The First Passage Time of Degradation Processes

THE FIRST PASSAGE TIME OF DEGRADATION PROCESSES

BY

CHENGWEI QIN, M.Sc., B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS & STATISTICS

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2017) McMaster University

(Mathematics & Statistics) Hamilton, Ontario, Canada

TITLE:	The First Passage Time of Degradation Processes
AUTHOR:	Chengwei Qin M.Sc. (National Central University) B.Sc. (Nankai University)
SUPERVISOR:	Prof. Narayanaswamy Balakrishnan
NUMBER OF PAGES:	xv, 116

To Nancy and my parents Jianping & Weiren

Abstract

The thesis gives perspectives for the first passage time (FPT) of degradation processes from both the parametric and nonparametric aspects. As an important reliability index of the manufactured products, $100p^{th}$ percentile of the FPT distribution is always required to provide by the market. If the assumed underlying process is misspecified and fails to fit the degradation data, the estimate of the reliability index will lose efficiency. The typically used degradation processes include the Wiener process, the gamma process, and the inverse Gaussian process because all of them are close under convolution. This property can help us to obtain the FPT in analytic forms. However, it is difficult to accurately fit the actual data with limited model selections, then it is important to discuss how to achieve the flexible selections of the underlying model.

To reduce the misspecification effects for the FPT density, the most straightforward way is to observe the failure times directly. For a good summary and set of references, see Prentice and Kalbfleisch (1979) and Kalbfleisch and Prentice (2002). However, to guarantee the sample size of failure time observation is large enough to make statistical inference, a large number of experiments are required which are too costly. That is the reason why people choose to reduce the experimental cost by alternatively observing the increments then the FPT density can be estimated based on the mathematical

properties of the selected underlying process. Hence, the current methods cannot simultaneously realize the low experimental cost and model robustness.

In this thesis, we firstly propose a novel parametric approach which can generalize the degradation processes only if the corresponding Laplace transforms exist. Then the Laplace transform of the FPT density function can be obtained in close-form and the survival probability can be computed through Laplace inversion. For many stochastic processes, their likelihood functions are intractable so the maximum likelihood estimate (MLE) of the parameters are unavailable to obtain. We estimate the parameters by generalized method of moments (GMM) which is a distance-based method. Specifically, the weighted convolution of two independent gamma processes incorporated with random effects is exemplified as the parametric underlying model, and it is motivated by the scenario of multiple sensors used for monitoring the degradation of the same critical component. Although the degradation processes generated from these sensors reflect the same degradation path, the corresponding scales and noise are significantly different. To find the unified degradation path, Hua *et al.* (2013) used the weight-averaged unified approach which determined the weights by the defined leadership scores. Then we develop the parametric model into a distribution-free model which can eliminate the misspecification effects caused by wrong process-type assumption. The theoretical Laplace transform of degradation process can be replaced by the empirical Laplace transform composed by the observed increments and the $100p^{th}$ percentile of the FPT distribution can be approximated by the empirical saddlepoint method. As one of the important applications in reliability engineering, the optimal design for degradation test is studied under both the parametric and nonparametric scenarios. To optimize the degradation test subject to the experimental cost not exceed the pre-specified

budget, the design factors in the experiment such as the number of test units, the number of measurements, the inspection frequency and the termination time are considered.

Keywords: First Passage Time, Lévy Subordinator, Saddlepoint Approximations, Laplace Inversion, Random Effects, Generalized Methods of Moments, Optimal Design.

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor, Professor Narayanaswamy Balakrishnan, for providing me the opportunity to pursue my academic dream, guiding me with his vast knowledge, unlimited patience and constant encouragement. He is always a great mentor and creates a nourishing environment throughout my time in McMaster University.

I would like to thank Professor Roman Viveros-Aguilera and Professor Traian Pirvu for being my doctoral supervisory committee members and giving me insightful suggestions which drives significant improvements for this thesis.

I would like to thank Prof. Viveros-Aguilera, Prof. Pirvu and Dr. Petar Jevtić as I had been the tutorial instructors for their courses including applied statistics and financial mathematics. I truly appreciate their generous time for the discussions and teaching me how to become a qualified teacher. I am also grateful to the collaboration works with Prof. Pirvu and Dr. Jevtić in financial mathematics and actuarial science. I really appreciate it that they showed me a bright interdisciplinary area of applying mathematics and statistics to solve financial and insurance problems.

I wish to express my gratitude to my girlfriend Xin Man who gives me confidence and supports whenever I was facing difficulties. I also want to extend my thankfulness to my officemates Yusuke, Diogo, Hanci, Alessandro, Jing, Sasha, Tyler and former

and current students of my supervisor Yanling, Dejing, Henry, Tian, Sandip, Mu, Kai, Xiaojun, Sayantee. Thank you for leaving me wonderful memories and making my life colourful.

Finally, I want to specially thank my family who have unconditionally supported me since I was born. I apologize for being away for so long.

Abbreviations

BS: Birnbaum-Sanders

CDF: cumulative distribution function

CGF: cumulant generating function

FPT: first passage time

GMM: generalized method of moments

K-S: Kolmogorov-Smirnov

MGF: moment generating function

MLE: maximum likelihood estimators

PDF: probability density function

SDE: stochastic differential equation

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Due to fatigue, wear or damages, the performance of a system will degrade over time. The failure happens when the accumulation of damage reaches a certain threshold level. Because of the requirements from markets, the $100p^{th}$ percentile of the FPT distribution is provided as an important information to reflect the reliability of the products. To model the degradation processes, there are three processes mainly used including Wiener process, gamma process and inverse Gaussian process. Discussions regarding Wiener process and its generalizations to approximate degradation processes can be seen in the literatures such as Doksum and Hóyland (1992), Doksum and Normand (1995), Whitmore (1995), Whitmore and Schenkelberg (1997), Padgett and Tomlinson (2004) and Wang (2010). But the two-directional Wiener process will cause difficulties on interpretation since it fails to possess the monotonicity property which is a typical feature for any degradation process in practice. Therefore, the monotonic processes such as gamma process and inverse Gaussian process are

the more appropriate selections to model degradation path. Gamma process is systematically studied in the area of reliability engineering by van Noortwijk (2009), Tseng *et al.* (2009) and Pan and Balakrishnan (2011). And the literatures regarding inverse Gaussian degradation process include Wang and Xu (2010), Ye and Chen (2014), Ye *et al.* (2014) and Peng (2015).

Knowing the lifetime distribution is very helpful for controlling the failure risk of manufacturing products but unfortunately only several type of continuous stochastic processes have the analytic density function so that we can find the close-form FPT density functions. For example, the FPT of Wiener process follows an inverse Gaussian distribution. Although the FPT of any monotonic processes can be approximated by Birnbaum-Sanders (BS) distribution for which only the first and second moments of the marginal distribution are required. there is always a significant bias for this method especially when the threshold is relatively small.

If the degradation process is assumed to have stationary, independent and positive increments, it can be identified as a monotonic Lévy process. The process itself and its FPT can also be called as Lévy subordinator and Inverse subordinator, respectively. In this regard, Yang and Klutke (2000) used Lévy process to characterize the properties of device lifetime distribution and Shu *et al.* (2015) proposed the cumulative degradation model with random jumps based on the characteristics of Lévy subordinators.

How to propose an appropriate model to capture the information from the actual dataset as much as possible is always a important issue. However, since there are only three types of stochastic processes potentially used in the current study, it is possible that none of them can fit the degradation data well. Once an incorrect model

is selected, the misspecification will be the major concern in degradation processes which may inevitably undermine the estimation accuracy for the survival probability and make the model inefficient. However, expanding the selection of degradation process cannot completely solve the misspecification problem which is one of the intrinsic flaws for the parametric models. Therefore, the nonparametric approaches can be more effective since it can eliminate the effects of misspecification caused by incorrect process-type assumption.

Assume the underlying process to be a Lévy process, the information is fully contained by its corresponding Laplace exponent. We can achieve the flexible selection of the underlying processes by using different Laplace exponent. But to fit a certain sample set, rather than characterizing the underlying process by the parametric exponent, it can be constructed as an empirical Lévy process in which the parametric Laplace exponent is replaced by the empirical Laplace exponent.

1.2 Models of Degradation Data and First Passage Time

In this section, we will briefly introduce the traditional parametric models used in analysing the degradation data.

1.2.1 General Degradation Path Model

Meeker and Escobar (1998) introduced the general degradation process by

$$X(t) = g(t, A),$$

where $A = (A_1, \cdots, A_n)$ is a random vector with positive components following distribution F_A and g is an increasing function and differentiable in t. For example, Levuliene (2002) suggested the tire wear data can be modelled by the linear degradation path model defined as

$$Z(t) = \frac{t}{A},$$

where A is a positive random variable.

Durham and Padgett (1997) proposed the cumulative damage approach to model the degradation data. The cumulative damage X_{n+1} after $n+1$ time intervals is related to the cumulative damage at the previous time X_n. The relation can be concluded as

$$X_{n+1} = X_n + D_n h(X_n),$$

where D_n denotes the damage incurred at the $(n+1)$st increment and $h(\cdot)$ is the damage model function loaded on the damage D_n. The newly increased damage $D_n h(X_n)$ is related to the damage accumulation at the previous time.

Park and Padgett (2005) generalized the cumulative damage model as

$$c(X_{n+1}) = c(X_n) + D_n h(X_n), \tag{1.1}$$

where $c(\cdot)$ is the damage accumulation function. Transfer this discrete process to a continuous process represented by

$$dc(X_u) = h(X_u) dD_u, \tag{1.2}$$

then the corresponding stochastic integral is given as

$$\int_0^t \frac{1}{h(X_u)} dc(X_u) = \int_0^t dD_u = D_t - D_0.$$

For example, suppose D_u is a Wiener process and $h(u) = 1$, $c(u) = u$, then the degradation process X_t is a Wiener process. The most commonly used degradation models include Wiener process, geometric Brownian motion, gamma process and inverse Gaussian process which can be represented by this general degradation process if the appropriate forms are selected.

1.2.2 Wiener Process with Drift

As mentioned before, if $h(u) = 1$ and $c(u) = u$, the discrete process (1.1) from 0 to t can be written as

$$X_t - X_0 = D_t,$$

where D_t is the Wiener process with drift $\alpha > 0$ and volatility $\sigma > 0$ following $N(\alpha t, \sigma^2 t)$.

It can be also written in the form of the stochastic differential equation (SDE)

$$dX_t = \alpha dt + \sigma dW_t,$$

where W_t is the Wiener process without drift.

If the initial point of the process is $X_0 = x_0$, then the pdf of X_t is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{(x-x_0-\alpha t)^2}{2\sigma^2 t}}.$$

It is well known that the FPT τ_C of Wiener process with a fixed threshold C defined as

$$\tau_C = \inf(t : X_t > C),\ C \geq 0,$$

follows an inverse Gaussian distribution with the pdf

$$g(\tau_C) = \frac{\sqrt{\lambda}}{\sqrt{2\pi\tau_C^3}} \exp\left(-\frac{\lambda(\tau_C - \mu)^2}{2\mu^2\tau_C}\right),$$

where $\mu = \frac{C-x_0}{\alpha}$ and $\lambda = \frac{(C-x_0)^2}{\sigma^2}$.

Then the probability of X_t not passing the threshold C up to t can be obtain by integrating the pdf of inverse Gaussian distribution as

$$P(\tau_C > t) = \Phi\left(\frac{C - x_0 - \alpha t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{-C + x_0 - \alpha t}{\sigma\sqrt{t}}\right) \cdot \exp\left(\frac{2\alpha(C - x_0)}{\sigma^2}\right).$$

1.2.3 Geometric Brownian Motion

The geometric Brownian motion is another good choice as a degradation process. If the dynamic of X_t is said to follow the geometric Brownian motion, its dynamic is given as an Itô type SDE

$$dX_t = \alpha X_t dt + \sigma X_t dW_t.$$

To write the SDE in the difference form as (1.1), we apply Itô lemma (see Appendix for details) and obtain the dynamic of the logarithm of X_t

$$d\ln(X_t) = (\alpha - \frac{\sigma^2}{2})dt + \sigma dW_t,$$

which gives the difference form as

$$\ln(X_t) - \ln(X_0) = (\alpha - \frac{\sigma^2}{2})t + \sigma W_t.$$

Then if let $c(u) = \ln(u)$, $h(u) = 1$, and D_t as Wiener process with drift $\mu - \frac{\sigma^2}{2}$, the general degradation model can be specified as the geometric Brownian motion.

Denote τ_C as the FPT of the geometric Brownian motion X_t to threshold C, the pdf of τ still follows an inverse Gaussian distribution

$$g(\tau_C) = \frac{\sqrt{\lambda}}{\sqrt{2\pi\tau_C^3}} \exp\left(-\frac{\lambda(\tau_C - \mu)^2}{2\mu^2\tau_C}\right),$$

where $\mu = \frac{\ln(C)-\ln(x_0)}{\alpha-\frac{\sigma^2}{2}}$ and $\lambda = \frac{(\ln(C)-\ln(x_0))^2}{\sigma^2}$.

