DEGRADATION PROCESSES AND RELATED RELIABILITY MODELS

JIN LU

Faculty of Management
McGill University
Montreal

July 1995

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Ph.D..

ISBN 0-612-12422-3

ACKNOWLEDGEMENTS

I wish to thank Professor Alex Whitmore for many enlightening discussions and many valuable comments on earlier drafts of the thesis, thesis proposal, and earlier theory papers that are now a part of this thesis; for financial support during my Ph.D. studies; and for his effort in obtaining the transistor data set. I also wish to thank Professors Richard Loulou, Brenda MacGibbon, and V. Seshadri for their valuable comments on my thesis proposal and earlier theory papers. I would like to express my gratitude to Dr. T.C. Denton and Dr. M. Chown of BNR Europe for providing the transistor data set.

I am grateful to fellow Ph.D. students: Claude Mathieu for his help in translating the thesis abstract to a French version; and Paul Berry and Li Jiang for their help in obtaining the stock price data set.

TABLE OF CONTENT

ABSTRACT

Reliability characteristics of new devices are usually demonstrated by life testing. When lifetime data are sparse, as is often the case with highly reliable devices, expensive devices, and devices for which accelerated life testing is not feasible, reliability models that are based on a combination of degradation and lifetime data represent an important practical approach. This thesis presents reliability models based on the combination of degradation and lifetime data or degradation data alone, with and without the presence of covariates. Statistical inference methods associated with the models are also developed.

The degradation process is assumed to follow a Wiener process. Failure is defined as the first passage of this process to a fixed barrier. The degradation data of a surviving item are described by a truncated Wiener process and lifetimes follow an inverse Gaussian distribution. Models are developed for three types of data structures that are often encountered in reliability studies, terminal point data (a combination of degradation and lifetime data) and mixed data (an extended case of terminal point data); conditional degradation data; and covariate data.

Maximum likelihood estimators (MLEs) are derived for the parameters of each model. Inferences about the parameters are based on asymptotic properties of the MLEs and on the likelihood ratio method. An analysis of deviance is presented and approximate pivotal quantities are derived for the drift and variance parameters. Predictive density functions for the lifetime and the future degradation level of either a surviving item or a new item are obtained using empirical Bayes methods. Case examples are given to illustrate the applications of the models.

RÉSUMÉ

Les caractéristiques de fiability des nouveaux équipements sont souvent estimées par des tests sur la durée de vie. Quand les données sur la durée de vie sont rares, comme c'est souvent le cas pour les équipements de haute fiability, les équipements dispendieux et ceux dont les tests accélérés sont impossibles, les modèles de fiability basés sur une combinaison de données sur la dégradation et de données sur la durée de vie représentent une application pratique importante. Cette thèse présente ces modèles de fiability où il y a une combinaison de données sur la dégradation et de données sur la durée de vie, et cela avec ou sans la présence des variables inter-reliées. Des techniques d'inférence statistique sont aussi développées.

Il est supposé que la dégradation suit un processus de Wiener. Le processus prend fin dès qu'une barrière fixe est franchie. Les données sur la dégradation de l'item survivant sont caractérisées par un processus tronqué de Wiener, et les durées de vie suivent une distribution inverse de Gauss. Des modèles sont développés pour trois types de structure de données, c'est-à-dire (i) les données sur les points terminaux (une combinaison entre les données sur la dégradation et ceux sur la durée de vie) et les données mixtes (une extension des données sur les points terminaux); (ii) les données conditionnelles de dégradation; et (iii) les données inter-reliées.

Les estimateurs du maximum de vraisemblance sont dérivés pour les paramètres de chacun des modèles. Les inférences statistiques sur les paramètres sont basées sur les propriétés asymptotiques des estimateurs du maximum de vraisemblance et sur la méthode des rapports de vraisemblance (likelihood ratio method). Une analyse de déviance est présentée et les valeurs pivotantes approximées sont dérivées pour les paramètres de tendance et de la variance. Les fonctions de densité utilisées pour prévoir la durée de vie et le niveau futur de dégradation de l'item survivant ou du nouvel item sont obtenues par les méthodes empiriques de Bayes. Des exemples sont présentés afin d'illustrer les applications du modèle.

CHAPTER 1 INTRODUCTION

Work on reliability models based on lifetime data extends back several decades. There is a large literature on this subject (reference books include Lawless, 1982; Nelson, 1982; Crowder, Kimber, Smith, and Sweeting, 1991; to name just a few). Several statistical distributions for lifetime data, such as the exponential, Weibull, extreme value, gamma, log-normal, and inverse Gaussian, and the statistical inference methods associated with these distributions, have been extensively studied. The application of the exponential distribution to lifetime analysis dates from the 1940's. An early influential paper on this topic is Davis (1952). The Weibull distribution is perhaps the most widely used. Many statistical methods developed for the Weibull distribution are now routinely used in life test and reliability work (Lawless, 1983). The inverse Gaussian distribution, as compared with the others, is a relative newcomer. It is particularly interesting since it arises as the first passage time distribution of a Wiener process which is a natural description of many physical processes. Since the early landmark papers on the statistical properties of the inverse Gaussian distribution by Tweedie (1957a,b), a substantial amount of work has been devoted to statistical methods and applications of the distribution. Details about the distribution can be found in Chhikara and Folks (1989) and Seshadri (1993).

The approach of reliability analysis based on lifetime data has its limitations in practical applications. Collecting a sample of lifetime data can be time consuming and costly. In engineering applications, observed lifetimes are often very long because of the high reliability of modern devices. Failure acceleration methods, including variable stress methods, have been developed to obtain a sample of lifetime data for devices in a relatively short time period. However, the acceleration mechanisms applied in such tests may not faithfully imitate the actual failure process that will prevail or no feasible acceleration mechanism may be available. Even with acceleration, few failures may be observed. Moreover, the devices under study may be very expensive and thus conducting a destructive life test of such devices becomes uneconomical or the consequences of destruction are disastrous and thus a life test becomes infeasible. For devices that are already in operation, such as the components of a telecommunication system or a nuclear power generator, the degradation levels of these components may be readily available from maintenance procedures while a sample of lifetime data under the same operating conditions is virtually nonexistent. The only practical approach in this type of situation would be to derive predictive inference and optimal maintenance policy based on degradation data.

The difficulty of collecting a sample of lifetime data becomes more evident in the health care area. Suppose one wishes to study the survival time of a patient who has received a certain medical treatment. Obtaining a sample of lifetime data can be extremely costly in human terms and time consuming also. On the other hand, a patient's vital signs, such as $CD4^{+}$ cell counts of AIDS patients (Lange, Carlin, and Gelfand, 1992), are well defined and readily measured. It would be ideal if one could draw inferences about survival time from measures of vital signs taken over time, instead of waiting years for a sample of lifetime data.

As been noted by Chown, Pullum, and Whitmore (1993), reliance on lifetime data alone is becoming less and less practical in engineering, and there exists a pressing need for reliability models that capture the degradation response of an item over time and statistical methods associated with those models. Nair (1988) states: "... Degradation data are a much richer source of information than time-to-failure data. The lack of statistical methods for analyzing them prevents users from exploiting this valuable source of information. ...." This thesis aims, in part, to provide a statistical tool for reliability analysis based on both degradation data and lifetime data.

Lifetime, degradation, and covariate data are the three building blocks of reliability data. Degradation data can be the measurements of a key indicator of a component's operational characteristic or the performance measurements of a system consisting of a group of components. The latter are sometimes called system data. Both degradation data and lifetime data can be obtained from the same reliability test. Consider, for example, observing a set of items on test and stopping observation at time t. One would often find that some items have already failed by time t while others remain operating. One could record both degradation measures over time for all items and lifetimes of the failed items. Also, the degradation process and the lifetime of an item are often affected by covariates, such as environmental or design factors. For example, as illustrated later, the degradation process of an electronic transistor accelerates when ambient temperature increases, which is a covariate affecting the degradation process of a transistor.

Reliability models based on lifetime data have been extensively studied. However, little work has been done to model the combination of degradation and lifetime data. Models based on degradation data alone have been studied mainly in the area of material fatigue analysis, although broader applications are beginning to appear (for example, Carey and Koenig, 1991; Lu and Meeker, 1993; Boulanger and Escobar, 1994). Modeling a combination of degradation and lifetime data is more effective than using lifetime data or degradation data alone when both types of data are available and covariates should be incorporated into the model whenever they exist. Neglecting one of these may result in incomplete information about the behavior of the items under study. Many items experience degradation before they fail. Models based on degradation data alone are important alternatives to lifetime models in many areas of application where failure data are not easily or economically available. This research develops reliability models based on either degradation data alone or a combination of degradation and lifetime data with and without covariates and explores statistical inference methods for each model.

The remainder of the thesis is organized as follows. Chapter 2 starts with a brief review of various approaches to modeling a degradation process with special attention being given to material fatigue and performance parameter drift, which are the main areas of degradation process modeling. Next, the main properties of a Wiener process are presented as it will be used in the thesis as the basic model for a degradation process. Finally, the basic types of data structures are presented and discussed. The most general data structure is defined as a mixture of degradation data and lifetime data, sometimes with covariate information. Several variations are given. The next three chapters are devoted to the development of reliability models for the data structures given in Chapter 2. Chapter 3 presents the density function of the level of a Wiener process conditioned on no first passage and then derives a reliability model for a terminal point data structure. The model is extended later to a mixed data structure. Next, pivotal quantities are found from which the approximate distributions of the maximum likelihood estimators (MLEs) are derived and a likelihood ratio test for sample path homogeneity is developed. Finally, the results of empirical Bayes analysis are presented. Predictive density functions for residual lifetime and future degradation level are developed. Chapter 4 presents a model for a conditional data structure. Chapter 5 extends the model for terminal point data to a covariate data structure that comprises lifetime, degradation, and covariate data. Inferences based on asymptotic normality and the likelihood ratio method are studied for each model. Applications of the models are illustrated in Chapter 6 by case examples that use real and simulated data. Finally, concluding remarks and some comments on further research are presented in Chapter 7.

CHAPTER 2 THE BASIC DEGRADATION MODEL AND ASSOCIATED DATA STRUCTURES

2.1 A Review of Degradation Models

The literature on statistical models for degradation is quite diverse. Representative work is found in two areas – models of material fatigue and models of performance parameter drift. Each of these areas is briefly reviewed in this section.

2.1.1 Material Fatigue

A considerable amount of work on degradation processes has concentrated on material fatigue analysis. There is a large body of literature on this subject (for example, Desmond, 1985 and 1987; Sobczyk,1987; Ditlevsen,1986). Sobczyk (1987) gives an overview of stochastic models for material fatigue damage. The following is a brief summary of some approaches to modeling the degradation of material strength based on Sobczyk (1987).

Degradation data can be modeled by a stochastic process in one or more dimensions, depending on the nature of the degradation mechanism. Material fatigue accumulation has been described by various kinds of Markov processes. To model fatigue accumulation by a continuous-time discrete-state Markov process, for instance, the random evolution of fatigue, $X(t)$, is assumed to be a one-dimensional process with $n+1$ states $E_0, E_1, ..., E_n = E^*$, where E_0 denotes an ideal state and E^* denotes a failure state. The degree of fatigue damage is usually characterized by the length of a "dominant crack" which, as it grows, eventually leads to failure. That is, the degradation process is the growth process of a dominant crack. Hence, states E_i, $i = 1, 2, ..., n-1$, represent the length of a dominant crack. This is a pure birth process with birth rates λ_i, $i = 0, 1, 2, ..., n$, and $\lambda_n = 0$. Let $P_i(t) = Pr\{X(t) = E_i\}$, then the transition equation of the pure birth process is the following.

$$P_i(t+\Delta t) = (1-\lambda_i \Delta t)P_i(t) + \lambda_{i-1}P_{i-1}(t)\Delta t + o(\Delta t) \quad \text{for} \quad i = 1, 2, ..., n$$
$$P_0(t+\Delta t) = (1-\lambda_0 \Delta t)P_0(t) + o(\Delta t)$$

The probability of failure at time t, $P_n(t)$, can be obtained by a recursion relation, which is the solution of the transition equations.

$$P_i(t) = \lambda_{i-1}\exp(-\lambda_i t)\int_0^t \exp(\lambda_i \tau)P_{i-1}(\tau)d\tau \quad \text{for} \quad i = 1, 2, ..., n$$
$$P_0(t) = \exp(-\lambda_0 t)$$

Fatigue damage can also be described by a Markov chain. Compared with the approach that has just been described, the new concept here is *duty cycle*, defined as a repetitive period of operation in the life of an item during which damage can accurs. It is assumed that fatigue damage occurs only at the end of each duty cycle. This process is then a Markov chain where time is measured by the number of elapsed duty cycles. As before, assume that there are $n+1$ states, $0, 1, 2, ..., n$, denoting the degree of material damage, and that state n represents failure, then the probability of failure can be computed from the transition matrix of this Markov chain in the standard manner.

It is often assumed that the growth of a crack is caused by the sequence of peaks of a random stress process. This leads to a shock model for fatigue damage in which the occurrence of peaks follows a Poisson process. There are many variations of this basic model, such as incorporating the phenomenon of crack growth retardation or acceleration (a decreasing or increasing rate of crack growth).

Crack growth can also be modeled by a continuous state stochastic process. In this type of model, the length of a dominant crack $L(t)$ at time t is governed by the following differential equation, often referred to as the Fokker-Planck- Kolmogorov equation.

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial l}[a(l,t)p(l,t;l_0,t_0)] + \frac{1}{2}\frac{\partial^2}{\partial l^2}[b(l,t)p(l,t;l_0,t_0)] \tag{2.1}$$

Here $a(l,t)$ and $b(l,t) > 0$ are drift and diffusion coefficients, respectively. Function $p(l,t;l_0,t_0)$ is the probability density function (*p.d.f.*) of a dominant crack length l at time t and satisfies the following conditions,

$$p(l,t;l_0,t_0) = 0 \quad \text{for} \quad l < l_0,$$
$$p(l,t;l_0,t_0) \geq 0 \quad \text{for} \quad l \geq l_0,$$

where $l_0 = L(t_0)$. Equation (2.1) is the diffusion equation of a Wiener process starting at l_0 at time t_0 if we set the coefficients $a(l,t)$ and $b(l,t)$ to be constants.

To relate the rate of crack growth to a covariate, such as the stress level applied to an item under study, the Paris law is often applied in fatigue damage analysis. The Paris law takes the following form.

$$\frac{dL(t)}{dt} = CK_{max},$$

where C is a constant and K_{max} is the maximum stress level an item has endured. A modified form of the Paris law is also often applied. The modifications are formulated based on the argument that the growth rate of a crack is related to the length of the crack, therefore, a varied rate of crack growth is assumed. Specifically,

$$\frac{\partial L(t)}{\partial t} = F(\Delta K, R)\xi(t),$$

where $\Delta K = (K_{max} - K_{min})\sqrt{\pi L(t)}$, $R = K_{min}/K_{max}$, K_{min} is the minimum stress level, and F is a given function. In addition, a multiplicative term $\xi(t)$ is introduced to represent the random fluctuation of external factors. For more details about the approaches summarized above, one may refer to Sobczyk(1987) and the cited references there.

2.1.2 Performance Parameter Drift

Degradation studies are also frequently carried out on the deterioration of one or more key performance parameters of an item. It is usually called the *parameter* *drift* problem in engineering. When the performance level drifts to a critical level or barrier that may be random or fixed, the item is no longer in good operational order, and therefore a "failure" is said to have occurred. For example, consider a telecommunication system in which a logical circuit is a key component. An important performance measure of a logical circuit is its *propagation delay*, the time required to respond to an input signal. Propagation delay increases over the life of a logical circuit as it degrades. When the delay reaches a certain level the logical circuit fails to function as required. The delay is also correlated with environmental conditions, such as temperature. Carey and Koenig (1991) present a case study about the reliability assessment of a logical circuit based on accelerated degradation data of propagation delay. They propose a nonlinear model with a regression structure to relate propagation delay to time. A parameter of this nonlinear model is assumed to be correlated with temperature (a covariate). The proposed model provides a good fit to their accelerated propagation delay data.

$$y_n - y_o = \theta(1 - \exp(-\sqrt{\lambda t_n})) + \varepsilon_n,$$

where $y_n - y_o$ is the increment of propagation delay over time interval (t_o, t_n), θ and λ are model parameters, and ε_n is a residual term with a normal distribution. The parameter θ is assumed to be related to temperature T according to the Arrhenius relationship

$$\log(\theta) = A - \frac{B}{kT} + \eta,$$

where A and B are parameters to be estimated, k is a known constant, and again η is a residual term having a normal distribution.

Engineering research on degradation phenomena is very extensive (see, for example, Leblebici and Kang 1993; Stucki 1994; Xiao and Bathias 1994, to mention a few). Most studies of degradation, however, concentrate on its physical characteristics without any reference to its statistical nature. Few have modeled degradation as a stochastic process. Carey and Koenig (1991) is one of the exceptions.

As may be seen from the preceding discussion and the examples given in Chapter 1, a variety of degradation processes are encountered in engineering applications. A degradation process can be the growth of a crack, the wear-out of material, the deterioration of a performance parameter such as propagation delay, or the electron-migration of metal in a transistor. Although degradation processes in different areas have different data structures and properties, they share the common feature that a random mechanism governs the degradation variables. This random mechanism is best represented by a stochastic process. Choosing a basic model for this fundamental stochastic process, we will be able to derive a series of models to accommodate a variety of data structures. The next two sections present the main properties of a Wiener process (or Brownian motion with drift), which is assumed to be the basic model for a degradation process in the thesis, and examine data structures that are often encountered in reliability studies.

2.2 The Wiener Process as a Degradation Model

A Wiener process $\{X(t), t \geq 0\}$ has the following three defining properties (Karlin and Taylor, 1975).

1. Every increment $X(t_i) - X(t_{i-1})$ for a time interval $(t_{i-1},\ t_i)$ is normally distributed with mean $\delta(t_i - t_{i-1})$ and variance $\nu(t_i - t_{i-1})$ where $\nu > 0$ is a fixed variance parameter and δ is a fixed drift parameter.

2. The increments for any set of disjoint time intervals are independent random variables having the distributions described in property 1.

3. $X(0) = 0$.

A Wiener process is a homogeneous process and has a continuous sample path with probability one. The increments $X(t_i) - X(t_{i-1})$ are independent of the past evolution of the process, that is,

$$P[X(t_i) \leq x_i | X(t_0) = x_0, X(t_1) = x_1, X(t_2) = x_2, ..., X(t_{i-1}) = x_{i-1}]$$
$$= P[X(t_i) \leq x_i | X(t_{i-1}) = x_{i-1}], \qquad i = 1, 2, ..., n, \tag{2.2}$$

for any $0 = t_0 < t_1 < t_2 < ... < t_n$.

Given condition $X(0) = 0$, the probability density function of $X(t)$ at $t > 0$ is the normal density function

$$\phi(x;t) = \frac{1}{\sqrt{2\pi\nu t}} \exp\left(-\frac{(x-\delta t)^2}{2\nu t}\right). \tag{2.3}$$

Under the same condition, the joint *p.d.f.* of $X(t_1)$, $X(t_2)$, ..., $X(t_n)$, $0 < t_1 < t_2 < ... < t_n$, is

$$f(x_1, x_2, ..., x_n) = \phi(x_1; t_1)\phi(x_2 - x_1; t_2 - t_1).....\phi(x_n - x_{n-1}; t_n - t_{n-1}). \tag{2.4}$$

Mathematical properties and inference methods for a Wiener process have been extensively studied. Selected references will be given throughout the thesis as appropriate.

A Wiener process, $\{X(t)\}$, is taken as the basic model of a degradation process. It is assumed that each item has its own degradation process which is independent of the others (for a given set of covariates). Items having the same design are assumed to have the same drift and variance parameters unless indicated otherwise. An item fails when its degradation process reaches a specified critical level for the first time. This critical level is referred to as a *barrier* and is denoted by a. The lifetime of an item corresponds to the first passage time, S, of the Wiener process to the barrier.

Wiener processes have been widely applied in both engineering and business situations. Many physical phenomena are described by Wiener processes. The matter of whether a Wiener process is a suitable model for degradation processes, however, deserves a few comments.

Often, degradation processes are monotonic, that is, degradation proceeds in only one direction, as in a wear-out process for example. The application of a Wiener process to this kind of degradation process is only approximate. However, when observed closely, the levels of many degradation processes vary bidirectionally over time as, for example, with the gain of a transistor or the extent of propagation delay. Other stochastic processes, such as gamma processes, may be considered if it is essential to represent degradation by a strictly monotonic process.

In this thesis, it is assumed that the degradation process of interest is a continuous process and, in many applications, this is a valid assumption. Where a degradation process is discrete and an approximation is not permitted, another type of stochastic process, such as a discrete-state Markov process, may be considered.

A Wiener process is a time homogeneous process but not all degradation processes have this property. For example, in reliability engineering, acceleration tests are often used to obtain lifetime and degradation data in a relatively short period of time. The stresses applied in these tests may be increased during the course of testing in order to bring about rapid failure. Because degradation parameters change as the stress level increases, the degradation process becomes time heterogeneous. As a second example, the physical mechanism that governs deterioration may tend to accelerate or decelerate degradation, as in crack propagation for instance, producing time heterogeneity. As will be discussed later, a transformation of the time scale in these situations can often convert a time heterogeneous degradation process to a time homogeneous process. In general, a time heterogeneous stochastic process, like the one described by the Fokker-Planck-Kolmogorov equation in (2.1), can be applied.

Modeling a degradation process by a Wiener process implies that the degradation process, given its current state, evolves to a future state independently of its past behavior. This is referred to as its *Markov property.* While the Markov property is a valid assumption in many applications, it does not always hold. For the latter case, several analytical methods, such as, the inclusion of supplementary variables and imbedding (Cox and Miller, 1965) may be considered.

If any practical degradation mechanism requires a stochastic model other than a Wiener process, one may find that the theoretical development in this thesis still provides useful guidelines although the technical details will differ in different applications.

2.3 Data Structures for Degradation and Lifetime Data

Reliability data are collected in various forms under many different conditions. In general, the data structure for a non-repairable item usually consists of lifetime, degradation, and covariate data. To illustrate a data structure, consider a random sample of m identical items on test. The data collection starts at time 0 and stops at time $t_n > 0$. Suppose that p of the m items fail before time t_n and $q = m - p$ items survive. The degradation levels (sample path levels) of the items are recorded at a set of fixed time points $0 < t_1 < t_2 < ... < t_n$. The values of the covariates may vary across groups of items and/or change over time. The resulting data consist of (1) sample path levels at times $t_1, t_2, ..., t_n$, (2) the first passage times of the p items that failed before the stopping time t_n, and (3) the values of the covariates.

2.3.1 Data Structure without Covariates

In the absence of covariates, the basic data structure, which will be referred to as the **mixed data** structure, has the following form:

for a failed sample path i, $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{in_i}, s_i)$ $\quad i = 1, 2, ..., p$

for a surviving sample path j, $\mathbf{x}_j = (x_{j1}, x_{j2}, ..., x_{jn})$ $\quad j = p+1, p+2.., p+q$

where x_{ik} is the level of sample path i at time t_k, x_{in_i} is the last observation of the sample path before failure for a failed item, and s_i is its failure time. Thus, the mixed data structure consists of two types of data: x_{ik}, the degradation data; and

s_i, the lifetime data. When all items survive to time t_n, the mixed data structure simplifies to

$$\mathbf{x_i} = (x_{i1}, x_{i2}, ..., x_{in}), \qquad i = 1, 2, ..., m.$$

These are **degradation data alone**. This case corresponds to a situation where lifetime data are not available. When all items have failed by time t_n and degradation data are not recorded, the mixed data structure simplifies to

$$\mathbf{s} = (s_1, s_2, ..., s_m).$$

These are **lifetime data alone** and represent a common form of data set considered in reliability literature.

Two variations of the mixed data structure are often of interest. In the first situation, the data structure consists of (1) the measurement x_{in} at the stopping time t_n only for the survivors and (2) the lifetime s_i of the failed items. In some applications, for example, the measurement of a degradation variable can be taken only when a test stops, thus, one would not be able to get degradation data before the test stops. Sometimes, even when the intermediate degradation levels can be measured, it may not be economical to do so because the cost of collecting the measurements may be excessive. In these cases, the mixed data structure reduces to the following.

$$\mathbf{x} = (x_{1n}, x_{2n}, ..., x_{qn}), \quad \mathbf{s} = (s_1, s_2, ..., s_p)$$

This type of data structure will be referred as **terminal point data** and is treated in Section 4.2.

The second variation arises when the structure consists of **censored lifetime data** only. In this case, one only knows that an item has either failed at time s or survived until the stopping time t_n. Thus, the data structure is

$$\mathbf{s} = (s_1, s_2, ..., s_p, t_n, t_n, ..., t_n),$$

where t_n is the censor time for items $p+1, p+2, ..., p+q = m$. Of course, if all items happen to have failed by time t_n, then, none of the lifetimes is censored and the data structure becomes lifetime data alone. As mentioned earlier, censored lifetime data, with and without covariates, have been studied extensively in the reliability literature. Figures 2.1a to 2.1d illustrate these various types of data structures.

In some applications, the data arise under restrictions or conditions. For example, one may have only the lifetime data for failed items $(s_1, ..., s_p)$ and have no knowledge of the number of survivors q. As another example, one may have degradation data only for the q survivors (x_i; $i = p+1, ..., p+q$), with no knowledge of the number or lifetimes of the p failed items. Both examples are important special cases representing situations of incomplete information, which often occurs with reliability data from field studies. The latter situation, which will be called **conditional degradation data**, is to be described by a truncated Wiener process and will be investigated in Chapter 4.

2.3.2 Data Structure with Covariates

In the presence of covariates, the measurements on these covariates are added to the data structure. A covariate can be one of the following three types. First, a covariate can vary across different items but remain fixed over time for each item. The variation may be controlled or may be random. For example, temperature can be controlled experimentally at one level for one group of items and at another level for another group of items; a design characteristic may differ among different groups of items; quality of workmanship and defects of material may vary randomly from one production batch to another. This type of covariate will be referred to as a **time fixed covariate**.

Second, a covariate can vary over time but the same variation is experienced by every item. This situation arises when all items are simultaneously subject to a stress level that changes only over time either randomly or by experimental control. For example, all items may be exposed to the same randomly varying level of temperature over time; a group of mechanical components may be subjected to the same experimentally controlled stress levels in an accelerated step-stress life test. This type of covariate will be referred to as a **time varying covariate**.

Third, a covariate can vary among groups of items and over time. Perhaps this is the type of covariate that is most often encountered in field studies. For example, automobiles are subjected to climate conditions that differ geographically and change over time. This type of covariate will be called a **mixed covariate**.

Time Fixed Covariates Assume now that we have k controlled covariates and/or random covariates that vary only across items. The measures on these variables are represented by a $(p+q) \times k$ matrix, denoted by Z.

$$Z = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1k} \\ z_{21} & z_{22} & \cdots & z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ z_{p1} & z_{p2} & \cdots & z_{pk} \\ \vdots & \vdots & \ddots & \vdots \\ z_{p+q,1} & z_{p+q,2}, & \cdots & z_{p+q,k} \end{pmatrix} \tag{2.5}$$

Usually, in engineering testing, items are grouped to form experimental blocks and each block is tested at a particular combination of covariate levels, which is experimentally controlled. In this case, each block corresponds to a subset of identical row vectors in the matrix Z. The transistor example discussed in Chapter 6 is one application of this test setting.

Time Varying Covariates Covariates can vary over time either randomly or by experimental control, but the same variation is experienced by every item. In the case of random variation over time, the degradation process $\{X(t)\}$ is accompanied by a group of covariate processes $\{Z_1(t), Z_2(t), ..., Z_k(t)\}$. If one takes measurements on the k covariates at n fixed time points, the result is an $n \times k$ covariate matrix Z,

$$Z = \begin{pmatrix} z_{1t_1} & z_{2t_1} & \cdots & z_{kt_1} \\ z_{1t_2} & z_{2t_2} & \cdots & z_{kt_2} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ z_{1t_n} & z_{2t_n}, & \cdots & z_{kt_n} \end{pmatrix} \tag{2.6}$$

where $z_{it_1}, z_{it_2}, ..., z_{it_n}$ correspond to the measurements on the ith covariate taken at time points $t_1, t_2, ..., t_n$, that is, $z_{it_j} = Z_i(t_j), i = 1, 2, ..., k, j = 1, 2, ..., n$.

