detailed discussion, see Nelson (2004) for example.

Constant stress loading is a time-independent test setting. Typically, $n_1, ..., n_m$ units are correspondingly tested at constant levels of accelerated stress $v_1, ..., v_m$ until failure or a censoring time. In general, $n_1 = ... = n_m$ but for optimal test plans, $n_1 > ... > n_m$ for $v_1 < ... < v_m$. Constant stress loads have several practical advantages. In particular,

(1) When in actual use, most units operate at constant use-level conditions. As a result, constant stress loads tend to mimic reality.

(2) Constant stress loads are simple and considerably easier to run. Specifically, maintaining the temperature at the same level in a thermal accelerated test is easier than having to change it with time.

(3) For some units, inference procedures for constant stress loads are well developed, empirically tested and computerised.

(4) If well-run, extrapolation from constant stress loads is more accurate than when stress is time dependent.

A disadvantage of constant stress tests is that they may need to run for a long time to yield enough failures, particularly at lower test stress levels.

In contrast, time-varying stress loads allow for a change in stress at different intermediate stages of the test. Commonly used time-varying stress loading schemes are step-stress and progressive stress loads. For $v_1 < ... < v_m$, step-stress loads take the form

$$V = \begin{cases} v_1 & \text{for} \quad 0 \leq t < t_1 \\ v_2 & \text{for} \quad t_1 \leq t < t_2 \\ \cdots & , \quad \cdots \\ v_m & \text{for} \quad t_{m-1} \leq t < t_m \end{cases} \tag{4.3}$$

where n units are placed on test at an initial lower level of stress v_1 until time t_1. Assume only $n_1 < n$ units fail during the test period $0 \leq t < t_1$. Unfailed units are then subsequently tested at increased stress levels for a period of time as in Equation 4.3 until all units fail. If n_2 units fail during the time interval $t_1 \leq t < t_2$ at a level of stress v_2 and so on until the remaining n_m units fail while being tested at v_m for the period $t_{m-1} \leq t < t_m$, then $n = n_1 + ... + n_m$.

In progressive stress tests on the other hand, test units are subjected to continuously increasing stress with time. A special case of progressive stress tests is a ramp test where stress is linearly increasing with time. The advantage of time-varying stress loads is that by design, they yield failures faster than constant stress loads assuming similar stress levels. Accordingly, the asymptotic theory of time-varying stress tests will consequently be a better estimation because of the many failures.

In terms of reliability estimation however, the standard error of estimates from test data is in general inversely proportional to the total time on test. Hence, estimates from time-varying stress tests are less accurate than from constant stress loads due to shorter test time. Further disadvantages of time-varying stress loads are linked to the design of the tests. Specifically

(1) Time-varying stress loads are difficult to maintain in practice and this may lead to additional experimental errors.

(2) Inference from time-varying stress loads require more complicated model assumptions. The model must properly account for the cumulative damage at successive stresses and at the same time, provide an estimate of unit lifetime under constant normal operating conditions.

(3) Time-varying stress loads induce failures at test stress levels far above the design stress level. Accordingly, the magnitude of the inevitable extrapolation error in these tests is higher than in constant stress loads.

In general, statistical methods for analysing censored data are readily available and computer packaged. As a result, it is not really necessary to force all test units to run to failure during life testing. For some tests, measured degradation data may also be collected during testing and utilised together with failure data to infer unit lifetime at use conditions. Besides, interest is often in lower percentiles of the lifetime distribution. Hence running all test units to failure adds very little information to lower distributional percentiles.

4.5 Mathematical description of the AFT model

The simplest ALT model particularly for units which degrade during testing is the accelerated failure time model (AFT model). Denote by v_0, the usual design stress experienced by a unit under normal operating conditions. Then the probability that a unit testing under test stress v

would survive till the moment t is the same as the probability that a unit operating under the use level stress v_0 would survive till the moment $g_v(t)$, called the transfer functional. That is, the transformed time $g_v(t)$ under v_0 is equivalent to t under the test stress v.

Now, let T_v be a non-negative absolutely continuous random variable corresponding to the failure time of the unit under test stress v. Obviously, the distribution of T_v depends on v. Thus, for any test stress v and $t \geq 0$,

$$S_v(t) = P\left(T_v \geq t\right) = P\left(T_{v_0} \geq g_v(t)\right) = S_{v_0}\left(g_v(t)\right) \tag{4.4}$$

where S_{v_0} is the baseline survival function. The AFT model, sometimes called additive accumulative damages model is defined on the basis of the properties of the functional $g_v(t)$ as follows. In the context of resource usage, the functional $g_v(t)$ may be taken as the amount of resource used until time t under stress v. Assuming stress is time-dependent, the rate of resource usage at the moment t is a function of the value $v(t)$ of the stress v at that moment and is given by the differential equation

$$\frac{d}{dt} g_v(t) = r[v(t)] \tag{4.5}$$

with the initial condition $g_v(0) = 0$. The unknown function r has the general interpretation of the popular failure rate model and can be estimated from test data. Thus, the AFT model is verified on the set of all admissible stresses E provided there exists a positive functional $r : E \to \mathbb{R}^+$ such that for any test stress $v \in E$ (Bagdonavicius, 1978), the relational function $g_v(t)$ satisfies 4.5. Integrating 4.5 with respect to t yields $g_v(t) = \int r[v(t)]dt$. Thus from 4.4 and in terms of survival functions, the AFT model takes the form

$$S_v(t) = S_{v_0}\left(\int_0^t r[v(s)]ds\right). \tag{4.6}$$

If the test stress is constant in time, then equation 4.6 reduces to

$$S_v(t) = S_{v_0}\left(r(v)t\right) \tag{4.7}$$

such that more severe test conditions shrink the time scale t by a factor $r(v)$ in the baseline survival distribution. Consequently for any two test stresses v_1 and v_2, the survival functions $S_{v_1}(t)$ and $S_{v_2}(t)$ will only differ in scale.

The baseline survival distribution S_{v_0} is often taken from a specified parametric family and if the functional $r()$ is also parameterised, then the corresponding AFT model is fully parametric. It becomes semi-parametric when either the functional $r()$ is parameterised and the baseline survival function S_{v_0} is completely unknown or vice versa. If both the functional $r()$ and the baseline survival function S_{v_0} are completely unknown, then the AFT model is nonparametric. In general however, parametric AFT models are used in practice though model choice is often guided by the assumed knowledge about the data.

4.6 Modeling under dependent random censorship

In this thesis however, the censoring variable X_1 and unit lifetime X_2 are competing to remove the unit from observation in a life test. As a result, test data comprise the time to removal of the unit from observation in a life test and the identity of the mode that actually removed the unit from observation in a life test. Clearly, subsurvival functions

$$S^*_{X_1}(t) = P\left(X_1 > t, X_1 < X_2\right)$$
$$S^*_{X_2}(t) = P\left(X_2 > t, X_2 < X_1\right)$$

and not true marginal survival functions $S_{X_j}(t) = P\left(X_j > t\right)$, $j \in (1, 2)$ are estimable from such observable competing risks data at each stress level unless the risks are independent. But since stochastic dependence is assumed between the censoring variable X_1 and unit lifetime X_2 on the premise of them being linked through the degradation process of the unit, the problem under consideration is that of dependent competing risks.

Given the event that the failure mode of interest removed the unit from observation in a life test and assuming continuity of $S^*_{X_1}(t)$ and $S^*_{X_2}(t)$ at $t = 0$,the conditional subsurvival functions

$$CS^*_{X_1}(t) = P\left(X_1 > t, X_1 < X_2 | X_1 < X_2\right) = \frac{S^*_{X_1}(t)}{S^*_{X_1}(0)}$$
$$CS^*_{X_2}(t) = P\left(X_2 > t, X_2 < X_1 | X_2 < X_1\right) = \frac{S^*_{X_2}(t)}{S^*_{X_2}(0)}$$

can be empirically obtained from the competing risks data at a stress level. Consequently, they may be important indicators when selecting the model that best fits competing risks data at a stress level.

Of interest in this investigation are the marginal survival functions $S_{X_j}(t) = P(X_j > t)$, $j \in (1, 2)$. In particular, the estimation problem is to extrapolate a use-level lifetime distribution from test data with the censoring variable removed. It is however well established, see for example Tsiatis (1975) that the true marginal survival functions are generally not identifiable from the observable dependent competing risks data alone. Additional simplifying assumptions are required and typically this entails restricting the joint survival function of the risk variables to a family of functions.

Carriere (1994) and Zheng and Klein (1995) generalised the identifiability of the true marginal survival functions to forms of dependency that are defined in terms of copulas. This obviously include the well-known result that the true marginal survival functions are identifiable when the risk variables are stochastically independent and corresponds to the independent copula. Carriere's model accommodates $j > 2$ competing risks whereas that of Zheng and Klein applies to $j = 2$ competing risks only. Lo and Wilke (2010) generalised the latter to $j > 2$ competing risks by pooling all other $k \neq j$ risks into a new risk variable.

The approach adopted in this investigation is as follows. Let $C(u_1, u_2)$ be a fixed copula that captures the stochastic dependence between unit lifetime and random censorship. Since unit lifetime and the censoring variable are competing to remove the unit from observation in a life test (competing risks scenario), $C(u_1, u_2)$ can not be estimated from test data since the data are incomplete. Accordingly, expert opinion is required to estimate $C(u_1, u_2)$. Assume that the estimated copula has continuous second-order partial derivatives with respect to $u_j \in (0, 1)$ and that the marginal dfs of unit lifetime and random censorship also exist at each test stress level. Further assume that $f^*_{X_j}(t) = -\frac{d}{dt}S^*_{X_j}(t)$ are continuous and denote by $h(x_1, x_2)$ the joint pdf of the competing risk variables X_1 and X_2. A straightforward calculation (see e.g. Bunea & Bedford, 2002) yields

$$F^*_{X_1}(t) \equiv P(X_1 \le t, X_1 < X_2) = \int_0^t \left(\int_{x_1}^{\infty} h(x_1, x_2) dx_2 \right) dx_1 \quad (4.8)$$
$$= F_{X_1}(t) - \int_0^t c_{u_1}(F_{X_1}(x_1), F_{X_2}(x_1)) f_{X_1}(x_1) dx_1$$

where

$$c_{u_1} = \frac{\partial C(u_1, u_2)}{\partial u_1}$$

is the first order partial derivative calculated in $(F_{X_1}(t), F_{X_2}(t))$. In the same way,

$$F^*_{X_2}(t) = F_{X_2}(t) - \int_0^t c_{u_2}(F_{X_1}(x_2), F_{X_2}(x_2)) f_{X_2}(x_2) dx_2 \quad (4.9)$$

where

$$c_{u_2} = \frac{\partial C(u_1, u_2)}{\partial u_2}$$

is also calculated in $(F_{X_1}(t), F_{X_2}(t))$. For any $u_1 \in \mathbf{I} = [0, 1]$, $\frac{\partial}{\partial u_2} C(u_1, u_2)$ exists for almost all $u_2 \in [0, 1]$ such that

$$0 \le \frac{\partial}{\partial u_2} C(u_1, u_2) \le 1. \quad (4.10)$$

Equation 4.10 is also true for $\frac{\partial}{\partial u_1} C(u_1, u_2)$ with u_1 and u_2 interchanging roles and the functions

$$u_1 \to c_{u_2}(u_1) \equiv \frac{\partial}{\partial u_2} C(u_1, u_2) \quad and \quad u_2 \to c_{u_1}(u_2) \equiv \frac{\partial}{\partial u_1} C(u_1, u_2)$$

are well-defined and almost everywhere non-decreasing on $[0, 1]$. Hence partial derivatives of a copula with respect to its arguments exists.

Put together and after rearranging, Equations 4.8 and 4.9 yield the non-linear system of differential equations

$$\begin{cases} [1 - c_{u_1}(F_{X_1}(t), F_{X_2}(t))] \, F'_{X_1}(t) = F^{*'}_{X_1}(t) \\ [1 - c_{u_2}(F_{X_1}(t), F_{X_2}(t))] \, F'_{X_2}(t) = F^{*'}_{X_2}(t) \end{cases} \quad (4.11)$$

with initial conditions $F_{X_1}(0) = F_{X_2}(0) = 0$. Given the estimated copula $C(u_1, u_2)$ and suitable functional forms for $F^*_{X_1}(t)$, $j \in (1, 2)$, the differential system in Equation 4.11 can be numerically solved for $F_{X_1}(t)$ and $F_{X_2}(t)$ at each test stress level. Consequently life data samples from these marginal distribution functions are easily generated at each test stress level. The problem of fitting a chosen lifetime distribution to life data is discussed next.

Several methods are available for fitting life data samples to a chosen lifetime distribution. In this thesis however, two parameter estimation methods namely maximum likelihood estimation (MLE) and median rank regression (MRR) are considered. Barring some exceptions, the former is the most robust and is thus the preferred parameter estimation method from a statistical point of view. On the other hand MRR is the preferred method in industry, specifically in reliability analysis. This is mainly because with MRR, data can be graphically displayed and parameters estimated by easily understood ordinary least squares method. Further, graphical representation of data also provides a basis for goodness-of-fit tests. Life data from accelerated tests are inherently censored. Hence MLE and MRR are discussed for censored test data assuming either the Weibull or the log normal distribution adequately describes test data.

4.7 Maximum likelihood estimation method

The method of MLE obtains the most likely values of the parameters for a chosen lifetime distribution that best describes test data. It has excellent asymptotic properties that make its use very attractive namely

(1) MLE parameter estimates converge to the right parameter values as sample size increases. Hence the method is asymptotically consistent.

(2) On average, MLE parameter estimates yield the correct parameter values for large samples and are therefore asymptotically unbiased.

(3) MLE parameter estimates are asymptotically normally distributed and this is the basis for the construction of confidence bounds to quantify parameter uncertainty.

(4) For large samples, MLE method produces minimum variance estimates and are therefore the most precise.

In addition, the method of MLE can also handle all kinds of test data including heavily censored test data where not even a single unit failure is observed in a life test. But when sample sizes for test data are very small, finite sample properties of the MLE parameter estimates would be less than optimal. Consequently the parameter estimates would also be biased.

4.7.1 Weibull distribution

Assume test data are adequately described by a $2-$parameter Weibull distribution. Denote by $\boldsymbol{\theta} = (\alpha, \beta)$ the unknown Weibull distribution parameters where $\alpha > 0$ is the scale parameter representing the characteristic life of the units and $\beta > 0$ is the shape parameter that determines the appearance of the Weibull pdf. The problem of estimating Weibull parameters has attracted significant attention in life testing and reliability theory in general (see e.g. Genschel and Meeker, 2010; Olteanu and Freeman, 2010 and the numerous references therein). The method of MLE obtains the Weibull parameter estimator $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta})$ which is the *highest ranked* of all possible $\boldsymbol{\theta}$ values given the observed test data at each test stress level by maximising the likelihood function

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{r} \left[\frac{\beta}{\alpha} \left(\frac{t_i}{\alpha} \right)^{\beta-1} \exp\left(- \left(\frac{t_i}{\alpha} \right)^{\beta} \right) \right] \prod_{i=r+1}^{n} \left[\exp\left(- \left(\frac{t_i}{\alpha} \right)^{\beta} \right) \right]$$

or equivalently the log of the likelihood $\mathscr{L}(\boldsymbol{\theta}) = \ln(L(\boldsymbol{\theta}))$ where r is the number of failures and $n - r$ is the number of right censored observations. Thus the likelihood function of the censored sample is the joint density of the n random variables and is a function of the unknown parameters. The ML estimates $\hat{\alpha}$ and $\hat{\beta}$ for α and β are solutions of

$$\frac{\sum_{i=1}^{n} t_i^{\hat{\beta}} \ln(t_i)}{\sum_{i=1}^{n} t_i^{\hat{\beta}}} - \frac{1}{r} \sum_{i=1}^{r} \ln(t_i) - \frac{1}{\hat{\beta}} = 0 \tag{4.12}$$

and

$$\hat{\alpha} = \left(\frac{1}{r} \sum_{i=1}^{n} t_i^{\hat{\beta}} \right)^{\frac{1}{\hat{\beta}}} \tag{4.13}$$

respectively. For more details, see for example Nelson (1982). In practice however, ML estimates of Weibull distribution parameters are obtained by numerical methods.

4.7.2 Lognormal distribution

Alternatively, assume the lognormal distribution with scale parameter μ' and shape parameter σ' provides a good description of test data. Then the log of test data, denoted by $T = lnX$ are normally distributed. That is $T \sim N(\mu, \sigma^2)$ where μ is the location parameter and σ is the scale parameter. The parameters of the normal and lognormal models are related through simple transformations $\mu' = e^{\mu}$ and $\sigma' = \sigma^{-1}$. Consequently MLE of the lognormal distribution can be recovered from the MLE of the normal distribution.

For a test sample of n units with r failures and $n - r$ censors, the likelihood function for the $2-$parameter normal distribution is

$$L = \prod_{i=1}^{r} \left[\frac{1}{\sigma} \phi \left(\frac{t_i - \mu}{\sigma} \right) \right] \prod_{i=r+1}^{n} \left[1 - \Phi \left(\frac{t_i - \mu}{\sigma} \right) \right] \tag{4.14}$$

where $\phi()$ and $\Phi()$ are the probability density and the cumulative distribution of the standard normal respectively. The parameter values μ and σ that maximise the likelihood function in Equation 4.14 or equivalently the log likelihood are the ML estimates $\hat{\mu}$ and $\hat{\sigma}$. Hence ML estimates $\hat{\mu'}$ and $\hat{\sigma'}$ for the lognormal distribution parameters μ' and σ' are obtained from ML estimates of the normal distribution parameters. Again, numerical methods are used in practice.

Denote by $\widetilde{\mu}$ and $\widetilde{\sigma}$, the mean and the standard deviation values of unit failure times. These values are not used as lognormal distribution parameters and are obtained through

$$\widetilde{\mu} = \exp \left(\mu' + \frac{1}{2} {\sigma'}^2 \right)$$

and

$$\widetilde{\sigma} = \sqrt{\left(e^{2\mu' + {\sigma'}^2} \right) \left(e^{{\sigma'}^2} - 1 \right)}$$

respectively. The life characteristic for the lognormal distribution is the median life and is given by $\kappa_{0.5} = e^{\mu'}$.

4.8 Median rank regression method

4.8.1 Introduction

The analytical motivation of MRR is the linearisation of the cdf of the lifetime distribution. This enables least squares regression analysis to be performed on the transformed test data. Consequently, distribution parameters are estimated through simple transformations of the estimated regression coefficients. Its implementation requires a nonparametric estimator of the failure probability $F(t_i)$. Typically, $\widehat{F}(t_i)$ is obtained by marking the median rank for each failure time at each stress level and hence the general name MRR. Different formulas for $\widehat{F}(t_i)$ have been proposed in the literature (see e.g. Skinner, Keats and Zimmer, 2001). The most popular approximation is Bernard's median rank estimator

$$\widehat{F}(t_i) = \frac{i - 0.3}{n + 0.4} \tag{4.15}$$

where i is the failure order number (FON), defined as the sequence number of that failure by age. The drawback of Bernard's median rank estimator is that FON is not defined at failure points after a suspension as follows.

Let $t_1 < t_2 < ... < t_s < ... < t_k < ... < t_n$ be a censored sample at a stress level. Assume for simplicity that the sample has a single suspension at time t_s. All unit failure times less than the suspension t_s have clearly defined failure order numbers. Thus t_1 is assigned a FON 1 since is the earliest failure by age, t_2 has a FON 2 and so on. For the suspension t_s however, if the unit was kept on test long enough, it could fail at any later time $t_l > t_s$ before or after the later failure time t_k. That is, either $t_l < t_k$ or $t_l > t_k$ and hence the FON at a failure point t_k becomes uncertain.