Hence, the probability the FPT is great than t is given by

$$P(\tau_C > t) = \Phi\left(\frac{\ln(C) - \ln(x_0 - (\alpha - \frac{\sigma^2}{2})t)}{\sigma\sqrt{t}}\right)$$
$$-\Phi\left(\frac{-\ln(C) + \ln(x_0) - (\alpha - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) \cdot \left(\frac{C}{x_0}\right)^{\frac{2(\alpha-\frac{\sigma^2}{2})}{\sigma^2}}.$$

1.2.4 Jump Process

We can define the continuous stochastic processes as following:

Definition 1.2.1 (Continuity in Time)

X is continuous at t_0 implies if $t \to t_0$, X_t converges to X_{t_0} in probability. X is continuous in time if this holds for all t_0.

Definition 1.2.2 (Continuous in Sample Paths)

A process X is continuous at t_0 if for almost all ω, $t \to t_0$ implies $X(t,\omega) \to X(t_0,\omega)$. A process is continuous in sample paths if, for almost all ω, $X(\cdot,\omega)$ is a continuous function.

Different with diffusion processes like the Wiener process or the geometric Brownian motion, a jump process is only continuous in time but discontinuous in sample paths. The gamma process and the inverse Gaussian process are two typical jump processes with only nonnegative increments. Consider the X_t is denoted as a jump process with $h(u) = 1$ and $c(u) = u$ in the general model. The cumulative damage at time t is given as

$$X_t - X_0 = D_t,$$

where D_t is assumed to be a gamma process or a inverse Gaussian process and initial point $X_0 = x_0$.

The marginal distribution of the gamma process is gamma distribution. The gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$ has pdf

$$f(x;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x),\ x > 0,$$

where $\Gamma(\cdot)$ is the gamma function.

Another definition of the gamma process is proposed by Berman (1981) who

considered that an event occurs at time zero and only every k^{th} event can be observed for a Poisson process with intensity rate $\lambda(t)$. This process is called non-homogeneous gamma process with rate function $\lambda(t)$ and shape parameter k. As a special case when $k = 1$, the process becomes a non-homogeneous Poisson process.

The gamma distribution also holds the additivity property. If X_i has a $Gamma(\alpha_i, \beta)$ distribution for $i = 1, \ldots, n$ such that all of the variables are independent and share the same rate parameter β, then their summation is given as

$$\sum_{i=1}^{n} X_i \sim Gamma\left(\sum_{i=1}^{n} \alpha_i, \beta\right).$$

With the additivity property, the gamma process can be constructed as a stochastic process $\{X_t, t \geq 0\}$ if

1. The increment $X_{t+s} - X_s$ follows a gamma distribution $Gamma(\alpha t, \beta)$ with shape parameter α and rate parameter β. The initial state $X_0 = 0$ with probability one.

2. If two time intervals $[t_1, t_2]$ and $[t_3, t_4]$are disjoint, the increments $X_{t_2} - X_{t_1}$ and $X_{t_4} - X_{t_3}$ are independent.

3. The gamma process is continuous in probability.

At any time point t, the position X_t has the distribution $Gamma(\alpha t, \beta)$.

The marginal distribution of an inverse Gaussian process follows the inverse Gaussian distribution with mean $\mu > 0$ and the shape parameter $\lambda > 0$. Its pdf is given by

$$f(x; \mu, \lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right), \ x > 0.$$

The inverse Gaussian distribution also holds the additivity property. If X_i has an inverse Gaussian distribution $IG(\mu\omega_i, \lambda\omega_i^2)$ for $i = 1, \ldots, n$ and all X_i are independent then

$$\sum_{i=1}^{n} X_i \sim IG\left(\mu \sum_{i=1}^{n} \omega_i, \lambda \left(\sum_{i=1}^{n} \omega_i\right)^2\right).$$

For $\omega_i = 1$, each $X_i \sim IG(\mu, \lambda)$, then

$$\sum_{i=1}^{n} X_i \sim IG(n\mu, n^2\lambda).$$

Wasan (1968) defined inverse Gaussian process as a stochastic process $\{X_t, t \geq 0\}$ with properties including

1. The increment $X_{t+s} - X_s$ follows an inverse Gaussian distribution $IG(\mu t, \lambda t^2)$ with mean μ and shape parameter λ. The initial state $X_0 = 0$ with probability one.

2. If two time intervals $[t_1, t_2]$ and $[t_3, t_4]$are disjoint, the increments $X_{t_2} - X_{t_1}$ and $X_{t_4} - X_{t_3}$ are independent.

3. The inverse Gaussian process is continuous in probability.

At any time point t, the position X_t has the distribution $IG(\mu t, \lambda t^2)$.

For the connection between the gamma process and inverse Gaussian process, Dufresne *et al.* (1991) showed the gamma process is the limit of a compound Poisson process with jump size conforming to a certain distribution and Ye and Chen (2014) showed that the inverse Gaussian process is also a limiting compound Poisson process with different jump size distribution.

The FPT densities of the gamma procecss and the inverse Gaussian process can be computed by the numerical Laplace inversion or approximated by the saddlepoint method. Park and Padgett (2005) also suggested to approximate the pdf by BS distribution if only the number of increments to threshold is very large. More details regarding the FPT of these two models will be further discussed in later chapters.

1.3 Laplace Transform

It is difficult to study the statistical properties of a stochastic process since it is common that the pdf for the sum of its increments is not available in close form. However, once the Laplace transform of the marginal distribution exists in the analytic form, we can always characterize the distribution of any point along the degradation paths by its Laplace transform.

Definition 1.3.1

The Laplace transform of $f(t)$, $t \geq 0$ *denoted by* $\mathcal{L}\{f(t)\}$ *is a statistical transform which transfers the variable from* t *to* s. *The Laplace transform can be written in a new function with variable* s

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t) dt.$$

As $f(t)$ is the function of time continuous on $[0, \infty)$, the Laplace transform to the frequency domain is one-to-one. The Laplace transform $F(s)$ can uniquely determine the function $f(t)$.

For example, let $f(t) = t$, then

$$F(s) = \int_0^\infty e^{-st} f(t)dt = \int_0^\infty e^{-st} t dt = \frac{1}{s^2}, \ s > 0.$$

The propositions of the Laplace transform will be discussed in Chapter 2.

1.4 Motivating Dataset

The light output for a laser device decreases due to degradation while it increases the operating current to maintain the constant level of light output. However, once the operating current is so high that reaches the failure threshold, the device will be considered to have failed. Some devices may fail suddenly with the possible reasons such as a hidden sudden change in the physical state of the device, manufacturing defects, the change of failure mode caused by high temperature and unexpected shocks happen to the device.

Suppose an experiment is conducted that a sample of lasers devices are tested at the temperature of 80°C and percentages increase in operating current are recorded. The data of this experiment is collected from Meeker and Escobar (1998) which contains 15 laser units, with measure frequency every 250 hours, and the experiment got terminated at 4000 hours. The dataset will be used throughout the thesis.

Figure 1.1 shows the plots of the 15 degradation paths. It can be seen that all the degradation processes can pass the dashed line which is the threshold assumed as 6. But if the threshold value becomes larger, the failure time of some units may not be observed up to the termination time 4000.

Even if all of the failure times of the units are observed, the sample size 15 is

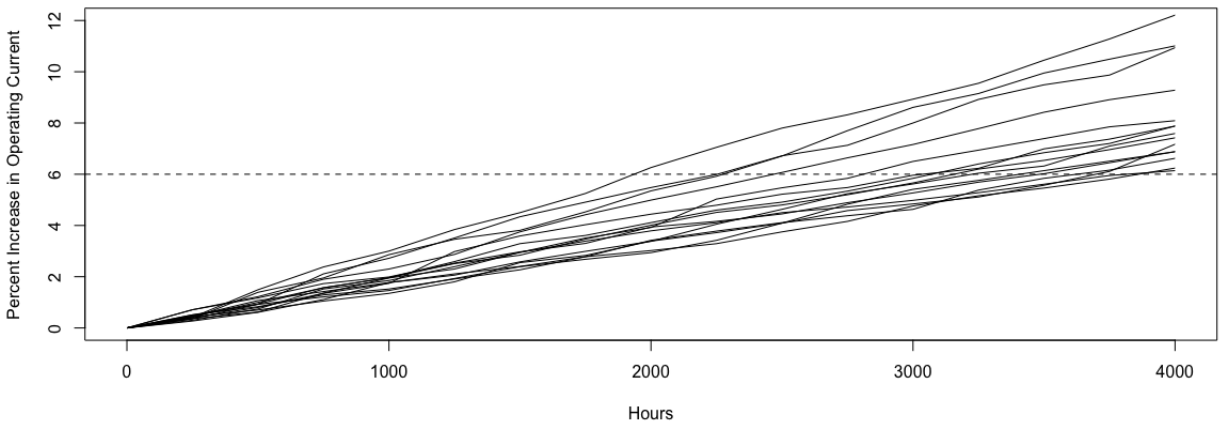

Figure 1.1: Degradation paths for laser data

still too small to make any inference for the density function or other properties of failure times. Conducting more times for the experiment will increase the economic burden so it is unreasonable to collect large number of failure time observations and then make inference. Compared with experimental cost, the measurement cost for the degradation paths are much cheaper then finding the statistical properties of failure time based on measured increments should be very beneficial to practice.

Therefore, rather than observing actual failure times, our idea is to alternatively observe the increments of single or a small number of test units. And all of the statistical inference towards the FPT will be based on the increments dataset and the threshold value.

1.5 Scope of the Thesis

The thesis evaluates the FPT distribution of the degradation processes through both parametric and nonparametric perspectives. In the parametric perspective, the study of more complicated processes can be realized in addition to the traditional degradation processes discussed by literatures in the past decades. Using the proposed methods, the close form of Laplace transform of the FPT density function can be derived and its inverse Laplace transform can be either obtained by numerical method or approximated by the saddlepoint method. Specifically, the weighted convolution of two independent gamma processes incorporated with random effects is exemplified as a 'complicated process' in Chapter 2. Since its likelihood function is intractable, the parameters can be alternatively estimated by GMM. For the nonparametric approach proposed in Chapter 3, we can make inferences to the FPT distribution without observing any failure time nor giving any process-type assumption, and the approximated FPT distribution only relies on the empirical Laplace transform of the underlying process. As an important application of the nonparametric method in reliability engineering, the optimal design for degradation test is discussed in Chapter 4. To optimize the degradation test, the design factors related to the experiment such as the number of test units, the number of measurements and the length of time interval are considered. Under the constraint that the experimental cost cannot exceed the pre-specified budget, the optimal design are determined by minimizing the bootstrap estimate of variance for the 100pth percentile of the FPT distribution.

Chapter 2

Parametric Evaluation of the First Passage Time of Degradation Processes

2.1 Introduction

This chapter discusses the parametric way to evaluate the first passage time (FPT) of the degradation process. As a generalization of the regular degradation process, the underlying process investigated in our study is assumed as weighted convolution of multiple degradation processes incorporated with random effects.

Consider a monotonic degradation process with stationary and independent increments, it is appropriate to assume it as an one-sided Lévy process and characterize it with Laplace exponent. Such a Lévy process can also be called as Lévy subordinator and its FPT is called as Inverse subordinator. In this regard, Yang and Klutke (2000) used Lévy process to characterize the properties of device lifetime distribution and

Shu *et al.* (2015) proposed the cumulative degradation model with random jumps based on the characteristics of Lévy subordinators.

By using Lévy subordinator, Eliazar and Klafter (2004) derived an analytic expression for the FPT density through the double Laplace transform and the n^{th} moment through the single Laplace transform. There are many methods to numerically obtain the FPT density by Laplace inversion as discussed by Veillette and Taqqu (2010) and Todorov and Tauchen (2012). The saddlepoint approximation is proposed by Daniels (1954) which can be used as an alternate for numerical inversion of the Laplace transform.

The weighted-convolution model is motivated by the scenario of multiple sensors utilized on monitoring a crucial component. Since the degradation information provided by a single sensor may be incomplete and unreliable, the multi-sensor monitoring can obtain more accurate reliability estimates. Although the degradation processes derived from these sensors reflect the same degradation path but their scales and noise are significantly different. To find the unified degradation path, Hua *et al.* (2013) used the weight-averaged unified approach which determined the weights by the defined leadership scores. Specifically, the degradation process generated from each sensor is modelled by gamma process which is widely used for the degradation path modelling in past studies such as Park and Padgett (2005), Tseng *et al.* (2009), Pan and Balakrishnan (2011).

Moreover, Chen and Ye (2016) mentioned that since the failure data in same database are often collected from many different data sources, it is necessary to investigate the heterogeneities by introducing random effects to model the variability. The random effects can be reflected by replacing the parameters of the underlying

process with the certain distributed variables. Lawless and Crowder (2004) and Tsai *et al.* (2012) incorporated random effects on gamma process, and Peng (2015) discussed the random effects when the underlying process is inverse Gaussian process.

Although the parameter estimation of degradation processes are commonly based on MLE, it is hard to find close-form likelihood function for many degradation processes without the important feature as gamma process or inverse Gaussian process that is closed under convolution. The likelihood function of the weighted convolution of gamma processes with random effects is intractable but its Laplace transform has the closed-form expression. Therefore, the parameters can be efficiently estimated by minimizing the distance between the parametric Laplace transform and empirical Laplace transform. This estimation approach can be realized through GMM. GMM was proposed by Hansen (1982) and the estimation methods of the empirical probability transforms were discussed by Carrasco and Florens (2000), Singleton (2001) and Yu (2004). A stochastic GMM was also proposed by Yin *et al.* (2011).