For the case of controlled variation over time, each covariate is preset at specified levels at a sequence of time points. The measurements of the covariates take the same form as matrix (2.6), except that each element is controlled rather than a random outcome.

Mixed Covariates When a covariate changes across items and over time, the covariate data structure is a combination of the previous two types. Suppose measurements are taken on k covariates at n fixed time points for each of m items, one will have m covariate matrices, $Z_1, Z_2, ..., Z_m$, corresponding to each of the m items, where each Z_i has the form of (2.6) with dimension $n \times k$.

2.4 Summary

This chapter first presents a brief review of degradation models, with special attention being given to material fatigue and performance parameter drift problems, which are the main concerns in the degradation modeling literature. Then, a Wiener process is presented as the basic model for a degradation process. Finally, several data structures of degradation and lifetime measures are discussed. The basic model needs to be expanded and modified in order to accommodate the conditions and restrictions implied in each type of data structure. Chapters 3 through 5 serve this purpose.

CHAPTER 3 A MODEL BASED ON DEGRADATION DATA AND LIFETIME DATA

Both lifetime and degradation data contain important information about the reliability properties of an item. In fact, a model based on both types of data may represent the only practical approach when failure experience is sparse as it will be with highly reliable items, new items, items in which failure is not easily accelerated, expensive items, etc. . In this chapter, a reliability model based on both degradation and lifetime data will be presented. The terminal point data structure will be studied first and the mixed data structure will then be studied as an extension of the terminal point case.

Recall that the degradation process of an item is assumed to follow a Wiener process and its lifetime, therefore, has an inverse Gaussian distribution. Notice, however, that the degradation data are restricted by the barrier at level a since an item fails when its sample path reaches the barrier, as shown in Figure 3.1. In other words, the sample paths are defined on $(-\infty, a)$ rather than on the whole real line as for a Wiener process. The distribution properties of a Wiener process presented earlier take no account of a restriction or barrier. It will be found later that for a sample path with a tight barrier, the effects of the barrier on statistical inference are significant and inferences based on a Wiener process without taking the effects of a barrier into consideration can be misleading. The degradation process of a surviving item, therefore, should be described by a Wiener process with the condition that $X(t)$ is restricted by a barrier at which the item fails. Such a process will be called a **truncated Wiener process** in this thesis. The following section presents several derivations of the density function of the level of a Wiener process given that it is restricted by a fixed barrier.

3.1 Density Function of a Truncated Wiener Process

The truncated Wiener process is denoted here by $\{X_t(\tau), 0 \leq \tau \leq t\}$ and has the following definition,

$$X_t(\tau) = X(\tau)|A,$$

where $[0,t]$ is a fixed observation period, $X(\tau)$ denotes a parent Wiener process with drift parameter δ and variance parameter ν, A denotes the restriction that $X(\tau) < a$ for all $0 \le \tau \le t$, and $a > 0$ denotes the level of the barrier.

The definition of the truncated process just presented will be referred to later as the **standard problem.** A number of other formulations can be converted to the standard problem by simple transformations. For example, a Wiener process with a negative barrier $a < 0$ is clearly the mirror image of the standard problem.

3.1.1 Derivation for the Standard Problem

The *p.d.f.* of the truncated random variable $X_t(t)$ for the standard problem is of interest here. In other words, we are interested in the derivation of the *p.d.f.* of the degradation level at time t of a surviving sample path. The density function can be derived in several ways. Four derivations are presented next. The derivations are of interest in their own right and each contributes insights into developments which follow later. The first two derivations are based on the following decomposition.

The probability density for the parent Wiener process $\{X(\tau), \tau \ge 0\}$ terminating at level x at time t is given by the normal density $\phi(x;t)$ in (2.3). Any sample path starting at $X(0) = 0$ and terminating at $X(t) = x$ either does not exit the barrier at $a > 0$ in the interval $(0,t)$ or exits the barrier in the interval. The former event is denoted here by A and the latter event by $\bar{A}$, the complement of A. Therefore, the density $\phi(x;t)$ consists of two components representing the density contributed by sample paths that do not exit, $f(x,A)$, and sample paths that do exit, $q(x,\bar{A})$, as follows.

$$\phi(x;t) = f(x,A) + q(x,\bar{A}) \tag{3.1}$$

Note that $f(x,A)$ and $q(x,\bar{A})$ are both joint *p.d.f.*s. With respect to x, $f(x,A)$

has support on $(-\infty, a)$ and $q(x, \bar{A})$ has support on the whole real line. The density function $f(x, A)$ can be derived by using (1) the inverse Gaussian distribution; (2) the method of images; (3) a partial differential equation method; and (4) Shepp's joint density function. The following four subsections describe the derivation by each method.

Inverse Gaussian Derivation Let S be the first passage time of the parent Wiener process $X(\tau)$ to the barrier $a > 0$, i.e.,

$$S = \inf[\tau | X(\tau) = a].$$

Then S has an inverse Gaussian distribution with the following *p.d.f.* (Chhikara and Folks, 1989; Seshadri, 1993).

$$h(s) = \frac{a}{\sqrt{2\pi\nu s^3}} \exp\left(-\frac{(a - \delta s)^2}{2\nu s}\right) \tag{3.2}$$

Note that when the drift δ is negative, the inverse Gaussian density function becomes defective (Whitmore, 1978). In this case, the sample path will cross the barrier with probability $\exp(2\delta a/\nu) < 1$. The extended density function when the drift is negative, $h(s; \delta < 0)$, has the following form.

$$h(s; \delta < 0) = \begin{cases} h(s) & \text{for} \quad 0 < s < \infty \\ 1 - \exp(2\delta a/\nu) & \text{for} \quad s = \infty \end{cases}$$

The following derivation is valid for both positive and negative drift, however.

For the outcome $(x, \bar{A})$ in $q(x, \bar{A})$ to occur, the sample path must exit the barrier at some time $s < t$ and then move to level x in the remaining time (s, t). Thus, the joint density function $q(x, \bar{A})$ in (3.1) is the product of the inverse Gaussian density $h(s)$ and normal density $\phi(x - a; t - s)$ integrated over $[0, t]$ with respect to s. That is,

$$q(x, \bar{A}) = \int_0^t h(s)\phi(x-a; t-s)ds = \phi(x;t)\exp\left(\frac{2a(x-a)}{\nu t}\right).$$

Accordingly, using equation (3.1), we have the desired *p.d.f.*

$$f(x, A) = \phi(x;t)\left(1 - \exp\left[\frac{2a(x-a)}{\nu t}\right]\right). \qquad (3.3)$$

Method of Images The density function $f(x, A)$ can also be derived by using a reflection principle. Consider the following case as shown in Figure 3.2. The Wiener sample path $X(\tau)$ crosses the barrier $a > 0$ at time s and reaches x at time t. Its image path $\tilde{X}(\tau)$ is defined as

$$\tilde{X}(\tau) = \begin{cases} X(\tau), & 0 \le \tau \le s \\ 2a - X(\tau), & s < \tau \le t. \end{cases}$$

The likelihood ratio of the image path relative to the original path is (Whitmore and Seshadri, 1987)

$$\Lambda(x) = \exp(\frac{2\delta(x-a)}{\nu}),$$

where $\Lambda(x)$ is independent of s.

Sample paths terminating at level $x < a$ either do not exit the barrier during the interval $(0, t)$ or exit the barrier and have image paths terminating at level $(2a - x)$. The former sample paths have total density $f(x, A)$ while the latter have total density $\phi(2a - x; t)/\Lambda(x)$. Again from equation (3.1), we have

$$f(x, A) = \phi(x;t) - \frac{\phi(2a-x;t)}{\Lambda(x)} = \frac{1}{\sqrt{2\pi\nu t}}\exp\left(-\frac{(x-\delta t)^2}{2\nu t}\right)\left(1 - \exp\left[\frac{2a(x-a)}{\nu t}\right]\right).$$

Partial Differential Equation Method Cox and Miller (1965) derive the *p.d.f.* of a Wiener process truncated by an absorbing barrier at $a > 0$. The density function, denoted by $p(x, t)$, is the solution of the partial differential equation

$$\frac{\partial p}{\partial t} = \frac{\nu}{2}\frac{\partial^2 p}{\partial x^2} - \delta\frac{\partial p}{\partial x} \qquad (x < a)$$

with boundary conditions

$$p(x;0) = \delta(x)$$

$$p(a;t) = 0 \quad t > 0$$

where $\delta(x)$ is the Dirac delta function. According to Cox and Miller, the solution is

$$p(x,t) = \frac{1}{\sqrt{2\pi\nu t}} \left(\exp\left[-\frac{(x-\delta t)^2}{2\nu t} \right] - \exp\left[\frac{2\delta a}{\nu} - \frac{(x-2a-\delta t)^2}{2\nu t} \right] \right).$$

By rearranging the right-hand side, $p(x,t)$ becomes

$$p(x,t) = \frac{1}{\sqrt{2\pi\nu t}} \exp\left(-\frac{(x-\delta t)^2}{2\nu t} \right) \left(1 - \exp\left[\frac{2a(x-a)}{\nu t} \right] \right).$$

The function $p(x,t)$ corresponds to $f(x,A)$.

Shepp's Joint Probability Density Function Shepp (1979) gives the following as the joint *p.d.f.* of the maximum value $Y = \max_{0 \le \tau \le t} X(\tau)$ and the terminal level $X = X(t)$ of a Wiener process.

$$g(y,x) = \frac{2(2y-x)}{\sqrt{2\pi\nu^3 t^3}} \exp\left(\frac{x\delta}{\nu} - \frac{\delta^2 t}{2\nu} \right) \exp\left(-\frac{(2y-x)^2}{2\nu t} \right)$$

By integrating $g(y,x)$ with respect to y over $[x,a]$, the joint *p.d.f.* of a Wiener process terminating at x and not exiting the barrier in the interval $(0,t)$ is obtained.

$$\begin{aligned} f(x,A) &= \int_x^a g(y,x)dy \\ &= \frac{1}{\sqrt{2\pi\nu t}} \exp\left(-\frac{(x-\delta t)^2}{2\nu t} \right) \left(1 - \exp\left[\frac{2a(x-a)}{\nu t} \right] \right) \end{aligned}$$

All four derivations necessarily result in the same joint *p.d.f.* for a Wiener process that does not exit the barrier in the interval $(0,t)$ and terminates at level $X(t) = x$. The conditional density function of x given no first passage is obtained by dividing the joint density $f(x,A)$ by the probability $P(A)$, where

$$P(A) = P(\max_{0 \le \tau \le t} X(\tau) < a) = P(S > t).$$

The probability $P(A)$ may be obtained from the following formula for the cumulative probability of an inverse Gaussian random variable,

$$P(S > t) = \Phi\left(\frac{a - \delta t}{\sqrt{\nu t}}\right) - \exp(\frac{2\delta a}{\nu})\Phi\left(-\frac{a + \delta t}{\sqrt{\nu t}}\right),$$

where Φ is the cumulative distribution function (c.d.f) of a standard normal variable. Thus, for $x < a$,

$$f(x|A) = \frac{f(x, A)}{P(A)} = \frac{1}{\sqrt{2\pi\nu t}P(A)} \exp\left(-\frac{(x - \delta t)^2}{2\nu t}\right)\left(1 - \exp\left[\frac{2a(x - a)}{\nu t}\right]\right). \tag{3.4}$$

Note that since $\int_{-\infty}^{a} f(x|A)dx = 1$, it follows that $P(A) = \int_{-\infty}^{a} f(x, A)dx$.

For a Wiener process with zero drift, the density function $f(x|A)$ in (3.4) becomes

$$f(x|A) = \frac{1}{\sqrt{2\pi\nu t}P(A)} \exp\left(-\frac{x^2}{2\nu t}\right)\left(1 - \exp\left[\frac{2a(x - a)}{\nu t}\right]\right) \tag{3.5}$$

where $x < a$ and the probability $P(A)$ has a special form, $P(A) = 1 - 2\Phi(-a/\sqrt{\nu t})$.

3.1.2 Cases Encompassed by the Standard Problem

A variety of problem settings can be transformed to the standard problem by appropriate transformations. Several of these settings are considered next.

Variable Stress Model The standard problem is one in which the rate of degradation of an item is represented by the parameter δ. This rate is constant and may be viewed as resulting from the application of a constant level of stress on the item. Thus, if the constant stress is applied for an interval of length t, the expected degradation will amount to δt. The stress may be produced by heat, voltage, humidity, vibration, or the like (Dasgupta and Pecht, 1991).

In variable stress tests, as noted earlier, the level of stress is changed during the test according to some fixed plan. Doksum and Hoyland (1992) propose a linear transformation of the time scale to convert the testing time r that has elapsed in a variable stress test to the effective operational time t of an item. A variable stress test can be represented as a standard problem by applying this or a similar transformation. For example, in a single step stress test, the level of stress is increased as a step function of time. Let us say, the stress increases from level y to level αy at time β, where $\alpha > 1$ is the acceleration coefficient, then the transformation has the following form.

$$t = h(r) = \begin{cases} r & r \leq \beta \\ \beta + \alpha(r - \beta) & r > \beta \end{cases}$$

The degradation process under variable stress, denoted by $\{D(r)\}$, is conveniently assumed to have the form

$$D(r) = X(h(r)),$$

where $\{X(t)\}$ is the Wiener process of the standard problem and $t = h(r)$ is a strictly increasing function of r which is either the linear form proposed by Doksum and Hoyland or some similar function that adjusts for the variable stress. If the function $h(r)$ is known, then the variable stress test results can be converted to the standard form by applying the transformation $t = h(r)$ to the time measurements in the data structure.

One implication of this transformation method is that the transformation $h(r)$ affects the mean and variance of a degradation increment in the same way. Specifically, in a time interval (r_1, r_2), the degradation increment $D(r_2) - D(r_1) = X\{h(r_2)\} - X\{h(r_1)\}$ and has the following distribution.

$$D(r_2) - D(r_1) \sim N\{\delta[h(r_2) - h(r_1)], \nu[h(r_2) - h(r_1)]\}$$

Stress-strength Model The stress-strength model is another case that can be transformed to the standard problem. Consider a stress process $\{X_2(t), t \geq 0\}$ and a strength process $\{X_1(t), t \geq 0\}$ that are Wiener processes (see Desmond, 1987). The strength level $X_1(t)$ acts as a random barrier for the stress process $X_2(t)$. When the stress first exceeds the strength then the item fails. Whitmore (1990) notes that the net strength process $\{W(t), t \geq 0\}$, where $W(t) = X_1(t) - X_2(t)$, is also a Wiener process with drift parameter $\delta = \delta_1 - \delta_2$ and variance parameter $\nu = \nu_1 + \nu_2$, where (δ_1, ν_1) and (δ_2, ν_2) are the respective drift and variance parameters of independent strength and stress processes. The net strength process $\{W(t), t \geq 0\}$ can be transformed to the standard problem by letting $W(0) - W(t)$ be the degradation process with δ and ν given above and a barrier $a = W(0) = X_1(0) - X_2(0)$.

A special case of this stress-strength model is one where the strength process is a deterministic linear function of time, i.e., $X_1(t) = X_1(0) + \delta_1 t$ (with $\nu_1 = 0$). This case corresponds to the standard problem with a barrier that is a linear function of time rather than a constant.

3.2 Inference for Terminal Point Data

This section presents statistical inference methods for terminal point data. It is assumed in this section that the barrier a is known but the drift parameter δ and the variance parameter ν must be estimated from the data. A comment about the case when a is also unknown is given later. The form of the terminal point data structure is repeated here for convenience. For simplicity, the second subscript n of x is suppressed and the termination time is represented by t without the subscript n.

$$\mathbf{x} = (x_1, x_2, ..., x_q), \quad \mathbf{s} = (s_1, s_2, ..., s_p)$$

The notation has the same meaning as in Chapter 2: x denotes the degradation level of a surviving item at time t and s denotes the lifetime of an item that fails before time t. Note that p, the number of failed items, is a binomial outcome with the following probability mass function.

$$\frac{(p+q)!}{p!q!} P(A)^p (1 - P(A))^q$$

Likelihood Function Since the sample paths are assumed to be mutually independent, the likelihood function of a sample of $p+q$ such items is the product of the individual likelihood functions of the failed and surviving sample paths. The density function of a sample path that failed at time s is given by $h(s)$ in (3.2). The density function of the level of a surviving sample path is given by $f(x, A)$ in equation (3.3). Therefore, the likelihood function of a sample of q surviving items and p failed items is

$$L(\delta, \nu) = \prod_{j=1}^{p} h(s_j) \prod_{i=1}^{q} f(x_i, A) = \prod_{j=1}^{p} \frac{a}{\sqrt{2\pi\nu s_j^3}} \exp\left(-\frac{(a-\delta s_j)^2}{2\nu s_j}\right) \times$$

$$\prod_{i=1}^{q} \frac{1}{\sqrt{2\pi\nu t}} \exp\left(-\frac{(x_i-\delta t)^2}{2\nu t}\right)\left(1-\exp\frac{2a(x_i-a)}{\nu t}\right). \tag{3.6}$$

Likelihood function (3.6) can be expressed in a simpler form with the following notation.

$$(d_i, t_i) = \begin{cases} (a, s_i) & \text{for a failed item} \\ (x_i, t) & \text{for a surviving item} \end{cases}$$

$$L(\delta, \nu) = \prod_{j=1}^{p+q} \frac{1}{\sqrt{2\pi\nu t_j}} \exp\left(-\frac{1}{2}\sum_{j=1}^{p+q}\frac{(d_j-\delta t_j)^2}{\nu t_j}\right) \prod_{j=1}^{p}\frac{a}{t_j} \prod_{i=1}^{q}\left(1-\exp\frac{2a(d_i-a)}{\nu t_i}\right) \tag{3.7}$$

Maximum Likelihood Estimators (MLEs) Let $d_S = \sum_1^{p+q} d_i$ and $t_S = \sum_1^{p+q} t_i$ represent the total accumulated degradation and total accumulated test time, respectively, then the corresponding MLEs for δ and ν are as follows,

$$\hat{\delta} = \frac{d_S}{t_S} \tag{3.8}$$

$$\hat{\nu} = \frac{1}{p+q}\left(\sum_{j=1}^{p+q}\frac{(d_i-\hat{\delta}t_i)^2}{t_i} - \sum_{i=1}^{q} K_i(\hat{\nu})\right) \tag{3.9}$$

where

$$K_i(\hat{\nu}) = \frac{4a(a-x_i)}{t\left(\exp\left[\frac{2a(a-x_i)}{\hat{\nu}t}\right]-1\right)}. \qquad (3.10)$$

The MLE of δ is the ratio of total accumulated degradation over total accumulated test time, which is an intuitive result. In the expression for the MLE of ν, the second term in parentheses would disappear if the Wiener process were not truncated, i.e., if the barrier were set at $a = \infty$. It is only this second term that yields estimates that differ from the standard estimates for the parameters of a Wiener process. Note that the function $K_i(\hat{\nu})$ depends on $\hat{\nu}$ but not on $\hat{\delta}$.

Looking now at the second order derivatives of $l = \ln(L(\delta,\nu))$, it is found that $\frac{\partial^2 l}{\partial\nu\partial\delta}$ is zero. Also, the second order derivative, $\frac{\partial^2 l}{\partial\delta^2}$, is negative everywhere. Thus, $\hat{\delta}$ corresponds to the maximum of the likelihood function. In addition, the second order derivative $\frac{\partial^2 l}{\partial\nu^2}$ exists for every pair of $\delta \in (-\infty,\infty)$, $\nu \in (0,\infty)$ and is negative for every root of equation (3.9). Since there must be a minimum between every two maxima, equation (3.9) must have a unique solution $\hat{\nu}$. Therefore, the MLEs $(\hat{\delta},\hat{\nu})$ correspond to the unique maximum of the likelihood function. Note that the maximum likelihood estimate of ν cannot lie at zero except in a degenerate case.

The MLEs are asymptotically independent because $\frac{\partial^2 l}{\partial\nu\partial\delta} = 0$. The estimated asymptotic covariance matrix of $\hat{\delta}$ and $\hat{\nu}$ is the following,

$$Cov(\hat{\delta},\hat{\nu}) = \begin{pmatrix} \frac{\hat{\nu}}{ts} & 0 \\ 0 & -\left[\frac{\partial^2 l}{\partial\nu^2}|_{\hat{\delta},\hat{\nu}}\right]^{-1} \end{pmatrix}, \qquad (3.11)$$

where

$$\frac{\partial^2 l}{\partial\nu^2}|_{\hat{\delta},\hat{\nu}} = -\frac{p+q}{2\hat{\nu}^2} - \sum_{i=1}^{q}\left(\frac{K_i(\hat{\nu})}{2\hat{\nu}^2}\right)^2 \exp\left(\frac{2a(a-x_i)}{\hat{\nu}t}\right). \qquad (3.12)$$

If the degradation data $\mathbf{x}$ are not included in the terminal point data structure then no more than p, q, and s are known. The structure then corresponds to the censored lifetime data structure defined in Chapter 2, that is,

$$\mathbf{s} = (s_1, s_2, ..., s_p, t, ..., t).$$

This case has been studied by Whitmore (1983) and Desmond (1982). The lifetimes s_i, $i = 1, 2, ..., p$, have an inverse Gaussian distribution with the *p.d.f.* $h(s)$ given in (3.2) and the q censored observations each has probability $P(A)$. The likelihood function, therefore, is

$$L(\delta, \nu) = \prod_{j=1}^{p} \frac{a}{\sqrt{2\pi\nu s_j^3}} \exp\left(-\frac{(a-\delta s_j)^2}{2\nu s_j}\right) [P(A)]^q. \tag{3.13}$$

The expression in (3.13) is obtained from (3.6) by integrating over x_i for each $i = 1, 2, ..., q$, which gives $P(A)$ in each case.

Inferences About Parameters Inferences about the parameters δ and ν can be derived based on the asymptotic normality of the MLEs or on the asymptotic properties of the likelihood ratio.

When the sample size $p + q$ is large, the MLEs of $\hat{\delta}$ and $\hat{\nu}$ are nearly independent and each has an approximate normal distribution with the estimated variance given in matrix (3.11). Inferences about the parameters can be derived from this asymptotic normal theory. For example, a $1 - \alpha$ confidence interval for δ is then $\hat{\delta} \pm z(\alpha/2)\hat{\sigma}_\delta$, where $\hat{\sigma}_\delta = \sqrt{\hat{\nu}/t_S}$ is the asymptotic standard deviation of $\hat{\delta}$ and $z(\alpha/2)$ is the $\alpha/2$ standard normal percentile. An asymptotic joint confidence region for δ and ν can also be derived in a corresponding manner.

The likelihood ratio method provides another approach to inference here. The likelihood ratio is defined as the ratio of the maximum likelihood under a reduced model to the maximum likelihood under a full model, where the reduced model is nested within the full model. The reduced model corresponds to a specified null hypothesis. A modified likelihood ratio will be used here whereby the nuisance parameters are estimated by the MLEs computed from the full model rather than from the reduced model. This modification is appropriate because in the case where the null hypothesis is true, the MLEs for this parameter calculated from the reduced and full model are asymptotically equivalent and in the case where the null hypothesis is false, the MLE calculated from the reduced model may be inconsistent.

Now, to illustrate the use of this method for hypothesis testing, assume that the following null hypothesis is of interest,

$$H_0 : \delta = \delta_0.$$

Then the modified likelihood ratio, denoted by Λ, is

$$\Lambda = \frac{L(\delta_0, \hat{\nu})}{L(\hat{\delta}, \hat{\nu})} = \exp\left(-t_S \frac{(\delta_0 - \hat{\delta})^2}{2\hat{\nu}}\right)$$

where $\hat{\delta}$ and $\hat{\nu}$ are the MLEs in (3.8) and (3.9), respectively. For a large sample size $p + q$,

$$-2\ln(\Lambda) = t_S \frac{(\delta_0 - \hat{\delta})^2}{\hat{\nu}} \sim \text{approx. } \chi^2(1). \tag{3.14}$$

Note that t_S here plays the role of the effective sample size. The preceding result can be rearranged to yield a confidence interval for δ of the form $\hat{\delta} \pm z(\alpha/2)\sqrt{\hat{\nu}/t_S}$, which was previously encountered in applying the asymptotic normality theory for MLEs. For this test, the modified likelihood ratio method produces a test statistic that is the same as the one obtained based on asymptotic normal theory. However, with the conventional likelihood method, the test statistics would not be identical.

Analysis of Deviance In addition to the two general inference methods just presented, an approximate analysis of deviance can also be derived for terminal point data. Let Q, Q_R, and Q_E define the following quadratic forms.

$$Q_R = t_S \frac{(\delta - \hat{\delta})^2}{\nu} \tag{3.15}$$

$$Q_E = \sum_{i=1}^{p+q} \frac{(d_i - \hat{\delta} t_i)^2}{\nu t_i} \tag{3.16}$$

$$Q = Q_R + Q_E = \sum_{i=1}^{p+q} \frac{(d_i - \delta t_i)^2}{\nu t_i}$$

Consider the following two limiting cases: (1) The barrier level approaches infinity ($a \to \infty$); (2) The termination time t approaches infinity ($t \to \infty$). In case (1), when the barrier level goes to infinity, the truncated Wiener process converges to a Wiener process and no failure occurs. Therefore, the terminal point data become a normal sample from $N(\delta t, \nu t)$. In case (2), when the termination time goes to infinity, all items fail and, therefore, the terminal point data become an inverse Gaussian sample from $IG(a/\delta, a^2/\nu)$. In both limiting cases (1) and (2), the quadratic forms are χ^2 distributed as follows,

$$Q_R \sim \chi^2(1), \qquad Q_E \sim \chi^2(p+q-1), \qquad Q \sim \chi^2(p+q),$$

and Q_R is independent of Q_E. Furthermore, Q_E and $(p+q-1)Q_R/Q_E \sim F(1, p+q-1)$ are pivotal quantities for ν and δ, respectively.

Now consider the terminal point data case with p failures and q survivors. Recall that the quadratic form Q appears in likelihood function (3.7) and that $Q_R \sim$ approx. $\chi^2(1)$, as shown in (3.14), based on asymptotic theory for MLEs. An interesting issue arises here concerning whether the chi-square distributional properties of Q_E and Q are preserved and whether Q_E and Q_R are independent or nearly so with terminal point data. This issue is taken up next.

Let Q_i denote a single term of Q, i.e.,

$$Q_i = \frac{(d_i - \delta t_i)^2}{\nu t_i}.$$

The Laplace transform of Q_i is

$$E(e^{-Q_i r}) = \int_{-\infty}^{a} \exp\left(-\frac{(x - \delta t)^2 r}{\nu t}\right) f(x, A)dx + \int_0^t \exp\left(-\frac{(a - \delta s)^2 r}{\nu s}\right) h(s)ds,$$

where $f(x, A)$ is given in (3.3) and $h(s)$ is given in (3.2). With some algebraic manipulation, the Laplace transformation of Q_i can be shown to have the following form,

$$E(e^{-Q_i r}) = (1 + 2r)^{-\frac{1}{2}} - \psi(r), \tag{3.17}$$

where

$$\psi(r) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\nu t}} \exp\left(\frac{2a(x-a)}{\nu t} - \frac{(2r+1)(x-\delta t)^2}{\nu t}\right)\left(1 - \exp\frac{4ar(x-a)}{\nu t}\right) dx.$$

Therefore, Q_i can have a $\chi^2(1)$ distribution when the term denoted by $\psi(r)$ is zero. Except in the two limiting cases, $\psi(r)$ is not identically zero and, hence, the exact distribution of Q_i requires further study.

Since the $Q_i, i = 1, 2, ..., p + q = m$, are mutually independent, the Laplace transform of Q is the product of the individual Laplace transforms of Q_i. Thus, the mean of Q has the form

$$E(Q) = m + m\psi'(0)$$

where

$$\psi'(0) = E[\frac{K(\nu)}{\nu}] = \frac{4a}{\sqrt{\nu t}} \exp(\frac{2a\delta}{\nu}) L(\frac{a + \delta t}{\sqrt{\nu t}}).$$

Here, $K(\nu)$ is the function defined in (3.10) and $L(\cdot)$ is the unit normal loss function, defined as

$$L(l) = \int_{l}^{\infty} (z - l) \exp(-\frac{z^2}{2}) dz.$$

Having taken the investigation this far analytically and taking account of the intended scope of the thesis, it was decided that further progress could only be made by numerical methods.