To ensure that the median rank is defined at all failure times for right censored test samples, the mean order number (MON) is used instead of the exact FON. See Wang (2004) for more details. The MON for the failure time t_k that is greater than age of suspension t_s is defined as the expected total number of failures before t_k assuming all units were kept on test and run to failure. That is

$$\widehat{F}(t_k) = \frac{E(FON_k) - 0.3}{n + 0.4} = \frac{MON_k - 0.3}{n + 0.4}. \tag{4.16}$$

The modification of median rank in Equation 4.16 is based on a statistical fundamental that unit failure times and suspensions have identical statistical properties since test units are a random sample from the same population. Denote by $\{t_i, \delta_i; i = 1, ..., n\}$, the ordered censored sample $t_1 < t_2 < ... < t_s < ... < t_k < ... < t_n$ where the censoring indicator

$$\delta_i = \begin{cases} 1 & \text{if} \quad t_i \quad \text{is} \quad \text{a} \quad \text{failure} \quad \text{time} \\ 0 & \text{if} \quad t_i \quad \text{is} \quad \text{a} \quad \text{censor} \quad \text{time.} \end{cases}$$

The total number of failures before age t_k is

$$FON_k = D_1 + D_2 + ... + D_n$$

where

$$D_i = \begin{cases} 1 & \text{if} \quad \text{unit} \quad \text{i} \quad \text{fails} \quad \text{before} \quad t_k \\ 0 & \text{if} \quad \text{unit} \quad \text{i} \quad \text{fails} \quad \text{after} \quad t_k. \end{cases}$$

But for a suspension at t_s before age t_k, D_i is random because the suspended unit could potentially fail before t_k ($D_i = 1$) with probability

$$P(D_i = 1) = \frac{\int_{t_s}^{t_k} f(t)dt}{1 - F(t_s)} = \frac{F(t_k) - F(t_s)}{1 - F(t_s)}$$

or after t_k ($D_i = 0$) with probability

$$P(D_i = 0) = \frac{\int_{t_k}^{\infty} f(t)dt}{1 - F(t_s)} = \frac{1 - F(t_k)}{1 - F(t_s)}.$$

Consequently, FON_k is a random variable and by definition

$$MON_k = E\left[FON_k\right] = E\left(D_1\right) + E\left(D_2\right) + ... + E\left(D_n\right). \tag{4.17}$$

It follows from Equation 4.17 that when $\delta_i = 0$, then $E(D_i) = \frac{F(t_k) - F(t_s)}{1 - F(t_s)}$ for $i = 1, ..., k$. On the other hand, when $\delta_i = 1$, then $E(D_i) = 1$ for $i = 1, ..., k$ and for $i = k + 1, ..., n$, $E(D_i) = 0$. Accordingly, the MON and hence the median rank estimator is completely defined for every unit failure time. Thus the MRR method for censored test data involves the following steps:

(1) Rank unit failure data from the smallest to the largest.

(2) For times-to-failure data only, use Equation 4.16 to assign median ranks

(3) Estimate distribution parameters using least squares analysis.

4.8.2 Weibull distribution

The cdf of the $2-$parameter Weibull distribution is linearised by taking the logarithm of the Weibull failure probability $F(t_i)$ twice. The resulting linear form is given by

$$\ln\left[-\ln\left(1 - F_w(t_i)\right)\right] = -\beta\ln(\alpha) + \beta\ln(t_i) \tag{4.18}$$

where the Weibull failure probability $F_w(t_i)$ is estimated from the median ranks. Equation 4.18 is in the standard linear form $y_i = a + bx_i$ where $y_i = ln\left[-ln\left(1 - F_w(t_i)\right)\right]$; $x_i = ln(t_i)$; $a = -\beta ln(\alpha)$ and $b = \beta$. Two kinds of regressions, namely regressing Y on X and regressing X on Y can be performed on the linear form in Equation 4.18. Correspondingly, these two regressions minimise the vertical and horizontal error sum of squares.

The scale with greater variability (see for example Berkson, 1950) is generally treated as the dependent variable. Unit failure times almost always exhibit larger error than median ranks. Hence unit failure times are treated here as the dependent variable. Consequently, the regression of X on Y is preferred. For more details, see Abernethy (1994). The best fitting equation for the regression of X on Y is the the straight line

$$x = \hat{a} + \hat{b}y. \tag{4.19}$$

Correspondingly, the equations for the regression coefficients are

$$\hat{a} = \bar{x} - \hat{b}\bar{y} \tag{4.20}$$

and

$$\hat{b} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}. \tag{4.21}$$

The Weibull parameters are recovered by writing Equation 4.19 in the form of Equation 4.18. This yields simple transformations $\hat{\alpha} = exp\left(\frac{\hat{a}}{\hat{b}} \times \frac{1}{\hat{\beta}}\right)$ and $\hat{\beta} = \frac{1}{\hat{b}}$ for obtaining the parameters of the Weibull model.

4.8.3 Lognormal distribution

The CDF of the lognormal distribution can be written as

$$F(t') = \Phi\left(\frac{t' - \mu'}{\sigma'}\right) \tag{4.22}$$

where $t' = ln(t)$ and t are the times-to-failure of the unit at a stress level. The parameters μ' and σ' as well as $\Phi()$ are as defined under MLE for the lognormal distribution. Rearranging Equation 4.22 yields an equivalent version

$$\Phi^{-1}\left[F(t')\right] = -\frac{\mu'}{\sigma'} + \frac{1}{\sigma'}t' \tag{4.23}$$

where Φ^{-1} is the inverse of the standard normal cdf. Comparing Equation 4.23 with the linear form $y_i = a + bx_i$ yields transformed probability axes $y_i = \Phi^{-1}\left[F(t_i')\right]$ where $F(t_i')$ is estimated from the median ranks and $x_i = ln(t_i)$.

Again, unit failure times are treated as the dependent variable because they exhibit larger error compared to median ranks. As a result, regressing X on Y is preferred. The best fitting regression equation remains exactly the same as in Equation 4.19 and the same applies to the equations for the regression coefficients. Simple transformations $\mu' = \frac{\hat{a}}{\hat{b}}\sigma'$ and $\sigma' = \hat{b}$ give the lognormal distribution parameters.

4.9 Discriminating between competing lifetime distributions

Deciding if a test sample is a realisation from a population with a particular lifetime distribution is an old but important problem in statistics. The problem is made even more difficult in accelerated reliability testing because test samples are generally small and test data are often right censored. Given test data at each test stress level, how will one discriminate between the Weibull and the lognormal distribution? From a purely statistical viewpoint, goodness-of-fit tests are used to discriminate between these two lifetime distributions.

Dumonceaux and Antle (1973) for example used the ratio of maximised likelihoods to formulate the hypothesis setting

$$H_0 : T \sim \text{Weibull}(\alpha, \beta) \ \text{ against } \ H_1 : T \sim \text{Lognormal}(\mu', \sigma') \tag{4.24}$$

for discriminating between the Weibull and the lognormal distributions. This setting considers model choice as a test of hypothesis where the distribution assigned to the null hypothesis is the preferred model to use. The test statistic for the test of hypothesis in Equation 4.24 is given by

$$TS_{MLR} = \frac{1}{\left(\sqrt{2\pi\sigma'^2 e}\right)\sqrt[n]{\prod_{i=1}^{n} t_i f_w(t_i)}} \tag{4.25}$$

where e is the exponent, n is the size of the test sample, t_i is the time to unit failure and $f_w()$ is the Weibull probability density function (pdf). The Weibull distribution is rejected in favour of the lognormal distribution if TS_{MLR} is greater than or equal to the tabulated critical value at a specified level of significance.

Alternatively, the lognormal distribution could be the preferred model. To allow this model choice, Dumonceaux and Antle (1973) also proposed the reverse hypothesis

$$H_0 : T \sim \text{Lognormal}(\mu', \sigma') \;\text{ against }\; H_1 : T \sim \text{Weibull}(\alpha, \beta). \tag{4.26}$$

Accordingly, the test statistic takes the form

$$TS_{MLR} = \left(\sqrt{2\pi\sigma'^2 e}\right)\sqrt[n]{\prod_{i=1}^{n} t_i f_w(t_i)}. \tag{4.27}$$

The lognormal distribution is rejected in favour of the Weibull distribution if TS_{MLR} is greater than or equal to the tabulated critical value at a specified level of significance.

Other methods that can be used to measure goodness-of-fit include the most powerful invariant (MPI) test due to Kent (1979). These methods are however very complex and hence less attractive than simpler statistical methods in practice. If the models were nested, that is, one model is a special case of the other, then they can be compared using likelihood ratio tests. Asymptotically, the test statistic takes the form

$$T_{LR} = -2\left(\mathscr{L}_s(\hat{\boldsymbol{\theta}}) - \mathscr{L}_g(\hat{\boldsymbol{\theta}})\right) \sim \chi^2_{p_g - p_s}$$

where $\mathscr{L}_s$ and $\mathscr{L}_g$ are log-likelihoods of the simpler and the general model respectively, p_s and p_g are the corresponding number of parameters in the models. There however does not seem to be a way to convert either of the Weibull or lognormal distributions to the other by fixing parameter(s). Consequently, these lifetime models are not nested.

Log likelihoods are also commonly used in statistical practice to compare models for the same data set provided the models being compared have the same number of parameters. If the number of parameters differ for the models in question, information-based criteria which extend log likelihood comparisons would be used instead. There are several information-theoretic approaches. But because of inherent small sample sizes in ALT, the corrected Akaike information criterion (AIC_c), defined as

$$AIC_c = -2\mathscr{L}(\hat{\boldsymbol{\theta}}) + 2K + \frac{2K(K+1)}{n-K-1}$$

where K is the number of estimated model parameters and n is the number of test units at a stress level. The model with minimum AIC_c value is better.

4.10 Comparison of the parameter estimation methods

An important question in terms of parameter estimation centers on the choice of the estimation method for the chosen lifetime distribution. A number of studies compare MLE and MRR methods for estimating Weibull parameters. Few studies include the lognormal distribution. Results from these studies are generally mixed because of study differences in terms of censoring schemes (data types), different censoring percentages and different evaluation criteria. Obviously, these factors lead to important differences when evaluating estimation methods. Emphasis here is on performance of MLE and MRR procedures in small samples under Type I censoring because few units are often available for testing and not all units are generally run to failure in life tests.

In cases where test samples are small and a high degree of censoring is apparent (Abernety, 2004), MLE estimates are known to be biased. On the other hand, let $t_1, ..., t_n$ be a test sample of size n. The MRR method uses corresponding order statistics $t_{(1)} \leq ... \leq t_{(n)}$. Accordingly, $t_i = t_{(i)}$ is the value assumed by the order statistic $T_{(i)}$. Hence $x_i = \ln(t_i)$ is interpreted as the value assumed by the log order statistic $X_{(i)}$. Similarly, the transformed value of the independent variable y_i is also considered as the value assumed by the (reduced) order statistic $Y_{(i)}$. As a result, the best fitting equation for the regression of X on Y in Equation 4.19 has the representation

$$X_{(i)} = a + bY_{(i)}.$$

Least squares regression of X on Y is justified as an estimation technique provided X_i are independent observations with constant variance. In this case however, $X_i = X_{(i)}$ is the log order statistic such that

$$Var\left(X_i\right) = b^2 Var\left(Y_{(i)}\right); \;\; Cov\left(X_i, X_j\right) = b^2 Cov\left(Y_{(i)}, Y_{(j)}\right), \;\; i \neq j.$$

Though essentially a least squares estimation method, MRR violates assumptions upon which the least squares estimation method is based. Accordingly, MRR is not statistically optimum as an estimation technique. Hence the estimates of the regression coefficients in Equations 4.20 and 4.21 are also biased.

In general however, the overall accuracy of an estimator cannot be properly evaluated unless both bias and precision as measured by the standard error of the estimator are considered. The MRR estimates $\hat{a}$ and $\hat{b}$ in Equations 4.20 and 4.21 respectively do not have minimum error variance since the assumptions of independent observations and constant variance are violated. On the other hand, ML estimates are often the most precise even in small test samples with few failures. Hence when these estimators are evaluated in terms of their overall accuracy, MLE almost always generally perform better than MRR in almost all practical situations.

Test data are almost always right censored. For such data, the MRR method only uses the location of the censored observation and not the exact time-to-censoring. As a result, it gives exactly the same results for cases with differing failure and suspension times as long as the order of unit states (failed or suspended) and hence failure order numbers are identical. MLE on the one hand would give different results for different cases because it uses the actual times-to-failure or suspension, not ranks. For highly reliable systems, few failures are often observed even under accelerated conditions. Consequently, the information contained in the suspended observations becomes very important. Hence MLE is attractive because it uses all information in the data.

4.11 Statistical methodology

The problem of identifying marginal distribution functions from dependent competing risks data has been studied in different contexts. Carriere (1994) for example transformed the differential system in Equation 4.11 into a system of difference equations and solving the problem recursively. The drawback of this approach for the study under consideration here is that the resulting marginal distribution functions $F_{X_1}(t)$ and $F_{X_2}(t)$ take only integer values. On the contrary, the censoring variable X_1 and unit lifetime X_2 are both supported on $[0, \infty)$. Zheng and Klein (1995) proposed an asymptotic copula graphic estimator that employs a bisection root-finding method to construct estimates $\widehat{F}_{X_1}(t)$ and $\widehat{F}_{X_2}(t)$ for the marginal distribution functions.

Studies of Kaishev, Dimitrova and Haberman (2007) and Dimitrova, Haberman and Kaishev (2013) directly solve the differential system in Equation 4.11 for $F_{X_j}(t)$, $j = 1, 2$ and any $t > 0$ in an actuarial context using `NDSolve`, a built-in function in `Mathematica`. Numerical integration methods implemented in `NDSolve` include Euler's method, the midpoint method and Runge-Kutta methods. The accuracy of numerical integration methods is measured by matching high terms with the Taylor expansion of the solution. Euler's method is the simplest but it has a local error of $O(h^2)$ and is thus first-order accurate. The midpoint method is second-order accurate and is also available through the more accurate Runge-Kutta methods.

Kaishev *et al.* (2007) and Dimitrova *et al.* (2013) assume a known copula and uses the data averaging spline interpolation method of De Boor (2001) to obtain functional forms of the observed occurrences of the competing risks variables. Bunea and Mazzuchi (2007) applied a life-stress relationship to conditional subsurvival functions at each test stress level and extrapolated the conditional subsurvival function at use-level conditions.

This investigation adopts an approach similar to that of Kaishev *et al.* (2007) and Dimitrova *et al.* (2013) but in an accelerated life testing setup. In addition, expert opinion is used to estimate the chosen copula model as opposed to just carrying-out a sensitivity analysis. Moreover, functional forms of the subdistribution functions are derived from the theory of stochastic processes. In Kaishev *et al.* (2007) and Dimitrova *et al.* (2013), cubic spline survival functions were fitted instead.

In particular, the methodology uses key results from chapters 2 and 3 as follows. The Frank copula is postulated in chapter 2 to adequately describe stochastic dependence between the censoring variable X_1 and unit lifetime X_2 and expert judgement is used to estimate its dependence parameter θ. The partial derivative with respect to u_1 of the chosen Frank copula in Equation 2.8 is given by

$$c_{u_1}^F(u_1, u_2) = \frac{\partial}{\partial u_1} C^F(u_1, u_2) = \frac{e^{-\theta u_1}\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1 + \left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}. \quad (4.28)$$

Clearly, $c_{u_1}^F(u_1, u_2)$ is defined for all $u_2 \in [0, 1]$. In particular, it is a strictly increasing function of u_2 for $u_1 \in [0, 1]$. Hence the partial derivative in Equation 4.28 is a conditional distribution function. Since the Frank copula is symmetric in u_1 and u_2, it follows that

$$c_{u_2}^F(u_1, u_2) = \frac{\partial}{\partial u_2} C^F(u_1, u_2) = \frac{e^{-\theta u_2}\left(e^{-\theta u_1} - 1\right)}{e^{-\theta} - 1 + \left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)} \quad (4.29)$$

is also a strictly increasing function of u_1 given the same set of parameters.

Suitable functional forms of the subdistribution functions are derived from the theory of stochastic processes in chapter 3. They are first passage time distributions of the Wiener maximum process and are shown in chapter 3 to be inverse Gaussian. Given the partial derivatives of the estimated Frank copula in Equations 4.28 and 4.29 and functional forms of $f_{X_j}^*$, $j = 1, 2$, the differential system in Equation 4.11 is numerically solved for $F_{X_1}(\cdot)$ and $F_{X_2}(\cdot)$ at each test stress level.

In practice however, there is no guarantee that both $F_{X_1}(\cdot)$ and $F_{X_2}(\cdot)$ will be non-defective for the chosen Frank copula and inverse Gaussian subdistribution functions. That is $P(X_j < \infty) = 1$ is not guaranteed for both risk variables. Unit lifetime X_2 is the object of estimation in this investigation. Saying that unit lifetime will be infinite with positive probability is obviously not plausible in practice since units cannot remain in test for ever. For more details on the issue of defective marginals in competing risks problems, see for example Bedford (2006).

But in this investigation, both X_1 and X_2 are assumed to be non-defective. Consequently, $F_{X_1}(\cdot)$ and $F_{X_2}(\cdot)$ are proper marginal distribution functions and are solutions of the differential system in Equation 4.11. The main object of estimation is the use-level lifetime distribution of the unit $F_{X_2}(\cdot)$. It is obtained from the ALT model as follows:

- Generate life data samples from numerical solutions of the unit lifetime distribution $F_{X_2}(\cdot)$ at each test stress level. Use information criteria to choose a lifetime distribution that best fits the generated life data samples.

- Based on an understanding of the physics of failure during testing or experience with similar life tests, choose an appropriate model that describes how a quantifiable measure of the assumed lifetime distribution changes with stress.

- Use the chosen life-stress relationship to extrapolate distribution percentiles from elevated test stresses to use-level conditions.

This identifies the use-level lifetime distribution of the unit from which reliability metrics such as warranty period, mean life etc. can be derived.

Chapter 5

Data analysis and results

5.1 Introduction

The analysis of test data on unit lifetime under random censorship is presented in this chapter. In particular, the extrapolation of a use-level unit lifetime distribution from test data with the censoring variable removed is illustrated. But accelerated test data where unit lifetime and random censorship are competing to remove the unit from observation in a life test are not readily available. This is almost always the case with real data from accelerated tests collected in internal research divisions of large companies. Test data are highly confidential (commercially sensitive) and are therefore difficult to access in general. Accordingly, statistical methods of analysing life test data are illustrated based on derived competing failure modes data that are analogous to unit lifetime and random censorship. These data are derived from an accelerated life test that also yielded competing failure modes.

5.1.1 Data description

The readily available and widely used ALT data when competing failure modes are acting is the Class-H insulation data collected from a temperature-accelerated life test of motorettes insulation. The test yielded three insulation failure modes namely Turn, Phase and Ground failures (Nelson, 2004; pp 393), each occurring on a separate part of the motorette. Ten motorettes were tested at high temperatures of $190°C$, $220°C$, $240°C$, $260°C$ and inspected periodically for failure. The recorded times (in hours) are midway between the failure time and the previous inspection time. Inspections were conducted frequently so that the error of rounding to the midpoint is minimal. The data are given in Table 5.1.