We define the model of weighted convolution gamma process with random effects in Section 2.2. And in Section 2.3 the Laplace transform of its FPT is derived in the analytic expression and two methods are discussed to invert the Laplace transform in Section 2.3. In Section 2.4, the the distance-based parameter estimation method is used due to intractable likelihood function. As an important indicator of reliability information, the estimated $100p^{th}$ percentiles based on numerical Laplace inversion and saddlepoint approximation are compared with the simulated percentiles in Section 2.5. A practical dataset is analyzed in Section 2.6.

2.2 Underlying Process

In this section, the definition of Lévy process and infinite divisible law are presented. Then we investigate the weighted convolution of gamma processes with random effects and derive its Laplace transform.

2.2.1 Lévy Process

The Lévy process is continuous in probability, has independent and identically distributed increments. The Poisson process and the Wiener process are fundamental examples of Lévy processes which are also the components of general Lévy processes as every Lévy process can be decomposed into a a deterministic drift, a Wiener process and a compound Poisson process. The strict definition of Lévy process is given as follow.

Definition 2.2.1 (Lévy Process)

A stochastic process X_t, $t \geq 0$ with initial value $X_0 = 0$ is called a Lévy process if it possesses the following properties:

1. *Independent increments: for every increasing sequence of times $t_0, \ldots, t_n$, the random variables X_{t_0}, $X_{t_1} - X_{t_0}$, $\ldots$, $X_{t_n} - X_{t_{n-1}}$ are independent.*

2. *Stationary increments: the law of $X_{t+h} - X_t$ does not depend on t.*

3. *Stochastic continuity: $\forall \epsilon > 0$,*

$$\lim_{h \to 0} \mathbb{P}\left(|X_{t+h} - X_t| \geq \epsilon\right) = 0$$

.

If a Lévy process is sampled at a regular time intervals 0, Δ, 2Δ, ..., we can write $S_n(\Delta) = X_{n\Delta} = \sum_{k=0}^{n-1} Y_k$ where $Y_k = X_{(k+1)\Delta} - X_{k\Delta}$ are i.i.d. random variables with the same distribution as X_Δ. Then the model $S_n(\Delta)$ can be specified by the sampling the Lévy process X at different frequencies. When $n\Delta = t$, for any $t > 0$ and $n \geq 1$, $X_t = S_n(\Delta)$ can be represented as a sum of n i.i.d. random variables with the same distribution as $X_{t/n}$. Then we can say X_t can be divided into n i.i.d. parts. The definition of infinitely divisible law is given by

Definition 2.2.2 (Infinite Divisibility)

A probability distribution F is said to be infinitely divisible if for any integer $n \geq 2$, there exists n i.i.d. random variables $Y_1, \ldots, Y_n$ such that $Y_1 + \cdots + Y_n$ has distribution F.

For example, if $X \sim N(\mu, \sigma^2)$ then we can write $X = \sum_{k=0}^{n-1} Y_k$ where Y_k are i.i.d. with $N(\mu/n, \sigma^2/n)$. Other common examples of infinitely divisible laws also include the gamma distribution, α-stable distribution and the Poisson distribution. On the other hand, given an infinitely divisible distribution F, with its n i.i.d. components, we can construct stochastic processes on a time grid with step size $1/n$ such that the distribution of X_1 is given by F. Then we have the proposition given in Tankov and Cont (2004).

Proposition 2.2.1 (Infinite Divisibility and Lévy Processes)

Let X_t, $t \geq 0$ be a Lévy process. Then for every t, X_t has an infinitely divisible distribution. Conversely, if F is an infinitely divisible distribution then there exists a Lévy process X_t such that the distribution of X_1 is given by F.

The Lévy process can be established by the Lévy-Khintchine formula (Bertoin

(1996))

$$\phi(\omega) = \mu\omega + \frac{1}{2}\sigma^2\omega^2 + \int_0^\infty \left(1 - e^{-\omega x}\right) \Pi(dx),$$

where μ is known as drift coefficient, σ is the volatility of the Brownian motion component, and Π is the measure satisfying $\int_0^\infty (1 \wedge x)\, \Pi(dx) < \infty$.

If a Lévy process is a pure jump process without drift or Brownian motion parts, and each jump has only positive direction, it can be called as subordinator.

Let $\{Z_t, t \geq 0\}$ be a Lévy subordinator starting from 0. It is characterized by its Laplace transform

$$\mathcal{L}(Z_t) = \mathbb{E}\left[e^{-\omega Z_t}\right] = e^{-t\phi(\omega)},\ \omega \geq 0, \tag{2.1}$$

where the notation $\mathcal{L}\{\cdot\}$ represents the Laplace transform. The function ϕ is called the Laplace exponent and is expressed by the Lévy-Khintchine formula for the one-sided Lévy process as

$$\phi(\omega) = \mu\omega + \int_0^\infty \left(1 - e^{-\omega x}\right) \Pi(dx).$$

where the Brownian motion part is removed.

For $t > s$, by writing $Z_{t+s} = Z_s + (Z_{t+s} - Z_s)$ and using the fact that $Z_{t+s} - Z_s$ is independent of Z_s, the Laplace transform is multiplicative.

Proposition 2.2.2

$L(Z_{t+s}) = \mathcal{L}(Z_{t+s} - Z_s) \cdot \mathcal{L}(Z_s) = \mathcal{L}(Z_t) \cdot \mathcal{L}(Z_s)$.

To model the degradation process, we only take the jump part of Lévy process which has independent and non-negative increments. Considering gamma jump

process Z_t, $t \geq 0$ where Z_t represents the degradation path for a testing unit at time t, then we have following properties

1. $Z_0 = 0$ with probability one;
2. $Z_t - Z_s \sim Gamma\left(\nu(t) - \nu(s), \beta\right)$ for all $t > s \geq 0$ where the shape parameter $\nu(t)$ is a non-decreasing, right-continuous, real-valued function for $t \geq 0$ with $\nu(0) = 0$ and $\beta > 0$ is scale parameter;
3. Z_t has independent increments.

To incorporate random effects to the model, the scale parameter β is assumed to follow a gamma distribution with parameters μ and λ and its pdf is given by

$$f(\beta) = \frac{\lambda^{\mu}}{\gamma(\mu)}\beta^{\mu-1}e^{-\lambda\beta}.$$

2.2.2 Weighted Convolution of Gamma Processes

If $Z_t^{(1)}$ and $Z_t^{(2)}$ are two independent Lévy processes, the weighted convolution can be written as

$$Z_t = (1-\rho)Z_t^{(1)} + \rho Z_t^{(2)}, \tag{2.2}$$

where ρ is the weight on $Z_t^{(2)}$. It can be seen that Z_t is also a Lévy process with corresponding Laplace transform as

$$\mathcal{L}\left(Z_t\right) = \mathbb{E}\left[e^{-\omega Z_t}\right] = \mathbb{E}\left[e^{-\omega(1-\rho)Z_t^{(1)}}\right] \cdot \mathbb{E}\left[e^{-\omega\rho Z_t^{(2)}}\right] = e^{-t\left(\phi^{(1)}((1-\rho)\omega)+\phi^{(2)}(\rho\omega)\right)},\ \omega \geq 0,$$

where $\phi^{(1)}(\cdot)$ and $\phi^{(2)}(\cdot)$ are Laplace exponents of $Z_t^{(1)}$ and $Z_t^{(2)}$ respectively. Then the Laplace exponent of the degradation process Z_t is $\phi(\omega) = \phi^{(1)}\left((1-\rho)\omega\right) + \phi^{(2)}\left(\rho\omega\right)$.

Specifically, $Z_t^{(1)}$ and $Z_t^{(2)}$ are assumed to be two independent gamma processes with random effects. The shape parameter function $\nu(t)$ is considered as a linear function $\nu(t) = \alpha t$. For a given time t, each process conditioned on β_i, $i = 1, 2$ follow gamma distributions as

$$Z_t^{(i)} \sim \text{Gamma}(\alpha_i t, \beta_i),\ i = 1, 2$$

with shape parameters $\alpha_i t$ and scale parameters β_i, $i = 1, 2$. Moreover, the scale parameters can be further assumed to follows gamma distribution as

$$\beta_i \sim \text{Gamma}(\mu_i, \lambda_i),\ i = 1, 2. \tag{2.3}$$

Then the Laplace transform of Z_t can be written as

$$\begin{aligned}
\mathbb{E}\left(e^{-\omega Z_t}\right) &= \mathbb{E}\left(\mathbb{E}\left(e^{-\omega Z_t}|\beta_1,\beta_2\right)\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(e^{-\omega(1-\rho)Z_t^{(1)}}|\beta_1\right)\cdot\mathbb{E}\left(e^{-\omega\rho Z_t^{(2)}}|\beta_2\right)\right) \\
&= \int_0^\infty \mathbb{E}\left(e^{-\omega(1-\rho)Z_t^{(1)}}|\beta_1\right) f(\beta_1)d\beta_1 \cdot \int_0^\infty \mathbb{E}\left(e^{-\omega\rho Z_t^{(2)}}|\beta_2\right) f(\beta_2)d\beta_2 \\
&= \int_0^\infty \left(1+\frac{\omega(1-\rho)}{\beta_1}\right)^{-\alpha_1 t} \frac{\lambda_1^{\mu_1}}{\Gamma(\mu_1)}\beta_1^{\mu_1-1}e^{-\lambda_1\beta_1}d\beta_1 \\
&\quad \cdot \int_0^\infty \left(1+\frac{\omega\rho}{\beta_2}\right)^{-\alpha_2 t} \frac{\lambda_2^{\mu_2}}{\Gamma(\mu_2)}\beta_2^{\mu_2-1}e^{-\lambda_2\beta_2}d\beta_2 \\
&= \frac{\Gamma(\mu_1+\alpha_1 t)}{\Gamma(\mu_1)}\left(\omega\lambda_1(1-\rho)\right)^{\mu_1} U(\mu_1+\alpha_1 t,\mu_1+1,\lambda_1\omega(1-\rho)) \\
&\quad \cdot \frac{\Gamma(\mu_2+\alpha_2 t)}{\Gamma(\mu_2)}\left(\omega\lambda_2\rho\right)^{\mu_2} U(\mu_2+\alpha_2 t,\mu_2+1,\lambda_2\omega\rho),
\end{aligned}$$

where the confluent hypergeometric Kummer U function is defined as

$$U(a,b,z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}dt.$$

We list some basic properties of confluent hypergeometric functions according to Slater (1960). There are two types of confluent hypergeometric functions, namely

$$M(a,b,z) = \sum_{n=0}^{\infty}\frac{(a)_n}{(b)_n}\cdot\frac{z^n}{n!},$$

where $(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = a(a+1)\ldots(a+n-1)$, $(a)_0 = 1$, and

$$U(a,b,z) = \frac{\pi}{\sin\pi b}\left\{\frac{M(a,b,z)}{\Gamma(1+a-b)\Gamma(b)} - z^{1-b}\frac{M(1+a-b,2-b,z)}{\Gamma(a)\Gamma(2-b)}\right\}.$$

The function $U(a,b,z)$ is analytic for all values of a, b and z, even when b is zero

or a negative integer. It can be represented as

$$U(a,b,z) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} dt,$$

for those values of a, b and z for which the integral exists.

The n^{th} derivative of $U(a,b,z)$ is given as

$$\frac{d^n}{dz^n} U(a,b,z) = (-1)^n (a)_n U(a+n, b+n, z).$$

The asymptotic behaviours for $U(a,b,z)$ as $z \to 0$ include

$$\begin{aligned} U(a,b,z) &\sim \frac{\Gamma(1-b)}{\Gamma(1+a-b)} \text{ if } b < 1, \\ &\sim -\frac{1}{\Gamma(a)} \left(\log(z) + \phi(a) - 2\gamma\right) \text{ if } b = 1, \\ &\sim \frac{\Gamma(b-1)}{\Gamma(a)} z^{1-b} \text{ if } b > 1, \end{aligned}$$

where $\phi(a) = \frac{\Gamma'(a)}{\Gamma(a)}$ and γ is Euler-Mascheroni constant defined as

$$\gamma = \lim_{n\to\infty} \left(-\log(n) + \sum_{k=1}^{n} \frac{1}{k} \right) = \int_1^\infty \left(\frac{1}{\lfloor x \rfloor} - \frac{1}{x} \right) dx.$$

Here, $\lfloor x \rfloor$ represents the floor function.

2.3 First Passage Time

The FPT of $\{Z_t, t \geq 0\}$ with threshold x is another process $\{\tau_x, x \geq 0\}$ which is also called as inverse subordinator. Its definition is given by

$$\tau_x = \inf(t : Z_t > x),\ x \geq 0;$$

evidently the probability of first passage time no earlier than t is simply

$$P(\tau_x > t) = P(Z_t < x). \tag{2.4}$$

The probability $\mathrm{P}(\tau_x > t)$ is called as the survival probability which represents the probability of an unit survives longer than t. The Laplace transform of the survival probability is given by the following proposition.