It is conjectured that the quantities Q and Q_E are approximately χ^2 distributed and Q_R and Q_E are nearly independent. If this is indeed the case, then the degrees of freedom of Q will approximately equal $E(Q)$, i.e.,

$$df \approx E(Q) = m + m\psi'(0). \tag{3.18}$$

Furthermore, the conjecture implies that Q and Q_E will have approximately $\chi^2(df)$ and $\chi^2(df-1)$ distributions, respectively.

From the two limiting cases discussed earlier, one knows that Q, Q_E, and Q_R have the conjectured properties when the drift parameter is very large and positive because then nearly all sample paths will exit the barrier before stopping time t. Likewise, the conjecture holds when the drift is very large and negative since, in this case, nearly all sample paths will survive to the stopping time t. The conjecture is most severely challenged when both exiting and surviving paths are common, which will occur when δ is in the neighbourhood of a/t.

A simulation study was conducted to provide empirical evidence for the sampling distributions of Q and Q_E. The distributional characteristics were studied for different parameter values and sample sizes $p+q$. The following simulation result is representative. The quantities Q, Q_E, and Q_R were simulated for $\delta = 0.2$, $\nu = 2$, $a = 10$, $t = 60$, and $m = p + q = 50$. The results were based on 1000 simulation trials. For these parameter values, $df = E(Q) = 61.5$ or 62 to the nearest whole number. Figures 3.3 (a), (b), and (c) show a $\chi^2(62)$ plot, a $\chi^2(61)$ plot, and a $\chi^2(1)$ for Q, Q_E, and Q_R, respectively. The plots indicate a good fit. Table 3.1 lists the summary statistics of the simulation results. The correlation coefficient of Q_R and Q_E is 0.009 which supports the conjecture of independence of the two quantities.

In summary, the combined analytical and numerical results suggest that, for $df = m + m\psi'(0)$,

$$Q_R \sim \text{approx. } \chi^2(1), \qquad Q_E \sim \text{approx. } \chi^2(df-1), \qquad Q \sim \text{approx. } \chi^2(df),$$

and Q_R is approximately independent of Q_E.

Statistical inferences about the parameters δ and ν can be drawn using these distributional properties for Q_R and Q_E. The quantity Q_E is an approximate pivotal quantity for ν. The ratio $(df-1)Q_R/Q_E$ is an approximate pivotal quantity for δ. The ratio has the following approximate F distribution.

$$F^* = (df-1)\frac{Q_R}{Q_E} = \frac{(df-1)(\hat{\delta}-\delta)^2 t_S}{\sum_1^{p+q}\frac{(d_i-\hat{\delta}t_i)^2}{t_i}} \sim \text{approx. } F(1, df-1) \tag{3.19}$$

Approximate interval estimates and tests for δ or ν can be developed in the usual manner using these distributional findings.

This investigation of analysis of deviance for terminal point data has left a number of interesting issues for future research. One issue concerns an analytical demonstration to support the numerical findings given above.

3.3 Inference for Mixed Data

The previous two sections have studied statistical inference for terminal point data. This section and the following section will extend the study to the mixed data structure. The data structure is defined in Chapter 2 and repeated here for convenience.

for a failed sample path i, $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{in_i}, s_i)$ $\qquad i = 1, 2, ..., p$

for a surviving sample path i, $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{in})$ $\qquad i = p+1, p+2.., p+q$

Likelihood Function Each sample path has n_i observations: $x_{i1}, x_{i2}, ..., x_{in_i}$ at a set of fixed time points, $0 < t_1 < t_2 < ... < t_{n_i}$, where $n_i = n$ for every surviving sample path. Notice that the degradation levels for a given sample path are not mutually independent. The following derivation gives the joint density function for $\mathbf{x}_i$ of sample path i. For simplicity of notation, the subscript i will be suppressed in the following derivation.

Let A_j denote the event that a sample path survives in time interval $[t_{j-1}, t_j]$, $\bar{A}_j$ denote the complement of A_j, and $A_{0n} = A_1 \cap A_2 \cap ..... \cap A_n$ denote no passage from time 0 to t_n. Based on the Markov property (2.2), we have $P(A_j \cap x_j | x_{j-1} \cap A_{j-1}) = P(A_j \cap x_j | x_{j-1})$. Therefore, the joint probability of x_1, $x_2, ..., x_n$ and A_{0n}, is

$$P(A_{0n} \cap x_1 \cap x_2 \cap ... \cap x_n) = P(A_1 \cap x_1)P(A_2 \cap x_2|x_1)......P(A_n \cap x_n|x_{n-1}).$$

Expressed alternatively in terms of density functions,

$$f(x_1, x_2, ..., x_n, A_{0n}) = f(x_1, A_1) f(x_2, A_2 | x_1) ...... f(x_n, A_n | x_{n-1}) \tag{3.20}$$

where $f(x_1, A_1)$ is defined by equation (3.3) and $f(x_j, A_j | x_{j-1})$, $j = 2, 3, ..., n$, can be obtained by replacing x with $\Delta x_j = (x_j - x_{j-1})$, t with $\Delta t_j = (t_j - t_{j-1})$, and a with $(a - x_{j-1})$ in equation (3.3). That is

$$f(x_j, A_j | x_{j-1}) = \frac{1}{\sqrt{2\pi\nu \Delta t_j}} \exp\left(-\frac{(\Delta x_j - \delta \Delta t_j)^2}{2\nu \Delta t_j}\right) \times \left(1 - \exp\left[\frac{2(a - x_{j-1})(x_j - a)}{\nu \Delta t_j}\right]\right). \tag{3.21}$$

Each surviving sample path will fail after t_n with a probability of $\int_{t_n}^{\infty} h(s) ds$, where $h(s)$ is the inverse Gaussian density function given by equation (3.2). Thus, the likelihood function of a surviving sample path i, denoted by $L_{qi}(\delta, \nu)$ is

$$L_{qi}(\delta, \nu) = f(x_{i1}, x_{i2}, ..., x_{in}, | A_{0n}) \int_{t_n}^{\infty} h(s) ds.$$

Note that $P(A_{0n}) = \int_{t_n}^{\infty} h(s) ds$ is the probability of no passage before time t_n. So,

$$L_{qi}(\delta, \nu) = f(x_{i1}, x_{i2}, ..., x_{in}, A_{0n}) \qquad i = p+1, p+2, ..., p+q.$$

The joint density function of a failed sample path can be derived as follows. Consider a sample path that starts at $X(0) = 0$ and ends at $X(s) = a$ with no passage in the time interval $(0, s)$. Now, denote the level of the sample path at time t_j, where $0 < t_j < s$, by x_j. Then, the joint density function of x_j and s is

$$f(x_j, s) = f(x_j, A_{0j}) h(s | x_j) \tag{3.22}$$

where $f(x_j, A_{0j})$ is the joint density function of $X(t_j) = x_j$ and no passage in the time interval $(0, t_j)$, given by (3.3); and $h(s|x_j)$ is the *p.d.f.* of the first passage time of a sample path starting at $X(t_j) = x_j$. The likelihood function of a failed sample path, $L_{pi}(\delta, \nu)$, can be readily derived from equation (3.20) and (3.22).

$$\begin{aligned} L_{pi}(\delta,\nu) &= f(x_{i1}, x_{i2}, ..., x_{in_i}, s_i) \\ &= f(x_{i1}, x_{i2}, ..., x_{in_i}, A_{0n_i}) h(s_i|x_{in_i}) \qquad\qquad i = 1, 2, ..., p \end{aligned}$$

Since each sample path evolves independently, the likelihood function for the full sample is

$$L(\delta,\nu) = \prod_{i=1}^{p} L_{pi}(\delta,\nu) \prod_{i=p+1}^{p+q} L_{qi}(\delta,\nu) = \prod_{i=1}^{p+q} f(x_{i1}, x_{i2}, ..., x_{in_i}, A_{0n_i}) \prod_{i=1}^{p} h(s_i|x_{in_i})$$

$$= \prod_{i=1}^{p} \frac{a - x_{in_i}}{\sqrt{2\pi\nu(s_i - t_{n_i})^3}} \exp\left(-\frac{[a - x_{in_i} - \delta(s_i - t_{n_i})]^2}{2\nu(s_i - t_{n_i})}\right) \times$$

$$\prod_{i=1}^{p+q} \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\nu\Delta t_j}} \exp\left(-\frac{(\Delta x_{ij} - \delta\Delta t_j)^2}{2\nu\Delta t_j}\right)\left(1 - \exp\left[\frac{2(a - x_{ij})(x_{i,j-1} - a)}{\nu\Delta t_j}\right]\right). \tag{3.23}$$

Again, to simplify the expression, let F and S be sets defined as:

$$F = \{(a - x_{in_i}, s_i - t_{n_i}); i = 1, 2, ..., p\},$$

$$S = \{(x_{ij} - x_{i,j-1}, t_{ij} - t_{i,j-1}); i = p+1, ..., p+q, j = 1, 2, ..., n_i\}.$$

Set F contains all paired increments that end in failure and S contains all other paired increments. Notation $(\Delta d, \Delta t)$ will represent a general element of these sets. Then the likelihood function can be rewritten as,

$$L(\delta,\nu) = \prod_{S,F} \frac{1}{\sqrt{2\pi\nu \triangle t}} \exp\left(-\frac{(\triangle d - \delta \triangle t)^2}{2\nu \triangle t}\right) \times$$

$$\prod_F \frac{\Delta d}{\Delta t} \prod_{i=1}^{p+q} \prod_{j=1}^{n_i} \left(1 - \exp\left[\frac{2(a - x_{ij})(x_{i,j-1} - a)}{\nu \Delta t_j}\right]\right). \tag{3.24}$$

Maximum Likelihood Estimators Letting $l = \log L(\delta, \nu)$ and setting the partial derivatives of l with respect to δ and to ν to zero, we obtain the maximum likelihood estimators of δ and ν as

$$\hat{\delta} = \frac{\sum_{S,F} \Delta d}{\sum_{S,F} \Delta t} = \frac{d_S}{t_S} \tag{3.25}$$

$$\hat{\nu} = \frac{1}{p + n_T} \left\{ \sum_{S,F} \frac{(\Delta d - \hat{\delta}\, \Delta\, t)^2}{\Delta t} - \sum_{i=1}^{p+q} \sum_{j=1}^{n_i} K_{ij}(\hat{\nu}) \right\} \tag{3.26}$$

where

$$K_{ij}(\hat{\nu}) = \frac{4(a - x_{ij})(a - x_{i,j-1})}{\Delta t_j \left(\exp\left[\frac{2(a - x_{ij})(a - x_{i,j-1})}{\hat{\nu} \Delta t_j}\right] - 1\right)} \tag{3.27}$$

and $n_T = \sum_{i=1}^{p+q} n_i$. For the same reasons stated in Section 3.2, $\hat{\delta}$ and $\hat{\nu}$ correspond to the unique maximum of the likelihood function. The estimated asymptotic covariance matrix of $\hat{\delta}$ and $\hat{\nu}$ is

$$Cov(\hat{\delta}, \hat{\nu}) = \begin{pmatrix} \frac{\hat{\nu}}{t_S} & 0 \\ 0 & -\left[\frac{\partial^2 l}{\partial \nu^2}|_{\hat{\delta},\hat{\nu}}\right]^{-1} \end{pmatrix}, \tag{3.28}$$

where

$$\frac{\partial^2 l}{\partial \nu^2}|_{\hat{\delta},\hat{\nu}} = -\frac{p + n_T}{2\hat{\nu}^2} - \sum_{i=1}^{p+q} \sum_{j=1}^{n_i} \left(\frac{K_{ij}(\hat{\nu})}{2\hat{\nu}^2}\right)^2 \exp\left(\frac{2(a - x_{ij})(a - x_{i,j-1})}{\hat{\nu}\, \Delta\, t_j}\right).$$

The off diagonal element is zero, indicating that $\hat{\delta}$ and $\hat{\nu}$ are asymptotically independent.

Notice that the MLE for δ, namely $\hat{\delta} = d_S/t_S$, and the estimated asymptotic variance of this MLE, $\hat{\nu}/t_S$, have the same form for both terminal point data and mixed data. Thus, the MLEs for δ for the two types of data sets have the same asymptotic efficiency. As for the MLE of ν, the asymptotic variance of $\hat{\nu}$ for the two types of data sets can be compared analytically and, as would be expected, the mixed data structure yields a smaller true asymptotic variance, provided that the mixed data set actually contains intermediate degradation readings so $n_T > q$. Thus, for any given set of sample paths, mixed data, in comparison to terminal point data, lead to a better estimate for ν but the same estimate for δ.

Inferences About Parameters Asymptotic theory for MLEs and the likelihood ratio method can be applied to the mixed data structure as they were to the terminal data structure with little modification. The analysis of deviance developed for terminal point data can be extended to the mixed data structure but the detailed technical development is left to future research.

3.4 Sample Path Homogeneity

Degradation processes for different items can follow a random process with different parameter values or even different random processes. Even for items with the same design, their degradation processes can be driven by different parameters because of the presence of manufacturing variability. This section develops two tests of sample path homogeneity: (1) variance parameter homogeneity and (2) drift parameter homogeneity, assuming that the degradation processes are Wiener processes.

First, a likelihood ratio test for variance parameter homogeneity is presented. As in Section 3.1, a modified likelihood ratio test is derived here. That is, the MLE of a nuisance parameter under the full model is used in calculating the maximum likelihood of both the reduced model under H_0 and the full model.

The null hypothesis is

$$H_0: \quad \nu_1 = \nu_2 = ... = \nu_{p+q} \tag{3.29}$$

while the drift parameters are not restricted to be the same. The corresponding likelihood ratio test statistic is

$$-2\ln(\Lambda) = \sum_{S,F} \ln\frac{\hat{\nu}}{\hat{\nu}_i} + \sum_{S,F}\left(\frac{1}{\hat{\nu}} - \frac{1}{\hat{\nu}_i}\right)\frac{(\Delta d - \hat{\delta}_i \Delta t)^2}{\Delta t} - 2\sum_{i=1}^{p+q}\sum_{j=1}^{n_i} \ln\frac{K_{ij}(-\hat{\nu}_i)}{K_{ij}(-\hat{\nu})} \tag{3.30}$$

where

$$\hat{\nu} = \frac{1}{p+n_T}\left\{\sum_{S,F}\frac{(\Delta d - \hat{\delta}_i \Delta t)^2}{\Delta t} - \sum_{i=1}^{p+q}\sum_{j=1}^{n_i} K_{ij}(\hat{\nu})\right\}, \tag{3.31}$$

$$\hat{\delta}_i = \frac{\sum_{S_i,F_i}\Delta d}{\sum_{S_i,F_i}\Delta t}, \tag{3.32}$$

and $K_{ij}(-\hat{\nu}_i)$ in (3.30) has the form of (3.27) with $\hat{\nu}$ being replaced by $-\hat{\nu}$. Notation S_i and F_i denote the subsets of S and F that contain the corresponding elements of sample path i. The MLE $\hat{\nu}_i$ is given as follows.

For failed sample path i

$$\hat{\nu}_i = \frac{1}{1+n_i}\left\{\sum_{S_i,F_i}\frac{(\Delta d - \hat{\delta}_i \Delta t)^2}{\Delta t} - \sum_{j=1}^{n_i} K_{ij}(\hat{\nu}_i)\right\}. \tag{3.33}$$

For surviving sample path i, the subset F_i is empty and in the above equation the term $1/(1+n_i)$ should be replaced by $1/n_i$.

It is well know that, for a large sample,

$$-2\ln(\Lambda) \sim \text{approx. } \chi^2(p+q-1).$$

If the null hypothesis (3.29) is accepted, naturally one would test next if the drift parameters are homogeneous. With the variance parameter assumed to be common, the null hypothesis of interest is the following.

$$H_0: \quad \delta_1 = \delta_2 = \ldots = \delta_{p+q}$$

The modified likelihood ratio test statistic, $-2\ln(\Lambda)$, has the following form and, for a large sample, is approximately χ^2 distributed.

$$-2\ln(\Lambda) = \sum_{i=1}^{p+q} \sum_{S_i F_i} \Delta t \frac{(\hat{\delta} - \hat{\delta}_i)^2}{\hat{\nu}} \sim \text{approx. } \chi^2(p+q-1) \tag{3.34}$$

where $\hat{\delta}$ and $\hat{\delta}_i$ are given by (3.25) and (3.32), respectively, and $\hat{\nu}$ is given in (3.31).

The test statistic for

$$H_0 : \delta_1 = \delta_2 = \ldots = \delta_{p+q}$$

without assuming a common variance parameter is

$$-2\ln(\Lambda) = \sum_{i=1}^{p+q} \sum_{S_i F_i} \Delta t \frac{(\hat{\delta} - \hat{\delta}_i)^2}{\hat{\nu}_i} \sim \text{approx. } \chi^2(p+q-1) \tag{3.35}$$

where $\hat{\delta}$ and $\hat{\delta}_i$ are given by (3.25) and (3.32) respectively and $\hat{\nu}_i$ are given in (3.33).

A simultaneous test for homogeneity of both drift and variance parameters can be derived in a similar manner. The details are omitted.

A final observation about terminal point data and mixed data with respect to sample paths homogeneity is worth noting. Parameter estimation based on terminal point data relies on the assumption of partial or complete sample path homogeneity because with m data points one can estimate at most m parameters. In contrast, with mixed data (containing some intermediate degradation readings), more parameters can be estimated; specifically, $p + n_T > m$ parameters can be estimated, where $n_T = \sum_{i=1}^{p+q} n_i$.

3.5 Inferences When Barrier a is Unknown

In some applications, both the barrier level a and the process parameters are unknown and must be estimated. For example, in engineering applications, the level of a performance parameter at which an electronic device fails may be unknown and have to be estimated from a time series of the parameter readings.

When a, δ and ν are unknown and both degradation and lifetime data are observed, the maximum likelihood estimation would proceed first by maximizing (3.7) or (3.24) with parameter a fixed and then by examining the profile likelihood function defined by $L[a, \hat{\delta}(a), \hat{\nu}(a)]$. Here $\hat{\delta}(a)$ and $\hat{\nu}(a)$ indicate the dependence of the parameter estimates on the chosen value for a. The profile likelihood may be maximized by a one-dimensional search over a. The feasible region for the estimate of a is $(\max(x_{ij}), \infty)$. See Cheng and Amin (1981) for a discussion of a related estimation approach.

3.6 Empirical Bayes Inference for Terminal Point Data

One objective of reliability analysis is to estimate the residual life of a surviving item or the lifetime of a new item. In this section, parameter estimation and predictive density functions for the future degradation level and for the lifetime of a surviving item or a new item are developed using Bayes methods. The results are derived for the terminal point data structure but the analysis can also be extended to mixed, conditional, and covariate data structures with no major conceptual alterations. The empirical Bayes implementation of the results is discussed after their presentation.

A normal prior distribution for the drift parameter is appropriate in many applications. In addition, this form of the prior is mathematically tractable because of its conjugacy with a Wiener process. A gamma prior distribution for the variance parameter would also seem to be a natural choice based on its conjugacy with a Wiener process. Thus, a normal distribution for δ and a gamma distribution for ν will be assumed in this section. First, the case where the variance parameter ν is known is considered. Then, a general case with unknown δ and ν is discussed.

3.6.1 Unknown δ and known ν

The data structure of terminal point data is repeated here for convenience.

$$\mathbf{x} = (x_1, x_2, ..., x_q) \qquad\qquad \mathbf{s} = (s_1, s_2, ..., s_p).$$

The joint density function of $\mathbf{x}$ and $\mathbf{s}$, $f(\mathbf{x}, \mathbf{s}|\delta)$, is given in (3.6). The drift parameter δ, as assumed, has a normal prior density function, $\xi(\delta|\theta)$, with mean θ and variance $\omega\nu$, where $\omega > 0$ is some specified multiple.

Bayes Posterior Distribution Given the prior density for δ just assumed, the posterior density of δ is

$$\xi(\delta|\mathbf{x}, \mathbf{s}, \theta) \propto f(\mathbf{x}, \mathbf{s}|\delta)\xi(\delta|\theta)$$

Some algebraic rearrangement gives

$$\xi(\delta|\mathbf{x}, \mathbf{s}, \theta) = \frac{1}{\sqrt{2\pi\nu c_1}} \exp\left(-\frac{(\delta - c_2)^2}{2\nu c_1}\right) \tag{3.36}$$

where

$$c_1 = \frac{\omega}{1 + \omega t_S} \qquad\qquad c_2 = \frac{\theta + \omega d_S}{1 + \omega t_S}$$

and, as defined earlier, $d_S = pa + \sum_1^q x_i$, $t_S = tq + \sum_1^p s_j$. The posterior density is normal, specifically, $\xi(\delta|\mathbf{x}, \mathbf{s}, \theta) \sim N(c_2, c_1\nu)$. The most probable value of δ, based on this posterior density function is c_2.

Estimation and Prediction The Bayes predictive density function of the first passage time of a surviving sample path with the current observation $X(t_n) = x_n$ is a mixture of the normal posterior density of δ in (3.36) and the inverse Gaussian density $h(s|x_n)$ relating to a first passage at time s in the future given the process starts at level x_n (below a) at time t_n.

$$h(s|\mathbf{x}, \mathbf{s}, \theta) = \int_{-\infty}^{+\infty} h(s|x_n, \delta)\xi(\delta|\mathbf{x}, \mathbf{s}, \theta)d\delta$$

$$= \frac{a - x_n}{\sqrt{2\pi\nu(s - t_n)^3[1 + c_1(s - t_n)]}} \exp\left(-\frac{[a - x_n - c_2(s - t_n)]^2}{2\nu(s - t_n)[1 + c_1(s - t_n)]}\right) \qquad s > t_n \tag{3.37}$$

Setting $x_n = t_n = 0$ in (3.37), one obtains the predictive density for the lifetime of a new item.

Note that (3.37) has a similar form to that of an inverse Gaussian density function. In fact, when the accumulated observation time t_S is large or ω is small, then the constant c_1 is small and $h(s|\mathbf{x}, \mathbf{s}, \theta)$ is approximately an inverse Gaussian density function with parameters (c_2, ν). Also note that in the limiting case where $\omega \to \infty$ (reflecting a diffuse prior for δ), the constants c_1 and c_2 take the limiting values $1/t_S$ and d_S/t_S, respectively.

A Bayes predictive interval for the failure time of a surviving item can be obtained accordingly. Let (L_1, L_2) denote the interval and γ the specified coverage of the interval. Then,

$$\gamma = \int_{L_1}^{L_2} h(s|\mathbf{x}, \mathbf{s}, \theta)ds.$$

Correspondingly, the empirical Bayes reliability function for a surviving item is given by the following integral.

$$R(t) = \int_{t}^{+\infty} h(s|\mathbf{x}, \mathbf{s}, \theta)ds \qquad t > t_n$$

When the sample size is large, $R(t)$ is approximated by the complement of an inverse Gaussian distribution function. An explicit form for $R(t)$ may be obtained from results in Whitmore (1986).

Now, consider the Bayes predictive density for a future degradation level. Note that one would not know at the moment of making the prediction whether the item will have failed by a future time point t ($t > t_n$). The quantity of interest, therefore, is the joint probability of surviving to time t and taking a particular degradation level $x < a$ at time t. Let A_{nt} denote the event of no passage in time interval $[t_n, t]$. Then, the predictive density of the future degradation level is the product of the normal posterior (3.36) and the *p.d.f.* of the level $X(t)$ of a truncated Wiener process given in (3.3) integrated with respected to δ $(-\infty < \delta < +\infty)$. The predictive density function for the future sample path level, $X(t) = x, (t > t_n)$, of a surviving item with the current sample path level $X(t_n) = x_n$ is then as follows.

$$f(x, A_{nt}|\mathbf{x}, \mathbf{s}, \theta) = \frac{1}{\sqrt{2\pi\nu\Delta t_n(c_1\Delta t_n + 1)}} \times$$
$$\exp\left(-\frac{(\Delta x_n - c_2\Delta t_n)^2}{2\nu\Delta t_n(c_1\Delta t_n + 1)}\right)\left(1 - \exp\left(\frac{2(a - x_n)(x - a)}{\nu\Delta t_n}\right)\right) \tag{3.38}$$

where $x < a$, $\Delta x_n = x - x_n$, and $\Delta t_n = t - t_n$. For a new item, the desired density is obtained by letting $x_n = t_n = 0$ in (3.38). As discussed earlier, when t_S is large or ω is small, then c_1 is small. As c_1 approaches zero, the density (3.38) approaches the form in (3.3) with the drift parameter replaced by c_2. On the other hand, as $\omega \to \infty$ (reflecting a diffuse prior), $c_1 \to 1/t_S$ and $c_2 \to d_S/t_S$. Then the density (3.38) has a similar form to (3.3) with the drift parameter replaced by d_S/t_S and the variance parameter inflated by a factor of $1 + \Delta t_n/t_S$.

Empirical Bayes Estimation and Prediction For an empirical Bayes implementation of the preceding results, it is assumed that prior parameter θ is unknown while ω is known. The MLE of θ can be obtained from the following likelihood function.

$$L(\theta) = \int_{-\infty}^{+\infty} f(\mathbf{x}, \mathbf{s}|\delta)\xi(\delta|\theta)d\delta$$

It can be shown that the MLE of θ is $\hat{\theta} = d_S/t_S = \hat{\delta}$. The empirical Bayes posterior density of δ therefore is

$$\xi(\delta|\mathbf{x}, \mathbf{s}, \hat{\theta}) = \frac{1}{\sqrt{2\pi\nu c_1}} \exp\left(-\frac{(\delta - \hat{c}_2)^2}{2\nu c_1}\right) \tag{3.39}$$

where

$$\hat{c}_2 = \frac{\hat{\theta} + \omega d_S}{1 + \omega t_S} = \frac{d_S}{t_S}.$$

Inference about the drift parameter δ can be made using the empirical Bayes normal posterior in (3.39). For example, the empirical Bayes point estimator of δ is $E(\delta|\mathbf{x}, \mathbf{s}, \hat{\theta}) = d_S/t_S$, which is the same as the maximum likelihood estimator of δ. Although the empirical Bayes estimator of δ corresponds to the MLE, the empirical Bayes interval estimator of δ will have a width that reflects the amount of prior information about δ as measured by the multiple ω in the prior variance.

Empirical Bayes confidence intervals, such as the one for δ based on (3.39), will often be too short or inappropriately centred because the interval takes no account of the sampling variability of $\hat{\theta}$. Methods have been proposed for adjusting these so-called "naive" confidence intervals. Carlin and Gelfand (1990), for example, propose a methodology for constructing bias-corrected intervals. The bias-corrected interval requires the sampling distribution of $\hat{\theta}$. This sampling distribution, however, cannot be found in a closed form here and the correction is not carried out.

3.6.2 Unknown δ and ν

When both parameters, δ and ν, are unknown, it is convenient to reparameterize by replacing ν with the precision parameter λ, defined as $\lambda = 1/\nu$. The precision parameter is assumed to have a gamma prior density function,

$$\xi_2(\lambda) = \frac{b^d \lambda^{d-1}}{\Gamma(d)} \exp(-b\lambda),$$

where $b > 0$ and $d > 0$. The drift parameter δ is assumed to have a conditional normal distribution, $\xi_1(\delta|\lambda)$, with mean θ and variance ω/λ.