$190°C$	Turn	Phase	Ground	$220°C$	Turn	Phase	Ground
1	7228	10511	10511+	11	1764	2436	2436
2	7228	11855	11855+	12	2436	2436	2490
3	7228	11855	11855+	13	2436	2436	2436
4	8448	11855	11855+	14	2436	2772+	2772
5	9167	12191+	12191+	15	2436	2436+	2436
6	9167	12191+	12191+	16	2436	4116+	4116+
7	9167	12191+	12191+	17	3108	4116+	4116+
8	9167	12191+	12191+	18	3108	4116+	4116+
9	10511	12191+	12191+	19	3108	3108	3108+
10	10511	12191+	12191+	20	3108	4116+	4116+
$240°C$	Turn	Phase	Ground	$260°C$	Turn	Phase	Ground
21	1175	1175+	1175	31	1632+	1632+	600
22	1881+	1881+	1175	32	1632+	1632+	744
23	1521	1881+	1881+	33	1632+	1632+	744
24	1569	1761	1761+	34	1632+	1632+	744
25	1617	1881+	1881+	35	1632+	1632+	912
26	1665	1881+	1881+	36	1128	1128+	1128
27	1665	1881+	1881+	37	1512	1512+	1320
28	1713	1881+	1881+	38	1464	1632+	1632+
29	1761	1881+	1881+	39	1608	1608+	1608
30	1953	1953+	1953+	40	1896	1896	1896

Table 5.1: *Class-H insulation failure mode data taken from Nelson (2004, pp. 393).*

Observe that each motorette has a recorded time-to-failure from each of Turn, Phase and Ground failure modes. This is because each failed part of the motorette was isolated electronically so that it could not fail again while the unit was kept on test and run to a second or third failure. But in a typical competing risks situation, the first occurring failure mode removes the unit from observation in a life test. Hence Table 5.1 contains pseudo-competing risks data.

To derive competing risks that are analogous to unit lifetime and random censorship from test data in Table 5.1, an understanding of the test purpose is required. The test purpose for the Class-H insulation data was three fold:

(1) To estimate the median life of the insulation system in motorettes at its use-level tempera-

ture of 180^0C.

(2) To determine the earliest occurring failure mode at the use-level temperature of 180^0C.

(3) To determine if redesign to remove the earliest failure mode would significantly improve the reliability of the motorette.

In his analysis of the Class-H insulation data Nelson (2004) found Turn failure to be the earliest failure mode at the design temperature of $180°C$. The reliability of the motorette was subsequently improved by eliminating Turn failure mode through a redesign of the motorette. Accordingly, the useful life of the redesigned motorette can only be ended by Phase or Ground failures in a life test. Hence the Class-H insulation data are reduced to two derived competing risks by classifying Turn failure mode as Risk 1 and treating the time to first failure from Phase or Ground failure modes as Risk 2. This translates to grouping Phase and Ground failure modes into a single mode and labeling it Risk 2.

The derived competing risks, namely Risk 1 and Risk 2 are obviously not degraded and critical failure modes respectively. However, Risk 2 is considered analoguous to critical failure mode in this investigation because its occurrence in a life test ends the useful life of the redesigned motorette. This makes Risk 2 the failure mode of interest. On the other hand, Risk 1 is considered analoguous to degraded failure mode because its occurrence would censor the failure mode of interest in a competing risk framework. Hence data on the derived censoring and unit lifetime variables X_1 and X_2 respectively at each test stress level are given in Table 5.2.

$190°C$	X_1	X_2	$220°C$	X_1	X_2	$240°C$	X_1	X_2	$260°C$	X_1	X_2
1	7228	10511	11	1764	2436	21	1175	1175	31	1632+	600
2	7228	11855	12	2436	2436	22	1881+	1175	32	1632+	744
3	7228	11855	13	2436	2436	23	1521	1881+	33	1632+	744
4	8448	11855	14	2436	2772	24	1569	1761	34	1632+	744
5	9167	12191+	15	2436	2436	25	1617	1881+	35	1632+	912
6	9167	12191+	16	2436	4116+	26	1665	1881+	36	1128	1128
7	9167	12191+	17	3108	4116+	27	1665	1881+	37	1512	1320
8	9167	12191+	18	3108	4116+	28	1713	1881+	38	1464	1632+
9	10511	12191+	19	3108	3108	29	1761	1881+	39	1608	1608
10	10511	12191+	20	3108	4116+	30	1953	1953+	40	1896	1896

Table 5.2: *Derived competing failure modes from the Class-H insulation failure mode data.*

The data on X_2 are obtained by taking the minimum of times to Phase and Ground failures.

A remark is necessary here. Treating times-to-failure from other modes as if they were times-to-failure from modes of interest is not new in the reliability literature. Bunea and Mazzuchi (2007) also reduced the same Class-H insulation failure mode data to two competing risks by classifying Turn failures as Risk 1 and grouping Phase and Ground failures into a single failure mode labeled Risk 2. They however did not attach any physical meaning to the derived competing risks.

In a different application, Dijoux and Gaudoin (2009) analysed times to electrical and mechanical component failure data of compressor units as corrective and preventive maintenance times respectively. Electrical component failures were treated as corrective maintenance times in the study because they are more expensive compared to cheaper mechanical component failures. For risk pooling in the biostatistical literature, see Kaishev *et al.* (2007) for example.

Recall that each of Turn, Phase and Ground failure modes occurred on a separate part of the motorette during testing. The part of the motorette that failed first (first occurring failure mode) was isolated while the unit was kept on test and run until the second or third failure mode occurred. This test design implies that the first occurring failure mode has no bearing on the occurrence of the remaining failure modes. As a result, failure data in Table 5.2 are times-to-failure from a specific mode as if the other failure modes were not acting. Nelson (2004), presumably the owner of the Class-H insulation data also assumed an independent competing risks model. Accordingly, probabilities of surviving a specific failure mode until the moment t (marginal survival functions) are identifiable from the competing failure modes data in Table 5.2.

Data analysis is in two parts. The first part uses real data in Table 5.2 to extrapolate a use-level unit lifetime distribution at the design temperature of $180°C$. However, these data are a special type of competing failure modes data in that independence is a consequence of experimental design, not the stochastic behaviour of the competing risks. It illustrates the well-known result that marginal survival functions of independent competing risks are identifiable. Consequently, it constitutes a simple case corresponding to knowing the independent copula. The second part of the analysis considers the more practical case where unit lifetime and random censorship are stochastically dependent. It uses simulated dependent competing risks data and combines numerical and theoretical results from Chapters 2, 3 and 4.

5.2 Test data analysis: Independent competing risks

The derived competing failure modes data in Table 5.2 are largely right censored. The only exceptions are degraded failures at test stresses of $190°C$ and $220°C$ which have complete samples. In addition, sample sizes are small. As has already been alluded to, for small test samples under Type I censoring, MLE has advantages over MRR method. Most importantly, test data in Table 5.2 have a number of unit failure and suspension times that are identical at each test stress level. Consequently, MRR method would assign different rank values to identical unit failure times. As a result, their corresponding failure probabilities as estimated by median ranks would also be different. But MLE uses actual failure and suspension times and is therefore preferred to MRR for these test data.

The scatter in unit life at each test stress level may in theory be described by any lifetime distribution. Table 5.3

$190°C$	X_1	X_2	$220°C$	X_1	X_2
$\hat{\alpha}$	9299.445	12610.309	$\hat{\alpha}$	2817.588	4020.505
$\hat{\beta}$	8.502	19.670	$\hat{\beta}$	7.398	2.952
LK value	−84.959	−36.715	LK value	−74.565	−54.440
$240°C$	X_1	X_2	$260°C$	X_1	X_2
$\hat{\alpha}$	1749.829	2585.271	$\hat{\alpha}$	1782.517	1311.883
$\hat{\beta}$	9.234	3.228	$\hat{\beta}$	8.183	2.636
LK value	−62.217	−27.427	LK value	−38.098	−68.849

Table 5.3: *ML estimates of Weibull parameters and log likelihood values for the derived competing failure modes data.*

contains ML estimates $\hat{\alpha}$ and $\hat{\beta}$ of the Weibull parameters α and β assuming the scatter in the derived competing failure modes data in Table 5.2 is adequately described by the $2-$parameter Weibull distribution. The 95% confidence intervals (CI) for the Weibull parameters α and β at each stress level and for each failure mode are given in `Appendix B`.

For the derived unit lifetime X_2 (variable of interest), the quantifiable life measure as measured by the Weibull scale parameter decreases with stress in Table 5.3. In theory, higher test stresses lead to early failures and hence a compression of unit lifetime. Consequently, results in Table 5.3 are consistent with theory. More imprortantly, the Weibull shape parameter, $\beta \approx 3$ for test data at all but the $190°C$ test stress. Hence the failure mechanism as represented by the Weibull shape parameter is largely the same, justifying subsequent extrapolation. The large shape parameter value at the $190°C$ stress level may be attributed to faulty testing or mishandled test units. In life testing studies, identifying causes of data peculiarities is more valuable than dropping suspicious data points and remodeling for example.

If the lognormal distribution is assumed to provide a good description of test data, the lognormal scale and shape parameters μ' and σ' are also estimated by the MLE method. Using the `Weibull++` software, the ML estimates $\hat{\mu'}$ and $\hat{\sigma'}$ for the derived competing failure mode data in Table 5.2 are given in Table 5.4. The 95% CIs for the lognormal scale and shape parameters μ'

$190°C$	X_1	X_2	$220°C$	X_1	X_2
μ'	9.071	9.432	μ'	7.863	8.141
σ'	0.144	0.086	σ'	0.182	0.372
LK value	-85.016	-36.784	LK value	-75.301	-53.292
$240°C$	X_1	X_2	$260°C$	X_1	X_2
μ'	7.409	7.783	μ'	7.444	6.981
σ'	0.144	0.484	σ'	0.198	0.417
LK value	-62.846	-27.185	LK value	-38.570	-68.178

Table 5.4: *ML estimates of lognormal parameters and log likelihood values for the derived competing failure modes data.*

and σ' at each stress level and for each failure mode are given in `Appendix C`.

The quantifiable life measure assuming the lognormal distribution best fits test data is median life $\kappa_{0.5} = e^{\mu'}$. Results in Table 5.4 show that median life for unit lifetime X_2 decreases with stress as suggested by theory. In addition, the failure mechanism as represented by the lognormal shape parameter, $\sigma' \approx 0.4$ at all but the $190°C$ test stress, justifying extrapolation.

Goodness-of-fit tests are used to choose between the the Weibull and lognormal distributions for these data. Since they are not nested (neither is a special case of the other), they can not be compared using likelihood ratio tests. Log likelihood values can be used since they both have the same number of parameters. Instead of using raw log likelihood values, AIC_c is used to discriminate between these two lifetime distributions. Table 5.5 gives AIC_c values calculated for each model with the derived competing failure modes data at each test stress level. The model with smaller AIC_c value is better.

$190°C$	X_1	X_2	$220°C$	X_1	X_2
Weibull	175.632	79.144	Weibull	154.844	114.595
Lognormal	175.747	79.282	Lognormal	156.317	112.298
$240°C$	X_1	X_2	$260°C$	X_1	X_2
Weibull	130.149	60.567	Weibull	81.911	143.412
Lognormal	131.406	60.084	Lognormal	82.853	142.069

Table 5.5: *AIC_c values for the Weibull and lognormal models calculated with the derived competing failure modes data.*

Based on the calculated AIC_c values in Table 5.5, the Weibull and the lognormal distributions largely fit the derived competing failure mode data in Table 5.2 equally well. This is consistent with the simulation results of Dumonceaux and Antle (1973) where the choice of either distribution is the test of hypothesis. The simulation results show that the power to discriminate between the Weibull and lognormal decreases with sample size. With sample sizes of 10 at each test stress level, either model can be used based on goodness-of-fit tests. For unit lifetime X_2, AIC_c values for the lognormal distribution are slightly less than for the Weibull distribution at all but the $190°C$ stress level. This suggests that the scatter in unit lifetime may be adequately described by the lognormal distribution.

The choice of the lifetime distribution will however be based on the analysis method, MLE in this case and the assumed life-stress relationship. Test data in Table 5.2 are from a temperature accelerated test of motorettes insulation. Accordingly, a life-stress relationship derived for temperature dependence is required. In particular, the Arrhenius relationship is preferred because it is based on the law of physics (nature). In addition, it is well-known (Nelson, 2004) to be a valid model in insulation work. Assuming the Arrhenius dependence of unit lifetime on temperature, log likelihood and AIC_c values for the Weibull, exponential and lognormal lifetime distributions are given in Table 5.6.

	Weibull	Exponential	Lognormal
LK-value	-192.9815	-205.9746	-190.663
AIC_c	391.667	414.449	387.040

Table 5.6: *LK and AIC_c values for the Weibull, exponential and lognormal lifetime distributions assuming the Arrhenius relationship for real test data.*

From Table 5.6, the lognormal distribution statistically has the best fit because it has the minimum AIC_c value. Hence the scatter in unit life at each test stress level is assumed to be described by the lognormal distribution. Consequently, the AFT model is assumed to be the Arrhenius-lognormal model where the quantifiable life measure is median life, $\kappa_{0.05} = e^{\mu'}$. By Equation 4.2,

$$\kappa_{0.05} = e^{\mu'} = \frac{1}{A} e^{\frac{E}{kV}}$$

such that $\mu' = \gamma_0 + \frac{\gamma_1}{V}$. Hence the pdf of the Arrhenius-lognormal model at stress V is given by

$$f(t, V) = \frac{1}{t\sigma' \sqrt{2\pi}} exp \left\{ -\frac{1}{2} \left(\frac{t' - \gamma_0 - \frac{\gamma_1}{V}}{\sigma'} \right)^2 \right\}. \tag{5.1}$$

5.3 Assessment of the Arrhenius-lognormal model

The best fitting Arrhenius-lognormal model assumes that the lognormal distribution adequately describes the scatter in unit life at all test stress levels. It also follows from Equation 5.1 that σ' is independent of stress. Since σ' has the interpretation of the lognormal shape parameter, the Arrhenius-lognormal model further assumes that the shape of the lifetime distribution is invariant to changes in stress.

The validity and accuracy of the reliability metrics derived from fitting the Arrhenius-lognormal model to test data largely depends on the extent to which the above assumptions are satisfied. In particular, lower percentile estimates at a stress are very sensitive to the assumed lifetime distribution and a shape parameter that varies with stress. Different methods for checking the assumptions of the Arrhenius-lognormal model are considered. These methods particularly check how well the assumed lognormal distribution fits test data and whether the distribution's shape parameter is constant across all test stress levels. These include graphical and numerical methods.

5.3.1 Graphical methods

Residual plots are important visual analysis tools when assessing the assumptions of the fitted model. They also help reveal any inadequacies in the assumed model as well as exposing outlying observations if any. The residual plot tool in ALTA 9.0 gives plots for the standardised and Cox-Snell residuals. At failure time T_i, the former are calculated by

$$\hat{\epsilon_i} = \frac{ln\,(T_i) - \hat{\mu'}}{\hat{\sigma'}} \sim N(0,1) \tag{5.2}$$

assuming the lognormal distribution adequately describes the scatter in unit life. Figure 5.1

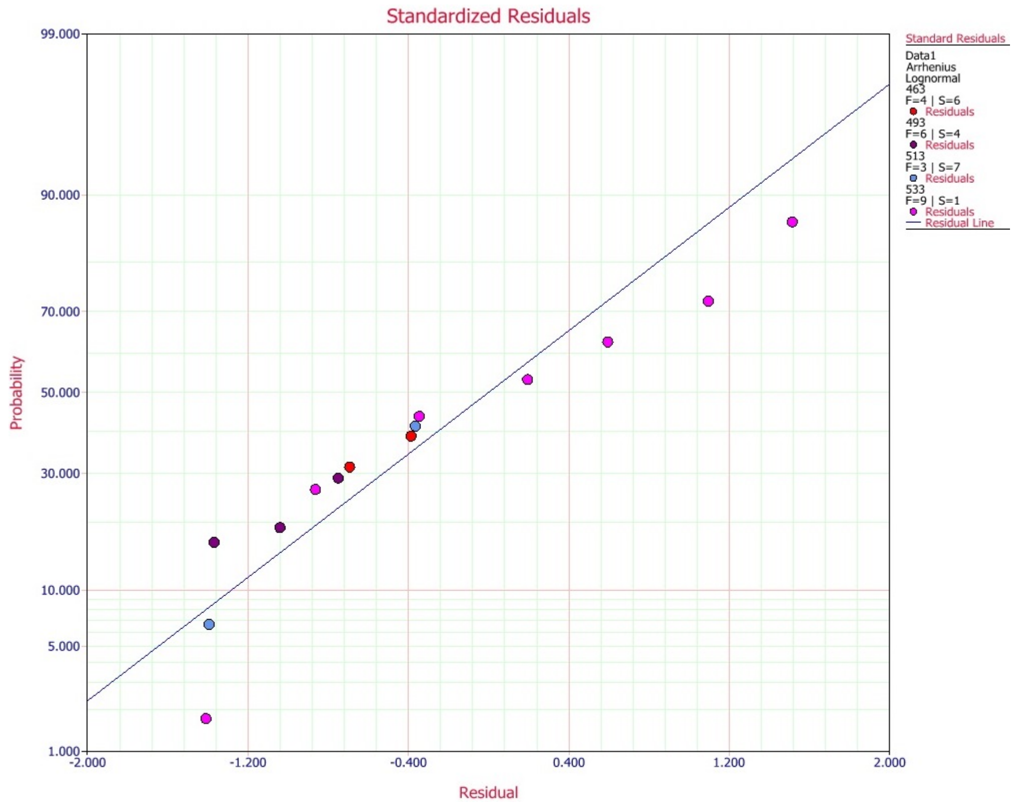


Figure 5.1: *Normal probability plot of the standardised residuals.*

gives a visual display of the standardised residuals on a normal probability plot. The lognormal distribution is considered to adequately describe the scatter in unit life if the standardised residuals roughly follow a straight line on a normal probability plot. It is clear from Figure 5.1 that the standardised residuals appear to fairly follow a straight line for unit failure data at all but the $260°C$ ($533K$) test stress level. In addition to faulty testing and mishandled units, the magnitude of acceleration may also be a factor since $260°C$ is double the use-level temperature.

Cox-Snell residuals on the other hand are calculated by

$$\hat{\epsilon}_i = ln\left[R\left(T_i\right)\right] \tag{5.3}$$

where $R\left(T_i\right)$ is the reliability value calculated at unit failure time T_i. Figure 5.2 is a visual display of the Cox-Snell residuals on an exponential probability paper.