Proposition 2.3.1

The Laplace transform of survival probability $P(\tau_x > t)$, with respect to x, is given by

$$\mathcal{L}\{P(\tau_x > t)\} = \frac{\exp\{-\phi(\omega)t\}}{\omega}, \tag{2.5}$$

Proof. If f is a smooth and bounded function and X is a non-negative random variable, then

$$\mathbb{E}\left[f(X)\right] = f(0) + \int_0^\infty f'(x)P(X > x)dx.$$

Apply this property to Laplace transform of Z_t, it yields

$$\mathbb{E}\left[\exp\{-\omega Z_t\}\right] = 1 - \int_0^\infty \omega \exp\{-\omega x\} P\left(Z_t > x\right) dx.$$

Using (2.1) and (2.4) gives

$$\exp\{-\phi(\omega)t\} = \omega \int_0^\infty \exp\{-\omega x\} P\left(\tau_x > t\right) dx,$$

and so

$$\int_0^\infty \exp\{-\omega x\} P\left(\tau_x > t\right) dx = \frac{\exp\{-\phi(\omega)t\}}{\omega}.$$

2.3.1 Inverse Laplace Transform

Let $\mathcal{L}^{-1}\{\cdot\}$ denote the operator of Inverse Laplace transform. Then the survival probability $P(\tau_x > t)$ can be obtained by inverting the Laplace transform as

$$P(\tau_x > t) = \mathcal{L}^{-1}\{\frac{\exp(-\phi(\omega)t)}{\omega}\}(x). \tag{2.6}$$

The Inverse Laplace transform is defined as

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{st} F(s) ds,$$

where the value γ is chosen such that the integration is done along the vertical line $Re(s) = \gamma$ and $F(s)$ is analytic in the region $Re(s) \geq \gamma$.

For many degradation processes, the close-form representations of (2.6) are unavailable due to the difficulties in the evaluation of the complex integrals but there are many numerical methods to approximate the inverse Laplace transform. In R and Matlab, the 'INVLAP' function was created to realize such computation based on the algorithm

proposed by Valsa and Brancik (1998). Once a underlying degradation process is specified, the numerical method can help us to find the true value of the survival probability.

To realize the computation, Valsa and Brancik (1998) approximate the exponential function by $e^{st} \approx E_c(st,a) = \frac{e^a}{2cosh(a-st)} = \frac{e^{st}}{1+e^{-2a}e^{2st}}$. When $a > \gamma t$, $\frac{e^{st}}{1+e^{-2a}e^{2st}}$ can be expanded into MacLaurin series $e^{st} + \sum_{n=1}^{\infty}(-1)^n e^{-2na}e^{(2n+1)st}$ such that

$$f(t) \approx f_c(t,a) = f(t) + \sum_{n=1}^{\infty}(-1)^n e^{-2na})f[(2n+1)t].$$

where the error, which is the sum term, can be controlled by choose the parameter a.

Expand $E_c(st,a)$ with respect to st by applying $\frac{1}{cosh(z)} = 2\pi\sum_{n=0}^{\infty}\frac{(-1)^n(n+1/2)}{(n+1/2)^2\pi^2+z^2}$ we have $E_c(st,a) = \pi e^a \sum_{n=0}^{\infty}\frac{(-1)^n(n+1/2)}{(n+1/2)^2\pi^2+(a-st)^2}$, therefore

$$f_c(t,a) = \frac{e^a}{2i}\int_{\gamma-i\infty}^{\gamma+i\infty} F(s)\pi e^a \sum_{n=0}^{\infty}\frac{(-1)^n(n+1/2)}{(n+1/2)^2\pi^2+(a-st)^2}ds.$$

By interchanging the integration and summation, and then integrating along the path of a semicircle with the infinite radius, we can obtain

$$f_c(t,a) = -\frac{e^a}{t}\sum_{n=0}^{\infty}(-1)^n \mathrm{Im}\left\{F\left[\frac{a}{t} + i(n+\frac{1}{2})\frac{\pi}{t}\right]\right\}$$

or

$$f_c(t,a) = \frac{e^a}{t}\sum_{n=0}^{\infty}(-1)^n \mathrm{Im}\left\{F\left[\frac{a}{t} + i(n-\frac{1}{2})\frac{\pi}{t}\right]\right\}.$$

2.3.2 Saddlepoint Approximation

Alternatively, rather than numerically computing the Laplace inversion, the saddlepoint approximation, which was initially introduced by Daniels (1954), can be used to approximate the inverse Laplace transform. This method provides a highly accurate approximation formula for any pdf or probability mass function of a distribution, based on the moment generating function (MGF) or other statistical transforms.

Given a close-form MGF $M(s)$, the saddlepoint approximation $\hat{f}(t)$ for the density is given by

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi K''(\hat{s})}} \exp[K(\hat{s}) - \hat{s}t], \tag{2.7}$$

where $K(s) = \log[M(s)]$ is the cumulant generating function (CGF) and $\hat{s}$ is the solution to the saddlepoint equation $K'(\hat{s}) = t$. The proof can be seen in Appendix.

Apply the equation (2.7), the Laplace Inversion of survival probability can be approximated by

$$\hat{P}(\tau_x > t) = \frac{1}{\sqrt{-2\pi\psi''(\hat{\omega})}} \exp\{-\psi(\hat{\omega}) + \hat{\omega}x\}, \tag{2.8}$$

where $\psi(\hat{\omega}) = -\log\left(\int_{-\infty}^{\infty} \exp\{-\hat{\omega}x\} P(\tau_x > t)dx\right)$ is the Laplace exponent of the survival probability and $\hat{\omega}$ satisfies $\psi'(\hat{\omega}) = x$.

Since the Laplace transform of survival probability is already known from Proposition 2.3.1, the Laplace exponent of survival probability $\psi(\omega)$ can be represented by the Laplace exponent of increment $\phi(\omega)$ as

$$\psi(\omega) = -\log\left(\frac{e^{-\phi(\omega)t}}{\omega}\right) = \phi(\omega)t + \log(\omega). \tag{2.9}$$

Its first- and second-order derivatives are given by

$$\begin{aligned}\psi'(\hat{\omega}) &= \phi'(\hat{\omega})t + \frac{1}{\hat{\omega}},\\ \psi''(\hat{\omega}) &= \phi''(\hat{\omega})t - \frac{1}{\hat{\omega}^2},\end{aligned}$$

and the estimate of $\hat{\omega}$ satisfies the equation

$$\phi'(\hat{\omega})t + \frac{1}{\hat{\omega}} = x.$$

By Equation (2.2) and Equation (2.3), the Laplace transform of the survival probability $P(\tau_x > t)$ is obtained as

$$\begin{aligned}\mathcal{L}\{P(\tau_x > t)\} &= \frac{\mathbb{E}\left(e^{-\omega Z_t}\right)}{\omega}\\ &= \frac{\Gamma(\mu_1+\alpha_1 t)}{\Gamma(\mu_1)}\left(\lambda_1(1-\rho)\right)^{\mu_1} U(\mu_1+\alpha_1 t, \mu_1+1, \lambda_1\omega(1-\rho))\\ &\quad \cdot\frac{\Gamma(\mu_2+\alpha_2 t)}{\Gamma(\mu_2)}\left(\lambda_2\rho\right)^{\mu_2} U(\mu_2+\alpha_2 t, \mu_2+1, \lambda_2\omega\rho)\\ &\quad \cdot\omega^{\mu_1+\mu_2-1}.\end{aligned}$$

The Laplace transform can be inverted by numerical Laplace inversion or using the saddlepoint approximation. To find the saddlepoint approximation, the first and

second-order derivatives of

$$\begin{aligned}
\psi(\omega) &= -\log\left(\mathcal{L}\{P(\tau_x > t)\}\right) \\
&= -\log\left(\frac{\Gamma(\mu_1+\alpha_1 t)}{\Gamma(\mu_1)}\left(\lambda_1(1-\rho)\right)^{\mu_1} U(\mu_1+\alpha_1 t, \mu_1+1, \lambda_1\omega(1-\rho))\right) \\
&\quad -\log\left(\frac{\Gamma(\mu_2+\alpha_2 t)}{\Gamma(\mu_2)}\left(\lambda_2\rho\right)^{\mu_2} U(\mu_2+\alpha_2 t, \mu_2+1, \lambda_2\omega\rho)\right) \\
&\quad -(\mu_1+\mu_2-1)\log(\omega)
\end{aligned}$$

need to be calculated.

Since $U(a,b,z)$ has derivative

$$\frac{d}{dz}U(a,b,z) = -aU(a+1,b+1,z),$$

the first and second-order derivatives of $\psi(\omega)$ are given by

$$\begin{aligned}
\psi'(\omega) &= -\frac{\mu_1+\mu_2-1}{\omega} + (1-\rho)\lambda_1(\mu_1+\alpha_1 t)\cdot\frac{U(\mu_1+\alpha_1 t+1, \mu_1+2, \omega(1-\rho)\lambda_1)}{U(\mu_1+\alpha_1 t, \mu_1+1, \omega(1-\rho)\lambda_1)} \\
&\quad +\rho\lambda_2(\mu_2+\alpha_2 t)\cdot\frac{U(\mu_2+\alpha_2 t+1, \mu_2+2, \omega\rho\lambda_2)}{U(\mu_2+\alpha_2 t, \mu_2+1, \omega\rho\lambda_2)}, \\
\psi''(\omega) &= \frac{\mu_1+\mu_2-1}{\omega^2} + (1-\rho)\lambda_1(\mu_1+\alpha_1 t)\cdot\frac{A}{U^2(\mu_1+\alpha_1 t, \mu_1+1, \omega(1-\rho)\lambda_1)} \\
&\quad \rho\lambda_2(\mu_2+\alpha_2 t)\cdot\frac{B}{U^2(\mu_2+\alpha_2 t, \mu_2+1, \lambda_2\omega\rho)},
\end{aligned}$$

where

$$\begin{aligned} A &= -\lambda_1(1-\rho)(\mu_1+\alpha_1 t+1)U(\mu_1+\alpha_1 t+2,\mu_1+3,\lambda_1\omega(1-\rho))U(\mu_1+\alpha_1 t,\mu_1+1,\lambda_1\omega(1-\rho)) \\ &\quad +\lambda_1(1-\rho)(\mu_1+\alpha_1 t)U^2(\mu_1+\alpha_1 t+1,\mu_1+2,\lambda_1\omega(1-\rho)) \\ B &= -\lambda_2\rho(\mu_2+\alpha_2 t+1)U(\mu_2+\alpha_2 t+2,\mu_2+3,\lambda_2\omega\rho)U(\mu_2+\alpha_2 t,\mu_2+1,\lambda_2\omega\rho) \\ &\quad +\lambda_2\rho(\mu_2+\alpha_2 t)U^2(\mu_2+\alpha_2 t+1,\mu_2+2,\lambda_2\omega\rho). \end{aligned}$$

Hence, the approximated survival probability can be obtained by Equation (2.8).

2.4 Parameter Estimation

Although the MLE approach is commonly used to estimate parameters, it is difficult to conduct under some certain circumstances. For example, there is no analytic-form density function for the convolution of two weighted beta prime distribution (gamma distribution with random effects), then its likelihood function does not have a tractable expression. But since there is a one-to-one relationship between the Laplace transform and the pdf of a random variable, an estimation method based on the distance between the parametric Laplace transform and the empirical Laplace transform is potentially an efficient way to estimate parameters. Feuerverger and McDunnough (1981a) showed that such approach can be realized by matching empirical characteristic function ϕ_n with parametric characteristic function ϕ_θ with a dynamic weight function and the parameters should satisfy the following equation

$$\int \omega_\theta(t)\left(\phi_n(t)-\phi_\theta(t)\right)dt = 0,$$

where the weighted function relies on the true parameters θ. Similarly, other empirical transforms including the empirical Laplace transform or empirical moment generating functions applied in estimation methods are also discussed by Braun *et al.* (2008).

2.4.1 Generalized Method of Moments

The GMM is proposed by Hansen (1982). Given the independent and identically distributed observations $X_1, \ldots, X_n$, there exists a function f such that $m(\theta) = \mathbb{E}\left(f(X_i;\theta)\right) = 0, i = 1, 2, \ldots, n$ where m is norm function. By the strong law of large numbers, the average $\hat{m}(\theta) = \frac{1}{n}\sum_i^n f(X_i;\theta)$ converges to $m(\theta)$ almost surely. The GMM method is looking for $\hat{\theta}$ such that $\hat{m}(\hat{\theta})$ is close to zero. The distance is defined as $\hat{m}(\theta)^T W \hat{m}(\theta)$ where W is a positive definite matrix. Schmidt (1982) considered the case $W = \Sigma^{-1}$ where Σ is asymptotic variance of $\hat{m}(\theta)$. The GMM estimator $\hat{\theta}$ is obtained by minimizing the distance function $\hat{m}(\theta)^T \Sigma^{-1} \hat{m}(\theta)$.