Let $f(\mathbf{x}, \mathbf{s}|\delta, \lambda)$ denote the joint density function of the terminal point data. Then, the posterior joint density function of δ and λ is

$$\xi(\delta, \lambda|\mathbf{x}, \mathbf{s}) = \frac{\xi_1(\delta|\lambda)\xi_2(\lambda)f(\mathbf{x}, \mathbf{s}|\delta, \lambda)}{g(\mathbf{x}, \mathbf{s})}, \tag{3.40}$$

where

$$g(\mathbf{x},\mathbf{s}) = \int_0^\infty \int_{-\infty}^\infty \xi_1(\delta|\lambda)\xi_2(\lambda) f(\mathbf{x},\mathbf{s}|\delta,\lambda)d\delta d\lambda$$

$$= \int_0^\infty F_0 \lambda^{\frac{p+q}{2}+d-1} \exp(-F_1\lambda) \prod_{i=1}^q \left[1 - \exp\left(\frac{2a(x_i - a)}{t}\lambda\right)\right] d\lambda, \tag{3.41}$$

$$F_0 = \frac{b^d a^p}{\Gamma(d)\sqrt{(2\pi)^{(p+q)}(1+\omega t_S)t^q}} \prod_{j=1}^p s_j^{-\frac{3}{2}},$$

and

$$F_1 = b + \sum_{j=1}^{p+q} \frac{d_j^2}{2t_j} + \frac{t_S\theta^2 - 2\theta d_S - d_S^2\omega}{2(1+\omega t_S)}.$$

The integrand in (3.41) can be expanded as a finite linear combination of gamma density functions and, hence, $g(\mathbf{x},\mathbf{s})$ can be expressed in a closed form. It is, however, a complicated expression in the prior parameters θ and ω and obtaining their empirical Bayes estimates will require the use of a numerical computer routine. Evaluating the predictive density functions $h(s|\mathbf{x},\mathbf{s})$ and $f(x, A_{nt}|\mathbf{x},\mathbf{s})$ is also analytically complex and requires numerical methods. Numerical simulation methods, such as Gibb's sampling, might be used as an alternative to numerical computation.

3.7 Summary

This chapter first presents derivations of the density function of a truncated Wiener process for the standard problem and then several variations of the standard problem are examined. Next, statistical inference methods based on terminal point data and mixed data are developed in reliability contexts where both degradation and lifetime data are available. The MLEs of the drift and variance parameters are obtained and found to be asymptotically independent. Inferences based on the asymptotic normality of the MLEs, a modified likelihood ratio method, and an approximate analysis of deviance are presented. Approximate pivotal quantities are found for both the drift and variance parameters. Tests of sample path homogeneity are also derived using the likelihood ratio method. Following these, the problem of estimating an unknown barrier is discussed briefly. Finally, an empirical Bayes analysis is developed for terminal point data. The predictive density functions for the lifetime and the future degradation level of a surviving item or a new item are presented. Two case examples that illustrate the applications of these methods are given in Chapter 6.

CHAPTER 4. INFERENCE FOR A CONDITIONAL RANDOM PROCESS

This chapter presents inference methods for conditional degradation data using the truncated Wiener process as the model for a conditional degradation process. In some applications, one can observe only the degradation process of a surviving item and no information about the failed item(s) is available. For example, consider a device that is already in service, such as an under-sea telecommunication link or a reactor in a chemical plant. The performance levels or degradation data of the device are readily available from routine maintenance procedures. It is, however, virtually infeasible or uneconomical to collect a sample consisting of the degradation data of both failed and surviving devices. In many cases, a failed device is replaced with one that is technologically advanced and hence has different design specifications. Moreover, the current device may be the first one operating in a certain environment or the only one of its kind. In these types of situations, it is desired to derive predictive inferences and optimal maintenance policies based on the degradation data of a device that is in service (i.e., a surviving item).

To discuss the related inference methods, the *conditional degradation data* structure, defined in Chapter 2 is taken as the representative form of the data. The plot in Figure 3.1 illustrates a sample path of a conditional random process that is restricted by a barrier. The conditional degradation data structure assumes that only a sample path that did not exit a barrier can be observed. Statistical inference for conditional data is then based on the observed levels of the process at a set of time points within an observation period, given that the path of the process within the period did not exit the barrier. Inference for the parameters of the conditional degradation process is discussed first. This is then followed by a detailed examination of the impact of conditioning on the parameter estimates.

4.1 Statistical Inference

Likelihood Function Consider a sample path that has not crossed a barrier at

a and measurements on sample path levels that have been taken at a set of time points. Let $\mathbf{x} = (x_1, x_2, ..., x_n)$ be n observations ($n > 2$) on the sample path, at the fixed time points $0 < t_1 < t_2 < ... < t_n$. As in Section 3.3, let A_i denote the event that there is no first passage to barrier a in time interval $[\, t_{i-1}, t_i]$ and let $\bar{A}_i$ denote the complement of A_i. The joint density function of $x_1, x_2, ..., x_n$ and A_{on} is given by equation (3.20). That is,

$$f(x_1, x_2, ..., x_n, A_{0n}) = f(x_1, A_1) f(x_2, A_2|x_1) ...... f(x_n, A_n|x_{n-1}).$$

Since all of the observations are conditioned on the sample path not exiting the barrier in the interval $[0, t_n]$, we have the following likelihood function for the n observations.

$$L(\delta, \nu) = f(x_1, x_2, ... x_n | A_{0n}) = f(x_1, x_2, ... x_n, A_{0n})/P(A_{0n})$$
$$= \frac{1}{P(A_{0n})} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\nu \triangle t_i}} \exp\left(-\frac{(\Delta x_i - \delta \triangle t_i)^2}{2\nu \triangle t_i}\right) \left(1 - \exp \frac{2(a - x_{i-1})(x_i - a)}{\nu \triangle t_i}\right) \tag{4.1}$$

where

$$P(A_{0n}) = \Phi\left(\frac{a - \delta t_n}{\sqrt{\nu t_n}}\right) - \exp(\frac{2\delta a}{\nu}) \Phi\left(-\frac{a + \delta t_n}{\sqrt{\nu t_n}}\right)$$

and $\Delta t_i = t_i - t_{i-1}$, $\Delta x_i = x_i - x_{i-1}$, $\Delta t_1 = t_1$, $\Delta x_1 = x_1$.

Maximum Likelihood Estimators Taking partial derivatives of the log-likelihood function $l = \ln L(\delta, \nu)$ with respect to δ and ν and setting these to zero gives the following equations. These may be solved simultaneously for the maximum likelihood estimators of δ and ν, denoted by $\hat{\delta}$ and $\hat{\nu}$.

$$\sum_{i=1}^{n} \left(\frac{\Delta x_i - \hat{\delta} \triangle t_i}{\hat{\nu}}\right) - \frac{1}{P(A_{0n})} \frac{\partial P(A_{0n})}{\partial \hat{\delta}} = 0,$$

$$\sum_{i=1}^{n} \left(-\frac{1}{2\hat{\nu}} + \frac{(\Delta x_i - \hat{\delta} \triangle t_i)^2}{2\hat{\nu}^2 \triangle t_i} - \frac{K_i(\hat{\nu})}{2\hat{\nu}^2}\right) - \frac{1}{P(A_{0n})} \frac{\partial P(A_{0n})}{\partial \hat{\nu}} = 0,$$

where

$$K_i(\hat{\nu}) = \frac{4(a-x_i)(a-x_{i-1})}{\Delta t_i \left(\exp\left[\frac{2(a-x_i)(a-x_{i-1})}{\hat{\nu}\Delta t_i}\right] - 1\right)}.$$

Furthermore, it can be shown that:

$$\frac{\partial P(A_{0n})}{\partial \hat{\delta}} = -\frac{2a}{\hat{\nu}} \exp(\frac{2a\hat{\delta}}{\hat{\nu}})\Phi\left(-\frac{a+\hat{\delta}t_n}{\sqrt{\hat{\nu}t_n}}\right)$$

$$\frac{\partial P(A_{0n})}{\partial \hat{\nu}} = \frac{2a\hat{\delta}}{\hat{\nu}^2} \exp(\frac{2a\hat{\delta}}{\hat{\nu}})\Phi\left(-\frac{a+\hat{\delta}t_n}{\sqrt{\hat{\nu}t_n}}\right) - \frac{a}{\sqrt{2\pi\hat{\nu}^3 t_n}} \exp\left(-\frac{(a-\hat{\delta}t_n)^2}{2\hat{\nu}t_n}\right).$$

The MLEs of δ and ν are therefore the roots of the following equations

$$\hat{\delta} = \frac{x_n}{t_n} + D_1, \tag{4.2}$$

$$\hat{\nu} = \frac{1}{n}\sum_{i=1}^{n} \frac{(\Delta x_i - \hat{\delta}\,\Delta\, t_i)^2}{\Delta t_i} + D_2, \tag{4.3}$$

where

$$D_1 = \frac{2a}{P(A_{0n})t_n} \exp(\frac{2a\hat{\delta}}{\hat{\nu}})\Phi\left(-\frac{a+\hat{\delta}t_n}{\sqrt{\hat{\nu}t_n}}\right), \tag{4.4}$$

$$D_2 = -\frac{1}{n}\sum_{i=1}^{n} K_i(\hat{\nu}) - \frac{2\hat{\nu}^2}{nP(A_{0n})} \frac{\partial P(A_{0n})}{\partial \hat{\nu}}. \tag{4.5}$$

It has not been confirmed that the roots of (4.2) and (4.3) define a unique global maximum but numerical analysis to date has produced no contrary evidence. Note that neither MLE can be expressed in a closed form.

The nature of the MLEs in (4.2) and (4.3) and the impact on them of the conditioning effect of the barrier are discussed in Section 4.2. First, however, inference results based on the MLEs are presented.

Inference About Parameters Inferences about the drift and variance parameters can be derived, as in Chapter 3, based on the asymptotic normality of the MLEs or using the modified likelihood ratio method.

The asymptotic bivariate normal distribution of $\hat{\delta}$ and $\hat{\nu}$ has a covariance matrix that is estimated by the inverse of the observed information matrix $\mathbf{I}(\hat{\delta}, \hat{\nu})$.

$$\mathbf{I}(\hat{\delta}, \hat{\nu}) = \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix} \tag{4.6}$$

where

$$I_{11} = \frac{t_n}{\hat{\nu}} - \left(\frac{t_n D_1}{\hat{\nu}}\right)^2 + \frac{1}{P(A_{0n})}\frac{\partial^2 P(A_{0n})}{\partial \hat{\delta}^2},$$

$$I_{22} = \frac{n}{2\hat{\nu}^2} + \sum_{i=1}^{n}\left(\frac{K_i(\hat{\nu})}{2\hat{\nu}^2}\right)^2 \exp\left(\frac{2(a-x_i)(a-x_{i-1})}{\hat{\nu}\Delta t_i}\right)$$

$$-\left(\frac{1}{P(A_{0n})}\frac{\partial P(A_{0n})}{\partial \hat{\nu}}\right)^2 + \frac{1}{P(A_{0n})}\frac{\partial^2 P(A_{0n})}{\partial \hat{\nu}^2} + \frac{2}{\hat{\nu}P(A_{0n})}\frac{\partial P(A_{0n})}{\partial \hat{\nu}},$$

and

$$I_{12} = \frac{t_n D_1}{\hat{\nu}^2} - \frac{1}{P(A_{0n})^2}\frac{\partial P(A_{0n})}{\partial \hat{\delta}}\frac{\partial P(A_{0n})}{\partial \hat{\nu}} + \frac{1}{P(A_{0n})}\frac{\partial^2 P(A_{0n})}{\partial \hat{\nu}\partial \hat{\delta}}.$$

The first order derivatives of $P(A_{0n})$ with respect to δ and ν are stated earlier and the second order derivatives can be derived readily. The MLEs, $\hat{\delta}$ and $\hat{\nu}$, cannot be expressed in closed forms and thus the covariance matrix must be evaluated numerically. In general, the MLEs are not asymptotically independent since the expected value of the second order partial derivative, $E(I_{12})$, is not identically zero at $(\delta, \nu) = (\hat{\delta}, \hat{\nu})$. Inferences based on the modified likelihood ratio method can also be developed as in Sections 3.2 and 3.3. No details will be given here, as the structure parallels that presented earlier.

Sampling Distribution of $\hat{\delta}$ Based on the form of the MLE for δ in (4.2), several interesting observations about the sampling distribution of $\hat{\delta}$ can be made.

Taking the value of ν to be given, we can derive the following equation, from the expression for the MLE $\hat{\delta}$ in equation (4.2).

$$x_n = \hat{\delta}t_n - \frac{2a}{P(A_{0n})}\exp\left(\frac{2a\hat{\delta}}{\nu}\right)\Phi\left(-\frac{a+\hat{\delta}t_n}{\sqrt{\nu t_n}}\right) \tag{4.7}$$

Here $P(A_{0n})$ is evaluated using $\hat{\delta}$ so that the right hand side of equation (4.7) is a function of $\hat{\delta}$ only. For the convenience of discussion, this function will be denoted by $\eta(\hat{\delta})$. Intuition suggests that $x_n = \eta(\hat{\delta})$ is a continuous increasing function of $\hat{\delta}$ for which x_n ranges over $(-\infty, a)$ as $\hat{\delta}$ ranges over $(-\infty, \infty)$. A mathematical proof of this result has proved to be elusive but the following exhaustive numerical examination supports this intuition.

To study equation (4.7), a reparameterization is found convenient. Let $\hat{b}_1 = \hat{\delta} t_n / a$ and $b_2 = a^2/\nu t_n$, then, equation (4.7) becomes

$$\frac{x_n}{a} = \hat{b}_1 - \frac{2}{P(A_{0n})} \exp(2\hat{b}_1 b_2) \Phi\left(-(1 + \hat{b}_1)\sqrt{b_2}\right) \tag{4.8}$$

The right hand side of (4.8) is a function of $\hat{b}_1$ and b_2 only. The numerical analysis is based on this equation. When parameters $\hat{b}_1$ and b_2 are both large, numerical computer routines for the normal distribution function Φ in (4.8) becomes unreliable, an approximation to the standard normal distribution function given in Abramowitz and Stegun (1967, p.933) is then used to calculate the tail area of the standard normal distribution. The numerical results are summarized in Figures 4.1 and 4.2.

Figure 4.1 shows a surface plot of x_n/a with respect to $\hat{b}_1$ and b_2. It can be seen that x_n/a is an increasing function of $\hat{b}_1$ for a wide range of b_2 values. The upper right corner of the surface is truncated because of numerical difficulty encountered when the exponential function causes overflow. Figure 4.2 plots x_n/a against $\hat{b}_1$ for $b_2 = 0.2$ and $b_2 = 8$, respectively, as illustrative cases. It is obvious that x_n/a approaches 1 from below as $\hat{b}_1$ increases.

By the technique of changing variables, the *p.d.f.* of $\hat{\delta}$, denoted by $g(\hat{\delta})$, is given by:

$$g(\hat{\delta}) = |J| f[\eta(\hat{\delta}) | A_{0n}] \tag{4.9}$$

where

$$J = \frac{\partial \eta(\hat{\delta})}{\partial \hat{\delta}}$$

and $f(\cdot|A_{0n})$ is the $p.d.f.$ in (3.4). Thus,

$$g(\hat{\delta}) = |J| \frac{1}{\sqrt{2\pi\nu t_n} P(A_{0n})} \exp\left(-\frac{(x_n - \delta t_n)^2}{2\nu t_n}\right)\left(1 - \exp(\frac{2a(x_n - a)}{\nu t_n})\right) \quad (4.10)$$

where

$$x_n = \hat{\delta} t_n - \frac{2a}{P(A_{0n})} \exp(\frac{2a\hat{\delta}}{\nu}) \Phi\left(-\frac{a + \hat{\delta} t_n}{\sqrt{\nu t_n}}\right). \quad (4.11)$$

Relation (4.10) provides a basis for studying distributional properties of $\hat{\delta}$, conditional on ν known. Note $P(A_{0n})$ in (4.11) is evaluated based on $\hat{\delta}$ while the one in (4.10) is evaluated based on δ. In general, the $p.d.f.$ $g(\hat{\delta})$ is unimodal, slightly right skewed and roughly centered on the value of δ. It can be shown that when the barrier $a \to \infty$, $\hat{\delta} \sim N(\delta, \nu/t_n)$, so $g(\hat{\delta})$ is a normal density.

A group of density curves are plotted in Figure 4.3 for different values of barrier level a, set at 5, 10, 20, and 40 respectively. The other parameters are fixed at $\delta = 0.2$, $\nu = 5$, and $t = 10$. The probability mass is roughly centered around the value of the drift parameter $\delta = 0.2$ and slightly right skewed. It can be seen that when the barrier gets tighter, the skewness increases. In other words, the distribution becomes more symmetrical as the barrier level increases and, as noted earlier, the distribution converges to a normal distribution when the barrier vanishes.

4.2 The Effects of a Barrier on MLEs

It is instructive to contrast the MLEs defined by (4.2) and (4.3) with those from sampling an unrestricted Wiener process. The joint density function of $x_1, x_2, ..., x_n$, from a Wiener process $\{X(\tau), \tau \geq 0\}$ with δ and ν, is

$$f(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\nu \triangle t_i}} \exp\left(-\frac{(\Delta x_i - \delta \triangle t_i)^2}{2\nu \triangle t_i}\right)$$

Letting $\hat{\delta}_n$ and $\hat{\nu}_n$ denote the maximum likelihood estimators of δ and ν for this model, it is well known that

$$\hat{\delta}_n = x_n / t_n \qquad\qquad \hat{\nu}_n = \frac{1}{n} \sum_{i=1}^{n} \frac{(\Delta x_i - \hat{\delta}_n \, \Delta \, t_i)^2}{\Delta t_i} \tag{4.12}$$

The estimators of δ and ν for a truncated Wiener process in (4.2) and (4.3) are both different from those of an unrestricted Wiener process in (4.12). The differences are seen by inspection to be the two terms, D_1 and D_2, given in (4.4) and (4.5), i.e.,

$$D_1 = \hat{\delta} - \hat{\delta}_n, \qquad\qquad D_2 = \hat{\nu} - \hat{\nu}_n.$$

Note that because $a > 0$ and $t_n > 0$, D_1 is positive for any value of $\hat{\delta}$ and $\hat{\nu}$. This finding implies that for conditional data one would tend to underestimate δ by using the estimator $\hat{\delta}_n$ when a barrier exists.

As the barrier gets tighter, the effects of the barrier on $\hat{\delta}$ become more significant. In fact, from the expression for D_1 in (4.4), it can be shown that D_1 increases steadily as the barrier level a decreases towards $\max(x_1, x_2, ..., x_n)$. The difference between $\hat{\nu}$ and $\hat{\nu}_n$, denoted by D_2, can be either positive or negative, depending on the particular sample path. It is not a monotonic function of a and simulation results, which are not reported here, show that D_2 tends to be relatively small under a wide range of conditions. Thus, the tightness of the barrier does not seem to have a significant impact on the estimator of ν.

On the other hand, one expects intuitively that, as the barrier moves away from the sample path, i.e., a moves away from $\max(x_1, x_2, , , , , x_n)$, the effects of the barrier will vanish. In fact, it is easy to show that the differences D_1 and D_2 do approach zero in the limit as a approaches ∞.

4.3 Summary

The truncated Wiener process describes a sample path with no first passage. This process is the appropriate model for the conditional degradation data structure. The maximum likelihood estimators of the drift and variance parameters $\hat{\delta}$ and $\hat{\nu}$ are derived for the this data structure. Inference based on asymptotic normality and the likelihood ratio method are discussed. When the barrier is tight, the results of statistical inference are found to be misleading without taking the effects of the barrier into consideration. A study of the MLEs shows that the estimator of δ is significantly affected by the barrier whereas that of ν is not.

The sampling distribution of $\hat{\delta}$ is developed when the variance parameter is given. The distribution converges to a normal distribution as the barrier level approaches infinity. The study of sampling distributions has been limited to the drift parameter while the variance parameter is assumed given. More study is needed to provide a complete analysis of the distributional properties of the MLEs of the drift and variance parameters of a truncated Wiener process.

The development of this chapter is based on a single sample path restricted by a fixed barrier. The results extend to multiple sample paths with no conceptual difficulty. A case example to illustrate the application of this model is given in Chapter 6.

CHAPTER 5. COVARIATE STRUCTURE

In many practical situations, such as the wear-out of a mechanical component or the deterioration of an electronic device, the random process changes in a manner that is correlated with one or more factors, that are called covariates. For example, a wear-out process is affected by temperature, material properties, and so on. While covariates differ from one situation to another, it is important to incorporate these covariates in reliability models. Much literature has been devoted to models for lifetime data with covariates. Lawless and Singhal (1980) give an exponential lifetime model where the response data depend on a vector of regression variables. Whitmore (1983) develops a regression method for censored lifetime data that follow an inverse Gaussian distribution. Some studies have formulated models that relate reliability to covariates. Guttman et al.(1988) derive a lower confidence bound for the reliability function of a classical stress-strength model where the strength and stress are correlated with explanatory variables. The Guttman model approaches the problem from the perspective of a static model assuming that strength and stress are independent and normally distributed. This chapter extends the lifetime data model in Whitmore (1983) to a data structure that comprises both lifetime and degradation data. The model presented here is developed for the terminal point data structure. The extension of this model to the mixed data structure is straightforward.

5.1 A Model for Time Fixed Covariates

Consider an item whose degradation process is assumed to be a Wiener process $\{X(t), t \geq 0\}$ having characteristics determined by k covariates Z_i, $i = 1, 2, ..., k$, which may vary across items but are fixed over time. By time t, this process either has crossed a barrier for the first time at $s < t$ (i.e., has failed) or has not crossed the barrier and terminates at degradation level $X(t) = x$ (i.e., has survived). Assume we have $p + q$ such items, then the resulting data structure, as described in

Chapter 2, consists of the levels of covariates represented by the matrix Z defined in (2.5), degradation data $\mathbf{x} = (x_1, x_2, ..., x_q)$, and lifetime data $\mathbf{s} = (s_1, s_2, ..., s_p)$. Again, p is the number of failed items and q is the number of the surviving items; s_i, $i = 1, 2, ..., p$, are the first passage times of the p failed items; t is the censoring time; x_i, $i = 1, 2, ..., q$, are the observed levels of the surviving items at the censoring time. The following matrix notation is convenient for mathematical derivations.

$$\mathbf{T} = \text{diag}(s_1, s_2, ..., s_p, t, ..., t), \quad \mathbf{X}' = (a, ...a, x_1, x_2, ..., x_q), \quad \mathbf{B}' = (\beta_1, \beta_2, ..., \beta_k)$$

In practice, covariates may influence the degradation process of an item by driving the process parameters in certain directions. They may affect the drift parameter δ, the variance parameter ν, or even the barrier level a. It will be assumed here that the relation between δ and the covariates is the dominate force driving the degradation process and that compared with it, the relations between the other parameters and the covariates are less significant. Of course, this assumption will need to be justified either theoretically or experimentally in particular applications. Here, for the simplicity of theoretical derivation, we consider the relation between δ and the covariates only and assume, in particular, that the drift parameters δ_i ($i = 1, 2, ..., p + q$) of the $p + q$ items are related to the covariates as follows

$$\Delta = \mathbf{ZB} \tag{5.1}$$

where $\Delta = (\delta_1, \delta_2, ..., \delta_{p+q})'$ and $\mathbf{Z}$ is the $(p + q) \times k$ matrix of covariate values.

5.2 Inference for Terminal Point Data with Covariates

Likelihood Function and Maximum Likelihood Estimators Following the pattern of Section 3.2, the likelihood function of the covariate data structure described above is

$$L(\mathbf{B}, \nu) = \prod_{j=1}^{p} \frac{a}{\sqrt{2\pi\nu s_j^3}} \prod_{i=1}^{q} \left(\frac{1}{\sqrt{2\pi\nu t_n}} \left(1 - \exp(\frac{2a(x_i - a)}{\nu t_n})\right)\right) \times$$

$$\exp\left(-\frac{1}{2\nu}(\mathbf{TZB}-\mathbf{X})'\mathbf{T}^{-1}(\mathbf{TZB}-\mathbf{X})\right).$$

The MLEs of parameters $\mathbf{B}$ and ν are the following.

$$\hat{\mathbf{B}} = (\mathbf{Z}'\mathbf{TZ})^{-1}(\mathbf{Z}'\mathbf{X}) \tag{5.2}$$

$$\hat{\nu} = \frac{1}{p+q}\left(\mathbf{X}'\mathbf{T}^{-1}\mathbf{X} - \mathbf{X}'\mathbf{Z}\hat{\mathbf{B}} - \sum_{i=1}^{q} K_i(\hat{\nu})\right) \tag{5.3}$$

where

$$K_i(\hat{\nu}) = \frac{4a(a-x_i)}{[\exp\frac{2a(a-x_i)}{\hat{\nu}t} - 1]t}.$$

It can be shown that the MLEs are the unique maximum of the likelihood function.

Inferences About Parameters Inferences about the drift and variance parameters can be conducted based on the asymptotic normality property of the MLEs, $\hat{\mathbf{B}}$ and $\hat{\nu}$, or the modified likelihood ratio method. Under the usual regularity conditions, $\hat{\mathbf{B}}$ and $\hat{\nu}$ have an asymptotic multivariate normal distribution with the following estimated covariance matrix.

$$Cov(\hat{\mathbf{B}},\hat{\nu}) = \begin{pmatrix} \hat{\nu}(\mathbf{Z}'\mathbf{TZ})^{-1} & 0 \\ 0 & -\left[\frac{\partial^2 l}{\partial \nu^2}|_{\hat{\delta},\hat{\nu}}\right]^{-1} \end{pmatrix} \tag{5.4}$$

Here,

$$\frac{\partial^2 l}{\partial \nu^2}|_{\hat{\delta},\hat{\nu}} = -\frac{p+q}{2\hat{\nu}^2} - \sum_{i=1}^{q}\left(\frac{K_i(\hat{\nu})}{2\hat{\nu}^2}\right)^2 \exp(-\frac{2a(x_i-a)}{\hat{\nu}t})$$

and $\hat{\nu}(\mathbf{Z}'\mathbf{TZ})^{-1}$ is the covariance matrix for $\hat{\mathbf{B}}$ (i.e., $\text{Cov}(\hat{\mathbf{B}}) = \hat{\nu}(\mathbf{Z}'\mathbf{TZ})^{-1}$). It can be seen that $\hat{\mathbf{B}}$ and $\hat{\nu}$ are asymptotically independent.

The quadratic form $(\hat{\mathbf{B}}-\mathbf{B})'\text{Cov}(\hat{\mathbf{B}})^{-1}(\hat{\mathbf{B}}-\mathbf{B})$ follows an approximate chi-square distribution with k degrees of freedom (*df*), based on the asymptotic theory for MLEs. Confidence intervals and tests can be constructed in the usual way. For example, a $1-\alpha$ confidence interval for β_i is $\hat{\beta}_i \pm z(\alpha/2)\hat{\sigma}_{ii}$, where $\hat{\sigma}^2_{ii}$ is the ith diagonal element of the covariance matrix for $\hat{\mathbf{B}}$ in (5.4) and $z(\alpha/2)$ is the $\alpha/2$

fractile of the standard normal distribution. When MLEs of the regression coefficients are normally distributed, the estimated drift parameters $\hat{\delta}_i = \mathbf{Z}_i\hat{\mathbf{B}}$, which are linear functions of the regression coefficients, are also normally distributed with mean $\mathbf{Z}_i E(\hat{\mathbf{B}})$ and variance $\mathbf{Z}_i \mathrm{Cov}(\hat{\mathbf{B}})\mathbf{Z}_i'$, where $\mathbf{Z}_i$ is the ith row of the $\mathbf{Z}$ matrix.

To illustrate inference based on the modified likelihood ratio method, consider hypotheses about the effects of some of the covariates on the parameters of the degradation process. Specifically, suppose we wish to test the null hypothesis

$$H_0 : \mathbf{B}_{(2)} = 0$$

corresponding to the partitions $\mathbf{B}' = [\mathbf{B}_{(1)}'|\mathbf{B}_{(2)}']$ and $\mathbf{Z} = [\mathbf{Z}_{(1)}|\mathbf{Z}_{(2)}]$, where $\mathbf{B}_{(2)}'$ is of dimension $r < k$. The likelihood ratio test statistic is then

$$-2\ln(\Lambda) = \frac{1}{\hat{\nu}}\mathbf{X}'\left(\mathbf{Z}(\mathbf{Z}'\mathbf{T}\mathbf{Z})^{-1}\mathbf{Z}' - \mathbf{Z}_{(1)}(\mathbf{Z}_{(1)}'\mathbf{T}\mathbf{Z}_{(1)})^{-1}\mathbf{Z}_{(1)}'\right)\mathbf{X}. \tag{5.5}$$

For a large sample, this test statistic is well approximated by a chi-square distribution, specifically

$$-2\ln(\Lambda) \sim \chi^2(r) \tag{5.6}$$

where r is the dimension of $\mathbf{B}_{(2)}$. The test statistic in (5.5) is a quadratic form in $\mathbf{X}$. A diagonal element of $\mathbf{T}$ will be either a failure time s with a probability of $1-P(A)$ or the censoring time t with a probability of $P(A)$. Each s_i in the diagonal matrix $\mathbf{T}$ follows an inverse Gaussian distribution that is truncated at time t and each x_i in matrix $\mathbf{X}$ is the level of a truncated Wiener process at time t and has the *p.d.f.* given by equation (3.4).