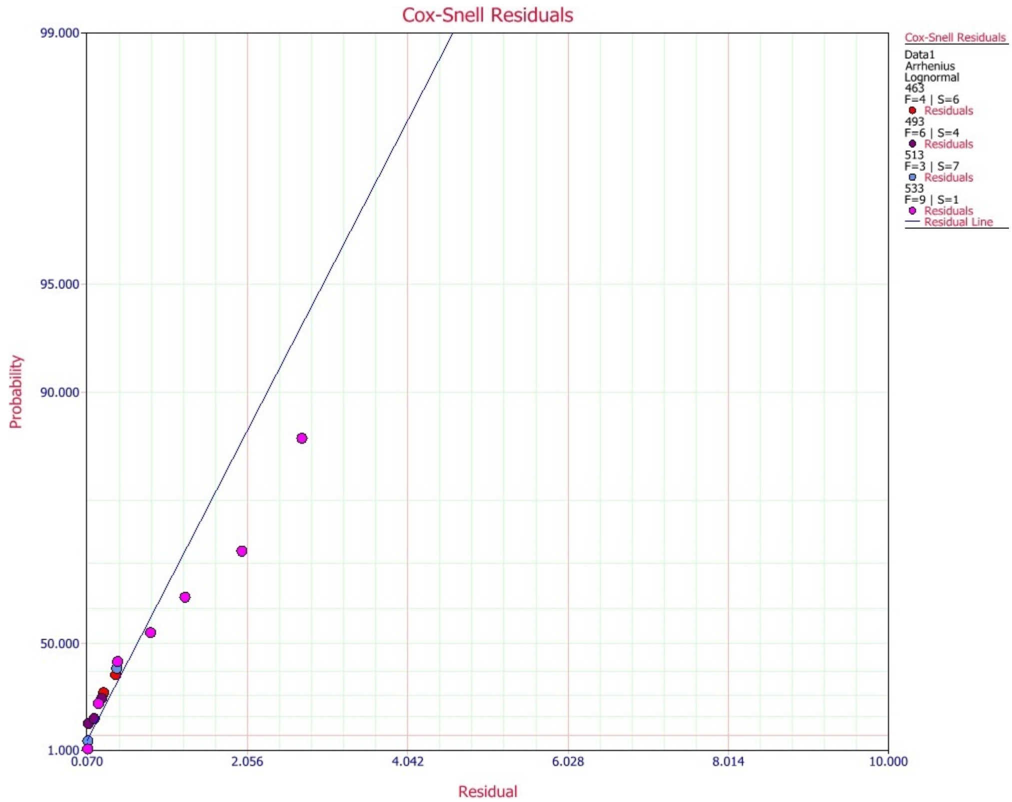


Figure 5.2: *Exponential probability plot of the Cox-Snell residuals.*

As with the standardised residuals, the Cox-Snell residuals in Figure 5.2 also appear to fairly follow a straight line for unit failure data at all but the $260°C$ ($533K$) test stress level.

Another graphical display that is useful when visually assessing how well the assumed lifetime distribution (lognormal in this case) fits test data is the probability plot. It is a plot of the cdf (unreliability) on linearised lognormal probability paper. In particular, fairly straight lognormal probability plots indicate that the scatter in unit life can be adequately described by a lognormal distribution. The probability plot for unit lifetime at all four test stress levels is shown in Figure 5.3.

There is a noticeable outlying point below the 10% failure probability at the $260°C$ ($533K$) test stress level in Figure 5.3. Potential causes of such peculiarities include mishandled units during testing. In practice, points in a probability plot are not necessarily expected to lie on the line for the assumed lifetime distribution to fit test data well. Hahn and Shapiro (1967) highlighted this in

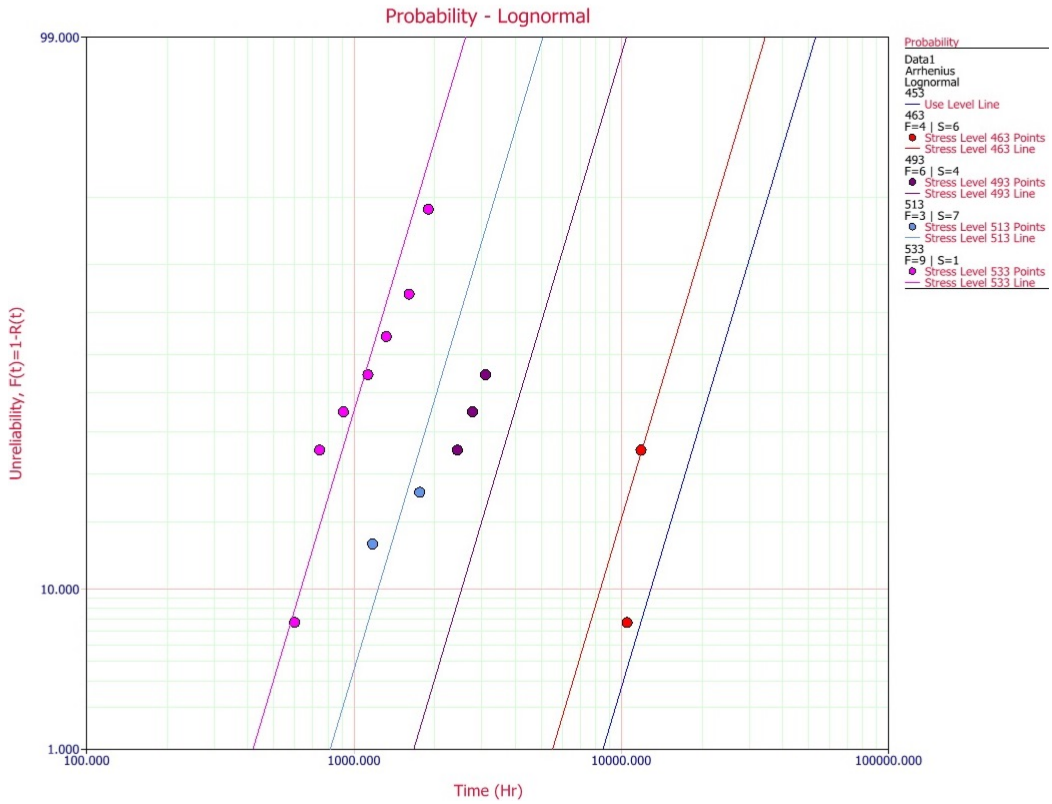


Figure 5.3: *CDF plot on linearised lognormal paper.*

a simulation study. Points in their probability plots for samples of 20 and as high as 50 simulated from a true distribution appear erratic.

Accordingly, points in a probability plot for real test data with small test samples will inevitably be erratic and only features of the plot that are striking may be taken to be properties of the population. Hence though erratic, points in the lognormal probability plot in Figure 5.3 are fairly straight. This suggests that the lognormal distribution may adequately describe the scatter in unit life at all test stress levels.

The probability plot is also useful when visually assessing the assumption of a common shape parameter for the assumed lifetime distribution across the various test stress levels. On the linearised lognormal probability paper, the shape parameter σ' is the slope of the line. If the failure mechanism as represented by the shape parameter σ' is independent of stress, plots of the cdf on linearised paper must be fairly parallel. The lines in Figure 5.3 are clearly parallel. They however have a compromise common slope obtained by a refined method as follows:

- Separately fit the lognormal distribution to test data at each of the j test stress levels. The fitted distributions appear as straight lines, each with its own separate slope.
- Obtain the common slope by weighting these separate slopes proportional to the number of units tested at a stress level.

Obviously data peculiarities at any test stress level will influence the value of the common slope.

5.3.2 Numerical methods

A more objective assessment of the assumption of a common lognormal shape parameter for test data across test stress levels utilises numerical methods. The likelihood ratio (LR) test described in Nelson (1990) is used in this investigation. Denote by T_{LR} the LR test statistic given by

$$T_{LR} = -2\left(\hat{\mathscr{L}}_0 - \left(\hat{\mathscr{L}}_1 + ... + \hat{\mathscr{L}}_j\right)\right) \tag{5.4}$$

where $\hat{\mathscr{L}}_1, ..., \hat{\mathscr{L}}_j$ are the maximum log likelihood values obtained by separately fitting the chosen lifetime distribution to test data at each of the j test stress levels. They are maximum log likelihood values for the unrestricted model. On the other hand, the maximum log likelihood value $\hat{\mathscr{L}}_0$ is obtained by fitting a model with a common shape parameter and a different scale parameter for each of the j test stress levels. Consequently, $\hat{\mathscr{L}}_0$ is the maximum log likelihood value for the restricted model.

If the true shape parameters of the assumed lifetime distribution do not differ at the j test stress levels, T_{LR} is asymptotically distributed as chi-square with $j-1$ degrees of freedom (Wilks, 1938). Otherwise, T_{LR} tends to assume larger values. Accordingly, the chi square test can be used as an approximate test of the hypotheses

H_0: The shape parameter is independent of stress level

H_1: The shape parameter is dependent on stress level.

The decision criterion is such that

- If $T_{LR} \leq \chi^2(\alpha; j-1)$, then the j shape parameter estimates are not statistically significantly different at the α level of significance.

- If $T_{LR} > \chi^2(\alpha; j-1)$, then the j shape parameter estimates are statistically significantly different at the α level of significance.

For the derived unit lifetime test data in Table 5.2, the maximum log likelihood values are $\hat{\mathscr{L}}_{463} = -36.784$, $\hat{\mathscr{L}}_{493} = -53.292$, $\hat{\mathscr{L}}_{513} = -48.446$, $\hat{\mathscr{L}}_{533} = -68.178$ and $\hat{\mathscr{L}}_{0} = -189.172$. Thus the value of the test statistic is

$$T_{LR} = -2\left(-189.172 - (36.784 + 53.292 + 48.446 + 68.178)\right) = -35.056.$$

At the 10% level of significance and for the $j = 4$ test temperatures, $\chi^2(0.1; 3) = 6.251$. Since the value of the likelihood ratio test statistic $T_{LR} = -35.056 < 6.251 = \chi^2(0.1; 3)$, H_0 cannot be rejected at the 10% level of significance. Consequently, the shape parameter estimates are not statistically significantly different at the 10% level of significance.

5.4 Assessment of the assumed life-stress relationship

The Arrhenius-lognormal model assumes that the relationship between the (transformed) quantifiable life measure, $\kappa_{0.5} = e^{\mu'}$ and stress V is linear. If this linearity assumption does not hold, extrapolation to low stresses (usually use-level conditions) will be difficult to justify. Consequently, the extrapolation will likely be inaccurate. The assumption of a linear life-stress relationship is assessed by both graphical and numerical methods.

5.4.1 Graphical methods

Life-stress plots are important visual analysis tools when assessing the linearity assumption of the life-stress relationship. They are obtained by plotting the quantifiable life measure against stress. Figure 5.4 shows the Arrhenius life-temperature plot for unit lifetime data.

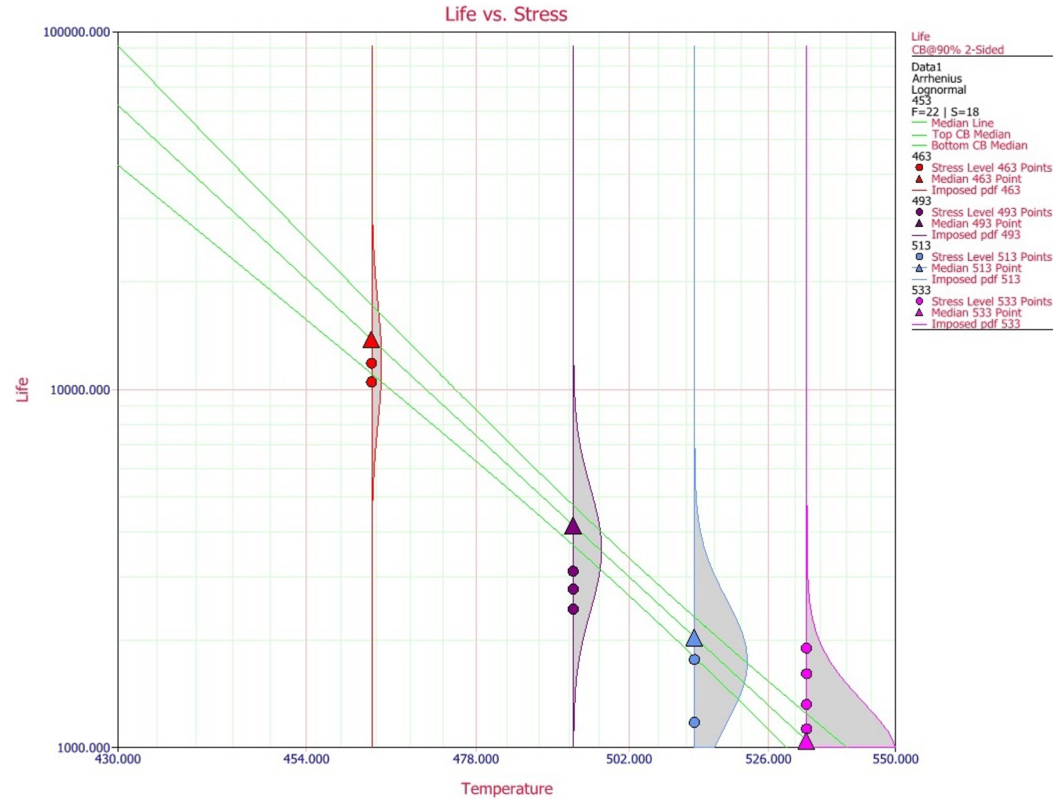


Figure 5.4: *Arrhenius life-temperature relationship plot for unit lifetime data.*

The relationship plot in Figure 5.4 shows that a straight line describes the Arrhenius dependence of life on temperature since the sample percentile line (middle line) is linear. This median line represents a path for extrapolating the quantifiable life measure from test stresses to use-level conditions. The linearity assumption seems to hold for these data since the dependence of unit life on temperature is failrly linear. Because of the small sample sizes, it is important to quantify uncertainty as much as possible. Accordingly, the Arrhenius life-temperature plot in Figure 5.4 also shows the upper and lower 90% confidence bounds on median life as represented by the top and bottom lines respectively. The imposed pdfs represent the distribution of test data at each test stress level.

5.4.2 Numerical methods

The LR test is also utilised to check linearity of the relationship between the quantifiable life measure, $\kappa_{0.5} = e^{\mu'}$ and stress V. The LR test statistic is given by

$$T_{LR} = -2\left(\hat{\mathscr{L}}_0 - \hat{\mathscr{L}}\right) \tag{5.5}$$

where $\hat{\mathscr{L}}$ is the maximum log likelihood value obtained by fitting the chosen lifetime distribution to test data assuming the specified time transformation function. For the derived unit lifetime test data, it is the maximum log likelihood value obtained by fitting the lognormal distribution to test data assuming the Arrhenius relationship whereas $\hat{\mathscr{L}}_0$ is defined as before. Accordingly, $\hat{\mathscr{L}}$ and $\hat{\mathscr{L}}_0$ are the maximum log likelihood values for the unrestricted and the restricted model respectively.

If the life-stress relationship is linear, T_{LR} is asymptotically distributed as chi-square with $j-1$ degrees of freedom. Otherwise, T_{LR} tends to assume larger values. As before, the chi square test is used as an approximate test of the hypotheses

H_0: Test data are consistent with a linear life-stress relationship

H_1: Test data are not consistent with a linear life-stress relationship.

The decision criterion is such that

- If $T_{LR} \leq \chi^2(\alpha; j-1)$, then test data are not statistically significantly different from a linear life-stress relationship at the α level of significance.
- If $T_{LR} > \chi^2(\alpha; j-1)$, then test data are statistically significantly different from a linear life-stress relationship at the α level of significance.

For the derived unit lifetime data in Table 5.2, $\hat{\mathscr{L}} = -190.663$ and $\hat{\mathscr{L}}_0 = -189.172$. Hence $T_{LR} = -2\,(-189.172 - (-190.663)) = -2.982$. At the 10% level of significance and for the $j = 4$ test temperatures, $\chi^2(0.1; 3) = 6.251$. Since $T_{LR} = -2.982 < 6.251 = \chi^2(0.1; 3)$, H_0

cannot be rejected at the 10% level of significance. Therefore there is insufficient evidence at the 10% level of significance to suggest that test data are statistically significantly different from a linear life-stress relationship.

5.5 Additional plots

Other useful plots include the reliability plot, failure rate plot, standard deviation plot, acceleration plot etc. Selected additional plots that are useful in accelerated life testing are presented in Figure 5.5.

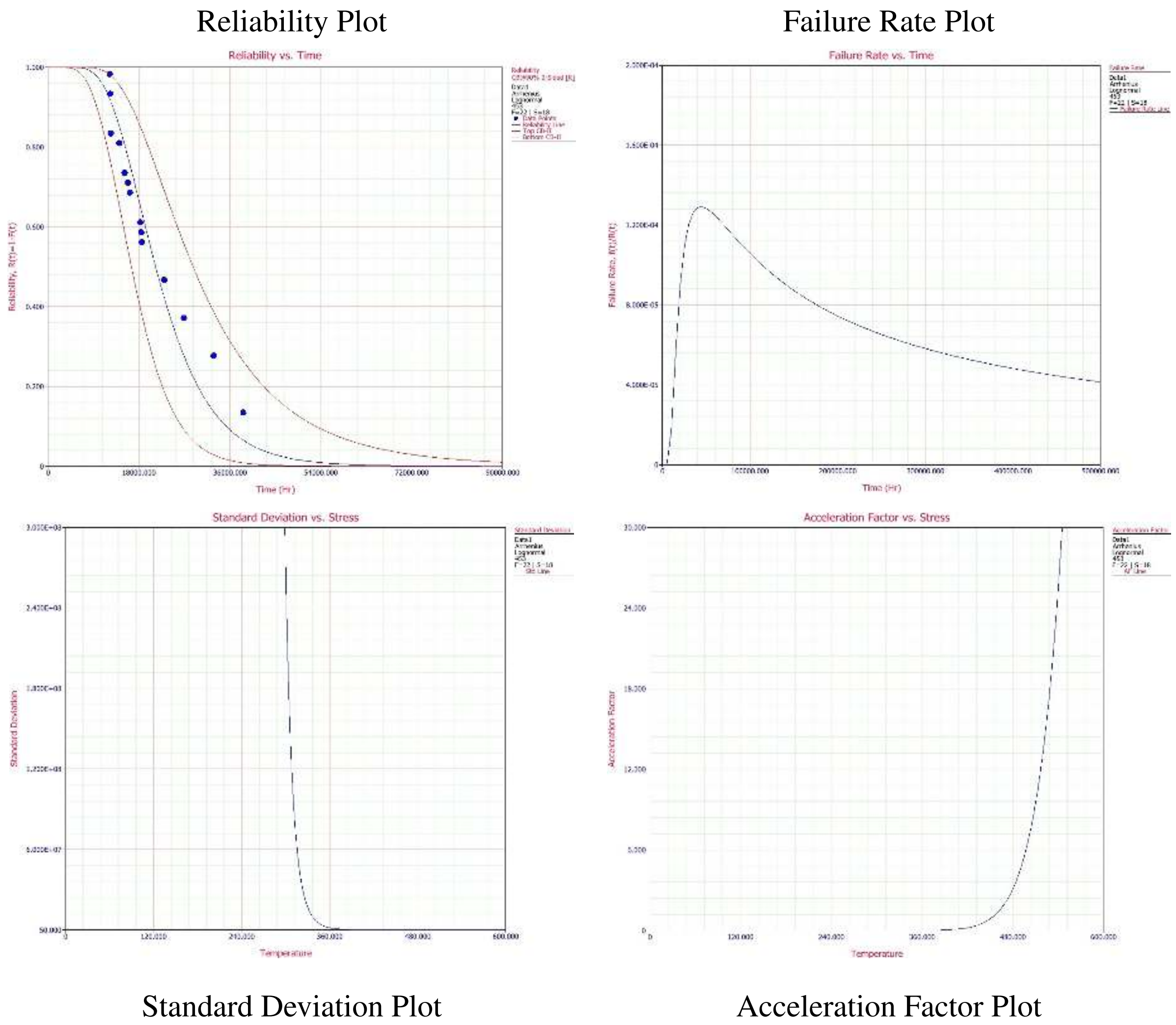


Figure 5.5: *Selected useful plots.*

The plot of the reliability against time, plotted here with confidence bounds is useful when one seeks reliability values for a given unit age and vice-versa. The failure rate against time plot gives the expected number of test units that would fail per unit time at a stress level. Both the reliability and the failure rate plots are at normal use conditions. On the other hand, the standard deviation plot shows how the data are spread at each stress level whereas the failure rate plot relates life at normal operating conditions to life at accelerated stresses.