If $L_\theta(\omega) = \mathbb{E}\left(e^{-\omega Y_1}\right)$ and $L_n(\omega) = \frac{1}{n}\sum_{i=1}^n e^{-\omega Y_i}$ are respectively the theoretical Laplace transform and the empirical Laplace transform where $Y_i = Z_{t_i} - Z_{t_{i-1}}$ is increment between t_{i-1} and t_i ($t_0, \ldots, t_n$ are equally spaced), the distance between the transforms converges by Gilvenko-Cantelli theorem

$$\sup_{\omega} |L_\theta(\omega) - L_n(\omega)| \to 0, \text{as } n \to \infty.$$

Selecting a set of discrete grids of $\omega = (\omega_1, \omega_2, \ldots, \omega_q)$ and letting $\mathbf{K}_\theta$ be a vector $(L_\theta(\omega_1), \ldots, L_\theta(\omega_q))^T$ and $\mathbf{K}_n$ be the vector of empirical Laplace transforms as $(L_n(\omega_1), \ldots, L_n(\omega_q))^T$, then the GMM estimator $\hat{\theta}$ is defined as

$$\hat{\theta} = \text{argmin}_\theta \left(K_n(\omega) - K_\theta(\omega)\right)^T \Sigma^{-1} \left(K_n(\omega) - K_\theta(\omega)\right), \tag{2.10}$$

where the asymptotic covariance matrix Σ has the elements

$$\begin{aligned}
&\mathrm{Cov}\left(L_\theta(u) - L_n(u), L_\theta(v) - L_n(v)\right) = \mathrm{Cov}\left(L_n(u), L_n(v)\right) \\
&= \mathrm{Cov}\left(\frac{\sum_{i=1}^n e^{-uY_i}}{n}, \frac{\sum_{i=1}^n e^{-vY_i}}{n}\right) \\
&= \frac{1}{n^2}\left(\mathbb{E}\left(\sum_{i=1}^n e^{-(u+v)Y_i}\right) + \sum_{i\neq j}\mathbb{E}\left(e^{-uY_i - vY_j}\right)\right) - \mathbb{E}\left(\frac{\sum_{i=1}^n e^{-uY_i}}{n}\right)\cdot\mathbb{E}\left(\frac{\sum_{i=1}^n e^{-vY_i}}{n}\right) \\
&= \frac{nL(u+v) + n(n-1)L(u)L(v)}{n^2} - L(u)L(v) = \frac{1}{n}\left(L(u+v) - L(u)L(v)\right). \qquad (2.11)
\end{aligned}$$

The first problem of the grids selection is to determine the value of q. If the dimension of the parameter space Θ is p, then the number of grids q should be selected no smaller than p (Yao and Morgan (1999)). The second problem is the values of $\omega_1, \ldots, \omega_q$ when q is determined. Schmidt (1982) suggested the grids should minimize the determinant of the asymptotic covariance matrix and they are close to each other. Feuerverger and McDunnough (1981b) showed that the asymptotic variance can be made arbitrarily close to the Cramer-Rao bound by selecting large value of q and making the grids sufficiently fine and extended. They also suggested the grids should be chosen with equal spacing, $\omega_k = \omega_0 + k\tau$, and the estimation will be efficient when $k \to \infty$ and $\tau \to 0$. However, when the grids are too fine, the covariance matrix becomes singular and the GMM estimator is unavailable to be computed. (Carrasco and Florens (2002)).

2.4.2 Estimation Algorithm and Procedures

It is difficult to simultaneously optimize the selection of grids and find the optimal parameters by (2.10). Before the proper grids are chosen to minimize $|\Sigma|$, the

parameter values need be known. Meanwhile, before θ is estimated, the selected grids should be used. To conduct these estimation procedures, it is necessary to separate the optimization into several stages.

First, rather than taking ω as a set of discrete points to determine the weight matrix, ω can be taken as a continuous variable. Considering the weight function as 1, the parameters are estimated by minimizing the integral of $(L_\theta(\omega) - L_n(\omega))^2$ with respect to ω.

Second, with parameter estimations obtained from the first step, the grids are optimally designed to minimize the determinant of asymptotic covariance matrix $|\Sigma|$. To reduce the computational burden, the number of grids q is set to be equal to the number of parameters p. To make them equally spaced as $\omega_k = \omega_0 + k\tau$, $k = 1, \dots p$, only the values of ω_0 and τ are searched.

Third, once the grids are determined, the GMM estimator can be obtained through (2.10). For all the optimization algorithms in this article, we applied Differential Evolution. Differential Evolution is an efficient global optimization algorithm proposed by Storn and Price (1997). It iteratively tries to improve the candidate solution with the certain fitness criterion.

Therefore, the three-step parameter estimations algorithm is given as following:

1 Find $\hat{\theta}_0$ such that $\hat{\theta}_0 = \text{argmin}_\theta \int_0^\infty (L_\theta(\omega) - L_n(\omega))^2 \, d\omega$;

2 Minimize $|\Sigma|$ by selecting the equally spaced $\omega_k = \omega_0 + k\tau$, $k = 1, \dots p$;

3 Use the selected grids to find GMM estimator $\hat{\theta}$ according to (2.10).

In the Section 2.6, the practical data is given by a multi-dimensional data set rather than a single-dimensional vector and the variations between test units are

considered. Then the details regarding how to determine $L_\theta(\omega)$ and $L_n(\omega)$ will be discussed later.

2.5 Simulation Studies

The random effects on $100p^{th}$ percentile of the FPT distribution are investigated when the underlying process is a the weighted convolution of gamma process incorporated with random effects under two different sets of pre-assumed parameters. The percentile of the FPT distribution can be obtained through approaches proposed in the previous sections and the results from both of the saddlepoint approximation and the numerical inverse Laplace transform are showed.

As the model previously assumed, the gamma convolution process can be written as

$$Z_t = (1-\rho)Z_t^{(1)} + \rho Z_t^{(2)},$$

where $Z_t^{(1)}$ and $Z_t^{(2)}$ are independent with respective weights $1-\rho$ and ρ.

As the gamma distribution is not closed under convolution when the rate parameters are different distributed random variables, the density function of any point along the degradation path is unavailable in analytic form.

Assume the $Z_t^{(1)}$ and $Z_t^{(2)}$ are two independent gamma processes with randomized scale parameters and Z_t is defined as the model in section 2.2 with two parameter sets $(\rho, \alpha_1, \mu_1, \lambda_1, \alpha_2, \mu_2, \lambda_2) = (0.8, 10, 2, 1, 5, 3, 1)$ and $(0.3, 8, 3, 1, 6, 2, 1)$.

In the Table 2.1 and Table 2.2, the raw simulation of the $100p^{th}$ percentile of the first passage time is taken as benchmark. Each simulated t_p is the $100p^{th}$ percentile

from 10,000 simulated first passage times and the mean and standard deviation of t_p are obtained after 1000 times of simulation. The percentiles of 90^{th}, 50^{th}, 10^{th} obtained by both numerical inverse Laplace transform and the saddlepoint approximation are presented under different threshold values of 5, 10 and 15. It can been seen that percentiles obtained by numerical inverse Laplace transform and the saddlepoint approximation are very close in values. All of the results obtained by numerical Laplace inversion fall inside the 95% confidence intervals of simulated percentiles and all of the results obtained by saddlepoint approximation fall inside the 95% condidence interval under 90^{th} and 50^{th} percentiles. For 10^{th} percentile, the saddlepoint approximations are slightly away from the 95% confidence interval. The confidence interval is constructed using the means and standard deviations given in the Table 2.1 and Table 2.2 as 'mean±1.96*standard deviation'.

Table 2.1: The $100p^{th}$ percentile of the first passage time obtained by numerical inverse Laplace transform, the saddlepoint approximation, mean and standard deviation of percentile from raw simulation (from top to bottom) with the underlying process as gamma convolution process parametrized as $(\rho, \alpha_1, \mu_1, \lambda_1, \alpha_2, \mu_2, \lambda_2) = (0.8, 10, 2, 1, 5, 3, 1)$ when threshold is assumed as 5, 10, 15 and p is assumed as 0.9, 0.5, 0.1.

	5	10	15
0.9	3.1594	6.1639	9.1664
	3.1577	6.1602	9.1608
	3.1538(0.0227)	6.1597(0.0450)	9.1624(0.0659)
0.5	1.6669	3.3033	4.9397
	1.6651	3.2987	4.9326
	1.6629(0.0123)	3.3006(0.0229)	4.9380(0.0335)
0.1	0.6871	1.4166	2.1481
	0.7478	1.5313	2.3175
	0.6995(0.0093)	1.4251(0.0195)	2.1515(0.0299)

In Figure 2.1, the $100p^{th}$ percentile t_p is obtained by the numerical Laplace

Table 2.2: The $100p^{th}$ percentile of the first passage time obtained by numerical inverse Laplace transform, the saddlepoint approximation, mean and standard deviation of percentile from raw simulation (from top to bottom) with the underlying process as gamma convolution process parametrized as $(\rho, \alpha_1, \mu_1, \lambda_1, \alpha_2, \mu_2, \lambda_2) = (0.3, 8, 3, 1, 6, 2, 1)$ when threshold is assumed as 5, 10, 15 and p is assumed as 0.9, 0.5, 0.1.

	5	10	15
0.9	2.6919	5.2461	7.7988
	2.6904	5.2430	7.7941
	2.6811(0.0204)	5.2393(0.0369)	7.7956(0.0565)
0.5	1.4264	2.8235	4.2205
	1.4247	2.8194	4.2141
	1.4220(0.0097)	2.8225(0.0206)	4.2189(0.0282)
0.1	0.5950	1.2280	1.8626
	0.6451	1.3215	2.0001
	0.6097(0.0079)	1.2346(0.0165)	1.8656(0.0240)

inversion. The counterpart by the saddlepoint approximation is not provided since their results are very close for higher percentile and the numerical Laplace inversion performs better for lower percentile. When p gets larger, t_p is increasing with a trend. The difference in t_p is also shown between the degradation models with various parameter values of μ_1 and μ_2 reflecting the variation of the convolution gamma process Z_t when the scale parameters follow the gamma distribution with different shape parameters. The other parameters of the underlying process in the subplot (a) and subplot (b) are fixed as $(\rho, \alpha_1, \lambda_1, \alpha_2, \lambda_2) = (0.8, 10, 1, 5, 1)$ and $(\rho, \alpha_1, \lambda_1, \alpha_2, \lambda_2) = (0.3, 8, 1, 6, 1)$, respectively.

2.6 Illustrative Data Analysis

In this section, we will analyze the reliability of the laser devices using the dataset introduced in Chapter 1. The underlying model is the convolution of multiple weighted

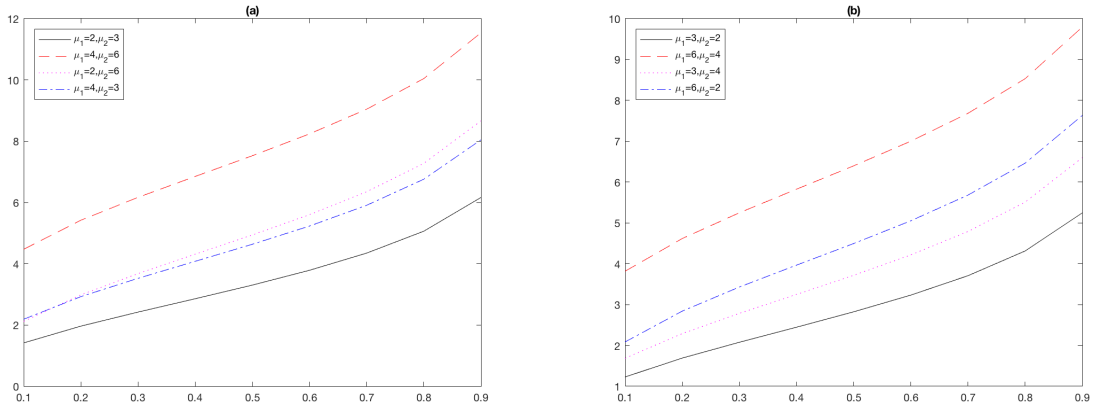

Figure 2.1: Plots of t_p for different shape parameters of β_1 and β_2 while other parameters are fixed as $(\rho, \alpha_1, \lambda_1, \alpha_2, \lambda_2) = (0.8, 10, 1, 5, 1)$ for subplot (a) and $(\rho, \alpha_1, \lambda_1, \alpha_2, \lambda_2) = (0.3, 8, 1, 6, 1)$ for subplot (b)

degradation processes can be explained as the laser device is monitored by two sensors in practice. The degradation paths derived from the sensors are different and assumed as two independent gamma processes. As the difference of experimental environments and data sources, the variations among the testing units are necessary to be considered and reflected by the randomization of scale parameters of the gamma processes of the two sensors. To unify the two degradation processes, the weight on each sensor is also required to be estimated.

As previously mentioned in Section 2, the Laplace transform of the marginal distribution of Z_t can be calculated as

$$\begin{aligned} L_\theta(\omega) &= \mathbb{E}\left(e^{-\omega Z_1}\right) = \mathbb{E}\left(\mathbb{E}\left(e^{-\omega Z_1}|\beta_1, \beta_2\right)\right) \\ &= \frac{\Gamma(\mu_1+\alpha_1)}{\Gamma(\mu_1)}\left(\omega\lambda_1(1-\rho)\right)^{\mu_1} U(\mu_1+\alpha_1, \mu_1+1, \lambda_1\omega(1-\rho)) \\ &\quad \cdot\frac{\Gamma(\mu_2+\alpha_2)}{\Gamma(\mu_2)}\left(\omega\lambda_2\rho\right)^{\mu_2} U(\mu_2+\alpha_2, \mu_2+1, \lambda_2\omega\rho). \end{aligned}$$

For the $n \times m$ data set with n test units and m measurements for each unit, the empirical Laplace transform of j^{th} column (each unit) can be written as $L_m^{(j)}(\omega) = \frac{\sum_{i=1}^{m} e^{-\omega Y_{ij}}}{m}$ where Y_{ij} is the increment from time $i-1$ to i for j^{th} unit. Since the degradation process for each unit follows gamma process conditioned on certain scale parameters, then $L_m^{(j)}(\omega) = \frac{\sum_{i=1}^{m} e^{-\omega Y_{ij}}}{m}$ for the test units $j = 1, \ldots, n$ are corresponded to the parametric Laplace transforms $\mathbb{E}\left(e^{-\omega Z_1} | \beta_1, \beta_2\right)$ with different parameters.