Analysis of Deviance The results of analysis of deviance presented in Section 3.2 can be extended to the covariate data structure. For the covariate structure, the respective counterparts of the quantity Q, Q_E, and Q_R defined in Section 3.2 are the following.

$$Q = \frac{1}{\nu}(\mathbf{TZB} - \mathbf{X})'\mathbf{T}^{-1}(\mathbf{TZB} - \mathbf{X}),$$
$$Q_R = \frac{1}{\nu}(\mathbf{B} - \hat{\mathbf{B}})'\mathbf{Z}'\mathbf{TZ}(\mathbf{B} - \hat{\mathbf{B}}),$$
$$Q_E = \frac{1}{\nu}(\mathbf{TZ\hat{B}} - \mathbf{X})'\mathbf{T}^{-1}(\mathbf{TZ\hat{B}} - \mathbf{X}) = \frac{1}{\nu}\mathbf{X}'(\mathbf{T}^{-1} - \mathbf{Z}(\mathbf{Z}'\mathbf{TZ})^{-1}\mathbf{Z}')\mathbf{X}.$$

Let Q_{E0} denote the counterpart of Q_E in the reduced model, i.e., under H_0 : $\mathbf{B}_{(2)} = 0$. Then, the test statistic given by equation (5.5) can be expressed as $-2\ln\Lambda = Q_{E0} - Q_E$ where $Q_{E0} - Q_E$ is

$$Q_{E0} - Q_E = \frac{1}{\hat{\nu}}\mathbf{X}'\left(\mathbf{Z}(\mathbf{Z}'\mathbf{TZ})^{-1}\mathbf{Z}' - \mathbf{Z}_{(1)}(\mathbf{Z}'_{(1)}\mathbf{TZ}_{(1)})^{-1}\mathbf{Z}'_{(1)}\right)\mathbf{X}.$$

Recall that this test statistic has an approximate $\chi^2(r)$ distribution. Furthermore, Q_R is approximately $\chi^2(k)$ based on asymptotic theory of MLEs. Following the rationale presented in Chapter 3, the following conjecture is certainly worth further study.

1) Q_R and Q_E are approximately independent,

2) Q is approximately $\chi^2(df)$ $(df > k)$, and

3) Q_E is approximately $\chi^2(df - k)$.

The degrees of freedom df in this conjecture is obtained by a direct analogy with (3.18) as

$$df = m + \sum_{i=1}^{m}\psi_i'(0),$$

where

$$\psi_i'(0) = E[\frac{K(\nu)}{\nu}] = \frac{4a}{\sqrt{\nu t}}\exp(\frac{2a\delta_i}{\nu})L(\frac{a+\delta_i t}{\sqrt{\nu t}}).$$

and

$$\delta_i = \mathbf{Z}_i\mathbf{B}.$$

The pivotal quantities presented in Section 3.2 may also be developed here. Q_E is an approximate $\chi^2(df-k)$ distributed pivotal quantity for ν and $(df-k)Q_R/(kQ_E)$ is an approximate $F(k, df-k)$ distributed pivotal quantity for $\mathbf{B}$. The technical details are omitted as the structure parallels to that of section 3.2.

5.3 Summary

This chapter presents statistical inference methods for the terminal point covariate data structure with covariates that vary among items but are fixed over time. The covariates are assumed to influence degradation processes through the drift parameter only and the variance parameter is unaffected. The drift parameter is assumed to be linearly related to the covariates. The MLEs of the regression coefficients and the variance parameter are derived and are found to be asymptotically independent. Statistical inference about the model parameters based on asymptotic normality and the likelihood ratio method are then presented. Analysis of deviance is also discussed.

This chapter has left a few issues to future research. One is that analytical work is required to prove the conjecture about analysis of deviance for the covariate data structure. Another is related to the assumption given in this chapter about the nature of the influence of covariates on a degradation process. It is certainly possible that covariates influence the variance parameter and barrier level in addition to the drift parameter. Finally, covariate data structures involving time varying covariates remain to be studied.

CHAPTER 6. CASE APPLICATIONS

This chapter presents four case examples to illustrate applications of the statistical methods developed in the preceding chapters for terminal point data, mixed data, conditional data, and covariate data. The first two sections employ simulated data to illustrate the methods for terminal point data with and without covariates. The remaining two sections employ real data to illustrate the methods for conditional and mixed data.

6.1 A Case for Terminal Point Data

To illustrate the analysis for terminal point data, as presented in Chapter 3, a numerical example is developed using simulated data. A set of 100 sample paths is simulated for a Wiener process with δ set at 0.2 and ν at 2. Of the 100 sample paths, 25 crossed a barrier at $a = 20$ before stopping time $t_n = 60$. The terminal point data set includes the failure times of the 25 failed sample paths and the degradation levels of the 75 surviving sample paths at time $t_n = 60$. The data are presented in Table 6.1. The MLEs of δ and ν are obtained from equations (3.8) and (3.9) using a numerical iteration procedure. For these data $d_S = 1045.102$ and $t_S = 5508.616$. The parameter estimates are

$$\hat{\delta} = 0.1897 \qquad \text{and} \qquad \hat{\nu} = 1.9853,$$

which, as expected, are close to the true parameter values specified in the simulation.

Using the inference methods developed in Chapter 3, the following results are obtained for this example.

The sample size of this example, $p+q = 100$, can be considered large. Hence, $\hat{\delta}$ and $\hat{\nu}$ are asymptotically independent and each has an approximate normal distribution with the variance being estimated according to matrix (3.11). Specifically, for this example, the estimated asymptotic variance of $\hat{\delta}$ is

$$Var(\hat{\delta}) = \hat{\nu}/t_S = 3.6040 \times 10^{-4}.$$

Interval estimates for δ can then be calculated for different confidence levels. For example, a 95% confidence interval for the drift parameter is

$$0.1525 \leq \delta \leq 0.2269.$$

A similar asymptotic interval estimate can be developed for the variance parameter. The estimated asymptotic variance of $\hat{\nu}$ is calculated using equation (3.12).

$$Var(\hat{\nu}) = 0.04995.$$

Then a 95% confidence interval for ν is

$$1.5473 \leq \nu \leq 2.4233.$$

The interval estimates for both δ and ν contain the respective true values of δ and ν. As they are asymptotically independent, the two confidence intervals will have a joint confidence level of approximately 90%.

Again, given that the sample size is large, the test statistic (3.14)

$$-2\ln(\Lambda) = t_S \frac{(\delta_0 - \hat{\delta})^2}{\hat{\nu}}$$

has an approximately $\chi^2(1)$ distribution. A likelihood ratio test for an hypothesis such as

$$H_0 : \delta = 0.2$$

can then be conducted based on this distributional property. The test statistic is $-2\ln(\Lambda) = 0.2944$, which corresponds to a p-value of 0.587. The large p-value supports the null hypothesis as would be expected because H_0, in fact, is a true hypothesis here. The test conclusion is also consistent with the fact that the 95% confidence interval spans $\delta = 0.2$.

A test for the same hypothesis can also be conducted based on the pivotal quantity (3.19), which was developed earlier from an analysis of deviance. The pivotal is repeated here for convenience.

$$F^* = (df-1)\frac{Q_R}{Q_E} = \frac{(df-1)(\hat{\delta}-\delta)^2 t_S}{\sum_1^{p+q} \frac{(d_i - \hat{\delta} t_i)^2}{t_i}} \sim \text{approx. } F(1, df-1)$$

The formula for the degrees of freedom df is given in (3.18). For this example, $Q_R = 0.2910$, $Q_E = 102.5807$, $df = 149.6$, $F^* = 0.4215$, and the corresponding p-value is 0.5172. Again, the large p-value supports H_0 as expected.

The pivotal quantity can also be used to construct an alternative interval estimate for δ. Using $F(1, 149.6, 0.95) = 3.904$ and values given earlier for $\hat{\delta}$, $\hat{\nu}$, t_S and Q_E, the 95% confidence interval is

$$0.1585 \le \delta \le 0.2209$$

Note that this interval is slightly tighter than the one calculated using MLE asymptotic normal theory and it includes the true value of δ as it should.

An empirical Bayes confidence interval for δ is also computed from the normal posterior distribution (3.36). It is assumed that $\omega = 3$ here. Then, from the sample information, we have

$$c_1 = \frac{\omega}{1+\omega t_S} = 1.8152 \times 10^{-4} \qquad \hat{c}_2 = \frac{\hat{\theta} + \omega d_S}{1+\omega t_S} = \hat{\delta} = 0.1897.$$

Since $\delta \sim N(c_2, c_1 \nu)$ and c_1 is very small, it can be seen that the posterior density in this example is highly concentrated on the mean value c_2. A 95% empirical Bayes confidence interval for this example is as follows.

$$0.1523 \le \delta \le 0.2271.$$

This interval is slightly wider than the intervals based on MLE asymptotic normal theory and the pivotal quantity from analysis of deviance.

The predictive density function $h(s|\mathbf{x}, \mathbf{s}, \hat{\theta})$ in (3.37) is plotted for this example. In addition to c_1 and c_2, the parameters of the density function have the following values,

$$a = 20, \quad t_n = 60, \quad \nu = 2, \quad q = 75, \quad p = 25, \quad \omega = 3.$$

The *p.d.f.* $h(s|\mathbf{x}, \mathbf{s}, \hat{\theta})$ is approximately inverse Gaussian as c_1 is very small. Figure 6.1 gives a plot of this density function with the above parameter values and $x_n = 0$. It is estimated, by computing the proportion of the area under the density curve, that this item will fail before time $t = 240$ with a probability of approximately 0.85, given that it has survived to time $t_n = 60$ with a degradation level of $x_n = 0$.

6.2 A Case for Terminal Point Data with Covariates

To illustrate applications of the covariate model described in Chapter 5, a sample of simulated terminal point data is analyzed. The simulation is designed to imitate the degradation of an electronic device. Two covariates, temperature (Z_1) and device design (Z_2), are experimentally manipulated. The temperature variable has five levels (0, 25, 50, 75, 100) and, for each temperature level, two designs are considered (denoted by levels 0 and 1, respectively). Thus, the experiment consists of $5 \times 2 = 10$ treatment combinations or experimental blocks. (The simulation is patterned after a real experimental setting that was used in a degradation test described in Section 6.4.)

Eight hundred sample paths are simulated; 80 sample paths in each experimental block. The relation between the drift parameter δ and the covariates is assumed to have the form

$$\delta = 0.1 + 0.008Z_1 + 0.1Z_2. \tag{6.1}$$

Thus, the drift parameter δ is set at 0.1 for the first block and an increment of 0.1 is added for every subsequent block. The drift parameter value δ for each treatment combination is therefore as follows.

		Temperature				
		0	25	50	75	100
Design	0	0.1	0.3	0.5	0.7	0.9
	1	0.2	0.4	0.6	0.8	1.0

The variance parameter ν is held constant to be consistent with the model assumption. The value $\nu = 2$ is chosen. The barrier level is set at 40 and stopping time at 60. All devices having sample paths that cross level $a = 40$ before time $t_n = 60$ are considered to have failed and their failure times s_j are recorded. The remaining devices are considered to have survived and their terminating sample path levels x_i at time $t_n = 60$ are recorded. The sample path levels x_i together with the barrier level $a = 40$ form an 800×1 vector $\mathbf{X}$ and the failure times together with the stopping time $t_n = 60$ form an 800×800 diagonal matrix $\mathbf{T}$. Both matrices are defined in Chapter 5. The 800×3 covariate matrix $\mathbf{Z}$ has the structure of (2.5) and contains the temperature and design levels with every 80 identical rows corresponding to one experimental block. The MLEs of the regression coefficients $\mathbf{B} = (\beta_0, \beta_1, \beta_2)'$ and the variance parameter ν and the estimated asymptotic variance covariance matrix are computed from equations (5.2),(5.3), and (5.4).

$$\hat{\mathbf{B}} = (0.1061682\ 0.0077912\ 0.0977416)'$$

$$\hat{\nu} = 1.82662$$

$$Cov(\hat{\mathbf{B}}, \nu) = \begin{pmatrix} 0.0001605 & -1.664 \times 10^{-6} & 0.000085 & 0 \\ -1.664 \times 10^{-6} & 3.5656 \times 10^{-8} & 5.1639 \times 10^{-8} & 0 \\ 0.000085 & 5.1639 \times 10^{-8} & 0.0001689 & 0 \\ 0 & 0 & 0 & 0.0073935 \end{pmatrix}$$

As expected, the MLEs of both the regression coefficients and the variance parameter are close to the true values specified in the simulation. The estimated linear relation between the drift δ and the covariates is

$$\hat{\delta} = 0.1061682 + 0.0077912Z_1 + 0.0977416Z_2.$$

Then the estimated drift parameters associated with the different combinations of covariate levels are

$$\begin{aligned} \hat{\Delta} &= (\hat{\delta}_1\ \hat{\delta}_2\ ...\ \hat{\delta}_{10})' \\ &= (0.1062\ 0.2039\ 0.3009\ 0.3987\ 0.4957\ 0.5935\ 0.6905\ 0.7882\ 0.8853\ 0.9830)'. \end{aligned}$$

Based on asymptotic normal theory, each of the $\hat{\delta}_i, i = 1, 2, ..., 10$, has an approximate normal distribution $N(\mathbf{Z}_i\mathbf{B},\ \mathbf{Z}_i Cov(\hat{\mathbf{B}})\mathbf{Z}_i')$, where $\mathbf{Z}_i$ is a row vector containing covariate levels for the ith block. From this distributional property, a set of 10 simultaneous confidence intervals for the drift parameters of the 10 experimental levels is computed.

	δ_1	δ_2	δ_3	δ_4	δ_5
Lower Bound	0.0735	0.1714	0.2752	0.3729	0.4722
Upper Bound	0.1388	0.2364	0.3267	0.4245	0.5192

	δ_6	δ_7	δ_8	δ_9	δ_{10}
Lower Bound	0.5695	0.6633	0.7603	0.8503	0.9472
Upper Bound	0.6174	0.7177	0.8162	0.9202	1.0188

The joint confidence coefficient is approximately 0.95 while each individual confidence interval is set at 0.995 level.

A test of the hypothesis $H_0 : \vec{\mathbf{B}} = (0.1\ 0.008\ 0.1)'$ is conducted. The asymptotic normal theory method and the likelihood ratio method developed in Chapter 5 give equivalent results. Based on the asymptotic normal theory, the following test statistic χ^* has an approximately $\chi^2(3)$ distribution under H_0.

$$\chi^* = (\hat{\mathbf{B}} - \mathbf{B})'[Cov(\hat{\mathbf{B}})]^{-1}(\hat{\mathbf{B}} - \mathbf{B}) = 1.7339. \tag{6.2}$$

The corresponding p-value is 0.182, as expected, H_0 cannot be rejected.

6.3 A Case for Conditional Data

In many economics and finance applications of stochastic processes, it is reasonable to assume that the time sequences under study, whether they are interest rates, stock prices, or the like, are restricted by a barrier or barriers. Applying a conventional stochastic process whose state space is defined on the whole real line to a restricted time sequence would not be appropriate and would give misleading results. For example, a stock's price, whether it increases or decreases in a time period, must remain above zero if a firm remains solvent. Much literature in finance have been devoted to searching for adjustments so that a Brownian motion process can capture the dynamics of stock price, since the Brownian motion process was first studied and applied as a stock price model by Bachelier in 1900 and the later adaptation of the process to changes of stock price (see, Cootner 1964 and Samuelson 1973). Because of the limited liability of stocks, literature in finance has adapted Brownian motion to describe the return of a stock by assuming that the logarithm of a stock price is normally distributed and follows a Brownian motion. The truncated Wiener process presented in Chapter 3 provides an alternative process that has properties that correctly model the limited liability of stocks. Instead of adapting the Wiener process to the logarithm of a stock price, the truncated Wiener process can be applied directly to a stock price that is restricted by a barrier placed at zero.

This section presents a finance application of the conditional data model, which was developed based on a truncated Wiener process. The application concerns a sequence of 100 daily returns of a randomly selected stock, which was listed on the New York Stock Exchange. Prices are calculated using the following equation: $y_i = (r_i+1)y_{i-1}$, where r_i and y_i represent return and price in period i, respectively. The initial price has been arbitrarily scaled to $y_0 = 5$. The price barrier level is set at zero. The stock price data are listed in Table 6.2 in time order.

The price sequence is transformed to correspond to the standard problem presented in Chapter 4. As the initial sample path level is $y_0 = 5$ and the barrier is below the initial level of the process, the transformation $X(t) = 5 - Y(t)$ is used, where $Y(t)$ denotes the original process. Thus, the transformed process $X(t)$ with $X(0) = 0$ and $a = 5$ corresponds to the standard problem described earlier.

The drift and variance parameters are estimated from equations (4.2) and (4.3) for the transformed sample path and then are converted to correspond to the original series $Y(t)$. The drift parameter $\hat{\delta}$ and the variance parameter $\hat{\nu}$ of the original series $Y(t)$ are as follows.

$$\hat{\delta} = -0.00196 \qquad\qquad \hat{\nu} = 0.02895.$$

The corresponding parameter estimates when the barrier is ignored can be calculated from (4.12) and are found to be

$$\hat{\delta}_n = -0.00090 \qquad\qquad \hat{\nu}_n = 0.02892.$$

Note that the difference between the values of the drift parameter $\hat{\delta}_n$ and $\hat{\delta}$ is significant whereas the variance parameter estimates are very close as anticipated by the theoretical analysis in Chapter 4. The rate of price change is underestimated when the barrier is ignored.

Inference for the drift and variance parameters can be derived from the asymptotic normal theory for MLEs and the likelihood ratio method. First, the results based on asymptotic normality of the MLEs are presented.

The covariance matrix of $\hat{\delta}$ and $\hat{\nu}$, which is the inverse of the information matrix (4.6), is estimated for this example.

$$Cov(\hat{\delta}, \hat{\nu}) = \begin{pmatrix} 0.2446 \times 10^{-3} & 0.5285 \times 10^{-4} \\ 0.5285 \times 10^{-4} & 0.2083 \times 10^{-4} \end{pmatrix}$$

Then setting the confidence level at $1 - \alpha = 0.95$, the individual confidence intervals for the drift and variance parameters are

$$-0.0326 \leq \delta \leq 0.0287,$$

$$0.0200 \leq \nu \leq 0.03790.$$

As noted in Chapter 4, the two MLEs are not asymptotically independent. Based on the asymptotic normality of the MLEs, the quantity $(\hat{\delta}-\delta \;\; \hat{\nu}-\nu)' Cov(\hat{\delta}, \hat{\nu})^{-1} (\hat{\delta}-\delta \;\; \hat{\nu}-\nu)$ has an approximate $\chi^2(2)$ distribution. A 95% confidence region for δ and ν is plotted in Figure 6.2.

From above results, one can see that the stock was not performing well. In fact, it decreased at an average rate of \$0.00196 per trading day over the time period observed.

6.4 A Case for Mixed Data with Covariates

The model for a mixed data structure is applied in this section to a problem related to the degradation of electronic transistors. An electronic transistor degrades and finally fails when its *gain*, a key performance measure of a transistor, falls to a level that makes it nonfunctional in the device where it is placed. The gain levels can be recorded at a set of time points and times to failure can also be observed in the course of laboratory experiments. Inferences about the reliability properties of the transistor are desired based on the set of degradation and life data collected in the experiment. The transistor degradation data used in this case application were made available through Dr. T. C. Denton, senior principal research engineer of BNR Europe Ltd., England. The data and application have been slightly disguised to protect their proprietary nature.

The transistor degradation data are listed in Table 6.3. Each transistor is referred to by an item number. The times shown are in hours. Ambient temperature and current were the experimental conditions. It can be seen that the degradation processes for gain are bidirectional. The transistor gain increases and decreases over time but tends to decline (degrade). This observation supports the remark made earlier about the plausibility of modeling degradation processes by a Wiener process. The following is some background information about the data set.

The experiment was conducted by varying both the ambient temperature and electric current levels applied to the transistors. A set of 20 transistors was divided randomly into five equal groups and each group was tested at a different temperature level. Each group is further divided randomly into two blocks consisting of two transistors each. At every temperature level, one block was tested at current level 1 and the other at current level 2. In this way, the 20 transistors were divided into 10 experimental blocks and the experiment was controlled at five different temperature levels and two electric current levels. The degradation processes were believed to be affected by both temperature and current levels applied. The temperature and current will be referred to as covariates.

The failure threshold for the gain of a transistor tends to vary slightly from one application to another. The barrier is arbitrarily set at 70 here, just to illustrate the application. That is, when gain falls to 70, a transistor is assumed to fail. Three of the 20 transistors failed at different times. The censoring times of the 17 surviving items are not equal while the standard model developed in Chapter 3 assumes a common censoring time t_n. Also, the initial gains vary from one transistor to another which is another departure from the standard model. The following transformations are used to convert the original data to the standard problem described by the model.

$$x_i = x_i^o(0) - x_i^o$$

$$a_i = x_i^o(0) - a^o$$

$$\delta_i = -\delta_i^o$$

$$t_i = t_i^o/1000$$

The notation on the right-hand side is for original data and $x_i^o(0)$ is the initial gain level of item i. The last transformation gives time measurements t_i in thousands of hours. Subsequent analyses are based on the transformed data unless pointed out otherwise.

Because of random variations in the manufacturing process, material quality, and other factors, transistors that have the same design specifications may have rather different performance levels and lifetimes even when they operate under the same conditions. In other words, these differences are contributed by factors other than the operating conditions or experimental conditions. These uncontrolled internal factors can interact with the experimental conditions and disguise the true effects of the covariates on the transistors' degradation processes. To rule out the possibility that the observed variations between transistors are caused by these uncontrolled factors, a statistical test of sample path homogeneity is necessary to check if transistors that operated under the same experimental conditions have statistically significant differences in their degradation processes. When sample path homogeneity has been established, one can analyze the effects of the covariates on the degradation processes. Regression analysis may be applied to reveal the statistical relations between the degradation process parameters and the covariates.

Sample Path Homogeneity Test Tests of sample path homogeneity for both the variance parameter and the drift parameter are conducted. For each experimental block, a test for a common variance parameter is performed first. Then, this is followed by a test for a common drift parameter.

The hypothesis for a common variance parameter is,

$$H_0 : \nu_i = \nu_j$$

where $(i, j) = (1,3), (2,4), ..., (18,20)$ denote the item numbers of the transistors in each experimental block. For this example, the test statistic (3.30) developed in Chapter 3 is approximately χ^2 distributed with $df = 1$. The test statistic is reproduced here for convenience.

$$-2\ln(\Lambda) = \sum_{S,F} \ln\frac{\hat{\nu}}{\hat{\nu}_i} + \sum_{S,F}(\frac{1}{\hat{\nu}} - \frac{1}{\hat{\nu}_i})\frac{(\Delta d - \hat{\delta}_i \bigtriangleup t)^2}{\Delta t} - 2\sum_{i=1}^{2}\sum_{j=1}^{n_i} \ln\frac{K_{ij}(-\hat{\nu}_i)}{K_{ij}(-\hat{\nu})}$$

The MLEs of the drift and variance parameters, $\hat{\delta}_i$ and $\hat{\nu}_i$, are based on each individual transistor's observed gains and calculated using equations (3.32) and (3.33). The common variance parameter estimates of each block are computed using equation (3.31). Columns 4, 6, and 7 of Table 6.4 list the above estimates. The test statistics are then calculated and listed in column 8 of Table 6.4. For each individual test, the α risk is controlled at 0.01 so that the ten separate tests have a joint α risk of about 0.1. Thus, H_0 is expected to be rejected in about 10% of the tests. In fact, H_0 is rejected in 1 of the 10 blocks based on $\chi^2(1; 0.99) = 6.63$. Block 3 (containing items 5 and 7) does not appear to have a common variance parameter. This can also be seen from the original data listed in Table 6.3. While transistor 5 has a large variation in gain level, the level of transistor 7 is relatively stable. More discussions about this block will follow later in this section.

Tests for a common drift parameter are conducted next. The hypothesis for a common drift parameter is

$$H_0 : \delta_i = \delta_j,$$

where $(i, j) = (1, 3), (2, 4), ..., (18, 20)$. Given that the two transistors in a block have a common variance parameter, the test statistic is (3.34), which is repeated here for convenience.

$$-2\ln(\Lambda) = \sum_{i=1}^{p+q} \sum_{S_i F_i} \Delta t \frac{(\hat{\delta} - \hat{\delta}_i)^2}{\hat{\nu}}$$

In this application, $p + q = 2$ so this statistic is approximately $\chi^2(1)$ distributed. The common drift parameter, $\hat{\delta}$, is computed based on equation (3.25), while $\hat{\delta}_i$, $i = 1, 2, ..., 20$, and $\hat{\nu}$ were estimated earlier using equations (3.32) and (3.31), respectively. Columns 5 and 9 of Table 6.4 list the $\hat{\delta}$ and the corresponding test statistics respectively. For all blocks except block 3, H_0 cannot be rejected.

For block 3, $H_0 : \nu_5 = \nu_7$ is rejected and the test for $H_0 : \delta_5 = \delta_7$ based on a common variance parameter is listed as inconclusive. To check if the drift parameters are affected by the uncontrolled factors, a test for $H_0 : \delta_5 = \delta_7$ without assuming a common variance parameter is then conducted. The test statistic is $-2\ln(\Lambda) = 0.01300$, calculated using (3.35). With reference to $\chi^2(1)$ distribution, this value is small and hence $H_0 : \delta_5 = \delta_7$ cannot be rejected. The remaining issue of the unequal variance parameters in block 3 will be addressed in the regression analysis that follows shortly.

To summarize, the effects of potential variations in the manufacturing process and other uncontrolled internal factors are statistically insignificant and, therefore, the changes in the estimated values of the drift and variance parameters are caused by the variations in temperature and current levels, except possibly for the variance parameters in block 3. Next, a regression analysis is carried out to establish the nature of the statistical relation between the experimental covariates and the parameters of the transistor degradation processes.

Regression Analysis In the covariate model developed in Chapter 5, the variance parameter is assumed to be comn n across blocks and it is also assumed that the covariates are linearly related to the drift parameter. Neither assumption is valid in this transistor case example. A modification to the covariate model, therefore, is necessary for the transistor data.

The individual estimates of $\hat{\delta}$ and $\hat{\nu}$ are plotted against the covariates levels in Figures 6.3 and 6.4. In the plots, circles denote the estimates for current level 1 and triangles for current level 2. The fitted curves are the regression functions developed later for this example. From the plots, it is clear that both sets of MLEs exhibit curvilinear relations with respect to the temperature level and that the current level has little effect on the estimates, especially, the drift parameter estimates. In addition, one can see that variation among the MLEs of ν at each temperature level appears to be considerable. As a consequence of this large variation, it becomes necessary to verify the assumption of a constant variance required by the regression analysis method. These two problems are addressed in the following discussions.

To accommodate the curvilinear relationships, transformations of the MLEs will be necessary. Knowledge of the engineering subject matter for this case example suggests that there should be an exponential relation between the parameters and the covariates. Exponential relations of the following general forms are plausible:

$$\delta = \beta_0 + \exp(\beta_1 + \beta_2 K^{r_1} + \beta_3 L K^{r_1}) \quad . \tag{6.3}$$

$$\nu = \alpha_0 + \exp(\alpha_1 + \alpha_2 K^{r_2} + \alpha_3 L K^{r_2}) \tag{6.4}$$

where K denotes the Kelvin temperature level measured in units of $100^o K$, L is a 0-1 indicator variable for the current level ($L = 0$, for current level 2), and β_0

and α_0 are shift parameters. Subject knowledge further suggests that the drift parameter should be zero when the temperature, measured on the Kelvin scale, reaches zero, which implies that the shift parameter $\beta_0 = -1$ and, also, that $\beta_1 = 0$. Unfortunately, the subject knowledge does not provide much guidance for the variance parameter. Considering the illustrative nature of this case example, a parallel structure is assumed here for the variance parameter. Thus, $\alpha_0 = -1$ and $\alpha_1 = 0$.