5.6 Extrapolating the use-level lifetime distribution

Based on results from the graphical and numerical assessment methods, the assumed Arrhenius-lognormal model adequately describes test data on unit lifetime X_2 in Table 5.2. Figure 5.6 is a pdf plot of the extrapolated use-level lifetime distribution of the redesigned motorette.

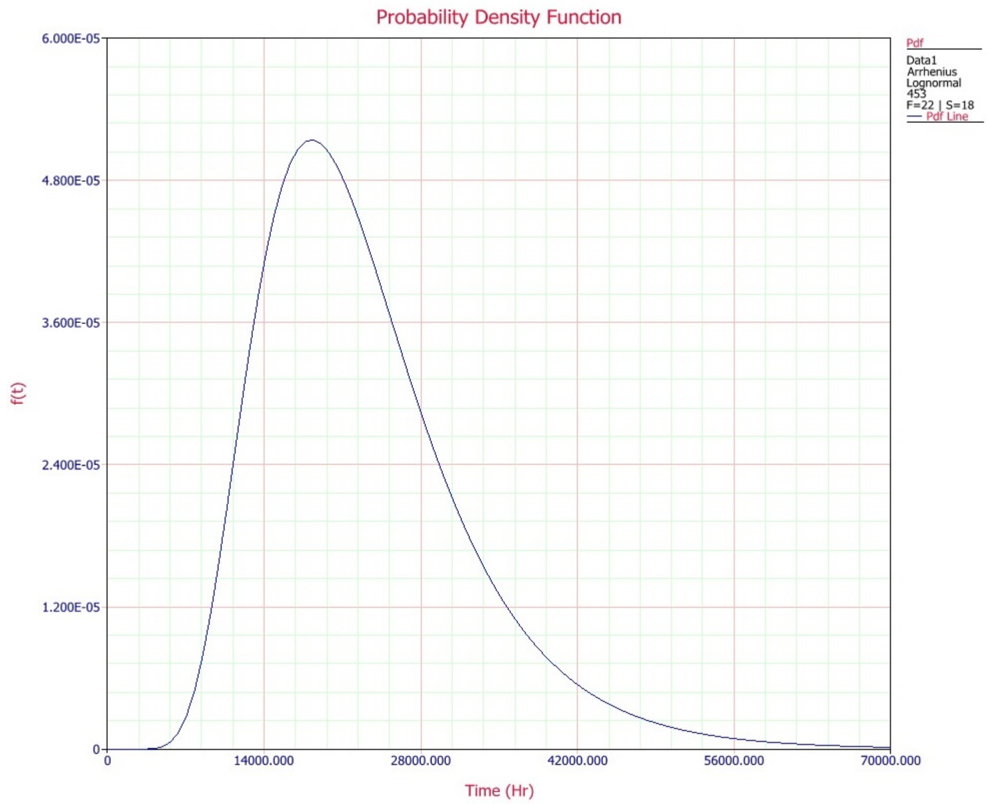


Figure 5.6: *The pdf plot of the extrapolated use-level lifetime distribution.*

It is obtained from the Arrhenius life-temperature plot in Figure 5.4 by extrapolating distribution percentiles from test data to use-level conditions.

The characteristics of the extrapolated use-level lifetime distribution which include distribution shape, skewness, mode etc. can be visualised from the pdf plot in Figure 5.6. But important quantities are reliability measures of the unit at normal operating conditions. Examples of measures of reliability include the quantifiable life measure, mean life, warranty time, conditional probability of failure etc.. These selected reliability measures depend on the estimated distribution parameters $\hat{\sigma'}_{453K} = 0.393536$ and $\hat{\mu'}_{453K} = 9.966209$ and their values are given in Table 5.7.

Use-level distribution characteristic	Estimate	90% Confidence Interval
Quantifiable Life Measure	212955	$(16319, 27787)$
Mean Life	23009	$(17516, 30226)$
Warranty Time	11147	$(8501, 14616)$
Conditional Probability of Failure	0.0873	$(0.0454, 0.1644)$

Table 5.7: *Selected measures of reliability at use-level conditions.*

The quantifiable life measure for the Arrhenius-lognormal model ($B50\%$ life) is the approximate time by which 50% of the population of units will fail. It is the median life and typically represents the lifetime of the unit. Accordingly, the lifetime of the redesigned motorette at normal operation conditions is 21295Hr. The test purpose was to achieve a median life of 20000Hr. The mean life is the average time to unit failure while the warranty time of 11147Hr is the time for a reliability of 0.95. It is also called the reliable life. Lastly, 0.0873 is the probability that the unit will fail within an additional time of 1000Hr given that it has successfully operated for 20000Hr. The uncertainty on these measures of reliability is quantified by the 90% confidence bounds.

5.7 Test data analysis: Dependent competing risks

A dependence structure between the censoring variable X_1 and unit lifetime X_2 is assumed at each stress level since degraded and critical failures are linked through the degradation process. But test data in Table 5.2 are not samples from a typical dependent competing risks situation because of the design of the life test. Accordingly, this part of the analysis uses test data from a simulation study. The problem of identifying marginal behaviour from dependent competing risks data is yet to be fully resolved. Typically, additional restrictions which cannot be tested from the observable competing risks data are imposed on the joint survival function in order to identify the marginals.

The approach adopted here is to assume that the dependence structure underlying the joint survival function of the competing risks is captured by a known copula model. Specifically, the Frank copula model is assumed to adequately describe the stochastic dependence between the censoring variable X_1 and unit lifetime X_2. To estimate the assumed Frank copula model, pairwise observations on (X_1, X_2) are generally required. In a competing risks situation however, only $Z = (X_1, X_2)$ along with the identity $j \in (1, 2)$ of the risk that achieved the minimum are observed at each test stress level. Hence test data are incomplete and expert opinion is required to estimate the assumed Frank copula model.

5.7.1 Simulation design: Dependent competing risks data

The Frank copula model that captures stochastic dependence between the censoring variable X_1 and unit lifetime X_2 was estimated from expert opinion in chapter 2. It therefore suffices to simulate observed occurrences (test samples) of X_1 and X_2 at each test stress level from the estimated Frank copula and derived competing risks data in Table 5.2 in two steps as follows:

(1) Fit life distributions to the derived competing risks data in Table 5.2 at each of the test stress levels.

(2) Generate bivariate outcomes (X_1, X_2) from the estimated Frank copula model using the fitted life distributions from the first step when inverting. In a competing risks situation,

only the minimum of X_1 and X_2 is observed. Hence obtain $Z = min(X_1, X_2)$ together with the identity $j \in (1, 2)$ of the mode that achieves the minimum.

This simulation design guarantees dependent competing risks samples at each test stress level. It does not simplify the analysis to any degree, neither does it significantly affect estimation results as follows:

- The adopted approach assumes a known copula model. Assuming test data on unit lifetime under dependent random censorship were available, expert opinion will still be required to estimate the assumed copula model since competing risks data are incomplete.
- The key factor when describing marginal behaviour from dependent competing risks data (Zheng and Klein, 1995) is a reasonable estimate of the stochastic dependence between competing risks, not the functional form of the copula.

Potential misspecification of the marginal distributions in the first step of the simulation design is a well-documented problem in copula modeling. On the basis of the calculated AIC_c values in Table 5.5, the Weibull and the lognormal distributions fit the derived competing failure mode data in Table 5.2 equally well. The Weibull distribution is preferred in this simulation study because of its further physical justification as a weakest link model. The general idea of simulating the full joint distribution first given by Genest (1987) and subsequently developed by Lee (1993) is adopted in this investigation. Assuming Weibull marginals, the algorithm is:

Algorithm 5.1: **Generating degraded and critical failure data using Frank's copula**

1. Generate independent uniform (0,1) random variables U_1 and U_2.
2. Set $X_1 = F_1^{-1}(U_1) = \alpha_1 \left(ln \frac{1}{1-U_1} \right)^{1/\beta_1}$ where α_1 and β_1 are the ML estimates of the Weibull scale and shape parameters for the degraded failure mode at a stress level.
3. Calculate X_2 as the solution to the equation

$$U_2 = e^{-\theta U_1} \left[\frac{e^{-\theta F_2(X_2)} - 1}{e^{-\theta} - 1 + (e^{-\theta U_1} - 1)(e^{-\theta F_2(X_2)} - 1)} \right].$$

That is calculate $X_2 = F_2^{-1}(U_{*2}) = \alpha_2 \left(ln \frac{1}{1-U_{*2}} \right)^{1/\beta_2}$ where $U_{*2} = -\frac{1}{\theta} ln \left[\frac{U_2 e^{-\theta} + e^{-\theta U_1}(1-U_2)}{U_2 + e^{-\theta U_1}(1-U_2)} \right]$. The parameters α_2 and β_2 are the ML estimates of the Weibull scale and shape parameters for the corresponding critical failure mode at a stress level. The parameter θ is the Frank copula parameter estimated from expert opinion.

4. Obtain $Z = min(X_1, X_2)$ and the identity of the cause that achieved the minimum.

This algorithm generates observed occurrences of degraded and critical failures in a dependent competing risks framework at each stress level. A remark is however necessary here. True acceleration alters the unit's operating conditions such that the failure causing mechanism is invariant to changes in stress but is accelerated at higher test stresses. Higher test stresses only alter the scale but not the failure mechanism as represented by the distribution's shape parameter. Accordingly, the shape parameter of the assumed lifetime distribution ought to be the same at all stress levels.

ML estimates of Weibull shape parameters in Table 5.3 slightly differ at all but the $190°C$ test stress level. These slight differences can be attributed to sampling error and varying degrees of censoring at different stress levels. With small test sample and few failures, external information is often required to supplement available data in practice. Typically, past experience or knowledge of the physics of failure is used to fix the value of the Weibull shape parameter. For more details on this approach, also called *Weibayes*, see Abernethy (2004) for example. In this thesis however, the compromise common slope obtained by fitting the life-stress model to test data estimates the Weibull shape parameter at all stress levels. Hence Weibull shape parameter values of 8 and 3 for degraded failure mode and critical failure mode respectively are used in the simulation study at all stress levels.

The `R code` in `Appendix D` generates degraded and critical failure times in a competing risks setting by implementing $\mathbf{Algorithm\ 5.1}$. The simulated competing risks data appear in `Appendix E` where column j indicates the identity of the cause that removed the unit from observation in a life test. That is, $j = 1$ if $X_1 < X_2$ and $j = 2$ if $X_2 < X_1$. Accordingly, the data in `Appendix E` would arise from a typical life test where unit lifetime is subject to dependent random censorship in a competing risks framework.

Test data in `Appendix E` are incomplete since only the minimum of the risk variables is observed. Hence only the subdistribution functions given by $F^*_{X_1}(t) = P(X_1 < t, X_1 < X_2)$ and $F^*_{X_2}(t) = P(X_2 < t, X_2 < X_1)$ can be estimated from these data, but not the true distribution functions $F_{X_1}(t) = P(X_1 < t)$ and $F_{X_2}(t) = P(X_2 < t)$. Functional forms of these estimable subdistribution functions $F^*_{X_j}(t)$, $j \in (1, 2)$ are derived in Chapter 3. In particular, observed occurrences of degraded and critical failure times are assumed to be distributed as inverse Gaussian since X_1 and X_2 have the interpretation of fist passage times of a degradation (stochastic) process to respective deterministic thresholds s_1 and s_2.

5.8 Numerical estimation of marginal survival functions

The target of estimation are the marginal survival (or distribution) functions of the censoring variable X_1 and unit lifetime X_2. They are solutions of the differential system in Equation 4.11. Input factors into this differential system are the estimated Frank copula model (Chapter 2 and derivatives of the subdistribution functions $F^*_{X_j}(\cdot)$ with respect to the time parameter t (chapter 3). The former are actually subdensity functions of the risk variables X_1 and X_2 and are denoted by $f^*_{X_j}(\cdot)$, $j \in (1, 2)$.

Suitable functional forms of $f^*_{X_j}(\cdot)$ are derived in Chapter 2. In particular, they are postulated to be inverse Gaussian with mean μ and scale λ. Table 5.8 contains the estimated parameters of the

$190°C$	X_1	X_2	$220°C$	X_1	X_2
μ	8835.214	7578.5	μ	2694.417	2033
λ	589669.5	402153.6	λ	87582.48	17499.47
$240°C$	X_1	X_2	$260°C$	X_1	X_2
μ	1682.429	3061.833	μ	NA	1187.421
λ	80450.86	2744.643	λ	NA	4519.269

Table 5.8: *Estimates of inverse Gaussian parameters for the simulated dependent competing risks data.*

inverse Gaussian distribution for the observed occurrences of the censoring variable X_1 and unit lifetime X_2 (simulated test data in `Appendix E`) at each test stress level. For both risks (X_1 and X_2), the inverse Gaussian distribution parameters μ and λ generally decrease with stress as expected. The are no parameter estimates for the censoring variable at $260°C$ stress level because the assumed distribution cannot be fitted to a single data point. That is, the censoring variable achieved the minimum once at $260°C$. Accordingly, simulated test data at the $260°C$ stress level will not be used further in the analysis.

Numerical solutions of survival functions of unit lifetime X_2 (variable of interest) at each stress level are presented in Figure 5.7.

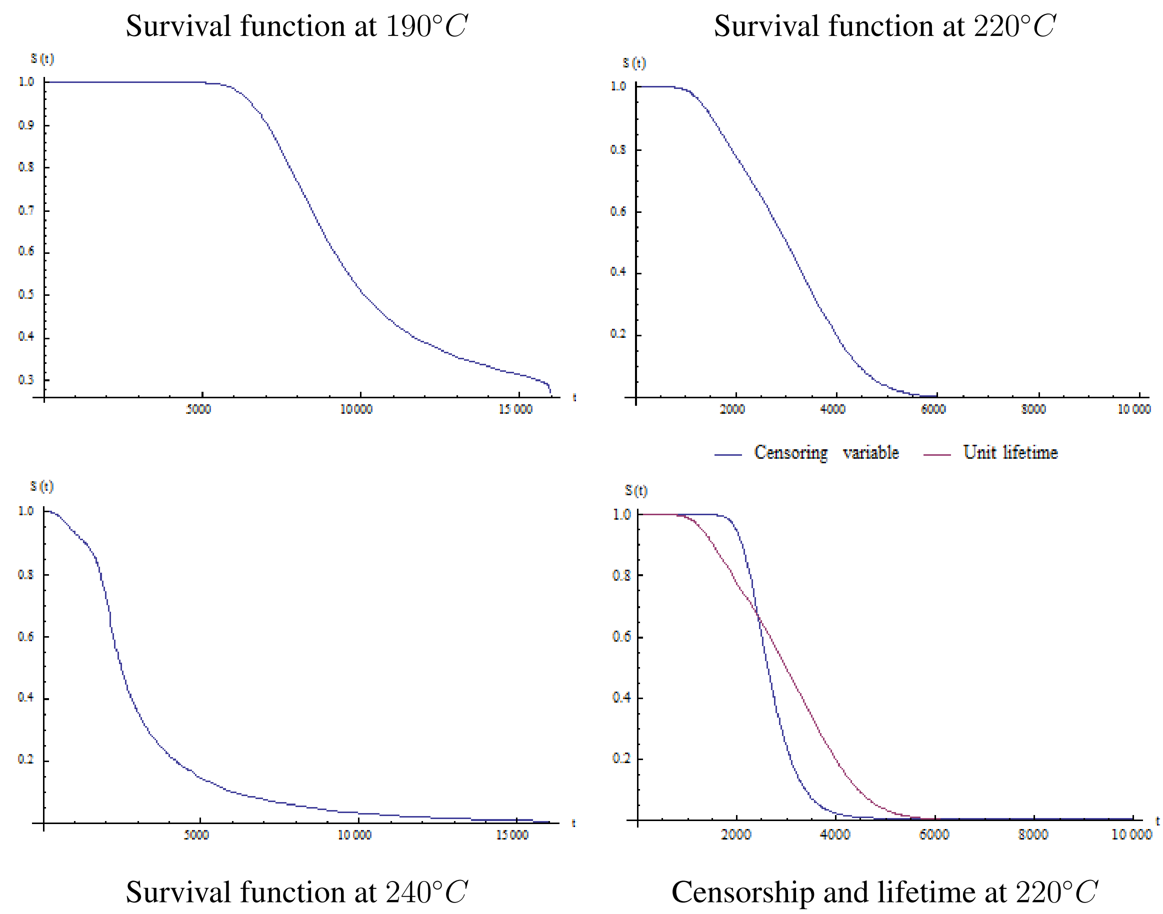


Figure 5.7: *Numerical solutions of survival functions of unit lifetime at different stress level. The bottom right plot shows both the censoring variable and unit lifetime .*

They are obtained by solving the non-linear differential system in Equation 4.11 using the Mathematica built-in function `NDSolve`. Mathematica gives numerical solutions of survival functions in the form of interpolation functions. They are complex objects that provide a smooth representation of the numerical solution. In practice however, there is no guarantee that the competing risks variables would both be non-defective for the estimated copula and subdistribution functions. In particular, unit lifetime X_2 may be defective as follows.

- The observable competing risks data are the pair $\{[Z = min(X_1, X_2)], j\}_{i=1}^{n}$. If the last observation(s) corresponds to the case(s) when X_1 achieved the minimum, then there may not be enough observed data to allow the tail of the marginal survival function of X_2 to be adequately estimated.
- A certain quantile of $F_{X_2}^{*}(\cdot)$ may only be reached at larger values of the time parameter t provided all quantiles of $F_{X_1}^{*}(\cdot)$ have been reached.

For practical purposes however, interest in accelerated testing is on early failures and hence lower percentiles of the lifetime distribution. Very high values of the time parameter may not have a defensible physical meaning in life tests. If the units are tested long enough, there will be a $t > 0$ where all units would fail by then. Admittedly however, the value of the time parameter t that will be considered too large for a specific investigation will obviously depend on that particular investigation. For example, interpolated survival functions in Figure 5.7 seem to tail off nicely for larger values of the time parameter t at all but the test temperature of of $190°C$. Real data on unit lifetime in Table 5.2 do not exceed $12191Hr$ at the test temperature of $190°C$. Since these data are used in the simulation, time parameter values in excess of $15000Hr$ may be deemed too large for this particular investigation.

Survival functions of unit lifetime in Figure 5.7 are obtained at test temperatures of $190°C$, $220°C$ and $240°C$. The target of estimation is the survival function of unit lifetime at the use-level temperature of $180°C$. Test data can be obtained from the numerical solutions in Figure 5.7 in different ways. The approach followed here is to first transform the interpolation functions to ordinary functions, find their corresponding inverses and substitute generated random uniform variates into the inverse functions. Sampled test data, together with the dependence of life on stress are utilised to extrapolate the use-level lifetime distribution of the unit.

5.9 Life-stress model for sampled test data on unit lifetime

Validity of the Arrhenius model is well established in insulation work (Nelson, 2004). Assuming the Arrhenius model, Table 5.9 shows likelihood and AIC_c values for the Weibull, exponential and lognormal lifetime distributions for sampled test data.