Since $L_\theta(\omega) = \mathbb{E}\left(\mathbb{E}\left(e^{-\omega Z_1} | \beta_1, \beta_2\right)\right)$, its counterpart empirical Laplace transform is obtained by the average of the empirical Laplace transform of each test unit

$$L_{mn}(\omega) = \frac{\sum_{j=1}^{n} L_m^{(j)}}{n} = \frac{\sum_{j=1}^{n} \sum_{i=1}^{m} e^{-\omega Y_{ij}}}{nm}.$$

Using the proposed three-step estimation algorithm proposed in section 3.2.2, the parameters are finally estimated as

$$(\hat{\rho}, \hat{\alpha}_1, \hat{\mu}_1, \hat{\lambda}_1, \hat{\alpha}_2, \hat{\mu}_2, \hat{\lambda}_2) = (0.8422, 8.5142, 23.4739, 3.8668, 3.0867, 18.9608, 1.9226),$$

and the grids selected in step 2 are $\omega_i = 0.0464 + 0.2513i$ where $i = 1, \ldots, 7$. Then estimated percentile $\hat{t}_p$ can be obtained by plugging these parameter estimations to (2.6).

Besides Laplace inversion and saddlepoint approximation, the $100p^{th}$ percentiles of the FPT obtained from the original data set are given in Table 2.3. When the values of threshold are taken as 2, 4, 6, the FPT for all of the 15 test units can be obtained by linear interpolation if the intercept of the degradation paths and thresholds are between two actual observations. In other words, if the two known points are given by the coordinates (x_0, y_0) and (x_1, y_1), the linear interpolant is the straight line between

the two points. For a value x in the interval (x_0, x_1), the value y along the straight line is given from the equation of slopes

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0},$$

and solve the equation for y in terms of x, gives

$$y = \frac{y_0(x_1 - x) + y_1(x - x_0)}{x_1 - x_0},$$

which is the linear interpolation in the interval (x_0, x_1).

From Table 2.3, the $100p^{th}$ percentiles t_p are given when $p = 0.9, 0.5, 0.1$ under the thresholds $2, 4, 6$. t_p obtained by numerical Laplace inversion, saddlepoint approximation and linear interpolation from measured degradation paths are very close. This illustrates that the proposed method performs well for the given dataset. To more accurately evaluate the performance of the approaches applied to the data set, a larger sample of experimental units are required.

2.7 Concluding Remarks

The parametric approach proposed in this chapter implements the scenario when there are multiple sensors to monitor the same component. It gives more freedom on the assumption of underlying processes which are mostly limited to gamma process or inverse Gaussian process in most studies. The Laplace inversion methods discussed in this chapter can help us to find the FPT even for the complicated degradation processes. Moreover, the parameters estimation method based on GMM can solve

Table 2.3: The $100p^{th}$ percentile of the first passage time for the laser data obtained by numerical inverse Laplace transform, the saddlepoint approximation and the $100p^{th}$ percentile revealed from the actual observations (from top to bottom) when the underlying process is assumed as gamma convolution process with estimated parameters $(\rho, \alpha_1, \mu_1, \lambda_1, \alpha_2, \mu_2, \lambda_2) = (0.8422, 8.5142, 23.4739, 3.8668, 3.0867, 18.9608, 1.9226)$, threshold is assumed as 2, 4, 6 and p is assumed as 0.9, 0.5, 0.1.

	2	4	6
	1338.53	2559.28	3781.05
0.9	1330.30	2556.35	3776.85
	1302.73	2455.53	3769.10
	992.85	2011.20	3010.68
0.5	1002.00	2175.53	2984.68
	1035.71	2018.02	3157.41
	720.53	1524.88	2327.73
0.1	697.00	1463.60	2237.90
	741.50	1464.95	2262.18

the problem when the likelihood function is mathematically intractable and the good performance is shown in the section of simulation and practical data analysis.

Although the gamma convolution process with random effects is taken as underlying process, more types of degradation processes can be studied by using the same approach such as weighted convolution of inverse Gaussian processes with random effects.

Both of the numerical Laplace inversion and saddlepoint approximation perform nice for median and right tail of FPT distribution and the numerical Laplace inversion performs better than saddlepoint approximation to capture the left-tail behaviour of FPT distribution. To improve the accuracy of the saddlepoint approximation, the higher-order terms of the $\psi(\omega)$ should be included.

Chapter 3

Nonparametric Evaluation of the First Passage Time of Degradation Processes

3.1 Introduction

This chapter discusses the nonparametric method of approximating the FPT density of degradation processes without observing any failure time or making any assumption on the type of process. Davison and Hinkley (1988) developed the empirical saddlepoint approximation by replacing the parametric Laplace exponent to the empirical Laplace exponent. Then the empirical saddlepoint approximation can estimate the density based on the sample set without being effected by the restraints of distributional assumptions. The parametric FPT density can be derived if the true underlying process is known which is taken as the standard to evaluate the performances of the nonparametric methods.

As a nonparametric benchmarks, BS distribution can approximate the FPT density by knowing the moments of the marginal distribution. The two-parameter BS distribution was originally proposed by Birnbaum and Saunders (1969a), Birnbaum and Saunders (1969b) as a failure time distribution. Park and Padgett (2005) showed that the failure time can be modelled by accelerated test versions of BS distribution driven by geometric Brownian motion and gamma process.

Another benchmark is to use a certain model to fit the degradation data. The most commonly used models are gamma process and inverse Gaussian process as we previously discussed. Select a certain underlying process then we can hypothesize a parametric process given the parameter θ and find the MLE $\hat{\theta}$ by the sample set. This approach is better when the correct parametric model is used. However, if the assumed model is incorrect, the nonparametric approach is superior which improves the model performance as directly estimating the Laplace transform by the empirical Laplace transform. Also, the nonparametric method will not be affected by misspecification. In this chapter, we will evaluate whether the nonparametric method can outperform the traditional parametric method of assuming the degradation process as a certain process and the nonparametric approximation as BS approach. The parametric method to obtain the FPT is the same as what discussed in Chapter 2.

The content of this chapter is organized as follow. We introduce the definition and properties of the empirical Laplace transform and develop the empirical saddlepoint approximation based on it to obtain the FPT density. Then we show that FPT can also be approximated by BS distribution and find its flaws by evaluating its performances by assuming the underlying process as gamma process and inverse Gaussian process. The accuracy criterion used is relative error and with its use

the saddlepoint approximation is compared with other two benchmarks when the underlying process is misspecified. The estimation and the corresponding standard error are obtained through bootstrap procedure. Some primary advantages of the proposed method are demonstrated through a simulation study. A laser degradation dataset is analyzed and the Kolmogorov-Smirnov goodness-of-fit test is used to test the adequacy of the model. Bootstrap confidence intervals are also constructed for the percentiles of the FPT distribution. Finally, some concluding comments are made.

3.2 Nonparametric Model of First Passage Time

This section describes two major nonparametric approaches to approximate the FPT density including the empirical saddlepoint approximation and BS distribution. We can obtain the Laplace inversion of Equation (2.6) by the empirical saddlepoint approximation in which the parametric Laplace transform is replaced by the empirical Laplace transform. On the other hand, BS approach only need the information of the first and second moments to approximate the FPT of monotonic degradation processes.

3.2.1 Empirical Laplace Transform

We will show that the empirical Laplace transform has the linearity, convergence and convolution properties.

Corresponding to the Laplace transform (2.1), the empirical Laplace transform is

defined as

$$L_n(\omega) = \int_0^\infty e^{-\omega x} dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} e^{-\omega X_i},$$

where F_n is the empirical distribution of the sample observations $X_1, \cdots, X_n$ defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x).$$

Suppose $\{X_i\}_{1\leq i\leq n}$ is a random sample from distribution F and $\{Y_j\}_{1\leq j\leq m}$ is a random sample from distribution G and their empirical distribution are respectively

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x),\ G_m(y) = \frac{1}{m}\sum_{j=1}^{m} I(Y_j \leq y),$$

and let $L^{F_n}(\omega)$ and $L^{G_m}(\omega)$ be the corresponding empirical transforms. Then the following lemma shows a series of propositions of the empirical transform which provides more insights into behaviours of the empirical transforms.

Lemma 3.2.1

1. *(Convolution) Suppose $L^{F_n \star G_m}(\omega)$ is the empirical Laplace transform of the convolution $F_n \star G_m$, then $L^{F_n \star G_m} = L^{F_n}(\omega) \cdot L^{G_m}(\omega)$.*

2. *(Linearity) Suppose A and B are constants, then the empirical Laplace transform of the linear combination of the two distributions $A \cdot F_n + B \cdot G_m$ is given by $A \cdot L^{F_n} + B \cdot L^{G_m}$.*

3. *(Convergence) L^{F_n} is an unbiased and consistent estimator of L^F.*

Proof. 1. The convolution of F_n and G_m is

$$\begin{aligned}
F_n \star G_m(z) &= \int_0^\infty F_n(z-y)dG_m(y) \\
&= \int_0^z F_n(z-y)dG_m(y) \\
&= \frac{1}{nm}\int_0^z \left[\sum_{i=1}^n I(X_i \le z-y)\right] \sum_{j=1}^m I(Y_j = y; 1 \le j \le m)dy \\
&= \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m \int_0^z I(X_i \le z-y)\cdot I(Y_j = y; 1 \le j \le m)dy \\
&= \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m I(X_i \le z - Y_j) \\
&= \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m I(X_i + Y_j \le z).
\end{aligned}$$

Then the Laplace transform of the convolution is

$$\begin{aligned}
L^{F_n \star G_m}(\omega) &= \int_0^\infty e^{-\omega z} d[F_n \star G_m(z)] \\
&= \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m \int_0^\infty e^{-\omega z} I(X_i + Y_j = z; 1 \le i \le n, 1 \le j \le m)dz \\
&= \frac{1}{nm}\sum_{i=1}^n \sum_{j=1}^m e^{-\omega(X_i+Y_j)} \\
&= \left(\frac{1}{n}\sum_{i=1}^n e^{-\omega X_i}\right)\left(\frac{1}{m}\sum_{j=1}^m e^{-\omega Y_j}\right) \\
&= L^{F_n}(\omega)\cdot L^{G_m}(\omega).
\end{aligned}$$

Hence, the Laplace transform of the convolution is the product of every single Laplace transform.

2. The Laplace transform of the empirical distribution function $A \cdot F_n + B \cdot G_m$ is

$$\begin{aligned} L^{A\cdot F_n + B\cdot G_m}(\omega) &= \int_0^\infty e^{-\omega z} d\left((A \cdot F_n + B \cdot G_m)(z)\right) \\ &= A\int_0^\infty e^{-\omega z} dF_n(z) + B\int_0^\infty e^{-\omega z} dG_m(z) \\ &= A \cdot L^{F_n}(\omega) + B \cdot L^{G_m}(\omega). \end{aligned}$$

Hence, the linearity holds.

3. As X_i are i.i.d. F-distributed, the expectation of L^{F_n} is

$$\begin{aligned} \mathbb{E}[L^{F_n}(\omega)] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n e^{-\omega X_i}\right] \\ &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}[e^{-\omega X_i}] \\ &= \mathbb{E}[e^{-\omega X_i}] \\ &= L^F(\omega). \end{aligned}$$

Then, the empirical Laplace transform L^{F_n} is an unbiased estimator of L^F.

By central limit theorem,

$$\sqrt{n}\left(\frac{\sum_{i=1}^n e^{-\omega X_i}}{n} - \mathbb{E}(e^{-\omega X})\right) \xrightarrow[n\to\infty]{d} N(0, \sigma^2(\omega)),$$

where

$$\sigma^2(\omega) = Var(e^{-\omega X}) = L^F(2\omega) - (L^F(\omega))^2$$

is a finite number. Then $\frac{\sigma^2(\omega)}{n}$ goes to zero when $n \to \infty$. Hence,

$$P(\lim_{n\to\infty} \sup |L^{F_n} - L^F| < \epsilon) = 1 \text{ for all } \epsilon > 0.$$

□

3.2.2 Empirical Saddlepoint Approximation

The parametric saddlepoint approximation requires the underlying Laplace exponent which may not be known. If there is no assumption made about the distribution of the increment, it is important to nonparametrically estimate the Laplace exponent by measuring the increment multiple times to get the observations. With these observations, substitution of the empirical Laplace exponent in place of unknown parametric Laplace exponent can be used to develop the empirical saddlepoint approximation.

The Laplace exponent $\phi(\omega)$ is replaced by the empirical Laplace exponent

$$\tilde{\phi}(\omega) = -\log\left(\frac{1}{n}\sum_{i=1}^{n} e^{-\omega Y_i}\right), \tag{3.1}$$

where $Y_1, \cdots, Y_n$ are the increments measured with equal-length time section.