Further, letting $\delta' = \ln(\delta + 1)$, $\nu' = \ln(\nu + 1)$, we have

$$\delta' = \beta_2 K^{r_1} + \beta_3 L K^{r_1} \tag{6.5}$$

$$\nu' = \alpha_2 K^{r_2} + \alpha_3 L K^{r_2} \tag{6.6}$$

Equations (6.5) and (6.6) are then the regression functions to be estimated based on the MLEs calculated for each combination of temperature and current levels.

Recall that the variations among the MLEs across temperature levels are relatively large, especially for the MLEs of the variance parameter. To check the possibility of violating the regression assumption of a constant error variance, a Hartley test for the equality of variance is conducted. Since the objective is to verify the assumption of the regression analysis, the test is conducted for the transformed parameters, namely $\hat{\delta}'$ and $\hat{\nu}'$.

Letting σ_i^2 denote the variance of the $\hat{\delta}'$ for $i = 1, 2, 3, 4, 5$, corresponding to the five temperature levels, the hypothesis takes the following form.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_5^2$$

The test statistic under H_0 is

$$H = \frac{\max\{\hat{\sigma}_i^2, i = 1, 2, \ldots, 5\}}{\min\{\hat{\sigma}_i^2, i = 1, 2, \ldots, 5\}} \sim H(1 - \alpha; \gamma, df)$$

For this example, $\gamma = 5$, corresponding to the five temperature levels, and $df = 3$. The level of significance is to be controlled at 0.01, so $H(0.99; 5, 3) = 151$. The maximum variance of $\hat{\delta}'$ is 0.178 and the minimum variance is 0.042. These occurred at the fourth temperature level and at the second temperature level, respectively. The test statistic is $H = 17.96$, which is less than the critical value of 151. That is, hypothesis H_0 cannot be rejected. A similar test is conducted for $\hat{\nu}'$. With the maximum variance of 0.712 occurring at the second temperature level and the minimum variance of 0.071 at the first temperature level, the test statistic is 100.56. Again, H_0 cannot be rejected. Having verified the assumption of a constant variance, we proceed to estimate the regression functions (6.5) and (6.6).

As the MLEs of δ and ν are asymptotically independent, the regression models (6.5) and (6.6) are estimated individually. The estimates of the exponent parameters, r_1 and r_2, are chosen so that the linearity of the respective regression functions is maximized. Using a numerical iteration method, they are found to be $r_1 = 11.695$ and $r_2 = 7.590$. To simplify the notation, let K_1 and K_2 denote K^{r_1} and K^{r_2}, respectively. The estimated regression functions are as follows.

$$\hat{\delta}' = 0.878 \times 10^{-6} K_1 + 0.039 L K_1 \tag{6.7}$$

$$\hat{\nu}' = 0.141 \times 10^{-3} K_2 + 0.393 L K_2 \tag{6.8}$$

Figures 6.5 and 6.6 show the fitted regression functions.

Recall that the sample path homogeneity test for the variance parameter reveals that items in block 3 do not share a common variance parameter. To determine whether this block has any significant effects on the regression function (6.8), a Cook's distance measure, D, is computed. Notice that, within block 3, item 5 is the one that has a larger departure while the other item (item 7) falls relatively in line with other blocks. Thus, the test is only performed for item 5. Specifically, for item 5, $D = 0.0317$ which corresponds to the 3rd percentile of an F(2, 18) distribution.

This result indicates that block 3 does not have significant effects on the regression function. Therefore, the block will not be removed from the regression analysis.

Next, t-tests of the hypothesis $H_0 : \beta_3 = 0$ and $H_0 : \alpha_3 = 0$ give p-values of 0.165 and 0.246, respectively. These results indicate that variables LK_1 and LK_2 have no statistically significant effects on $\hat{\delta}'$ and $\hat{\nu}'$ and, hence, that the two current levels (represented by variable L) have no differential effect on degradation. After removing the variables, the regression functions become,

$$\hat{\delta}' = 0.916K_1$$

$$\hat{\nu}' = 0.152K_2.$$

Expressing the above regression functions in the form of (6.3) and (6.4), we have the following functions describing the relation between the gain degradation process and the ambient temperature level K measured in units of $100^o K$.

$$\delta^o = 1 - \exp(0.916 \times 10^{-6} \times K^{11.695})$$

$$\nu = \exp(0.152 \times 10^{-3} \times K^{7.590}) - 1$$

Here the drift parameter δ^o is the gain differential per 1000 hours (note, $\delta^o = -\delta$). Similarly, the variance parameter ν is a measurement of gain level variation per 1000 hours.

The disguised nature of the data makes it difficult to give a precise interpretation of the results. Yet, one can see that a higher level of temperature increases the degradation rate and variation of transistor gain, as expected. What is interesting here is that only the temperature level has significant effects on the degradation rate and variation, hence, on the lifetime of a transistor. The effects of the current level are insignificant.

Before closing this section, three observations should be noted. First, the overall pattern of the estimates of the variance parameter appears to be disturbed by the estimates of block 3. It appears that block 3 is an outlier, although it does not have significant effects on the relation between the gain levels and the experimental conditions. This outlier may represent an unlikely event as there were 10 hypotheses tested simultaneously. This outlier, on the other hand, may represent an important factor that has not been captured by the regression function, as the results of the sample path homogeneity test have indicated. However, a firm conclusion regarding this issue can be drawn only when a thorough investigation based on the engineering subject knowledge has been conducted.

Second, a close study of the MLEs for δ and the covariates reveals that, at higher temperature levels, the drift parameter is sharply higher than at lower temperature levels. From Figures 6.3 and 6.5, one can see that the estimates of δ at the three lower temperature levels appear to be in one cluster while those at the two higher temperature levels appear in another cluster. A step regression function might provide a better fit to the data. It is possible that a different degradation mechanism becomes a dominating force when the temperature level increases above 75^oC. To confirm or to reject this hypothesis, a larger sample that provides more measures at the intermediate temperature levels will be necessary.

Third, it has been found that the degradation rates of several transistors decrease in a short initial period of time and then remain relatively stable over time. It is conceivable that there exists one or more time varying covariates that cause the degradation rate to change over the initial period of time. While the temperature and current levels are fixed over time for a group of transistors, one or more unobserved covariates, which may be related to the underlining physical processes inside a transistor, may vary over time. A time scale transformation or inclusion of time in the regression model should be considered if subject knowledge supports this conclusion.

6.5 Summary

This chapter presents four case examples for the models and related inference methods developed earlier. The main objective of the case examples is to illustrate the applications of the models and methods. The fit of the models to the data has not been tested for the cases reported in Sections 6.3 and 6.4. Sections 6.1 and 6.2 present simulation studies for the terminal model and the covariate model so goodness of fit is not an issue.

In Section 6.1, the terminal data model developed in Chapter 3 is illustrated using a simulated example. First, point estimates for the drift and variance parameters are calculated. Next, confidence intervals for the parameters are constructed based on the asymptotic normality of the MLEs and the modified likelihood ratio method. Then, tests based on asymptotic normal theory, the modified likelihood ratio method, and the pivotal quantity derived from the analysis of deviance are also conducted and the test results are consistent with the interval estimates. Finally, limited predictive inferences based on the empirical Bayes results derived in Chapter 3 are presented.

A simulated example for the covariate model is given in Section 6.2. The regression coefficients are estimated and confidence intervals for the drift and variance parameters at each covariate level are computed. Also, a test for the regression function coefficients is conducted.

The conditional model is applied to a sequence of stock price data and the analysis is presented in Section 6.3. MLEs for the drift and variance parameters are estimated and a joint confidence region for the parameters is plotted.

Finally, a real case example of transistor degradation processes is analyzed in Section 6.4. First, sample path homogeneity tests are conducted for the drift and variance parameters to determine if there are any factors other than the two known covariates (temperature and current) affecting the gain levels observed. Then, a regression analysis is carried out to establish a statistical relation between the covariates and the parameters of the transistor degradation process. It is found that the experimental ambient temperature is the major factor affecting the degradation processes of the transistors.

CHAPTER 7. CONCLUSION AND DISCUSSION

Traditional reliability models are based on lifetime data alone. When lifetime data are not available, which is often the case with highly reliable items, expensive items, and items for which accelerated life testing is not feasible, reliability models based on degradation data become an important and, perhaps, the only approach to reliability analysis. When both lifetime and degradation data are available, the degradation data contain important information about the future behavior of surviving items and new items. In both cases, evaluating reliability based on degradation or the combination of degradation and lifetime data has practical importance.

Main Contributions This research has two main contributions. One is the application of a truncated Wiener process to modeling the degradation process of a surviving item. The other is the development of reliability models and inference methods based on a combination of degradation and lifetime data with and without covariates.

Particularly, the following subjects are studied in the thesis. Chapter 2 first presents a brief review of degradation modeling in engineering, which is found to be mainly concentrated on material fatigue analysis and the parameter drift problem of electronic devices. This is then followed by a discussion of several reliability data structures. The structures include terminal point data (a combination of degradation and lifetime data), mixed data (a general case of terminal point data), conditional degradation data, and covariate data.

In Chapter 3, the derivations of the probability density function for a truncated Wiener process are presented first. The density function is derived for a standard problem that is defined as a Wiener process with a positive, fixed and known barrier. Several variations of the standard problem are examined and simple transformations are given to convert some of them to the standard problem. Next, taking the inverse Gaussian distribution as the lifetime distribution for failed items, the likelihood function for a sample of terminal point data is formulated. Maximum likelihood estimators (MLEs) are obtained for the drift and variance parameters. Inference results for the parameters are derived using asymptotic normal theory and a modified likelihood ratio method. In addition, an approximate analysis of deviance is developed, which yields pivotal quantities for the parameters. Then the inference is extended to a mixed data structure. The problem of sample path homogeneity is discussed and likelihood ratio tests for sample path homogeneity are developed. Using an empirical Bayes procedure, predictive density functions for the lifetime and the future degradation level of either a surviving item or a new item are obtained and estimations based on the posterior density function are also presented.

In Chapters 4 and 5, inference results for conditional and covariate data structures are developed. MLEs are obtained for both data structures. Estimates for parameters of each data structure are derived based on asymptotic normality and the modified likelihood ratio method. For the conditional data structure, the sampling distribution for the MLE of the drift parameter with known variance parameter is studied. The density function of the sampling distribution of the drift parameter is found to be slightly skewed to the right with skewness decreasing as the barrier level approaches infinity. An illustrative set of density curves are numerically calculated and plotted for different values of the barrier level. For the covariate data structure, an analysis of deviance is presented and pivotal quantities, similar to those for terminal point data, are also given.

Finally, in Chapter 6, applications of the models are illustrated by case examples using real data and simulated data. Two real data examples and two simulation studies are presented. The major case application involves experimental data on transistor degradation. In this application, transistors of a given design are observed to degrade under experimental controls on the ambient temperature and applied current level (the covariates). The results of the sample path homogeneity tests show that, excluding the effects of the experimental covariates, the transistors' degradation patterns can be considered identical. In other words, the transistors' degradation processes and lifetimes would not be statistically different from one transistor to another, if they were tested under the same experimental conditions. It is also found that degradation is controlled only by the level of ambient temperature. Time to failure largely depends on the temperature level and the initial gain level. The current level has no statistically significant effects on either the drift parameter or the variance parameter of the degradation process. Specifically, the drift and variance parameters are exponential functions of temperature while the relations between the parameters and the current level are not statistically significant. Another real data example is developed for the conditional degradation process or truncated Wiener process. In this example, a sequence of daily stock prices was analyzed. The results show that the stock had a negative drift parameter (the rate of price change) during the time period observed. The drift parameter is underestimated if the barrier at zero is ignored. Two simulation studies are conducted for terminal point data and covariate data. Inference results based on the methods derived earlier in the thesis are found to be consistent with the true parameter values prescribed in the simulations. Point and interval estimations, hypothesis tests, and predictive inferences are demonstrated in the two simulated examples.

Some Open Research Questions This research has only touched the surface of many deeper questions. A number of issues have been left to be dealt with in subsequent research. The followings are some of them.

It is assumed in this thesis that a global degradation variable exists and is measurable. It is well known that the degradation of a physical item is caused by the progress of the underlying physical and/or chemical processes within the item and is affected by external environmental factors which accelerate or delay the progress of these processes. In many applications, these processes can be represented only approximately by a single variable. Thus, extending the models and inference methods into a multivariate framework would be a good next step. Moreover, since the underlining processes are determined by the properties of an item and, in many applications, subject matter knowledge about these properties is incomplete, an understanding of the processes can only be obtained from empirical data, either experimental data or field data. Thus, a model or a computer system containing the model that can interact with experimental scientists or field experts and is capable of learning from large sets of data would be a practical tool in many industrial and business applications.

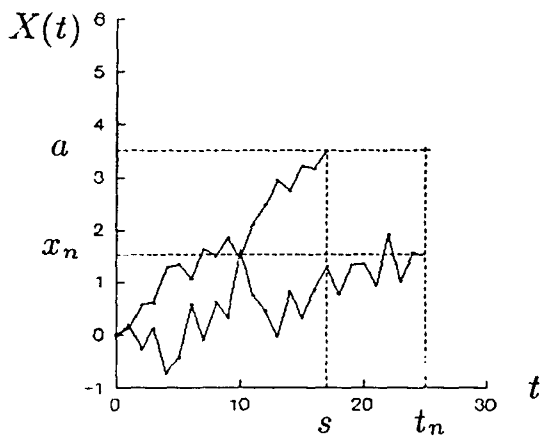
In this thesis, several variations of the standard problem described in Chapter 3 have been discussed and some still need to be studied. One variation is to study a degradation process drifting toward a random barrier. In many applications, the exact location of a barrier level at which a device fails to function properly is unknown or actually is a random variable itself. Another variation requiring further investigation is a degradation process accompanied by one or more time-varying random covariate processes.

A number of mathematical problems require further research. It has been observed that the modified likelihood ratio method produces the same test statistics as the asymptotic normal theory, for the tests studied in Chapters 3 and 5, while the conventional likelihood method produces test statistics that would be different from that based on asymptotic normal theory. This suggests that the modified likelihood ratio method is more appropriate than the conventional likelihood ratio method, at least for the tests studied in the two chapters. More studies are needed to investigate this issue. Another issue is the analytical structure of the analysis of deviance and the theoretical verification of the conjecture about this structure that was given in Chapter 3.

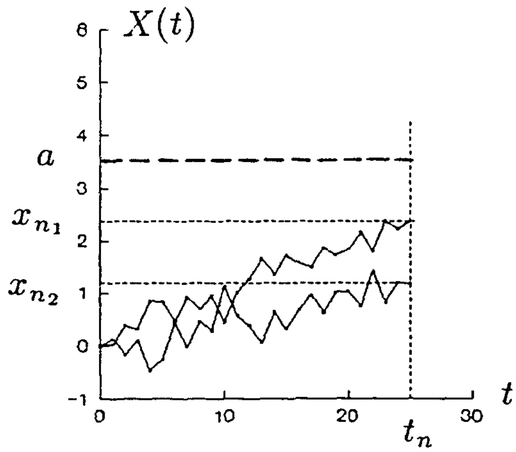
The application context assumed in this thesis is mainly industrial engineering. The methodology, however, may be applied to a number of economic and business settings with few modifications, as illustrated by the stock price application in Chapter 6. The methodology may also be applied to clinical trial data and data from other medical applications.

FIGURES AND TABLES

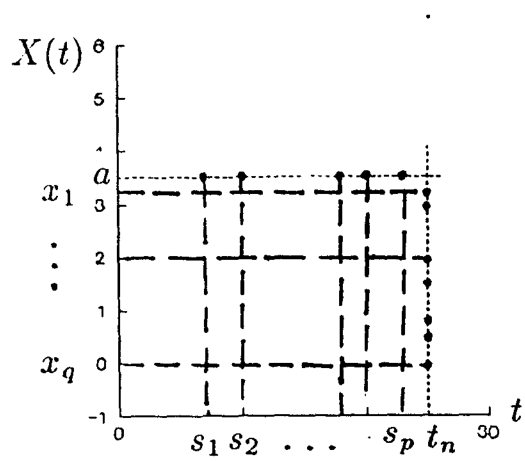
Figure 2.1 Illustrations of Types of Data Structures



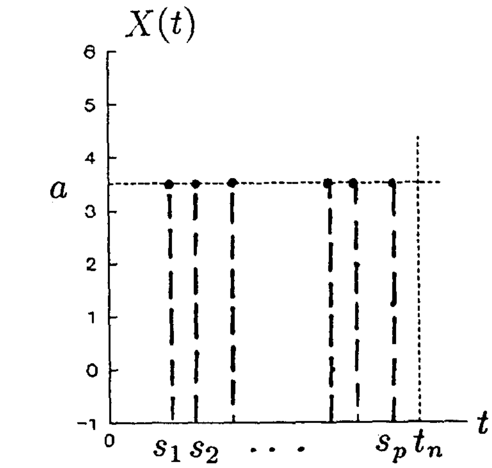
(a) Mixed Data Structure



(c) Degradation Data Alone



(b) Terminal Point Data



(d) Lifetime Data Alone($p = n$) or Censored Lifetime Data ($p < n$)