	Weibull	Exponential	Lognormal
LK-value	-252.5685	-280.297	-253.0792
AIC_c	508.851	563.094	511.873

Table 5.9: *LK and AIC_c values for the Weibull, exponential and lognormal lifetime distributions assuming the Arrhenius relationship for sampled unit lifetime data.*

The Weibull distribution has minimum AIC_c and is thus chosen to describe the scatter in the sampled unit lifetime data at each test stress level. Consequently, the ALT model for the sampled test data is assumed to be the Arrhenius-Weibull model where the quantifiable life measure is the Weibull scale parameter α. By Equation 4.2,

$$\alpha = \frac{1}{A} e^{\frac{E}{kV}} = C e^{\frac{B}{V}}$$

and the pdf of the Arrhenius-Weibull model at stress V is given by

$$f(t, V) = \frac{\beta}{Ce^{\frac{B}{V}}} \left(\frac{t}{Ce^{\frac{B}{V}}} \right)^{\beta - 1} exp \left(- \left(\frac{t}{Ce^{\frac{B}{V}}} \right) \right). \tag{5.6}$$

5.10 Assessment of the Arrhenius-Weibull model

The chosen Arrhenius-Weibull model assumes that

- The Weibull distribution adequately describe the scatter in the sampled unit lifetime data at each test stress level.

- The Weibull shape parameter β does not change with stress and the relationship between the Weibull scale parameter α (quantifiable life measure) and temperature (stress) is linear..

These life-stress model assumptions are checked by both graphical and analytical methods.

5.10.1 Graphical methods

Standardised residuals when the assumed Weibull distribution is fitted to sampled unit lifetime data are calculated from

$$\hat{\epsilon}_i = \hat{\beta}\left[ln\left(T_i\right) - ln\left(\hat{\alpha}(V)\right)\right] \tag{5.7}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the estimated Weibull scale and shape parameters respectively. They are a sample from the type III smallest extreme value distribution with zero mean. Accordingly, Figure 5.8 is a plot of the standardised residuals on a smallest extreme value probability paper.

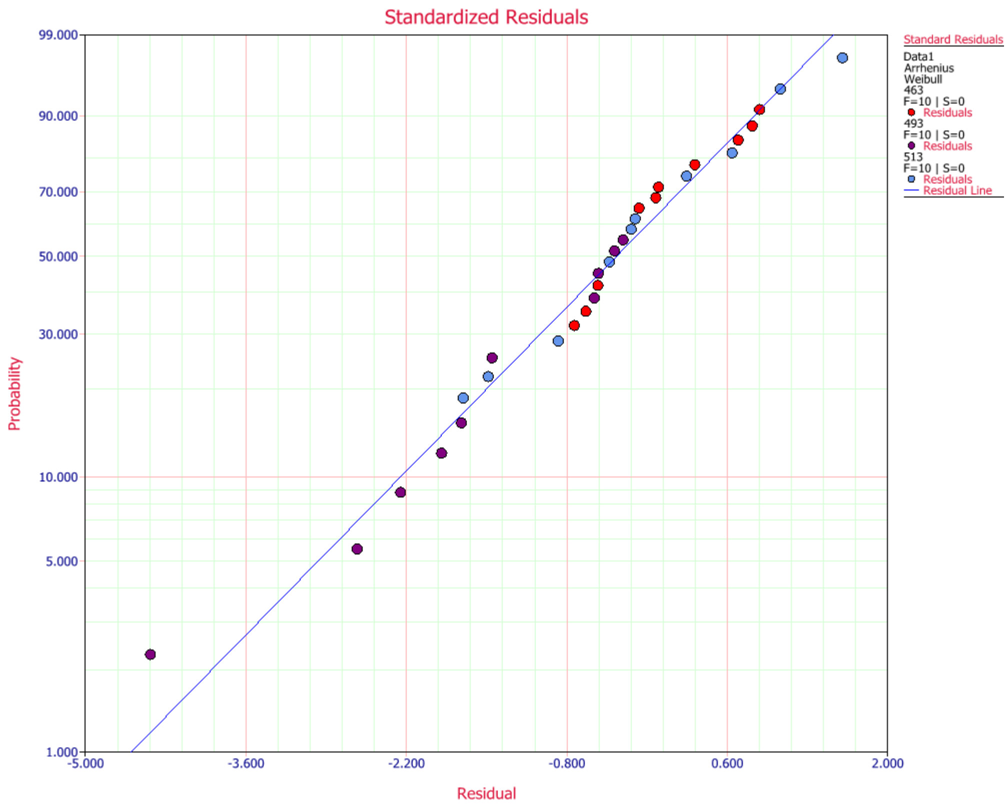


Figure 5.8: *Normal probability plot of the standardised residuals for sampled unit lifetime data.*

The standardised residuals in Figure 5.8 appear to follow a straight line on a smallest extreme value probability paper though they are some outlying data points. This suggests that the Weibull distribution is an adequate model for the sampled unit lifetime data at all stress levels. On the other hand, Figure 5.9 is a plot of the Cox-Snell residuals on an exponential probability paper.

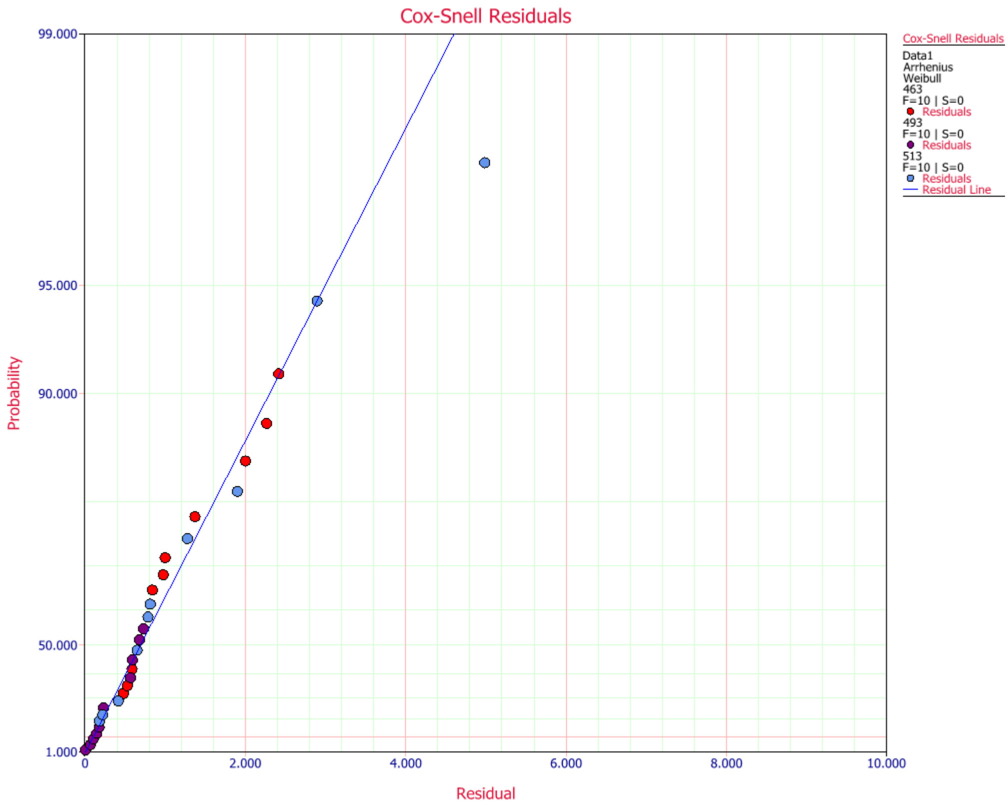


Figure 5.9: *Exponential probability plot of the Cox-Snell residuals for sampled unit lifetime data.*

As with standardised residuals, Cox-Snell residuals in Figure 5.9 fairly follow a straight line for sampled unit lifetime data at all test stress levels. Consequently, the scatter in sampled test data at all test stresses may be described be described by the Weibull distribution.

The adequacy of the assumed Weibull distribution is visually assessed further by means of the probability plot. Fairly straight Weibull probability plots indicate adequate model fit. The probability plot for sampled unit lifetime data is given in Figure 5.10.

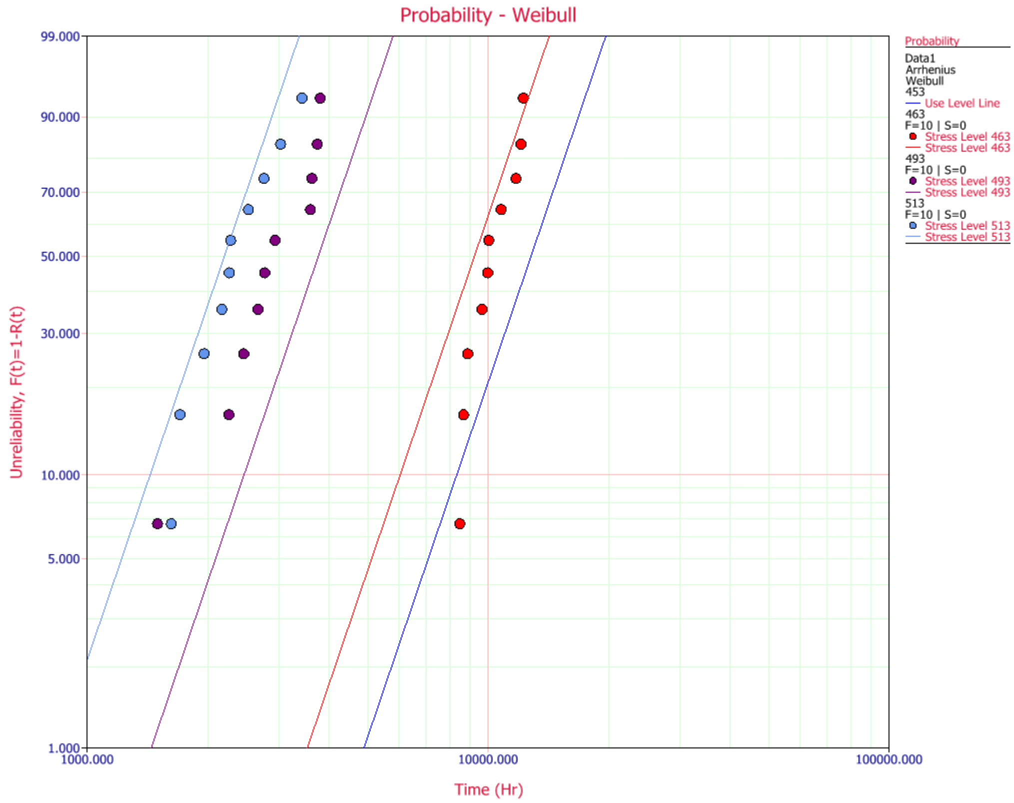


Figure 5.10: *CDF plot on linearised lognormal paper for sampled unit lifetime data.*

Though erratic, points in the Weibull probability plot in Figure 5.10 are fairly straight. This suggests that the scatter in the sampled unit lifetime data can be described by the Weibull distribution. In addition, the CDF plots on linearised Weibull paper in Figure 5.10 are fairly parallel, albeit with a compromise common slope. This implies that the failure mechanism does not necessarily change with stress as required under true acceleration.

In order to extrapolate a use-level unit lifetime distribution from the sampled unit lifetime data at the different test stress levels, the Arrhenius-Weibull model assumes that the scale parameter (quantifiable life measure) α linearly changes with temperature (stress). This linearity assumption is visually assessed by life-stress plots of the sampled unit lifetime data in Figure 5.11.

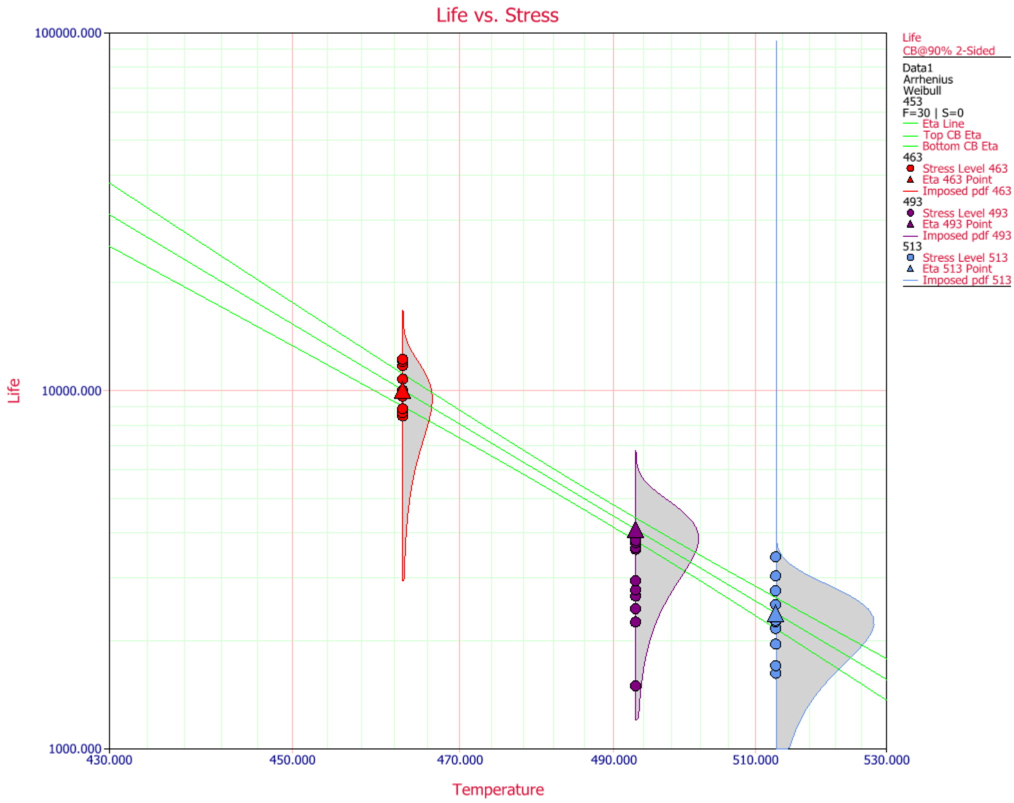


Figure 5.11: *Arrhenius life-temperature relationship plot for sampled unit lifetime data.*

The relationship between the quantifiable life measure, α and temperature in Figure 5.11 is clearly linear based on the linear sample percentile (middle) line. Accordingly, the linearity assumption seems to hold for these data. The linear life -stress relationship allows distribution percentiles to be extrapolated from test stresses to use-level conditions. This yields the survival (or distribution) function of the unit at design stress.

5.10.2 Numerical methods

The LR test is utilised to objectively assess the assumptions of a common Weibull shape parameter across the sampled unit lifetime data at the different stress and a linear life-stress relationship. The hypotheses when testing if the Weibull shape parameter depends on stress level are:

H_0: The shape parameter is independent of stress level

H_1: The shape parameter is dependent on stress level.

For the sampled unit failure time data in `Appendix F`, the maximum log likelihood values are $\hat{\mathscr{L}}_{463} = -86.370$, $\hat{\mathscr{L}}_{493} = -79.750$, $\hat{\mathscr{L}}_{513} = -77.435$ and $\hat{\mathscr{L}}_0 = -245.223$. Hence the value of the test statistic is

$$T_{LR} = -2\left(-245.223 - (-86.370 - 79.750 - 77.435)\right) = 3.336.$$

At the 10% level of significance and for the $j = 3$ test temperatures, $\chi^2(0.1; 2) = 4.605$. Since the value of the likelihood ratio test statistic $T_{LR} = 3.336 < 4.605 = \chi^2(0.1; 2)$, H_0 cannot be rejected at the 10% level of significance. Hence the Weibull shape parameter estimates are not statistically significantly different at the 10% level of significance.

When testing if the life-stress relationship is linear, the hypothesis are

H_0: Sampled unit lifetime data are consistent with a linear life-stress relationship

H_1: Sampled unit lifetime data are not consistent with a linear life-stress relationship

For the data in Appendix E, $\hat{\mathscr{L}} = -251.569$ $\hat{\mathscr{L}}_0 = -245.223$. Consequently, the value of the test statistic is

$$T_{LR} = -2\left(-245.223 - (-251.569)\right) = -12.692$$

whereas $\chi^2(0.1; 2) = 4.605$. Since $T_{LR} = -12.692 < 4.605 = \chi^2(0.1; 2)$, H_0 cannot be rejected at the 10% level of significance. Therefore there is insufficient evidence at the 10% level of significance to suggest that the sampled unit lifetime data are statistically significantly different from a linear life-stress relationship.

5.11 Extrapolating the use-level survival function

Following the graphical and numerical assessment results, the Arrhenius-Weibull model appears to adequately describe the sampled unit lifetime data. The extrapolated survival function of the unit at the use-level temperature of $180°C$ is given in Figure 5.12.

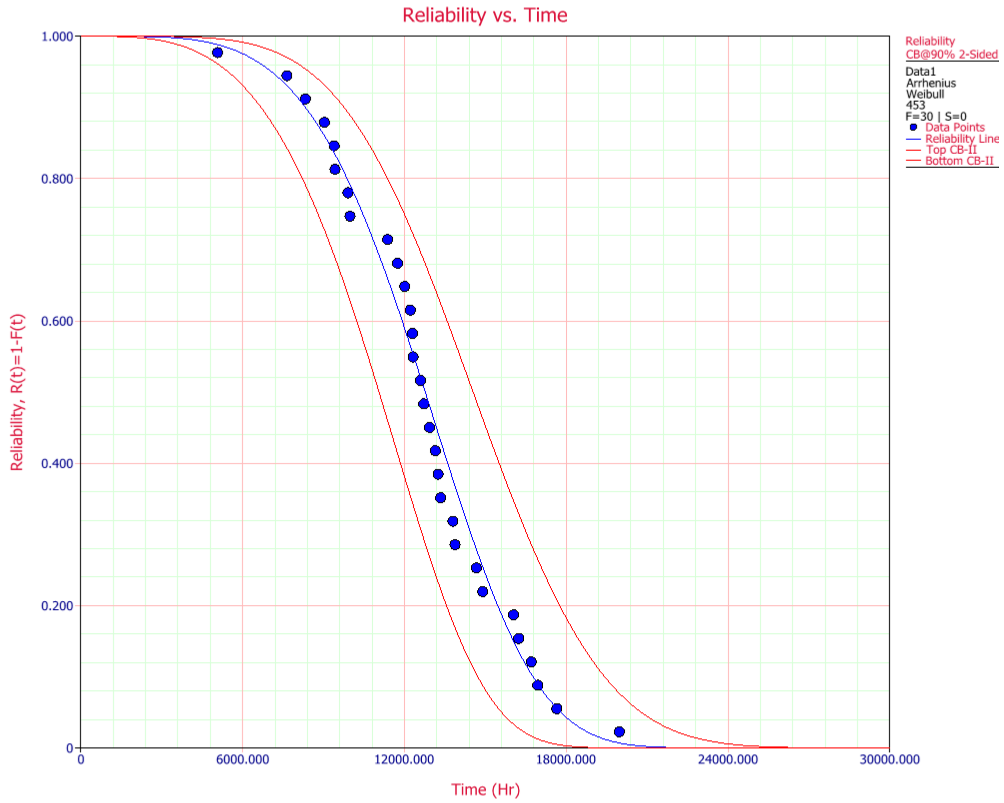


Figure 5.12: *The survival function of the unit at the use-level temperature.*

Estimated Weibull distribution parameters at the use-level temperature of $180°C$ are $\hat{\alpha} = 13876$Hr and $\hat{\beta} = 4.411$ where the former estimates the lifetime of the redesigned unit. Consequently, important measures of reliability are derived. Selected important reliability measures calculated from the extrapolated use-level survival function are given in Table 5.10