However, the empirical Laplace transform may result in ill-posed problem when inverting this Laplace transform in numerical methods. More details regarding ill-posed problem can be referred to Appendix. To avoid this problem, the empirical saddlepoint approximation method is taken as an alternate to the numerical inversion of empirical Laplace transform.

The empirical Laplace exponent of survival probability is equal to

$$\tilde{\psi}(\hat{\omega}) = \tilde{\phi}(\hat{\omega})t + \log(\hat{\omega}),$$

and its first- and second-order derivatives are given by

$$\begin{aligned} \tilde{\psi}'(\hat{\omega}) &= \tilde{\phi}'(\hat{\omega})t + \frac{1}{\hat{\omega}}, \\ \tilde{\psi}''(\hat{\omega}) &= \tilde{\phi}''(\hat{\omega})t - \frac{1}{\hat{\omega}^2}, \end{aligned}$$

where the derivatives of the empirical Laplace exponent of increment are calculated as

$$\begin{aligned} \tilde{\phi}'(\hat{\omega}) &= \frac{\sum_{i=1}^n Y_i e^{-\hat{\omega}Y_i}}{\sum_{i=1}^n e^{-\hat{\omega}Y_i}}, \\ \tilde{\phi}''(\hat{\omega}) &= \frac{(\sum_{i=1}^n Y_i e^{-\hat{\omega}Y_i})^2 - (\sum_{i=1}^n Y_i^2 e^{-\hat{\omega}Y_i})(\sum_{i=1}^n e^{-\hat{\omega}Y_i})}{(\sum_{i=1}^n e^{-\hat{\omega}Y_i})^2}. \end{aligned}$$

The estimate of $\hat{\omega}$ satisfies the equation

$$\tilde{\psi}'(\hat{\omega}) = \frac{\sum_{i=1}^n Y_i e^{-\hat{\omega}Y_i}}{\sum_{i=1}^n e^{-\hat{\omega}Y_i}} t + \frac{1}{\hat{\omega}} = x,$$

so that the survival probability can be approximated as

$$\tilde{P}(\tau_x > t) = \frac{1}{\sqrt{-2\pi\tilde{\psi}''(\hat{\omega})}} \exp\{-\tilde{\psi}(\hat{\omega}) + \hat{\omega}x\}. \tag{3.2}$$

3.2.3 The Birnbaum-Saunders Distribution Approach

Park and Padgett (2005) discussed the BS distribution as a nonparametric approach

for the FPT distribution through the use of the Central Limit Theorem. It is unnecessary to know the empirical Laplace exponent as the empirical saddlepoint approximation, the BS approach only requires sample mean and variance.

Suppose the lifetime X follows a two-parameter BS distribution, $BS(\alpha, \beta)$, with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$. The cumulative distribution function (CDF) is given by

$$F(x; \alpha, \beta) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)\right], \ x > 0 \tag{3.3}$$

where $\Phi(\cdot)$ is the standard normal CDF; the PDF of x corresponding to the CDF in (3.3) is

$$f(x; \alpha, \beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\left[\left(\frac{\beta}{x}\right)^{1/2} + \left(\frac{\beta}{x}\right)^{3/2}\right] \exp\left[-\frac{1}{2\alpha^2}\left(\frac{x}{\beta} + \frac{\beta}{x} - 2\right)\right], \ x > 0.$$

The relationship between BS distribution and mixture inverse Gaussian distribution is discussed by Balakrishnan *et al.* (2009). Consider the inverse Gaussian distribution denoted by $IG(\alpha, \beta)$, with pdf $f(x) = \left(\frac{2\pi\alpha^2 x^3}{\beta}\right)^{-1/2} \exp\left(-\frac{(x-\beta)^2}{2\alpha^2\beta x}\right)$, $x, \alpha, \beta > 0$, Jorgensen *et al.* (1991) propose the mixture of inverse Gaussian distribution with its complementary reciprocal which is defined as

$$X_1 \sim \text{IG}(\alpha, \beta) \text{ and } X_2^{-1} \sim \text{IG}(\frac{\alpha}{\beta^2}, \frac{1}{\beta}),$$

where X_2 is called complementary reciprocal of X_1 and the new distribution is given

by

$$X = \begin{cases} X_1 \text{ with probability } 1-\rho \\ X_2 \text{ with probability } \rho, \end{cases}$$

X is called mixture inverse Gaussian distribution.

When $\rho = \frac{1}{2}$, the mixture inverse Gaussian distribution defined above can be written in the form of BS distribution.

The approximation of FPT by BS distribution can be justified as below if the degradation paths are assumed as discrete process. Let N_x denote the number of increments when the process reaches the threshold x and $\{Z_k, k \geq 0\}$ be the strictly increasing process having the identical and independent non-negative increments $Y_i = Z_i - Z_{i-1}$, $i = 1, 2, ...$, with $Z_0 \equiv 0$.

As in the case of continuous processes, the FPT and the number of increments satisfy the relation

$$\mathbb{P}(N_x > n) = \mathbb{P}(Z_n < x) = \mathbb{P}(\sum_{i=0}^{n} Y_i < x).$$

By Central Limit Theorem, when n is large, the probability of the FPT is given by

$$\begin{aligned} \mathbb{P}(N_x < n) &= 1 - \mathbb{P}(\sum_{i=0}^{n} Y_i < x) \approx 1 - \Phi\left(\frac{x - \mathbb{E}[\sum_{i=0}^{n} Y_i]}{\sqrt{Var(\sum_{i=0}^{n} Y_i)}}\right) \\ &= \Phi\left(\sqrt{\frac{x\mathbb{E}[Y_1]}{Var(Y_1)}}\left(\sqrt{\frac{n}{x/\mathbb{E}[Y_1]}} - \sqrt{\frac{x/\mathbb{E}[Y_1]}{n}}\right)\right). \end{aligned}$$

Upon setting the parameters

$$\alpha = \sqrt{\frac{\mu_2 - \mu_1^2}{x\mu_1}} \text{ and } \beta = \frac{x}{\mu_1}, \tag{3.4}$$

where μ_n represents the n^{th} moment of the increment Y_1, the FPT can be taken as BS distribution with parameters α and β with CDF as in (3.3).

With moments μ_1, μ_2 and threshold x, the mean and variance of the FPT N_x are given by

$$\begin{aligned}
\mathbb{E}[N_x] &= \frac{x}{\mu_1} + (\frac{\mu_2}{2\mu_1^2} - \frac{1}{2}), \\
\text{Var}(N_x) &= (\frac{\mu_2 - \mu_1^2}{\mu_1^3})x + (\frac{5\mu_2^2}{4\mu_1^4} - \frac{\mu_2}{2\mu_1^2} + \frac{5}{4}).
\end{aligned}$$

Smith (1959) gave the approximated mean and variance of the renewal process N_x as below when x is large

$$\begin{aligned}
\mathbb{E}[N_x] &\approx \frac{x}{\mu_1} + \frac{\mu_2}{2\mu_1^2}, \\
\text{Var}(N_x) &\approx (\frac{\mu_2 - \mu_1^2}{\mu_1^3})x + (\frac{5\mu_2^2}{4\mu_1^4} - \frac{\mu_2}{2\mu_1^2} - \frac{2\mu_3}{3\mu_1^3}),
\end{aligned}$$

where $\mu_r = E[Y_1^r]$ $(r = 1, 2, ...)$ and Y_1 is a renewal.

Comparing the means and variances of the BS distribution and those of the renewal process, it can be seen that the BS distribution tends to have smaller expectation but larger variance than the renewal process if the moments of each increment are finite. For both the BS distribution and the renewal process, the ratios $\mathbb{E}[N_x]/x$ and $Var(N_x)/x$ converge to $\frac{1}{\mu_1}$ and $\frac{\mu_2 - \mu_1^2}{\mu_1^3}$, respectively, when $x \to \infty$.

Bertoin *et al.* (1999) and Lageras (2005) also discussed the relations between

renewal process and Lévy subordinators. Each subordinator has the corresponding inverse subordinator which represents its FPT process. From Proposition 3 of Lageras (2005), if the mean of some specific subordinator μ_1 is finite, then the expectation of the FPT τ_x, $E[\tau_x] \sim \frac{x}{\mu_1}$ as $x \to \infty$.

If the degradation path is the summation of the increments, the FPT to a threshold level x follows BS distribution when x is much greater than each increment. However, the performance of BS approach under small threshold level deserves discussions. In the following parts, we will look at the bias between the estimated FPT obtained by BS approach and the true FPT based on numerical Laplace inversion. Besides evaluating the FPT by BS approach, the results of parametric saddlepoint approximation are also presented. The two exemplified models are gamma and inverse Gaussian process which are often directly used to model the degradation data. The expressions of survival probability in Laplace transform for these two processes will be obtained as well as the corresponding estimated forms derived from the parametric saddlepoint approximation. The parameters of BS distribution can be expressed in terms of the parameters of the underlying degradation processes (namely, gamma and inverse Gaussian).

Gamma Process

Gamma process is a degradation process with positive gamma-distributed increments. The parameters of the gamma distribution are, respectively, α_G as shape parameter and β_G as rate parameter.

Suppose the marginal distribution of gamma process has PDF

$$f(x) = \frac{\beta_G^{\alpha_G}}{\Gamma(\alpha_G)} x^{\alpha_G - 1} e^{-\beta_G x},\ x > 0;$$

its Laplace exponent is given by

$$\phi(\omega) = \alpha_G \log(1 + \frac{\omega}{\beta_G}).$$

Then, the survival probability in the Laplace transform, with respect to x, under the gamma process can be expressed as

$$\mathcal{L}\{P(\tau_x > t)\} = \frac{\exp\{-\alpha_G \log(1 + \frac{\omega}{\beta_G})t\}}{\omega},$$

and the corresponding Laplace exponent is

$$\psi(\omega) = \alpha_G \log(1 + \frac{\omega}{\beta_G})t + \log(\omega).$$

To approximate the inverse Laplace transform by the saddlepoint method, the first- and second-order derivatives are given by

$$\begin{aligned} \psi'(\omega) &= \frac{\alpha_G t}{\beta_G + \omega} + \frac{1}{\omega}, \\ \psi''(\omega) &= -\frac{\alpha_G t}{(\beta_G + \omega)^2} - \frac{1}{\omega^2}, \end{aligned}$$

and then find $\hat{\omega}$ through

$$\psi'(\hat{\omega}) = \frac{\alpha_G t}{\beta_G + \hat{\omega}} + \frac{1}{\hat{\omega}} = x.$$

Using the results above, the survival probability can be approximated by the parametric saddlepoint approximation as shown in Equation (2.8).

Alternatively, if the FPT is approximated by the BS distribution, its parameters α and β can be estimated from the first two moments of the marginal distribution of the underlying degradation process. The first and second moments of gamma distribution are given by

$$\mu_1 = \frac{\alpha_G}{\beta_G} \quad \text{and} \quad \mu_2 = \frac{\alpha_G^2 + \alpha_G}{\beta_G^2}.$$

Using Equation (3.4), the parameters of the BS distribution can then be expressed, using these two moments and threshold x, as

$$\begin{aligned} \hat{\alpha} &= (\beta_G x)^{-\frac{1}{2}}, \\ \hat{\beta} &= x\alpha_G^{-1}\beta_G. \end{aligned} \tag{3.5}$$

Then, for a fixed value of x, the survival probability $\hat{P}(\tau_x > t)$ can be directly approximated by the CDF of BS$(\hat{\alpha}, \hat{\beta})$ as

$$\hat{P}(\tau_x > t) = 1 - \hat{F}_{BS}(t), \tag{3.6}$$

where $\hat{F}_{BS}(t)$ is the CDF of BS$(\hat{\alpha}, \hat{\beta})$.

As an important quantity of interest, the $100p^{th}$ percentile of the FPT distribution,

t_p, defined by

$$P\left(\tau_x > t_p\right) = 1 - p, \tag{3.7}$$

can be considered.

Suppose $\theta = t(F)$ is the parameter of interest if the true underlying distribution F is known. Then $\hat{\theta}_n = t(\tilde{F})$ represents the nonparametric estimate of θ and this estimate is also consistent. Moreover, $\hat{\theta}^W = t(F^W)$ is defined as the estimate of θ if the underlying distribution F is incorrectly assumed as F^W.

Finding an accurate estimate for the percentiles purely by raw simulation of the empirical FPT will be very time-consuming and computationally burdensome. However, if the marginal distribution of the underlying process is exactly known, the expression of the Laplace transform of survival probability can be found by using Equation (2.8). The numerical Laplace transform methods for an analytic form have been discussed extensively in computational mathematics. The results obtained through such numerical methods will tend to be quite accurate. For this reason, the numerical Laplace inversion is taken as the true value of percentiles.

The visualized comparisons between different approaches under gamma process with certain parameters are given in Figure 3.1 and Figure 3.2. In Figure 3.1, the plots of 90^{th} percentile, t_{90}, are shown for different threshold values. Taken as true values, the red plot is obtained through Laplace inversion. The black dashed plot almost coincides with the red plot for both pairs of parameters which shows that the saddlepoint approximation is very effective in this case. On the other hand, the BS approach, which is in blue color and dotted type, performs worse than the saddlepoint approximation as it is clearly away from the true value. Figure 3.2 shows the plots

for various percentiles with threshold fixed as 10. For small p values, the saddlepoint approximation has larger bias than the BS approach, but coincides with the red plot when p is greater than 0.3. The bias of BS approach becomes larger when p grows.