Figure 3.1 Sample Path of a Truncated Wiener Process

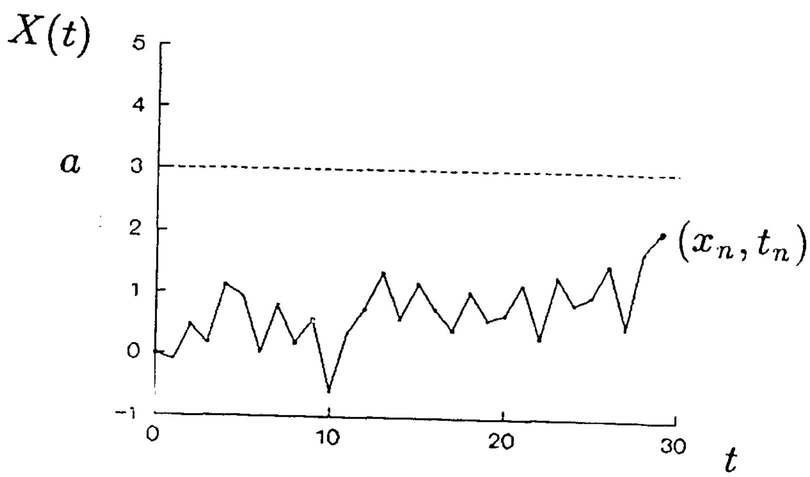


Figure 3.2 Sample Path of a Wiener Process

and Its Image Path

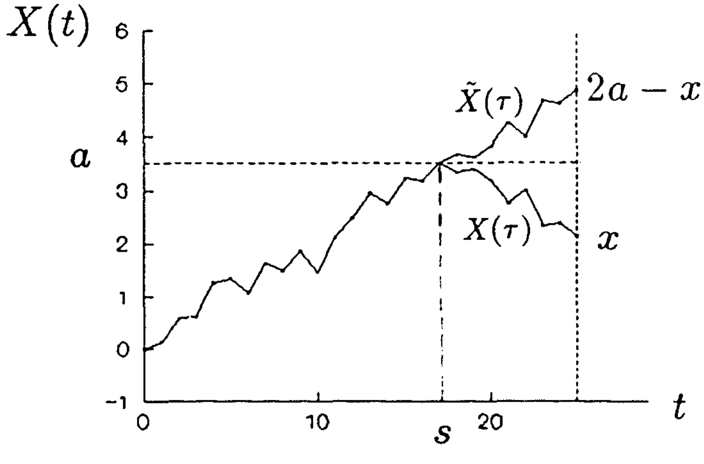


Figure 3.3 χ^2 Plots for Quantities Q, Q_R, and Q_E

Based on Simulated Data

(a) $\chi^2(62)$ Plot for Q

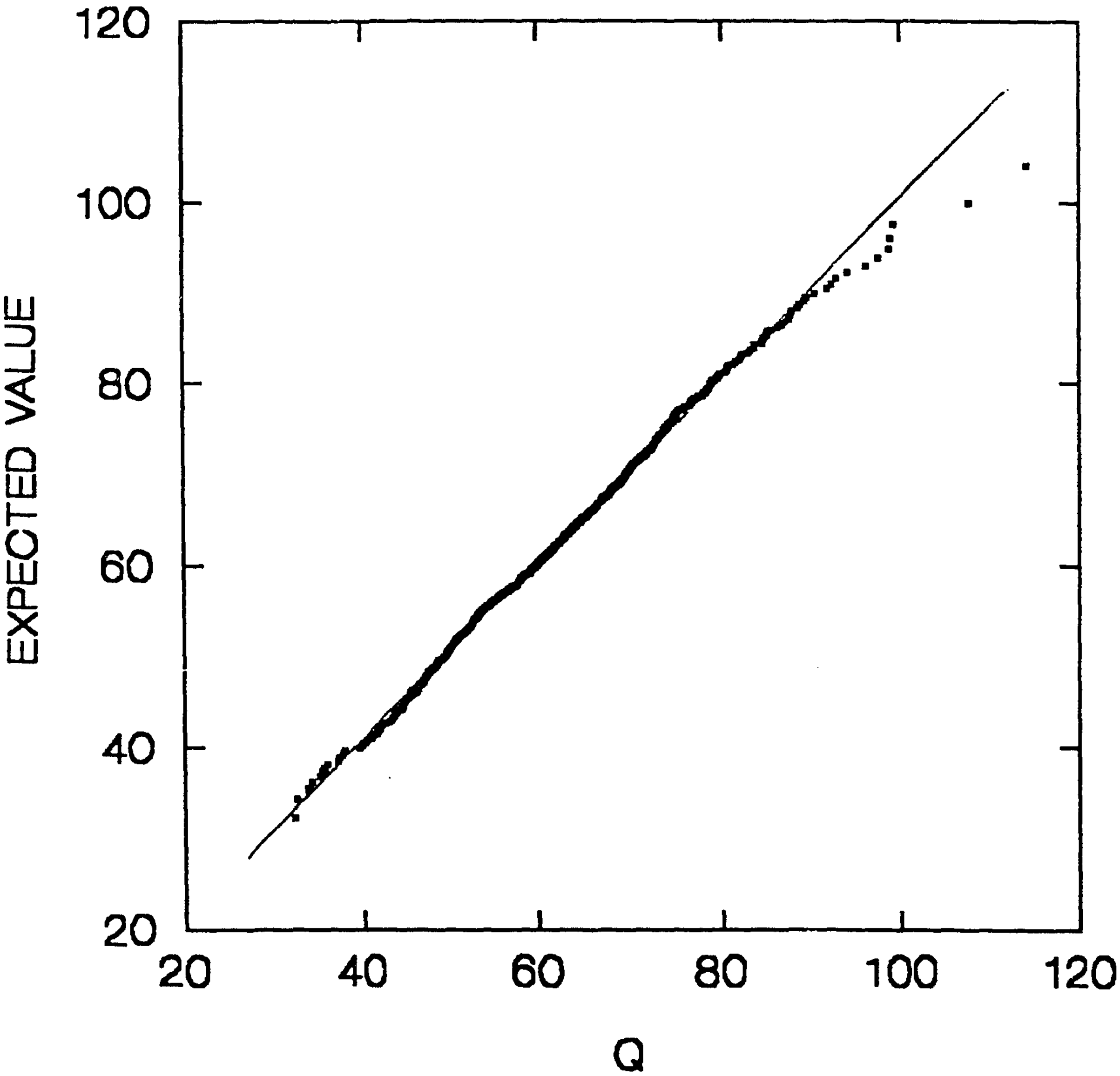


Figure 3.3 Contd. (b) $\chi^2(61)$ Plot for Q_E

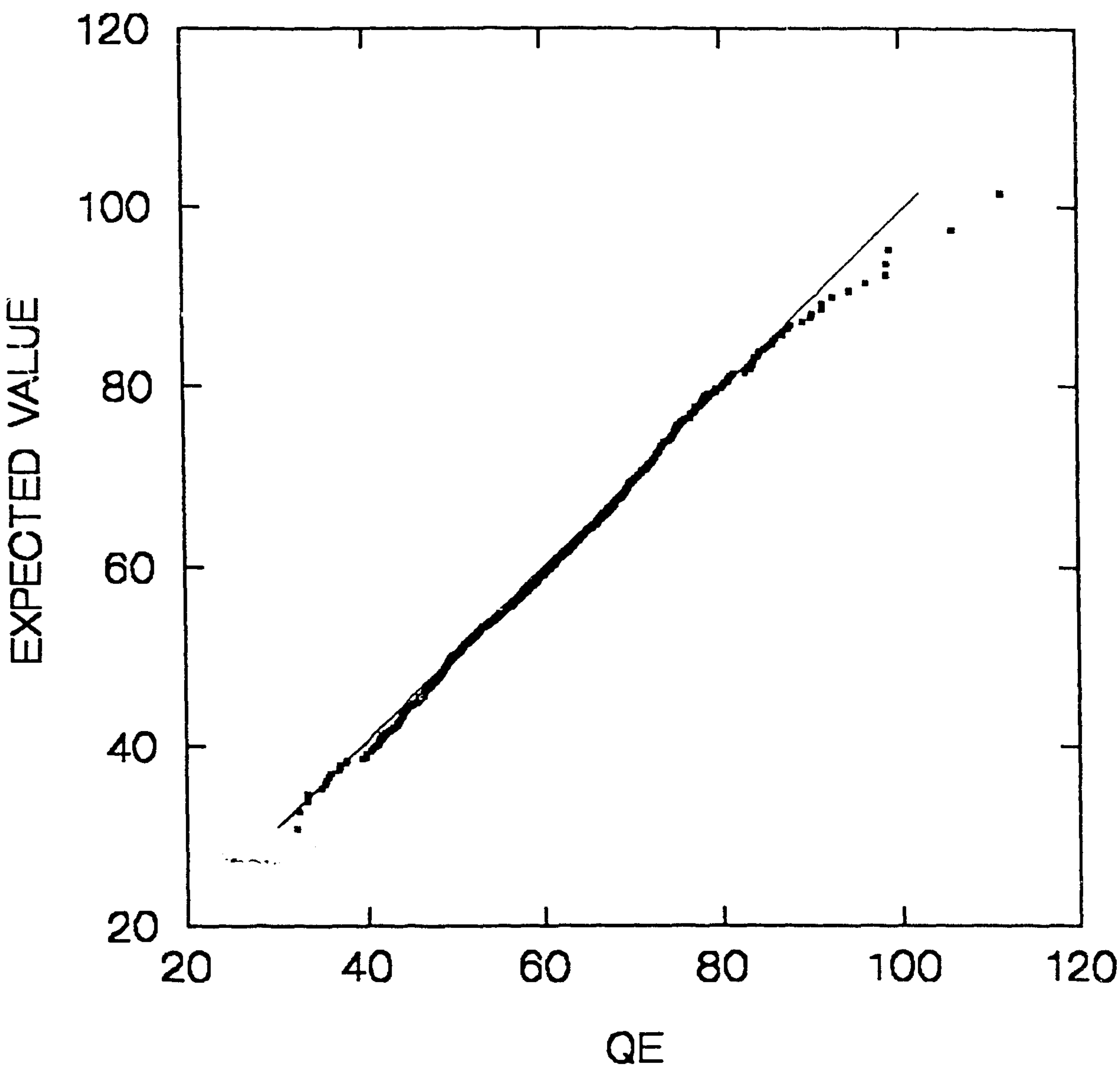


Figure 3.3 Contd. (c) $\chi^2(1)$ Plot for Q_R

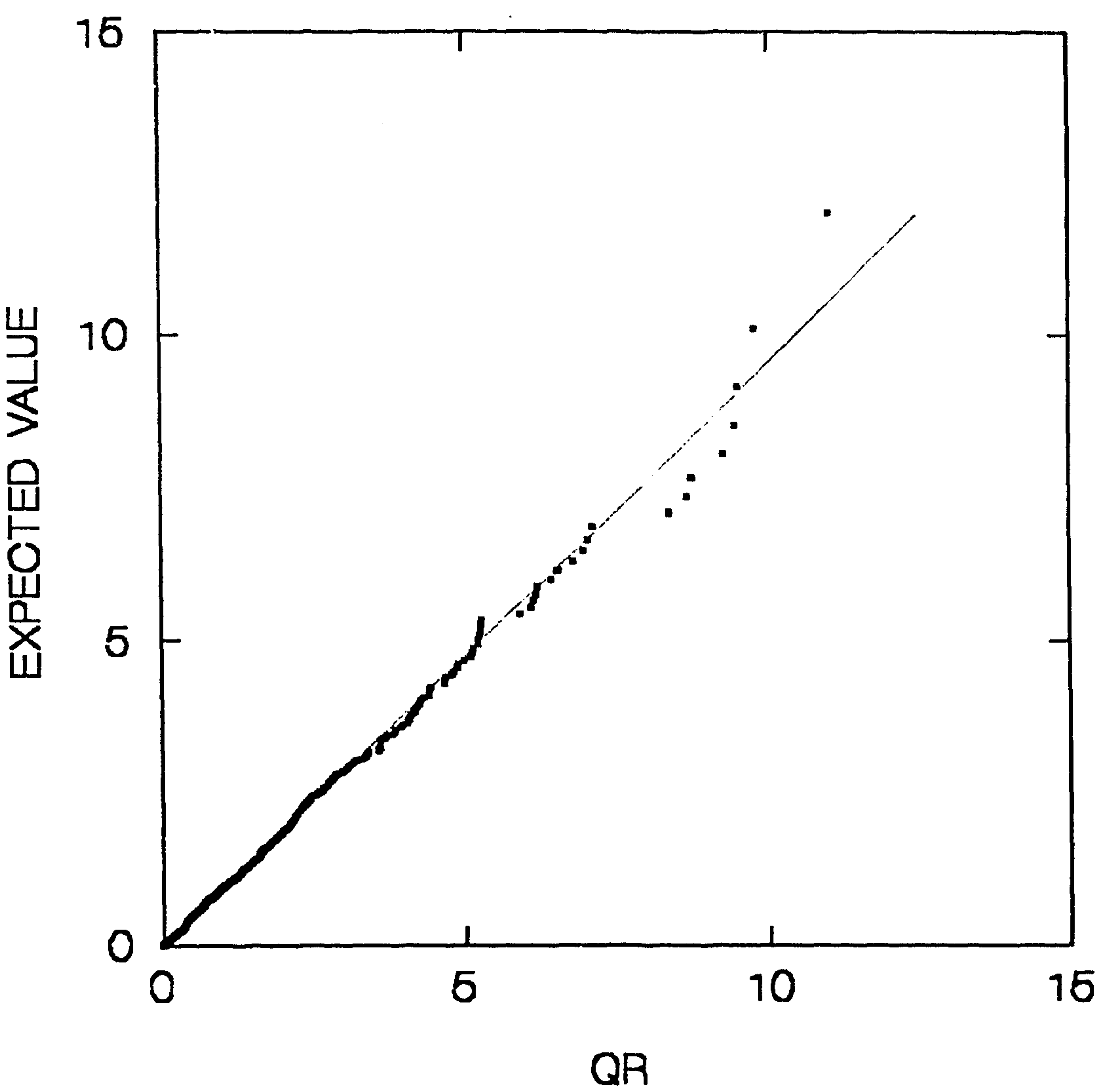


Figure 4.1 A Plot of Equation (4.8), x_n/a with Respect to $\hat{b}_1$ and b_2

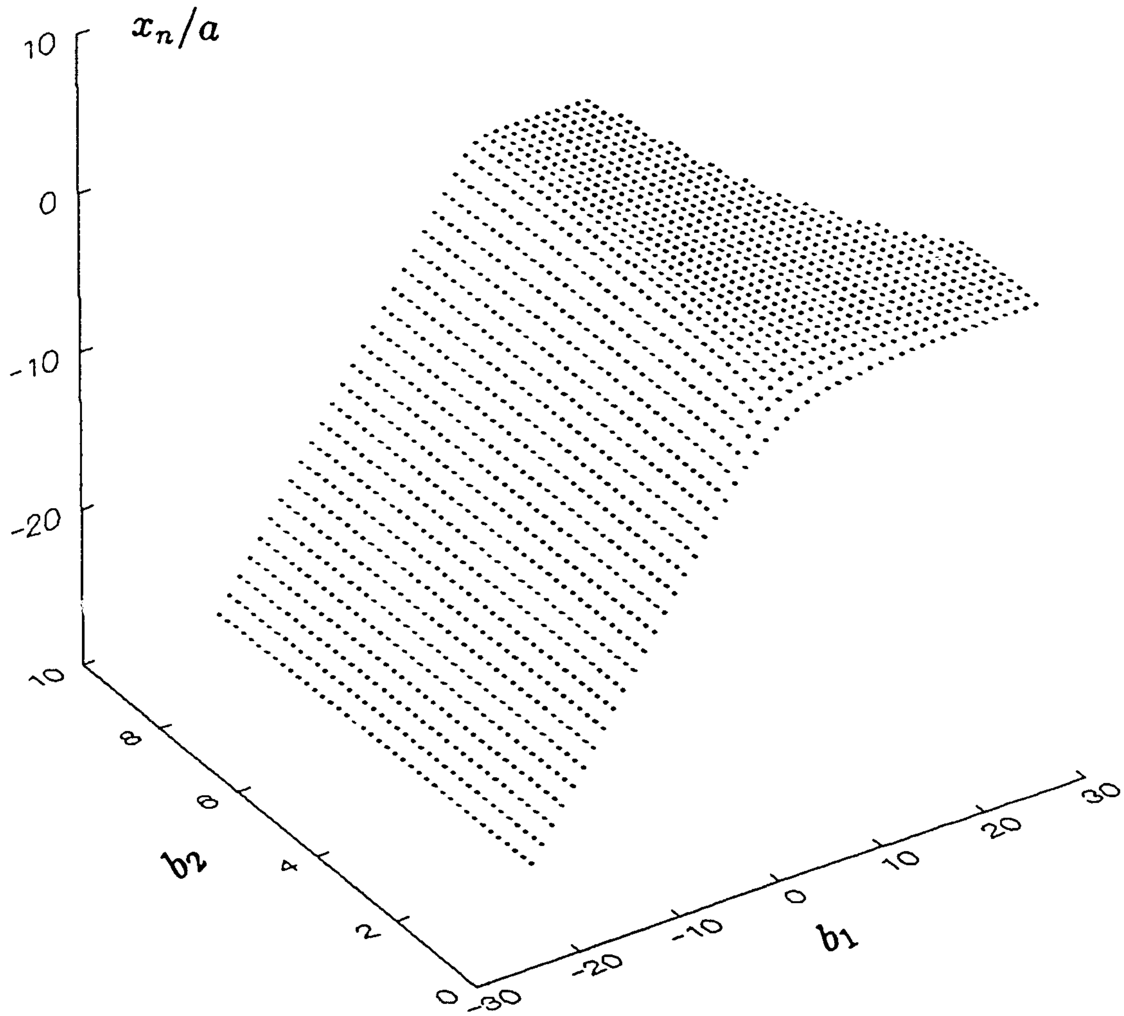


Figure 4.2 A Plot of Equation (4.8), $x_n/a \to 1$ as $b_1 \to \infty$

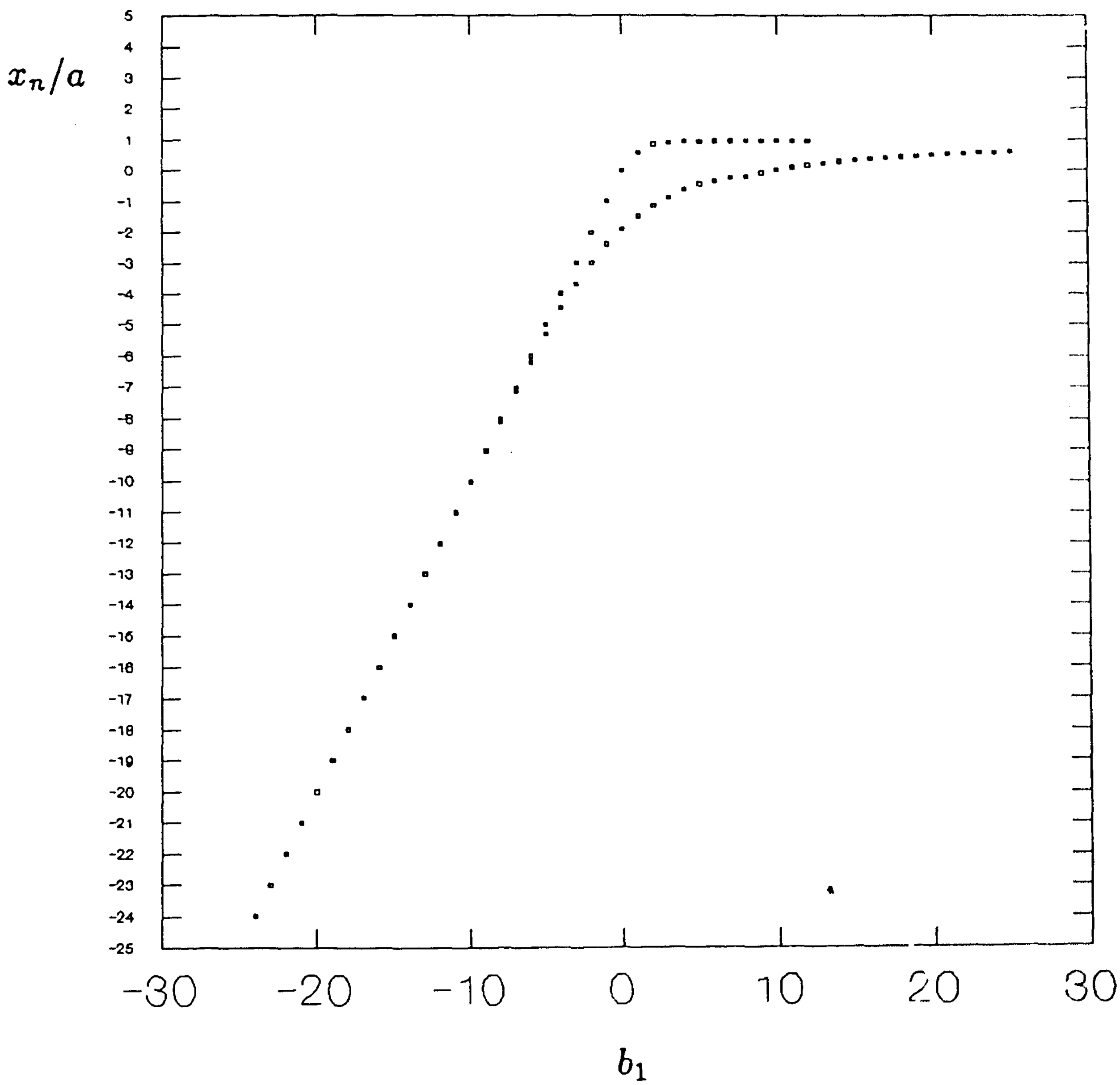


Figure 4.3 Density Curves of $\hat{\delta}$–Sampling Distribution for Different Values of a, the Barrier Level

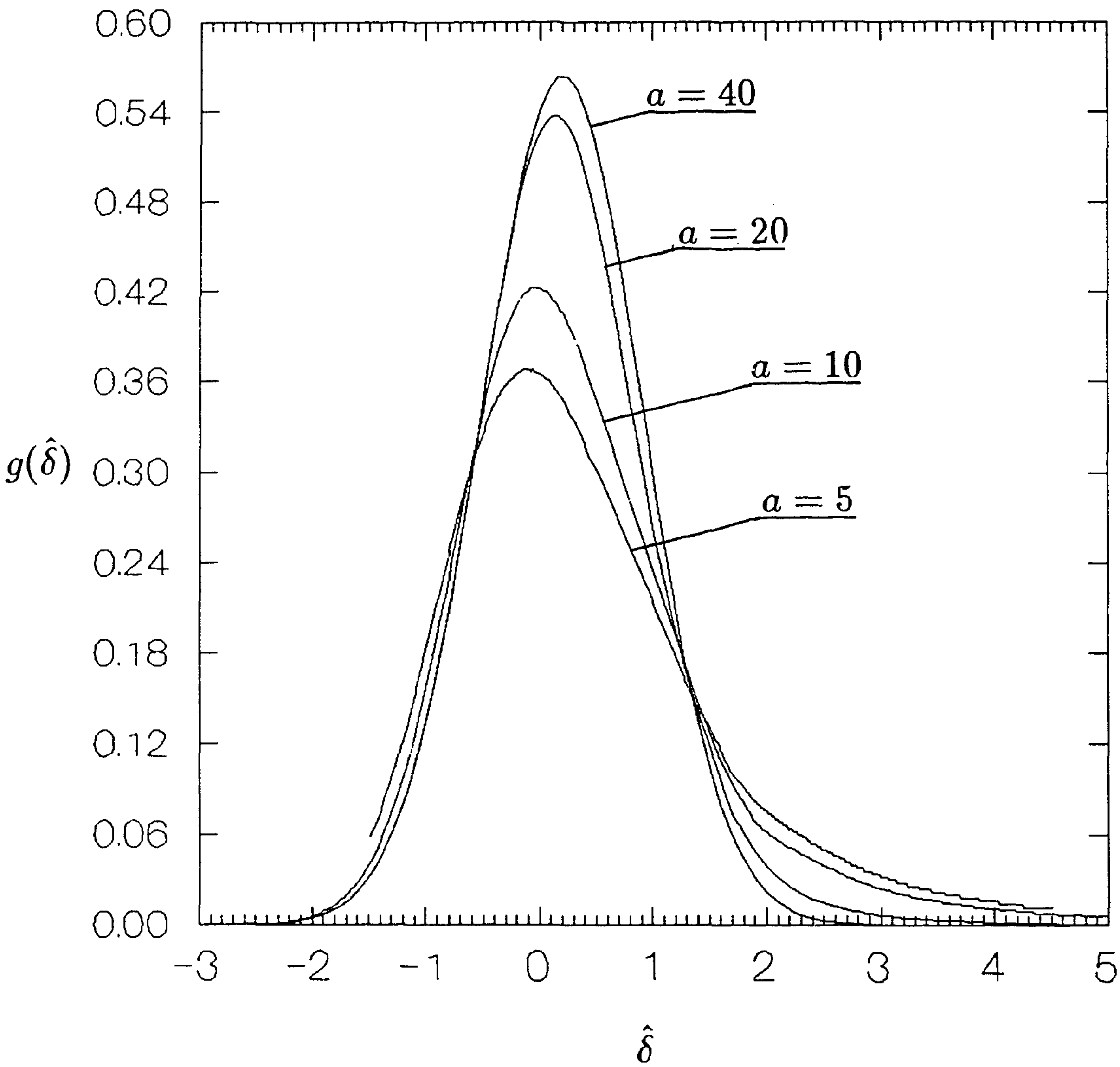


Figure 6.1 A Curve of Predictive Density Function $h(s|\mathbf{x}, \mathrm{s}, \hat{\theta})$ of an Item Surviving at $X(60) = 0$

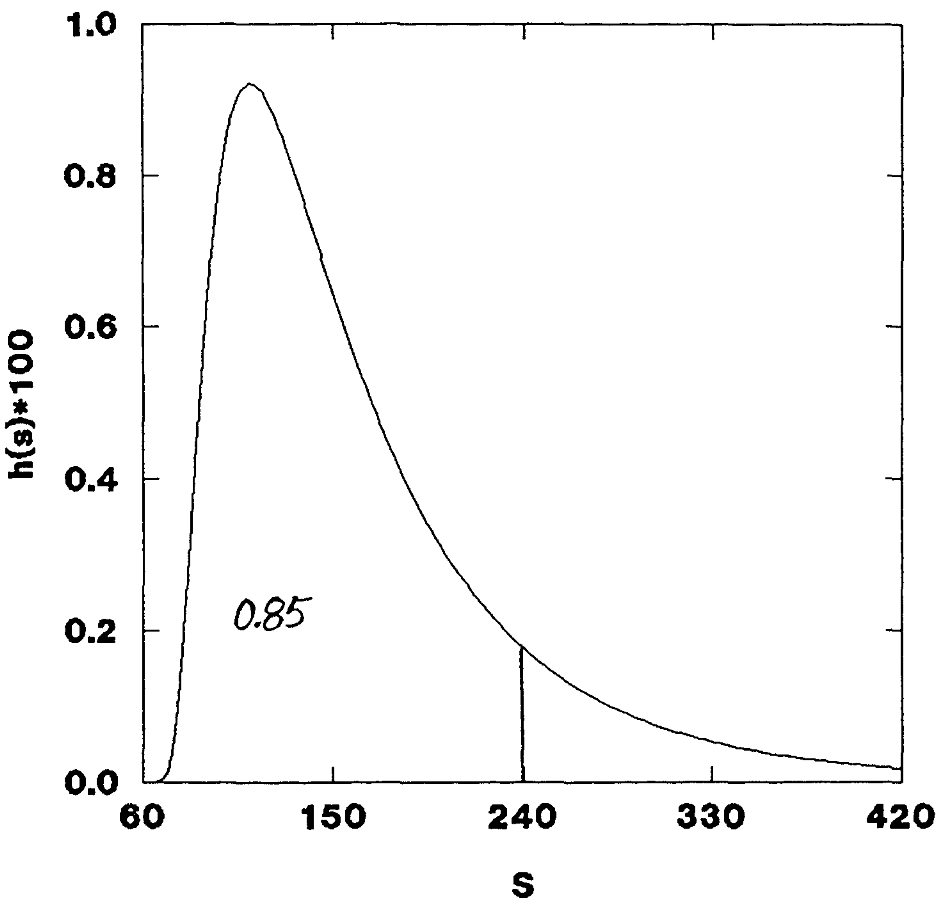


Figure 6.2 A 95% Confidence Region for δ and ν

Based on a Sequence of 100 Daily Stock Returns

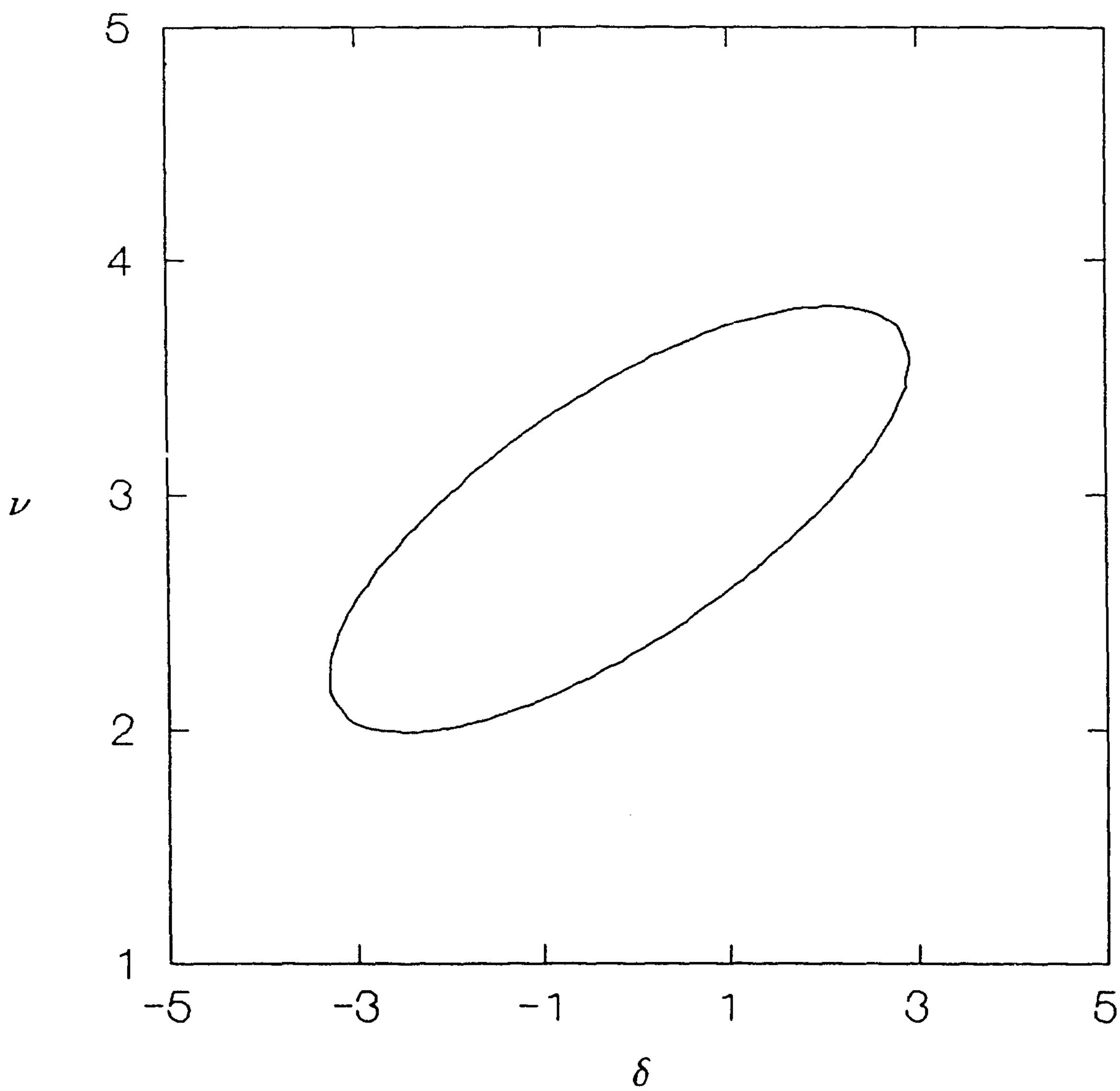


Figure 6.3 A Plot of $\hat{\delta}$ and K, the Temperature Level, of the Transistor Degradation Data Example

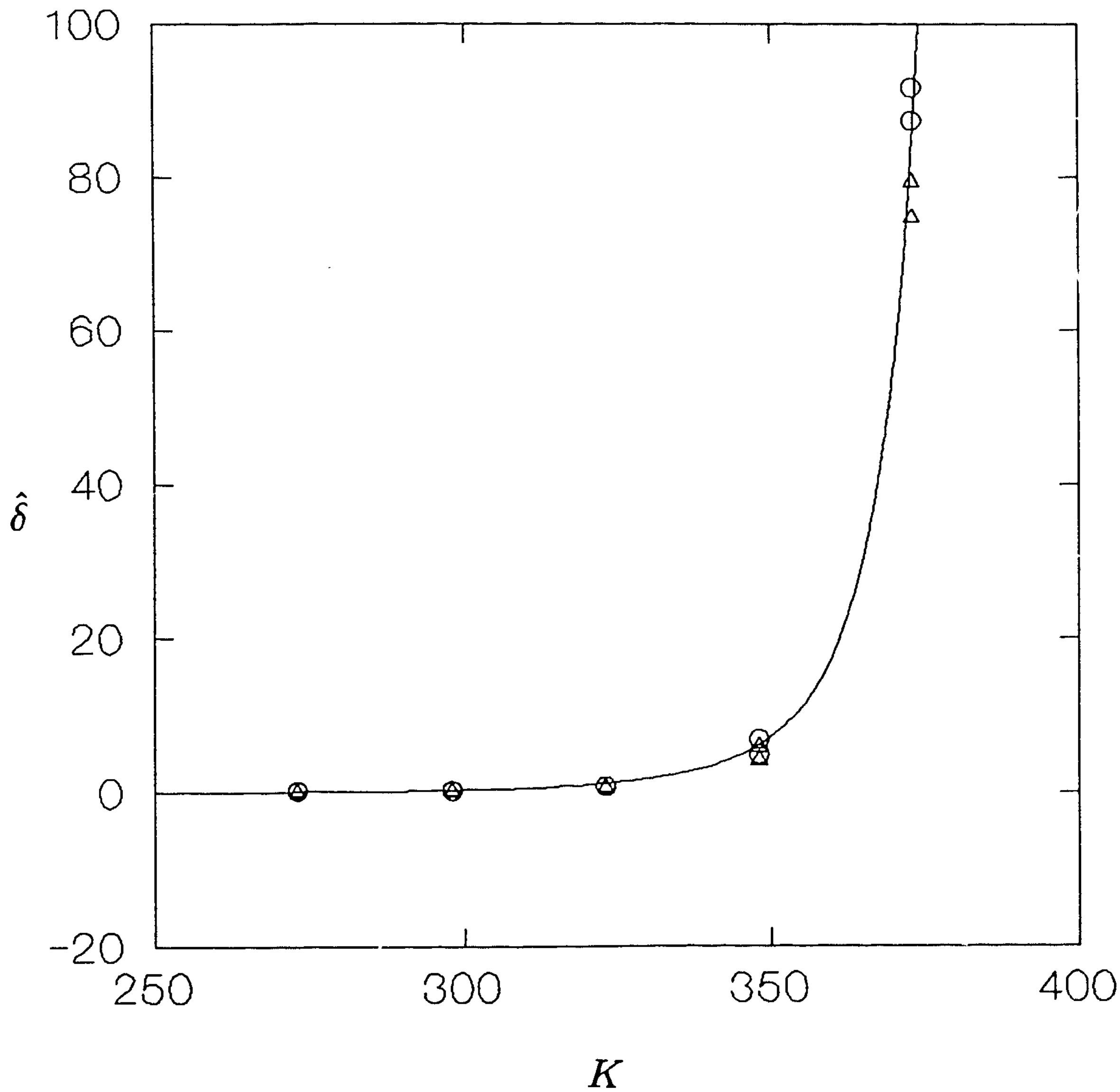


Figure 6.4 A Plot of $\hat{\nu}$ and K, the Temperature Level, of the Transistor Degradation Data Example

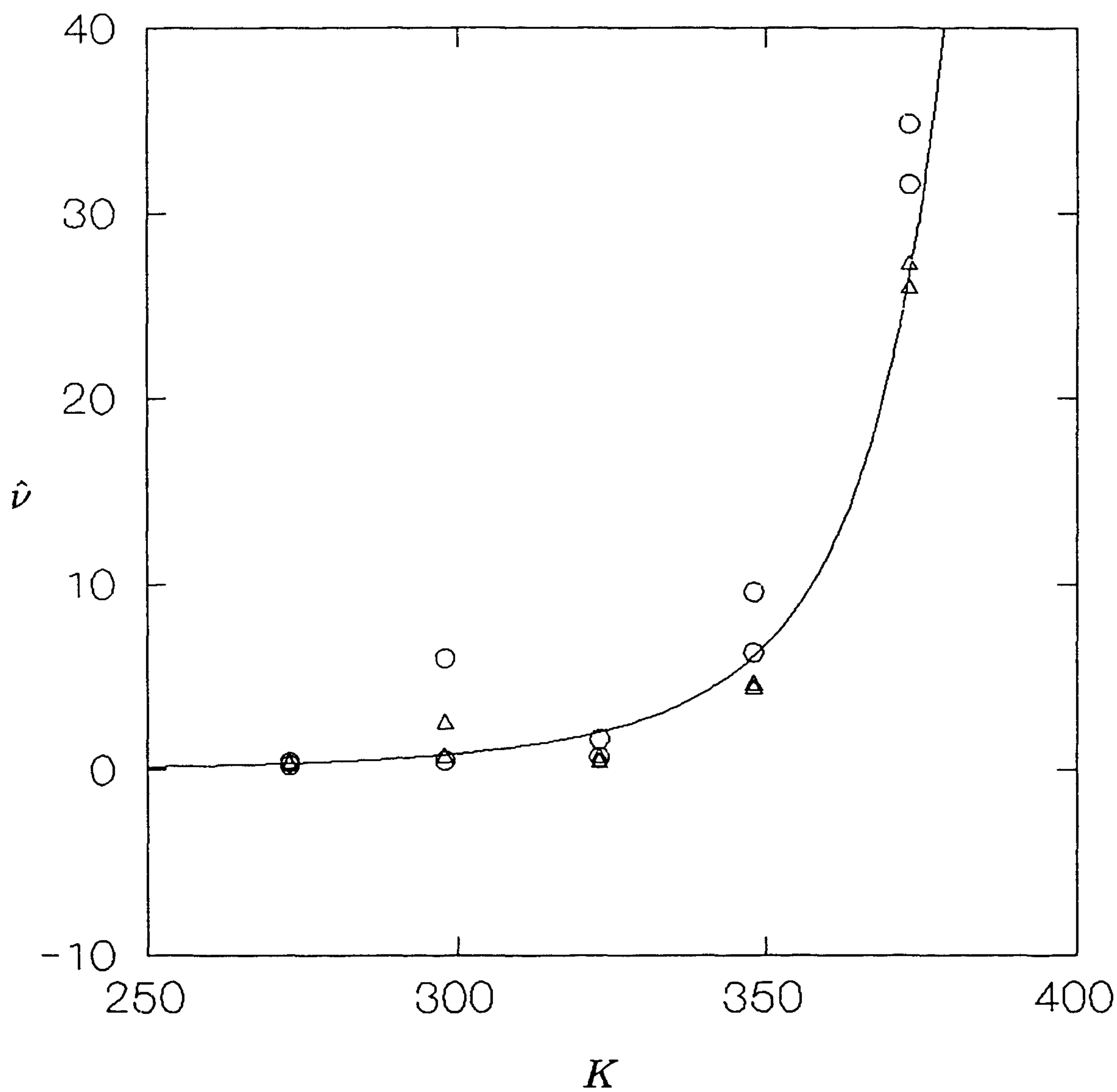


Figure 6.5 A Plot of the Transformed $\hat{\delta}'$ and K_1 of the Transistor Degradation Data Example