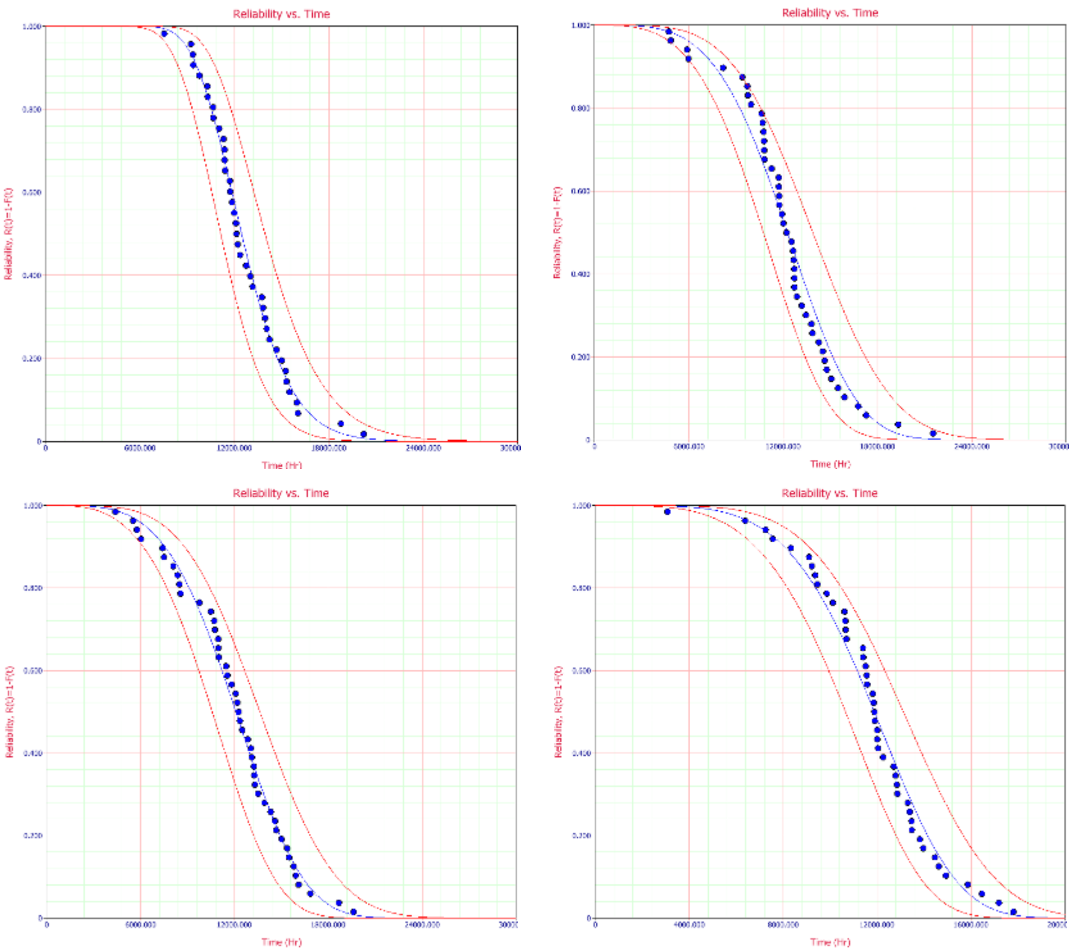
	Reliability measure	90% Confidence limits
$B50\%$ Life	12770.354Hr	$(11177.662, 14589.988)$
Mean life	12648.872Hr	$(11065.936, 14458.241)$
Warranty time	7077.222Hr	$(5736.233, 8731.700)$

Table 5.10: *Selected reliability measures at use-level temperature.*

5.12 Sensitivity analysis

The adopted copula-based competing risks methodology largely depends on the elicited rank correlation and hence, the estimated copula model parameter. This section presents the sensitivity of the extrapolated survival function of the unit at the use-level temperature of $180°C$ with respect to different degrees of stochastic dependence between the risk variables. In particular, the differential system in Equation 4.11 is further solved for values of the copula dependence parameter θ that correspond to values of Kendall's τ equal to 0.25, 0.5, 0.75 and 0.9. The resulting survival functions of the unit at the use-level temperature of $180°C$ are given in Figure 5.13.

Survival function: $\tau = 0.25, \theta = 2.371$ Survival function: $\tau = 0.5, \theta = 5.736$



Survival function: $\tau = 0.75, \theta = 14.138$ Survival function: $\tau = 0.9, \theta = 38.281$

Figure 5.13: *Use-level survival functions of the unit assuming different degrees of dependence.*

Results in Figure 5.13 show the difference that the elicited degree of dependence between the censoring variable and unit lifetime makes when estimating the lifetime distribution of the unit at use-level conditions. The extrapolated use-level survival functions in Figure 5.13 reveal an apparent shift to the left as the strength of rank correlation and hence stochastic dependence between the risks increases. Thus for strong positive rank correlation, removal of the censoring variable leads to poorer survival with respect to unit lifetime. This is made more clearer by looking at the estimated measures of reliability contained in Table 5.11.

	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.9$
$B50\%$ Life	12456Hr	12270Hr	12092Hr	11881Hr
Mean life	12704Hr	12190Hr	12011Hr	11745Hr

Table 5.11: *Sensitivity of estimated reliability measures to different degrees of dependence.*

Poor survival with respect to the remaining failure mode assuming strong positive rank correlation makes sense as follows. Assuming strong stochastic dependence between the failure modes, the remaining failure mode will continue to operate in the same way as the removed mode. As a result, cause removal will not significantly improve survival with respect to the remaining modes. Of interest in this investigation is quantifying reliability characteristics of the redesigned unit from the extrapolated use-level lifetime distribution. Results in Table 5.11 show slight differences in selected reliability measures for different rank correlation values. In particular, $B50\%$ Life is approximately 12000Hr for the considered four rank correlation values.

This somewhat surprising result of slight differences in reliability measures for different degrees of dependence was also obtained by Meeker, Escober and Hong (2009). In their analysis however, they assumed a bivariate lognormal model for the competing risks and estimated model parameters by ML estimation. To the contrary, this investigation assumes a copula model. The practical implication of this result is as follows: When estimating the survival function of the unit at use conditions, one may use a degree of dependence between the censoring variable and unit lifetime that is believed to be realistic to admit.

Chapter 6

Summary, conclusions and suggestions for future research

6.1 Research summary

Industrial units generally fail from different failure modes and these are often present in ALT. When mode and lifetime information are available, competing risks theory provides the appropriate model for analysing failure data. In the case of life tests where unit lifetime is subject to random censorship, two failure modes namely degraded failure and critical failure are distinguished at each stress level.

A simplifying assumption that is often made when analysing competing risks data is that the risks act independently. This ensures identifiability of the marginal (and hence joint) survival functions. But since degraded and critical failures are linked through the degradation process, the investigated problem is that of modeling dependent competing risks. More general copula methods are preferred to classical families of multivariate distributions and their parametrisation by rank correlation is used to estimate the copula model using expert opinion. In particular, the Frank copula dependence parameter $\hat{\theta} = 2.8405$ corresponding to an assessed rank correlation $\hat{\tau} = 0.29317$ was obtained. This result makes sense for degrading units during testing since:

- Degraded failures are expected to occur close to critical failures. Hence the agreement between their rankings is largely expected to be positive.
- For any two independent draws from a population of test units, if it turns out in a life test that the lifetime of one unit is longer than for the other, it does not necessarily follow that its corresponding censoring time will also be higher. Hence the agreement between the rankings of degraded and critical failures is not expected to be perfect.

Functional forms (or models) of the observed occurrences of degraded and critical failures (competing risks data) are derived from a stochastic process point of view. Stochastic processes that are commonly used in reliability and life testing studies as degradation models are the Wiener and gamma processes. Both satisfy the Markov property and are therefore Markov processes. In addition, the adopted degradation modeling viewpoint assumes that the underlying failure causing process is not fully observable. Together with Wiener and gamma processes being Markov processes, this motivated the modeling framework of hidden Markov processes, a more general statistical and structural approach.

Hidden Markov processes are bivariate stochastic processes with a hidden failure causing process that can only be observed through another process called the observation (marker) process. Emphasis was placed on selecting a suitable probability structure for the doubly stochastic process model satisfying the following:

- Both the underlying failure causing process and the marker process are Markov processes.
- Sample paths of the underlying failure causing process are restricted strictly monotone increasing functions in order to account for the irreversible accumulation of damage that leads to unit failure in a life test. On the other hand, sample paths of the marker process are only restricted to continuous functions.
- The marker process must be useful in terms of tracking progress of the underlying failure causing process. Hence the marker and the failure causing process must be linked in a natural way.

A bivariate stochastic process model with these properties (Whimore *et al.*, 1998) is obtained by describing the latent failure causing process by the Wiener maximum process and the marker process by the Wiener process. Time to unit failure was estimated by first passage times of the failure causing Wiener maximum process to deterministic failure thresholds and hence no additional degradation data was necessarily required. The first passage time of the Wiener maximum process to a failure threshold coincides with that of the Wiener process to the same failure threshold. It is well-known (Chhikara and Folks, 1989) to be distributed as inverse Gaussian. Consequently the inverse Gaussian distribution is postulated as the probability model for the observed occurrences of degraded and critical failures in a life test. Contributions of the observable competing risks data to the likelihood function are derived and ML estimators are obtained.

Statistical modeling of life data from accelerated tests relied on the popular ALT model which combines the lifetime distribution and the life-stress relationship. The former is assumed to come from a specified parametric family while the choice of the latter was based on theory and literature on similar tests. In particular, the Weibull and lognormal distributions are motivated as appropriate models for life data from accelerated tests. Early unit failures in life tests imply that test data are generally positively skewed. Thus the Weibull and lognormal distributions are popular in life testing studies partly because they have positively skewed frequency curves.

In addition, a physical motivation for the Weibull distribution as a model for test data stems from its interpretation as a limiting distribution for minima. This makes it an acceptable model for the first occurring failure in situations where there are competing failure modes as is the case in this investigation. On the other hand, physical motivation for the Lognormal distribution is based on the central limit theorem. But more importantly, interest in accelerated testing is largely in estimating lower percentiles of the lifetime distribution and both the Weibull and lognormal distributions flexibly fit test data over their lower tails. Goodness-of-fit tests, in particular AIC_c was used to discriminate between these two lifetime distributions where the model with minimum AIC_c value is better

Maximum likelihood estimation and median rank regression are discussed as statistical methods for fitting test data samples to the chosen lifetime distribution. Results from studies that compare MLE and MRR methods for estimating Weibull and lognormal distributions are generally mixed because of study differences. Their performance in small samples under Type I censoring is considered since few units are often tested and test data are right censored. Both estimation procedures are shown to be biased for such data but MLE is preferred in terms of the overall accuracy since ML estimators have minimum error variance.

Extrapolation of the use-level unit lifetime distribution is based on the life-stress relationship which describes how a quantifiable life measure of the assumed lifetime distribution changes with stress. Test data are from a temperature accelerated test of motorette insulation. The Arrhenius relationship is chosen firstly because it is based on the laws of physics and secondly because of its well established validity in insulation work (Nelson, 2004).

The analysis of test data on unit lifetime under dependent random censorship is complicated by

- Test data not being readily available. Hence the analysis relied on test data from a simulation study.
- Marginal behaviour of the competing risks variables not generally identifiable. Consequently, parametric restrictions had to be placed on the joint behaviour of the competing risks.

In particular, stochastic dependence between unit life and the censoring variable is described by the Frank copula model estimated using expert opinion at each stress level. Given the estimated copula model, test data identify the marginal behaviour of the competing risks at each test stress level. Emphasis is on the marginal behaviour of unit lifetime since it is the variable of interest. Based on AIC_c and assuming the Arrhenius relationship, the scatter in simulated test data is adequately described by the Weibull distribution. Consequently the Arrhenius-Weibull model is the assumed life-stress relationship for these data. It is assessed for goodness-of-fit using both graphical and numerical methods. The chosen Arrhenius dependence of life (as measured by the Weibull scale parameter) on temperature (stress) is utilised to extrapolate a use-level unit lifetime distribution from simulated test data.

Based on the extrapolated unit lifetime distribution, a number of reliability measures of the redesigned motorette are derived. These include the lifetime of the unit as estimated by the Weibull scale parameter, $B50\%$ life, mean life, warranty time etc. But the adopted modeling approach largely depends on the elicited rank correlation and hence the estimated copula model. Accordingly, sensitivity of the derived measures of reliability to rank correlation values ranging from $\tau = 0.25$ to $\tau = 0.9$ yielded the following results and conclusion:

- The stronger the agreement between the rankings of unit lifetime and the censoring variable leads to poorer survival with respect to the failure mode of interest (unit lifetime in this case) when the other mode is removed.
- Derived reliability measures from the extrapolated use-level unit lifetime distribution differ slightly for a wide range of rank correlation values. This somewhat surprising result led to the conclusion that a degree of dependence that is believed to be realistic to admit is the important factor when estimating marginal survival functions from dependent competing risks data. Consequently prior knowledge or experience with similar tests may be useful factors in the analysis.

6.2 Suggestions for future reasearch

The modeling framework adopted in this investigation pertains to first passage times of the underlying failure causing degradation process to failure thresholds. Consequently the measurable variable is unit failure time and is estimated by obtaining crossing times of the degradation process to these failure thresholds. Other problems worthy of further study are as follows:

- In addition to failure times, test samples may also include observations of increments of the degradation process. That is, degradation increments are assumed to be observable during testing provided the underlying failure causing process has not exceeded failure thresholds. This modeling framework is useful particularly for highly reliable units where failure times may not be observed even under accelerated conditions. Inference procedures for these and other related models are described in Kahle and Lehmann (1998) for example, though not necessarily in an accelerated testing context.

- For some units, field operating conditions are often highly variable whereas accelerated tests are carefully controlled. To construct a model that relates the two conditions, field data from warranty returns for example may be required in addition to data from life tests. Among the possible models that relate these two conditions is the use rate model. It applies to cases where failure mode(s) depends on the use-rate, which intern, varies enormously among units in the product population. Otherwise, the reliability-based methodology for relating the two conditions is required and would typically involve the following steps. (1) Develop ALTs that yield the same failure mode(s) as in field performance. The failure mode(s) must also be driven by the same failure mechanism. This can be achieved through physical failure mode analysis and if there are discrepancies between the two conditions as field data become available, testing procedures may be modified so that failure modes agree. (2) If ALT mimics field conditions, field performance would be directly estimated by testing performance. Since field and testing performance generally differ, a bias correction factor may be incorporated into the reliability-based methodology. The shift factor, possibly deterministic, may also be incorporated to account for the highly variable field operation conditions and other unobservable variables. The latter implies that a random model error term must also be incorporated.

- Given the small sample sizes in ALT, the Bayesian approach may be worthy considering in future research.

Appendix A

The R code for generating an estimated value of Kendall's τ from the assessed concordance probability.

```
simwei=function(n,a,b,p,e)\{
x2=rweibull(n,shape=a,scale=b)
y2=rep(0,n)
z1=rexp(n,e)
z2=rexp(n,e)
count=1
repeat\{
if(count==(n+1)) break
y=rweibull(1,shape=a,scale=b)
if(y>x2[count])\{
y2[count]=y
count=count+1
\}
\}


x1=p*x2+z1
y1=p*y2+z2
k=0
```

```
for(i in 1:n)\{
if(y1[i]>x1[i]) k=k+1
\}
prob=k/n
tau=2*prob-1
list(x2=x2,y2=y2,z1=z1,z2=z2,x1=x1,y1=y1,n=n,k=k,prob=prob, tau=tau)
\}
sim1=simwei(n=75,a=3,b=1,p=0.85,e=1)
```

Repeating the simulation 1000 times

```
simN=1000
output=c(0,0,0)
for(s in 1:simN)\{
out=simwei(n=75,a=3,b=1,p=0.85,e=1)
output=rbind(output,c(out$n,out$k,out$prob,out$tau))
\}
output=output[-1,]
colnames(output)=c("n","k","prob","tau")
output=as.data.frame(output)
x=output$tau
```

Appendix B

The 95% confidence intervals for Weibull parameter α and β at each test stress level and for each failure mode.

X_1			X_2		
$190°C$	Estimate	95% CI	$190°C$	Estimate	95% CI
α	9299.445	$(8510.697; 10102.115)$	α	12610.309	$(12081.471; 14471.543)$
β	8.502	$(4.898; 13.198)$	β	19.670	$(6.453; 43.913)$
$220°C$	Estimate	95% CI	$220°C$	Estimate	95% CI
α	2817.588	$(2543.305; 3100.454)$	α	4020.505	$(3106.157; 6360.134)$
β	7.398	$(4.198; 11.742)$	β	2.952	$(1.307; 5.375)$
$240°C$	Estimate	95% CI	$240°C$	Estimate	95% CI
α	1749.829	$(1613.825; 1904.822)$	α	2585.271	$(1916.236; 9853.710)$
β	9.234	$(5.035; 14.725)$	β	3.228	$(0.853; 7.969)$
$260°C$	Estimate	95% CI	$260°C$	Estimate	95% CI
α	1782.517	$(1615.519; 2182.962)$	α	1311.883	$(986.133; 1759.732)$
β	8.183	$(3.393; 14.903)$	β	2.636	$(1.459; 4.207)$

Appendix C

The 95% confidence intervals for the lognormal scale and shape parameters μ' and σ' at each test stress level and for each failure mode.

X_1			X_2		
$190°C$	Estimate	95% CI	$190°C$	Estimate	95% CI
μ'	9.071	$(8.977; 9.165)$	μ'	9.432	$(9.371; 9.579)$
σ'	0.144	$(0.093; 0.229)$	σ'	0.086	$(0.046; 0.234)$
$220°C$	Estimate	95% CI	$220°C$	Estimate	95% CI
μ'	7.863	$(7.744; 7.983)$	μ'	8.141	$(7.886; 8.530)$
σ'	0.182	$(0.118; 0.290)$	σ'	0.372	$(0.218; 0.813)$
$240°C$	Estimate	95% CI	$240°C$	Estimate	95% CI
μ'	7.409	$(7.311; 7.514)$	μ'	7.783	$(7.425; 9.073)$
σ'	0.144	$(0.096; 0.251)$	σ'	0.484	$(0.228; 1.681)$
$260°C$	Estimate	95% CI	$260°C$	Estimate	95% CI
μ'	7.444	$(7.305; 7.689)$	μ'	6.981	$(6.695; 7.283)$
σ'	0.198	$(0.115; 0.450)$	σ'	0.417	$(0.277; 0.729)$

Appendix D

The R code for generating observed occurrences of degraded and critical failure times in a competing risks setting.

```
u1=runif(20, min=0, max=1)
u2=runif(20, min=0, max=1)

v=(-1/2.8405)*log((u2*exp(-2.8405)+exp(-2.8405*u1)*
(1-u2))/(u2+exp(-2.8405*u1)*(1-u2)))

x1=alpha1*(log(1/(1-u1)))^(1/8)
x2=alpha2*(log(1/(1-v)))^(1/3)
y=cbind(x1,x2)
min=apply(cbind(x1,x2),1,min)
out=x1-x2
```

out[$out > 0$]="x2"

out[out==0]="same"

out[$out < 0$]="x1"

Appendix E

Simulated test data on unit lifetime under dependent random censorship.

$190^\circ C$	$min(X_1, X_2)$	j	$220^\circ C$	$min(X_1, X_2)$	j	$240^\circ C$	$min(X_1, X_2)$	j	$260^\circ C$	$min(X_1, X_2)$	j
1	8322	1	21	2362	2	41	915	2	61	1399	2
2	9491	1	22	2663	1	42	11751	2	62	1264	2
3	6329	1	23	1728	1	43	1548	1	63	1274	2
4	10434	1	24	2681	1	44	1826	1	64	198	2
5	8341	1	25	2341	2	45	1855	2	65	1545	2
6	7686	2	26	3006	1	46	1419	1	66	1197	2
7	9482	1	27	3090	1	47	1688	1	67	1436	2
8	6085	2	28	3019	1	48	2058	1	68	1246	2
9	8853	1	29	1881	2	49	1357	1	69	1921	1
10	8003	1	30	2696	1	50	1799	1	70	1450	2
11	8862	1	31	1979	2	51	1504	1	71	1672	2
12	8008	2	32	2441	1	52	1422	2	72	1194	2
13	8724	2	33	847	2	53	1855	1	73	1239	2
14	9455	1	34	2019	2	54	1871	1	74	782	2
15	8542	2	35	3065	1	55	1226	1	75	928	2
16	7640	1	36	3100	1	56	1472	2	76	1495	2
17	6426	2	37	2043	1	57	956	2	77	1111	2
18	9284	1	38	2314	2	58	1929	1	78	1385	2
19	9383	1	39	2521	2	59	1808	1	79	811	2
20	9814	1	40	2801	1	60	1666	1	80	935	2

Appendix F

Sampled data on unit lifetime from the numerical solutions at test temperatures of $190°C$, $220°C$ and $240°C$.

$190°C$	X_2	$220°C$	X_2	$240°C$	X_2
1	8161	16	3866	31	3434
2	8460	17	2665	32	2760
3	8897	18	1498	33	4320
4	12072	19	3634	34	492
5	13623	20	2255	35	1704
6	8492	21	3815	36	1622
7	11738	22	4983	37	2279
8	12912	23	4397	38	2523
9	10032	24	2457	39	1957
10	9776	25	3752	40	2261
11	12249	26	2945	41	3036
12	10776	27	917	42	2166
13	8687	28	3604	43	5625
14	9649	29	1323	44	5749
15	13745	30	2772	45	627

References

Aalen, O. O and Gjessing, H. (2001), Understanding the shape of the hazard rate: a process point of view, *Statistical Science*, vol. 16, pp. 1–22.

Abdel-Hamid, M. (1975), A gamma wear process, *IEEE Transactions on Reliability*, vol. 24, n.2, pp. 152–153.

Abernethy, R.B. (1994), *The New Weibull Handbook*, Gulf Publishing, Texas, USA.

Abernethy, R.B. (2004), *The New Weibull Handbook*, North Palm Beach, FL.

Anjos, U., Kolev, N. and Tanaka, N. (2005), Copula associated to order statistics, *Brazilian Journal of Probability and Statistics*, vol. 19, pp. 111–123.

Bagdonavicius, V. (1978), Testing the hypothesis of the additive accumulation of damages, *Journal of Statistical Planning and Inference*, vol. 126, pp. 535–551.

Barlow, R.E. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, Inc.

Bedford, T. (2006), Copulas, degenerate distributions and quantile tests in competing risk problems, *Journal of Statistical Planning and Inference*, vol. 136, pp. 1572–1587.

Bedford, T. and Alkali, B.M. (2009), Competing risks and opportunistic informative maintenance, *Journal of Risk and Reliability*, vol. 223, pp. 363–372.

Bedford, T. and Cooke, R.M. (2001), *Probabilistic Risk Analysis: Foundations and Methods*, Cambridge University Press.

Beichelt, F. (2006), *Stochastic Processes in Science, Engineering and Finance*, Chapman & Hall/CRC.

Berkson, J. (1950), Are there two regressions, *Journal of the American Statistical Association*, vol. 45, pp. 164–180.

Bhattacharrya, G.K. and Fries, A. (1982), Fatigue failure models - Birnbaum-Saunders vs. inverse Gaussian, *IEEE TRANSACTIONS ON RELIABILITY*, vol. 31, no.5, pp. 439–441.

Birnbaum, Z. W. and Saunders, S. C. (1969), A new family of life distributions, *Journal of Applied Probability*, vol. 6, pp. 319–327.

Bunea, C. and Bedford, T. (2002), The Effect of Model Uncertainty on Maintenance Optimization, *IEEE TRANSACTIONS ON RELIABILITY*, vol. 51, no.4, pp. 486–493.

Bunea, C. and Mazzuchi, T. A. (2007), Accelerated Life Tests: Analysis with Competing Failure Modes. In “Encyclopedia of Statistics in QUALITY and RELIABILITY”. Vol **1**, pp. 30–43. Edited by F. Ruggeri, R. S. Kenett, F. W. Faltin, John Wiley and Sons Ltd, England.

Carriere, J.F. (1994), Dependent Decrement Theory, *Transactions of Society of Actuaries*, vol. 46, pp. 45–65.

Chhikara, R. S. and Folks, J. L. (1989), *The Inverse Gaussian Distribution: Theory, Methodology and Applications*, Marcel Dekker, New York, NY.

Christer, A. (2002), A review of delay time analysis for modelling plant maintenance, In: Osaki, S. (Ed.), Stochastic Models in Reliability and Maintenance. Springer, Berlin, pp. 89–124.

Cinlar, E. (1972), Markov additive processes II, *Z Wahrscheinlichkeitstheorie verw. Geb.*, vol. 24, pp. 95–121.

Clemen, R.T., Fischer, G.W., and Winkler, R.L. (2000), Assessing dependence: Some experimental results, *Management Science*, vol. 46, pp. 1100–1115.

Clemen, R.T. and Reilly, T. (1999), Correlations and copulas for decision and risk analysis, *Management Science*, vol. 45, no.2, pp. 208–224.

Conover, W.J. (1999), *Practical nonparametric statistics*, John Wiley and Sons, New York.

Cooke, R.M. (1991), *Experts in uncertainty*, Oxford University Press.

Cooke, R. (1996), The design of reliability databases, part II: competing risk and data compression, *Reliability Engineering System Safety*, vol. 51, pp. 209–223.

Cooke, R. and Bedford, T. (2002), Reliability databases in perspective, *IEEE TRANSACTIONS ON RELIABILITY*, vol. 51, no.3, pp. 294–310.

Cooke, R.M. and Goossens, L.H.J. (2000), Procedures guide for structured expert judgement in accident consequence modelling, *Radiation Protection Dosimetry*, vol. 90, no.3, pp. 303–309.

Crowder, M. (2001), *Classical competing risks*, Chapman & Hall/CRC

David, H. A. and Moeschberger, M. L. (1978), *The Theory of Competing Risks*, London: Griffin.

De Boor, C. (2001), *A Practical Guide to Splines*, Springer, New York.

Dijoux, Y. and Gaudoin, O. (2009), The alert-delay competing risks model for maintenance analysis, *Journal of Statistical Planning and Inference*, vol. 139, no.5, pp 1587–1603.

Dimitrova, D.S., Haberman, S. and Kaishev, V.K. (2013), Dependent competing risks: cause elimination and its impact on surviaval, *Insurance: Mathematics and Economics*, vol. 53, pp. 464–477.

Doksum, K.A. (1991), Degradation models for failure time and survival data, *CWI Quarterly, Amsterdam*, vol. 4, pp. 195–203.

Dudley, R.M. (1989), *Real analysis and probability*, Wadsworth, Inc., Belmont, California.

Dumonceaux, R. and Antle, C.E. (1973), Discrimination between the lognormal and the Weibull distributions, *Technometrics*, vol. 9, pp. 119–129.

Durante, F. and Sempi, C. (2010), Copula theory: an introduction. In: F. Durante, W. Hardle, P. Jaworski, T. Rychlik (eds.) Workshop on Copula Theory and its Applications, Springer, Dortrecht.

Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

Escarela, G. And Carriere, J.F. (2003), Fitting competing risks with an assumed copula, *Statistical Methods in Medical Research*, vol. 12, pp. 333–349.

Escober, L. A. and Meeker, W. Q, (2006), A Review of Accelerated Test Models, *Statistical Science*, vol. 21, no.4, pp. 552–577.

Fang, K.T., Kotz, S. and Ng, K.W. (1990), *Symmetric multivariate and related distributions*, Chapman & Hall, London.

Gail, M. (1975), A review and critique of some models used in competing risk analysis, *Biometrics*, vol. 31, pp. 209–222.

Garthwaite, P.H., Kadane, J.B. and O'Hagan, A. (2005), Statistical methods for eliciting probability distributions, *Journal of the American Statistical Association*, vol. 100, no.470, pp. 680–701.

Genest, C. (1987), Frank's family of bivariate distributions, *Biometrika*, vol. 74, pp. 549–555.

Genest, C. and Favre, A.-C. (2007), Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, vol. 12, no.4, pp. 347–368.

Genest, C. and Remillard, B. (2006), Discussion of "Copulas: Tales and facts", *Extremes* vol. 9, pp. 27–36.

Genschel, U. and Meeker, W.Q. (2010), A comparison of maximum likelihood and median-rank regression for Weibull estimation. *Quality Engineering*, vol. 22, pp. 236–255.

Gigerenzer, G. (1991), How to make cognitive illusions disappear: Beyond heuristics and biases, *European Rev. Social Psych.*, vol. 2, pp. 83–115.

Gigerenzer, G., Hoffrage, U. and Kleinbolting, H. (1991), Probabilistic mental models: A Brunswikian theory of confidence, *Psychological Rev.*, vol. 83, pp. 506–528.

Gumbel, E.J. (1958), *Statistics of extremes*, Columbia Univ. Press, New York, USA.

Hahn, G.J. and Shapiro, S.S. (1967), *Statistical models in engineering*, Wiley, New York, USA.

Hove, H. (2013), On the use of expert opinion to characterise the joint behaviour of competing risks in industrial accelerated life testing, In: Yeomans, J.S., Montemanni, R. and Norlander, T.E. (eds), *Lecture Notes in Management Science*, Vol. 5, pp. 33–38, Tadbir

Kahle, W. (1994), Simultaneous confidence regions for the parameters of damage processes, *Statistical Papers*, vol. 35, pp. 27–41.

Kahle, W. and Lehmann, A. (1998), Parameter estimation in damage processes: Dependent observations of damage increments and first passage time, In: Kahle, W. *et al.* (eds), *Advances in Stochastic Models for Reliability, Quality and Safety*. Boston: Birkhauser, pp. 139–152.

Kahle, W. and Lehmann, A. (2010), The Wiener process as a degradation model: Modeling and parameter estimation, In: Nikulin, M.S., Limnios, N., Balakrishnan, N., Kahle, W. and Huber-Carol, C. (eds), *Advances in Degradation Modeling: Applications to reliability, survival analysis and finance*, Boston: Birkhauser, pp. 127–145.

Kahle, W. and Wendt, H. (2004), On a cumulative damage processe and resulting first passage times, *Applied Stochastic Models in Business and Industry*, vol. 20, pp. 17–26.

Kaishev, V.K., Dimitrova, D.S. and Haberman, S. (2007), Modelling the joint distribution of competing risks survival times using copula functions, *Insurance: Mathematics and Economics*, vol. 41, pp. 339–361.

Kent, J.E.A. (1979), An efficient proceedure for selecting among five reliability models, *PhD Thesis*, North Carolina State University, Releigh, USA.

Kurowicka, D. And Cooke, R. (2006), *UNCERTAINTY ANALYSIS with High Dimensional Dependence Modeling*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd.

Lawless, J. (2003), *Statistical Models and Methods for Lifetime Data*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc.

Lee, J. (1993), Generating random binary deviates having fixed marginary distributions and specified degrees of association, *The American Statistician*, vol. 47, pp. 209–215.

Lemeshko, B.Y., Lemeshko, S.B., Akushkina, K.A., Nikulin, M.S. and Saaidia, N. (2010), Inverse Gaussian model and its applications in reliability and survival analysis, In: Rykov, V.V., Balakrishnan, N., and Nikulin, M.S. (eds), *Mathematical and Statistical Models and Methods in Reliability*. Birkhauser, pp. 433–453.

Lindqvist, B. H. and Skogsrud, G. (2009), Modeling of dependent competing risks by first passage times of Wiener processes, *IIE Transactions*, vol. 41, pp. 72–80.

Lindqvist, B.H., Stove, B. and Langseth, H. (2006), Modelling of dependence between critical failure and preventive maintenance: the repair alert model, *Journal of Statistical Planning Inference*, vol. 136, no.5, 1701–1717.

Ling, C.-H. (1965), Representation of associative functions, *Publ. Math. Debrecen*, vol. 12, pp. 189–212.

Liu, C.-C. (1997), A comparison between the Weibull and lognormal distribution models used to analyse reliability data, *PhD. Thesis*, University of Nottingham, Nottingham, UK.

Lo, S.M.S. and Wilke, R.A. (2010), A copula model for dependent competing risks, *Journal of the Royal Statistical Society*, vol. 59, no.2, pp. 359–376.

Lu, C.J. and Meeker, W.Q. (1993), Using degradataion measures to estimate a time-to-failure distribution, *Technometrics*, vol. 35, pp. 161–174.

Lu, C.J., Meeker, W.Q. and Escober, L.A. (1996), A comparison of degradation and failure-time analysis methods for estimation a time-to-failure distribution, *Statistica Sinica*, vol. 6, pp. 531–546.

Mann, N.R., Schafer, R.E. and Singpurwalla N.D. (1974), *Methods for Statistical Analysis of Reliability and Life data*, J. Willey, New York.

Marshal, A.W. and Olkin, I. (1967), A multivariate exponential distribution, *Journal of the American Statistical Association*, vol. 62, pp. 30–44.

Marshal, A.W. and Olkin, I. (1988), Families of multivariate distributions, *Journal of the American Statistical Association*, vol. 83, pp. 834–841.

McCullagh, P. And Nelda, J.A. (1983), *Generalized Linear Models*, Chapman & Hall, London.

McNeil, A.J., Frey, R. and Embrechts, P. (2005), *QUANTITATIVE RISK MANAGEMENT: Concepts, Techniques and Tools*, Princetown University Press, Princetown and Oxford.

McNeil, A.J., Neslehova, J. (2009), Multivariate Archimedean copulas, d-monotone functions, ℓ_1- norm symmetric distributions, *Annals of Statistics*, vol. 37, no.5B, pp. 3059–3097.

Meeker, W. Q., Escober, L. A. and Hong, Y. (2009), Using accelerated test results to predict product field reliability, *Technometrics*, vol. 51, no.2, pp. 146–161.

Nelsen, R.B. (1999), *An Introduction to Copulas*, Springer: New York.

Nelsen, R.B. (2005), Dependence modeling with Archimedean copulas, In: Proceedings of the Second Brazilian Conference on Statistical Modeling in Insurance and Finance, pp. 45–54.

Nelson, W. (1982), Applied life data Analysis, *Wiley Series in Probability and Mathematical Statistics*, Wiley, New York.

Nelson, W. (2004), Accelerated Testing: Statistical Models, Test Plans, and Data Analysis, *Wiley Series in Probability and Statistics*, Wiley, New Jersey.

Oakes, D. (1989), Bivariate survival models induced by frailties, *Journal of the Amer-ican Statistical Association*, vol. 84, pp. 487–493.

Olteanu, D. and Freeman, L. (2010), The evaluation of median-rank regression and maximum likelihood estimation techniques for a two-parameter Weibull distribution, *Quality Engineering*, vol. 22, pp. 256–272.

Padgett, W.J. and Tomlinson, M.A. (2004), Inference from accelerated degradation and failure data based on Gaussian process model, *Life Data Analysis*, vol. 10, pp. 191 - 206.

Park, C. and Padgett, W. J. (2005), Accelerated degradation models for failure based on geometric Brownian motion and Gamma processes, *Lifetime Data Analysis*, vol. 11, pp. 511–527.

Pan, Z., Zhou, J. and Zhao, P. (2010), Joint accelerated failure mode modeling of degradation and traumatic failure times, *Proceedings of the World Congress on Engineering.*, vol. 10, no.3, June 30 - July 2, London, U.K.

Prentice, R., Kalbfleisch, J., Peterson, A., Flournoy, N., Farewell, V. and Breslow, N. (1978), The analysis of failure times in the presence of competing risks, *Biometrics*, vol. 34, pp. 541–554.

Qu, X., Zhou, J. and Shen, X. (2010), Archimedean copula estimation and model selection via l_1- norm symmetric distribution, *Insurance: Mathematics and Economics*, vol. 46, pp. 406–414.

Renyi, A. (1959), On measures of dependence, *Acta. Math. Acad. Sci. Hungar.*, vol. 10, pp. 441–451.

Sato, K. (1999), Levy Processes and Infinitely Divisible Distributions, *Cambridge University Press*, United Kingdom.

Schrodinger, E. (1915), Zur theorie de fall- und Steigversuche an Teilchen mit Brownscher Bewegung, *Physikalische Zeitschrift*, vol. 16, pp. 289–295.

Schweizer, B and Sklar, A, (1983), *Probabilistic Metric Spaces*, Ney York: North-holland.

Si, X.-S., Wang, W., Hu, C.-H. and Zhou, D.-H. (2011), Remaining useful life estimation - A review on the statistical data driven approaches, *European Journal of Operations Research*, vol. 213, pp. 1–14.

Singpurwalla, N.D. (2006), *Reliability and Risk: A Bayesian perspective*, John Wiley & Sons, New Jersey.

Singpurwalla, N.D. (2006b), On competing risks and degradation processes, The second Erich Lehmann Symposium - Optimality, Institute of Mathematical Statistics - Monograph Series, vol. 49, (ed. J. Rojo), pp. 362–377.

Sklar, A. (1973), Random variables, joint distribution functions, and copulas, *Kybernetika*, vol. 9, no.6, pp. 449–460.

Skinner, K.R., Keats, J.B. and Zimmer, W.J. (2001), A comparison of three estimators of the Weibull parameters, *Quality Reliability Engineering International*, vol. 35, pp. 219–237.

Tsiatis, A. (1975), A nonidentifiability aspect of the problem of competing risks, *Proceedings of National Academy of Sciences USA*, vol. 72, no.1, pp. 20–22.

Tweedie, M. C. K. (1957a), Statistical properties of inverse Gaussian distributions, I, *The Annals of Mathematical Statistics*, vol. 28, no.2, pp. 289–304.

van Noortwijk, J.M. (2009), A survey of the application of gamma processes in maintenance, *Reliability Engineering and Systems Safety*, vol. 94, pp. 2–21.

van Noortwijk, J.M., Dekker, R., Cooke, R.M. and Mazzuchi, T.A. (1992), Expert Judgment in Maintenance Optimization, *IEEE TRANSACTIONS ON RELIABILITY*, vol. 43, no.3, pp. 427–432.

Wang, W. (2004), Refined rank regression method with censors, *Quality Reliability Engineering International*, vol. 20, pp. 667–678.

Whitmore, G.A., Crowder, M.J. and Lawless, J.F. (1998), Failure inference from a marker process based on a bivariate Wiener model, *Lifetime Data Analysis*, vol. 3, pp. 27–45.

Whitmore, G.A. and Schenkelberg, F. (1997), Modelling accelerated degradation data using Wiener diffusion with a time scale transformation, *Lifetime Data Analysis*, vol. 4, pp. 229–251.

Wilks, S.S. (1938), The large-sample distribution of the likelihood ratio test for testing composite hypothesis, *The Annals of Mathematical Statistics*, vol. 9, no.1, pp. 60–62.

Zhao, W. and Elsayed, E.A. (2004), An accelerated life testing model involving performance degradation, *RAMS.*, pp. 324 - 329.

Zheng, M. and Klein, J.P. (1995), Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika*, vol. 82, no.1, pp. 127–138.

Zografos, K. (2000), Measures of multivariate dependence based on a distance between Fisher information matrices, *Journal of Statistical Planning and Inference*, vol. 89, pp. 91–107.