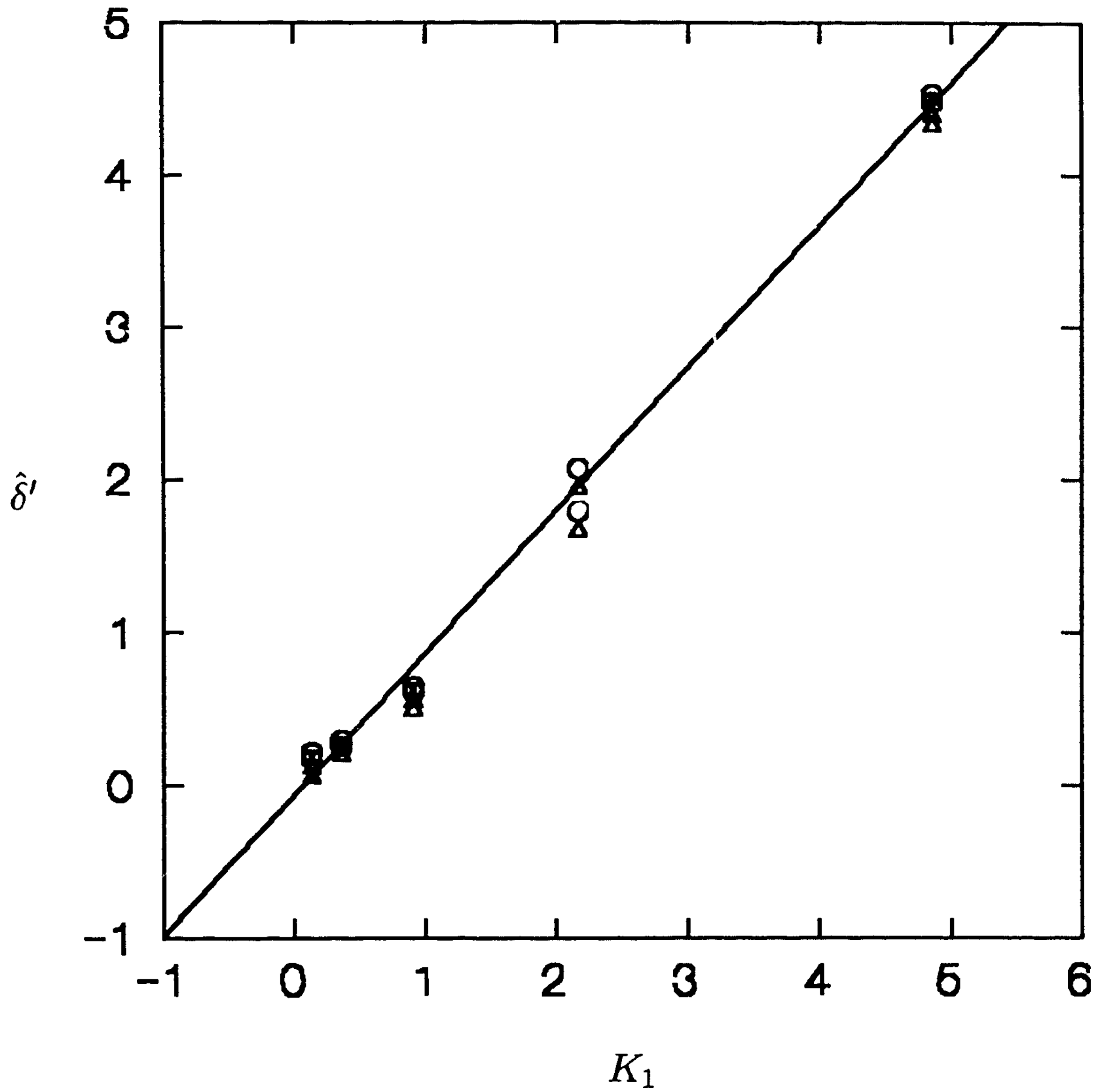


Figure 6.6 A Plot of the Transformed $\hat{\nu}'$, and K_2 of the Transistor Degradation Data Example

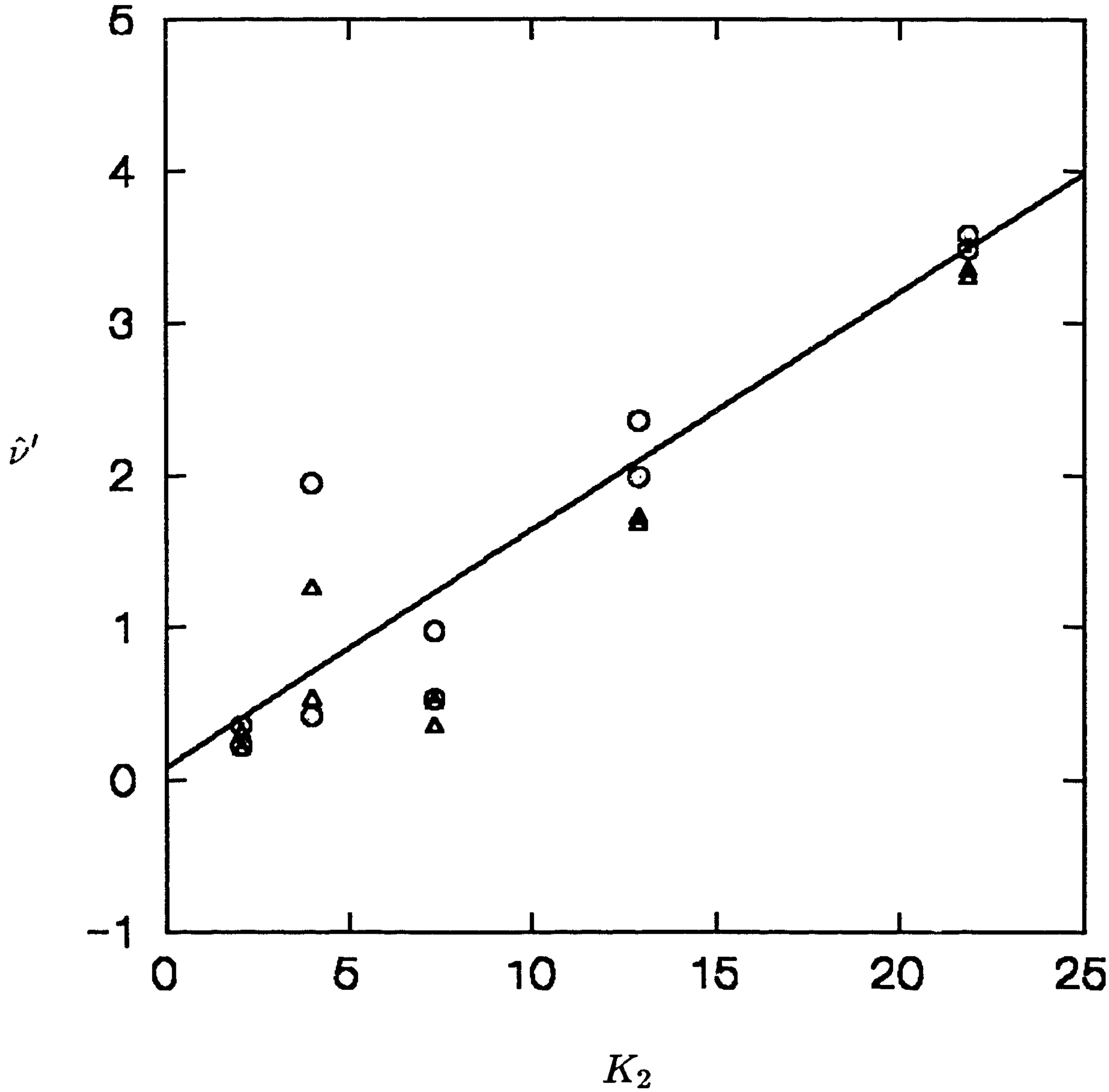


Table 3.1 Descriptive Statistics of Quantities Q, Q_E, and Q_R Based on Simulated Data

	Q	QE	QR
Number of cases	1000	1000	1000
Mean	61.504	60.470	1.034
Variance	127.035	124.613	2.135

Correlation coefficient of QE and QR= 0.009

Table 6.1 Simulated Terminal Point Data for Case Example 6.1

(a) Degradation Levels of Surviving Items

at Termination Time $t_{T} = 60$ (75 Cases)

12.14213	15.62855	17.69043	-2.44341	5.11325	15.62297	14.91234
15.56237	8.68442	1.25276	6.98332	11.81844	16.70418	2.12379
.67851	10.74363	8.08719	8.76551	7.22935	17.64452	7.40810
12.77163	15.62129	16.23109	6.85663	16.94567	4.79681	14.92917
1.58111	16.13156	-.76809	5.29686	9.65507	2.89318	12.26031
6.50739	17.98543	-6.09736	6.37003	1.06473	-1.13387	-5.44475
14.34272	14.84135	5.95957	8.71392	8.14419	7.34078	3.75003
8.23961	10.96491	-1.77257	-6.29091	-8.38495	3.78228	8.80557
13.05949	2.09561	11.86486	-8.48722	-4.50586	-10.3007	9.81885
5.75186	8.42572	15.65565	16.48809	6.11814	17.39534	2.44143
17.33543	-5.77587	7.17196	5.12634	.17986		

(b) Lifetimes of Failed Items (25 cases)

17.10966	54.91068	23.37911	45.80112	57.46943	32.81833
49.93590	25.50655	50.55935	52.92153	45.47594	30.20828
30.97360	38.45463	30.13472	43.61834	39.79428	47.12710
48.82117	34.85248	33.14257	46.19419	34.66541	52.74341
41.99863					

Table 6.2 Stock Prices Used in Case Example 6.3

5.000000	2.820256	4.615042	2.597682	3.275383
4.671046	2.820285	4.504636	2.575482	3.297677
4.539468	2.621990	4.469190	2.508880	3.074840
4.605254	2.445732	4.181613	2.730910	3.063719
4.745434	2.644047	3.517817	2.642112	2.985740
4.635568	2.842336	3.053189	2.642118	2.807521
3.646902	2.908449	3.008994	2.575473	2.651559
3.405240	3.106783	2.964752	2.442258	2.673832
3.273407	3.062730	2.832021	2.428674	2.629231
3.163527	3.371204	2.898403	2.428674	2.651518
3.097631	3.864221	3.208150	2.852033	2.651517
3.493081	3.842154	3.097525	3.743286	2.540113
3.668893	3.709680	2.986898	3.542736	2.495553
3.405283	3.555136	2.854144	3.654145	2.495553
3.273441	3.886397	2.743544	3.609613	2.428682
3.141632	3.908472	2.478051	3.654168	2.406411
2.878007	4.040920	2.686513	3.297673	2.384117
2.952469	4.107150	2.619908	3.253115	2.384126
2.886349	4.327985	2.619908	3.186263	2.228160
2.776191	4.504653	2.686494	3.230829	1.849366

Table 6.3 Degradation Data of Electronic Transistors Under Different Temperature and Current Levels

	Current Level			
	1	2	1	2
	Temperature 0°C			
Time	Item 1	Item 2	Item 3	Item4
0	100	93.3	91.9	85.9
165	99.6	93.6	92.1	86.7
450	98.4	92.8	91.3	86.2
801	98.4	92.8	91.0	85.9
1300	98.2	92.3	90.9	85.4
1800	97.9	92.2	90.2	85.4
2300	97.7	92.3	90.3	85.7
3000	97.5	91.9	90.1	85.4
3690	97.1	91.7	89.7	85.1
4504	96.7	91.6	89.7	85.3
5345	97.1	91.7	89.5	85.3
6255	96.7	91.5	89.3	85.0
7300	97.1	91.7	89.6	85.2
8500	97.2	91.8	89.6	81.5
9825	96.5	91.2	89.2	84.8
11500	97.1	91.6	89.7	85.2
14160	97.2	91.8	89.9	85.3
16315	96.0	91.3	88.8	85.0

Table 6.3 Contd.

Temperature 50°C				
Time	Item 9	Item 10	Item 11	Item 12
0	90.9	85.3	87.6	82.1
50	90.3	85.2	86.5	82.2
115	90.1	84.9	86.3	82.0
180	89.9	84.7	86.1	81.6
250	89.6	84.5	86.1	81.1
320	89.6	84.5	85.7	81.5
420	89.3	84.2	85.6	81.2
540	89.1	84.1	85.4	81.0
630	89.0	84.0	85.4	81.0
720	89.1	84.1	85.0	80.9
810	88.5	83.7	84.8	80.7
875	88.4	83.6	84.6	80.6
941	88.5	83.7	84.8	80.6
1010	88.3	83.3	84.3	80.2
1100	87.7	82.9	83.9	79.8
1200	87.5	82.7	83.6	79.7
1350	87.0	82.5	83.5	79.6
1500	87.1	82.6	83.2	79.7
1735	86.9	82.6	82.9	79.0
1896	86.5	82.3	82.6	79.1
2130	86.4	82.6	83.0	79.6
2460	85.9	81.9	82.4	79.1
2800	85.4	81.6	81.4	78.6
3200	85.2	81.4	80.0	78.3
3400	84.6	81.3	80.4	78.1
4600	83.8	80.1	79.3	76.7
5650	83.7	80.2	79.7	77.1
6600	83.7	80.3	79.8	77.2
7800	82.3	79.8	77.0	76.5
8688	82.5	78.1	76.7	75.2
10000	82.3	77.9	78.4	75.6

Temperature 75°C				
Time	Item 13	Item 14	Item 15	Item 16
0	78.6	74.9	81.1	76.7
25	76.9	73.9	79.8	75.6
65	76.6	73.1	79.2	75.2
130	76.1	72.6	79.1	75.0
250	75.5	72.0	78.4	74.4
420	74.6	71.3	77.5	73.6
609	73.5	70.4	76.7	72.9
818	72.6	69.6	76.1	72.4
1004	71.7	68.8	75.5	71.9
1240	69.7	67.2	74.9	71.4

Table 6.3 Contd.

Temperature 25°C				
Time	Item 5	Item 6	Item 7	Item 8
0	107	98.8	87.1	82.3
50	109	100	87.0	82.7
120	108	99.3	87.0	82.2
210	107	98.6	86.2	81.6
300	107	99.0	86.4	81.8
400	107	98.8	86.4	81.7
540	107	98.8	86.3	81.6
730	107	98.7	86.1	81.5
950	107	98.6	85.6	81.1
1260	107	98.5	85.7	81.2
1740	106	98.1	85.4	81.0
2350	106	97.8	85.1	80.7
3000	106	97.9	85.3	80.8
4000	105	97.6	84.7	80.4
5200	104	96.6	84.1	79.8
6450	104	96.4	83.7	79.5
8030	104	96.1	83.2	79.2
10000	104	96.6	83.6	79.5

Temperature 100°C				
Time	Item 17	Item 18	Item 19	Item 20
0	101	93.8	100	93.1
15.5	99.1	91.9	98.0	91.2
40	98.0	91.0	97.3	90.4
80	95.9	89.1	95.0	88.7
150	90.1	84.2	89.7	84.2
300	73.5	69.9	73.8	70.7

Table 6.4 Maximum Likelihood Estimates of δ and ν Based on Transistor Degradation Data and the Results of Sample Path Homogeneity Test

(1) Block #	(2) Item #	(3) a	(4) $\hat{\delta}_i$	(5) $\hat{\delta}$	(6) $\hat{\nu}_i$	(7) $\hat{\nu}$	(8) Test statistics Ho: $\nu_i=\nu_j$	(9) Test statistics Ho: $\delta_i=\delta_j$
1 Temp.=0 Current=1	1 (i=1)	30.0	0.2452	0.2176	0.4251	0.3372	1.1950 Ho accepted	0.0737 Ho accepted
	3 (j=3)	21.9	0.1900		0.2494			
2 Temp.=0 Current=2	2 (i=2)	23.3	0.1226	0.0889	0.2215	0.2859	0.8865 Ho accepted	0.1296 Ho accepted
	4 (j=4)	15.9	0.0552		0.3503			
3 Temp.=25 Current=1	5 (i=5)	37.0	0.3000	0.3250	6.0188	3.2727	20.8211 Ho rejected	Inconclusive
	7 (j=7)	17.1	0.3500		0.5225			
4 Temp.=25 Current=2	6 (i=6)	28.8	0.2200	0.2500	2.4836	1.5833	6.6245 Ho accepted	0.0114 Ho accepted
	8 (j=8)	12.3	0.2800		0.6831			
5 Temp.=50 Current=1	9 (i=9)	20.9	0.8600	0.8900	0.6982	1.1753	5.2624 Ho accepted	0.0153 Ho accepted
	11 (j=11)	17.6	0.9200		1.6601			
6 Temp.=50 Current=2	10 (i=10)	15.3	0.7400	0.6950	0.4175	0.5321	1.5727 Ho accepted	0.0761 Ho accepted
	12 (j=12)	12.1	0.6500		0.6627			
7 Temp.=75 Current=1	13 (i=13)	8.6	6.9355	5.9677	9.5996	7.9763	0.3923 Ho accepted	0.2911 Ho accepted
	15 (j=15)	11.1	5.0000		6.3172			
8 Temp.=75 Current=2	14 (i=14)	4.9	5.9902	4.9563	4.5776	4.4530	0.0089 Ho accepted	0.2755 Ho accepted
	16 (j=16)	6.7	4.2742		4.3317			
9 Temp.=100 Current=1	17 (i=17)	31.0	91.6667	89.5000	31.5576	33.1844	0.0120 Ho accepted	0.0849 Ho accepted
	19 (j=19)	30.0	87.3333		34.8111			
10 Temp.=100 Current=2	18 (i=18)	23.8	79.3333	77.0000	25.8913	26.5313	0.0029 Ho accepted	0.1231 Ho accepted
	20 (j=20)	23.1	74.6667		27.1713			

REFERENCE

Abramowitz,M., and Stegun,I.(1967), *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables* (Applied Mathematics Series, No.55), Washington, DC: National Bureau of Standards.

Anderson,R.M., and May,R.M.(1992), 'Understanding the AIDS Pandemic', *Scientific American*, May, p-58.

Boulanger, M. and Escobar, L.A.(1994), 'Experimental Design for a Class of Accelerated Degradation Tests', *Technometrics*, 36, 260-272.

Carey, M.B. and Koenig, R.H.(1991), 'Reliability Assessment Based on Accelerated Degradation: A Case Study,' IEEE Transactions on Reliability, 40(5), 499-506.

Carlin,B.P., and Gelfand,A.E.(1990), 'Approaches for Empirical Bayes Confidence Intervals,' *Journal of the American Statistical Association*, 85(409), 105-114.

Cheng,R.C.H., and Amin,N.A.K.(1981), 'Maximum Likelihood Estimation of Parameters in the Inverse Gaussian Distribution, With Unknown Origin,' *Technometrics*, 23(3), 257-263.

Chhikara,R.S., and Guttman,I.(1982), 'Prediction Limits for the Inverse Gaussian Distribution,' *Technometrics*, 24(4), 319-324.

Chhikara,R.S., and Folks,J.L.(1989), *The Inverse Gaussian Distribution, Theory, Methodology, and Applications*, New York: Marcel Dekker.

Cho,D.C., and Frees,E.W.(1988), 'Estimating the Volatility of Discrete Stock Prices,' *The Journal of Finance*, XLIII(2), 451-466.

Chown,M., Pullum,G.G., and Whitmore,G.A.(1994), *Reliability in Communication Technology*, England: Chapman & Hall.

Cootner, P.H. (ed) (1964), *The Random Character of Stock Market Prices*, Cambridge, Massachusetts:The M.I.T. Press.

Cox, D.R.(1972), 'Regression Models and Life-Tables' (with discussions), *J. Roy. Statist. Soc. Ser. B*, 34, 187-220.

Cox,D.R., and Miller,H.D.(1965), *The Theory of Stochastic Processes*, New York: Chapman and Hall.

Crowder, M.J., Kimber, A.C., Smith, R.L., and Sweeting,T.J. (1991), *Statistical Analysis of Reliability Data*, London, Chapman and Hall.

Dasgupta,A. and Pecht,M.(1991),' Material Failure Mechanisms and Damage Models,' *IEEE Transactions on Reliability*, 40(5), 531-536.

Davis,D.J.(1952), 'An Analysis of Some Failure Data,' *Journal of the American*

Statistical Association, 47, 113-150.

DeGroot, M.H.(1969), *Optimal Statistical Decisions*, New York: McGraw-Hill.

Desmond, A.F.(1987), 'An Application of Some Curve-Crossing Results for Stationary Stochastic Processes to Stochastic Modelling of Metal Fatigue,' *Applied Probability Processes, and Sampling Theory, 51-63*, MacNeill,I. and Umphrey, G.J. (eds.), D. Reidel Publishing Company.

Desmond, A.F.(1985), 'Stochastic Models of Failure in Random Environments,' *The Canadian Journal of Statistics,* 13(2), 171-183.

Ditlevsen,O.(1986), 'Random Fatigue Crack Growth - A First Passage Problem,' *Engineering Fracture Mechanics*, 23(2), 467-477.

Doksum,K.A., and Hoyland,A.(1992), 'Models for Variable-Stress Accelerated Life Testing Experiments Based on Wiener Processes and the Inverse Gaussian Distribution,' *Technometrics*, 34(1), 74-82.

Guttman,I.; Johnson,R.A.; Bhattacharyya,G.K. and Reiser,B.(1988), ' Confidence Limits for Stress-Strength Models With Explanatory Variables,' *Technometrics*, 30(2), 161-168.

Gydesen,H.(1984), 'A Stochastic Approach to Models for the Leaching of Organic Chemicals in Soil,' *Ecological Modelling*, 24, 191-205.

Howlader,H.A.(1985), 'Approximate Bayes Estimation of Reliability of Two Parameter Inverse Gaussian Distribution,' *Communications in Statistics*, A, 14(4), 937-946.

Karlin,S., and Taylor,H.M.(1975), *A First Course in Stochastic Processes*, 2nd edition, New York:Academic Press.

Lawless,J.F.(1983), 'Statistical Methods in Reliability,' (with discussions), *Technometrics*, 25(4), 305-335.

Lawless,J.F. (1982), *Statistical Models and Methods for Lifetime Data*, New York:John Wiley

Lawless,J.F., and Singhal,K.(1980), 'Analysis of Data from Life-Test Experiments Under an Exponential Model,' *Naval Research Logistics Quarterly*, 27, 323- 334.

Lange,N., Carlin,B.P., and Gelfand,A.E. (1992), 'Hierarchical Bayes Models for the Progression of HIV Infection Using Longitudinal CD4+ Counts,' *Journal of the American Statistical Association*, 87, 615-626.

Leblebici, Y. and Kang, S.M. (1993), 'Modeling and Simulation of Hot-Carrier-Induced Device Degradation in MOS Circuits,' *IEEE Journal of Solid-State Circuits*, 28(5), 585- 595.

Letac,G., Seshadri,V., and Whitmore,G.A.(1985), 'An Exact Chi-squared Decomposition Theorem for Inverse Gaussian Variates,' *Journal of the Royal Statistical Society*, B, 47(3), 476-481.

Lu, C.J. and Meeker, W.Q.(1993), 'Using Degradation Measures to Estimate a Time-to- Failure Distribution,' *Technometrics*, 35(2), 161-174.

Nair, V.N.(1988), 'Discussion of: Estimation of Reliability in Field Performance Studies (Kalbfleisch,J.D. and Lawless,J.F.), ' *Technometrics*, 30, 379-383.

Nelson,W.(1982), *Applied Life Data Analysis*, New York: John Wiley & Sons.

Reiser,B. and Guttman,I.(1986), 'Statistical Inference for Pr(Y¡X): The Normal Case', *Technometrics*, 28(3), 253-257.

Samuelson, P. (1973), 'Mathematics of Speculative Price', *SIAM Review*, 15(1), 1-43.

Seshadri,V. (1993), *The Inverse Gaussian Distribution - A Case Study in Exponential Families*, Oxford Science Publications.

Shepp,L.A.(1979), 'The Joint Density of the Maximum and Its Location for a Wiener Process With Drift,' *Journal of Applied Probability*, 16, 423-427.

Sinha,S.K.(1986), 'Bayesian Estimation of the Reliability Function of the Inverse Gaussian Distribution,' *Statistics & Probability Letters*, 4, 319-323.

Sobczyk,K.(1987), 'Stochastic Models for Fatigue Damage of Materials,' *Adv. Appl. Prob.*, 19, 652-673.

Stucki,F. (1994), 'Injection of a Minimal Space Charge as Mechanism for the Initial Phase of Electrical Polymer Degradation,' *IEEE Transactions on Dielectrics and Electrical Insulation*, 1(2) 231-234.

Taksar,M.I.(1984), 'Storage Model with Discontinuous Holding Cost,' *Stochastic Processes and Their Applications*, 18, 291-300.

Tweedie,M.C.K.(1957a), 'Statistical Properties of Inverse Gaussian Distributions I,'. *Annals of Mathematical Statistics*, 28, 362-377.

Tweedie,M.C.K.(1957b), 'Statistical Properties of Inverse Gaussian Distributions II,' *Annals of Mathematical Statistics*, 28, 696-705.

Viveros,R.(1991),'Combining Series System Data to Estimate Component Characteristics,' *Technometrics*, 33(1), 13-23.

Weerahandi,S. and Johnson,A.(1992), 'Testing Reliability in a Stress-Strength Model When X and Y are Normally Distributed,' *Technometrics*, 34(1), 83-91.

Whitmore,G.A.(1990), ' On the Reliability of Stochastic Systems: A Comment,' *Statistics & Probability Letters*, 10, 65-67.

Whitmore, G.A., and Seshadri,V.(1987), 'A Heuristic Derivation of the Inverse Gaussian Distribution,' *The American Statistician*, 41(4), 280-281.

Whitmore,G.A.(1986), 'Normal-gamma Mixtures of Inverse Gaussian Distributions,' *Scand. J. Statist.*, 13, 211-220.

Whitmore,G.A.(1983), 'A Regression Method for Censored Inverse-Gaussian Data,' *The Canadian Journal of Statistics*, 11(4), 305-315.

Whitmore,G.A.(1978), in the Discussion of 'The Inverse Gaussian Distribution and Its Statistical Application – A Review' by J.L.Folks and R.S. Chhikara. *J.R. Statist. Soc. B.*, 40(3), 285.

Xiao,J.Y. and Bathias,C.(1994), 'Strength Prediction and Damage Mechanisms of Glass/Epoxy Woven Laminates with Circular Holes', *Fatigue & Fracture of Engineering Material & Structures*, 17(4), 411-428.