Inverse Gaussian Process

Inverse Gaussian distribution with mean parameter μ_{IG} and shape parameter λ_{IG} has the PDF

$$f(x) = \left[\frac{\lambda_{IG}}{2\pi x^3}\right]^{1/2} \exp\left\{\frac{-\lambda_{IG}(x-\mu_{IG})^2}{2x\mu_{IG}^2}\right\}.$$

Its Laplace exponent is given by

$$\phi(\omega) = \frac{\lambda_{IG}}{\mu_{IG}}\left[\sqrt{1+\frac{2\mu_{IG}^2\omega}{\lambda_{IG}}}-1\right].$$

The Laplace transform of the survival probability based on the inverse Gaussian process is derived from (2.8) as

$$\mathcal{L}\{P(\tau_x > t)\} = \frac{\exp\left\{-\frac{\lambda_{IG}}{\mu_{IG}}\left[\sqrt{1+\frac{2\mu_{IG}^2\omega}{\lambda_{IG}}}-1\right]t\right\}}{\omega},$$

and the Laplace exponent of this survival probability is

$$\psi(\omega) = \frac{\lambda_{IG}}{\mu_{IG}}\left[\sqrt{1+\frac{2\mu_{IG}^2\omega}{\lambda_{IG}}}-1\right]t + \log(\omega).$$

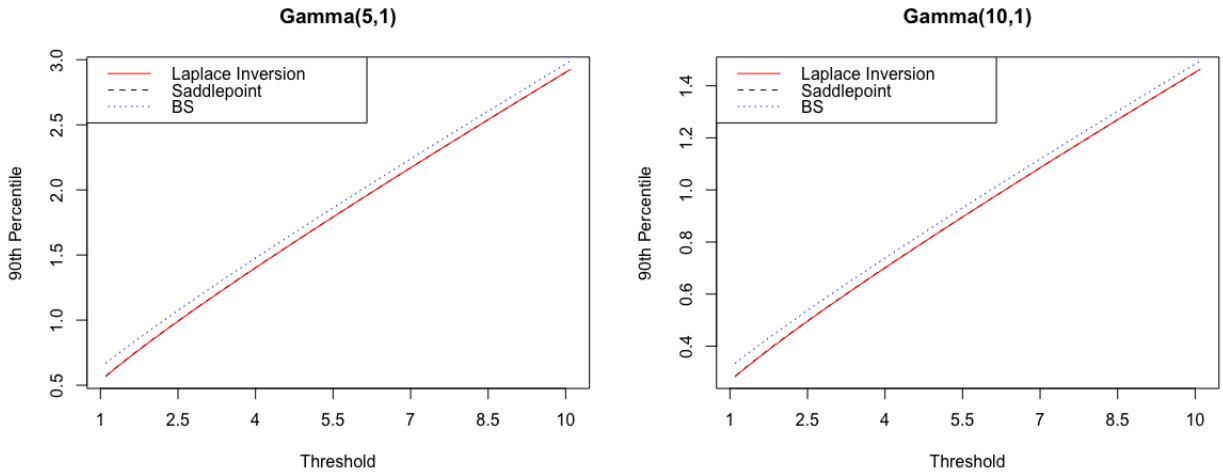

Figure 3.1: Comparison of the 90^{th} percentiles at certain threshold values if the underlying process is known as gamma with parameters (5,1) and (10,1) obtained by Laplace inversion, the saddlepoint approximation, and the BS approach

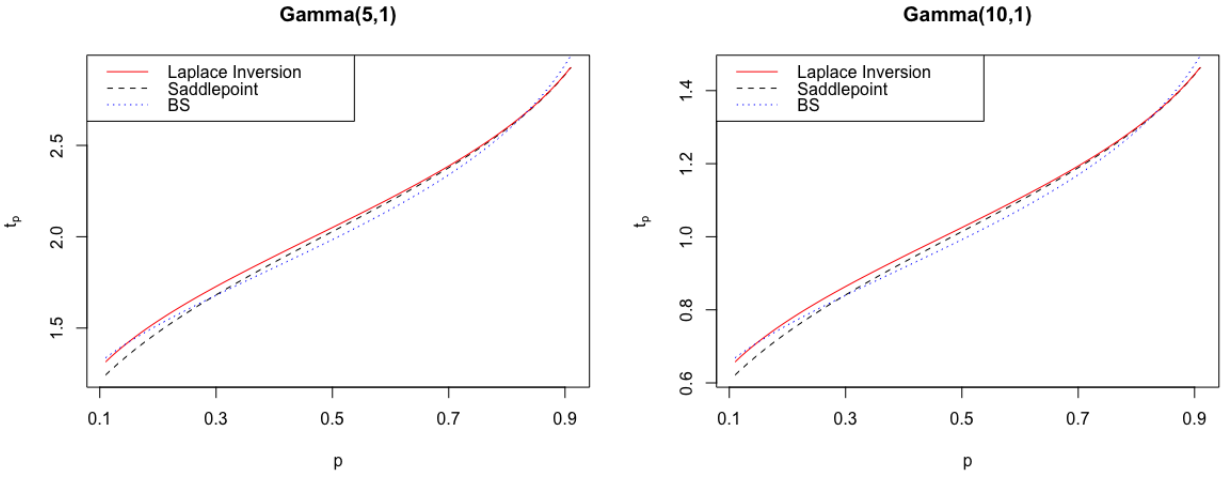

Figure 3.2: Comparison of the $100p^{th}$ percentiles with threshold fixed as 10 if the underlying process is known as gamma with parameters (5,1) and (10,1) obtained by Laplace inversion, the saddlepoint approximation, and the BS approach

The first- and second-order derivatives of $\psi(\omega)$ are calculated as

$$\begin{aligned}\psi'(\omega) &= \mu_{IG}\left(1+\frac{2\mu_{IG}^2\omega}{\lambda_{IG}}\right)^{-\frac{1}{2}}+\frac{1}{\omega},\\ \psi''(\omega) &= -\frac{\mu_{IG}^3}{\lambda_{IG}}\left(1+\frac{2\mu_{IG}^2\omega}{\lambda_{IG}}\right)^{-\frac{3}{2}}-\frac{1}{\omega^2},\end{aligned}$$

and $\hat{\omega}$ satisfies the equation

$$\mu_{IG}\left(1+\frac{2\mu_{IG}^2\hat{\omega}}{\lambda_{IG}}\right)^{-\frac{1}{2}}+\frac{1}{\hat{\omega}}=x.$$

With these results, the survival probability can be approximated by the saddlepoint method by using equation (2.8).

If the BS approach is used instead, the first two moments of the inverse Gaussian distribution, given by

$$\mu_1=\mu_{IG} \quad \text{and} \quad \mu_2=\frac{\mu_{IG}^3}{\lambda_{IG}}+\mu_{IG}^2,$$

are needed.

With these two moments, the corresponding parameters of the BS distribution are estimated as

$$\begin{aligned}\hat{\alpha} &= \mu_{IG}(\lambda_{IG}x)^{-\frac{1}{2}},\\ \hat{\beta} &= x\mu_{IG}^{-1},\end{aligned} \tag{3.8}$$

and the survival probability under the inverse Gaussian process, $\hat{P}(\tau_x > t)$, can then be approximated by Equation (3.6) when $F_{BS}(t)$ is the CDF of $BS(\hat{\alpha}, \hat{\beta})$.

Figure 3.3 and Figure 3.4 show the plots when the underlying processes are selected as inverse Gaussian processes with parameters $(15, 10)$ or $(25, 10)$. Figure 3.3 fixes the percentile as 0.9, and in this case the saddlepoint approximation displays the superiority for all the threshold values compared to the BS approach as the bias is always much smaller. In Figure 3.4, the saddlepoint approximation is seen to be much better than the BS approach for the two sets of parameters when p is greater than 0.3 if the threshold is fixed as 10. Although both the saddlepoint approximation and the BS approach have larger bias than the case when the underlying process is gamma process, the bias for the saddlepoint approximation is still considerably smaller than the BS approach. The saddlepoint approximation coincides with the red plot when p gets large and its bias concentrates only around small p values. On the contrary, the plot of the BS approach completely diverges for large p and the trend is essentially different from the plot of true values. From these results, it can be seen that the overall bias of the saddlepoint approximation is much smaller than the BS approach.

Once the assumption of the underlying process is correct, the parametric methods such as Laplace inversion or parametric saddlepoint approximation can undoubtedly capture more information than BS approach which only requires the first two moments. Then, it is understandable that the parametric method can perform better than BS distribution. Despite this, its large bias under small threshold level still makes people reluctant to approximate FPT density by using BS approach. Therefore, it is important to investigate whether the empirical saddlepoint approximation can outperform BS approach under nonparametric scenarios.

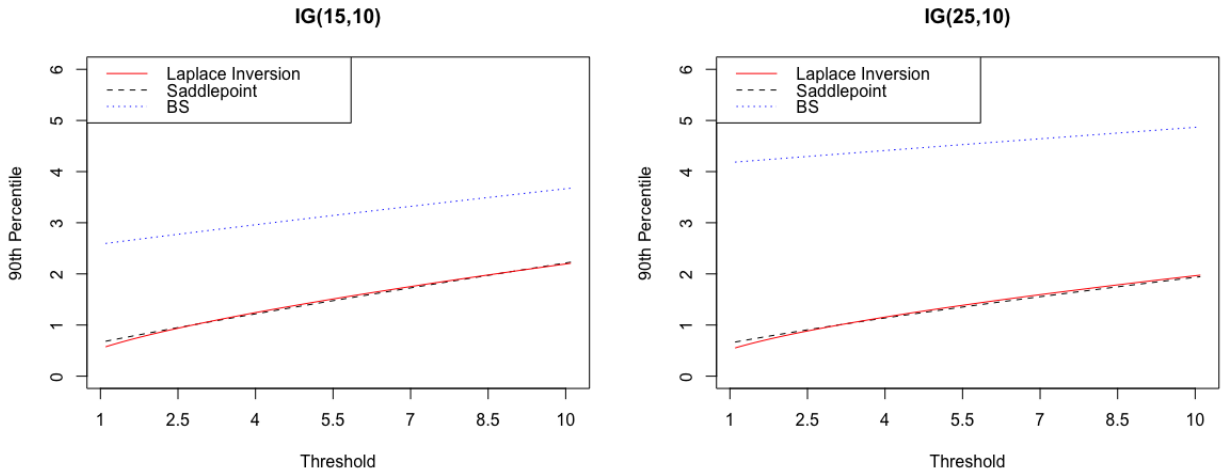

Figure 3.3: Comparison of the 90^{th} percentiles at certain threshold values if the underlying process is known as inverse Gaussian with parameters (15,10) and (25, 10) obtained by Laplace inversion, the saddlepoint approximation, and the BS approach

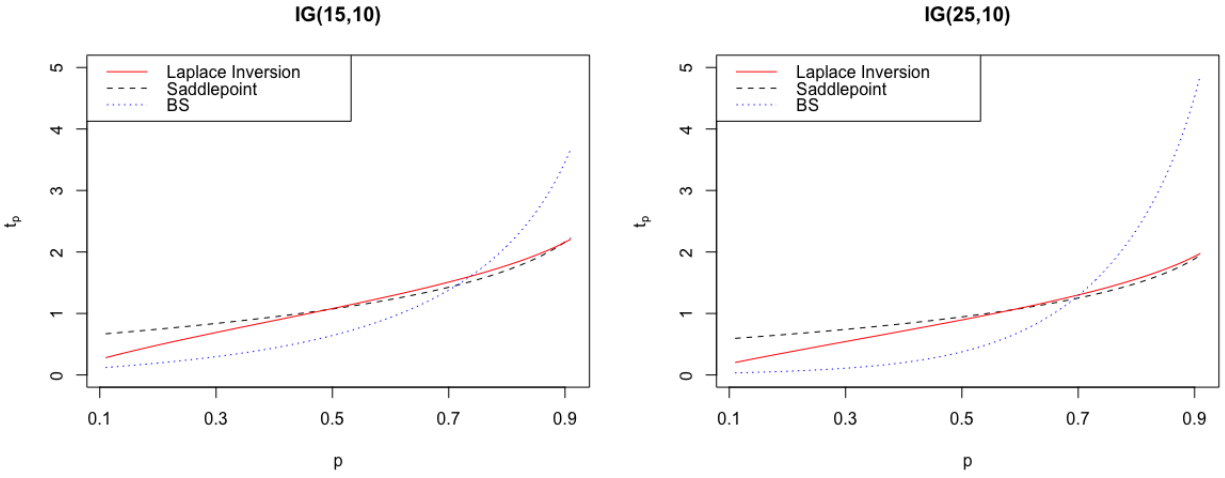

Figure 3.4: Comparison of the $100p^{th}$ percentiles with fixed threshold as 10 if the underlying process is known as inverse Gaussian with parameters (15,10) and (25, 10) obtained by Laplace inversion, the saddlepoint approximation, and the BS approach

3.3 Model Evaluation

Tsai *et al.* (2011) discussed the case when a set of observations following a gamma process is incorrectly assumed as Wiener process and showed that the estimation results in a large bias due to this misspecification. Thus, the type of underlying process is failed to be specified but only assumed as nondecreasing Lévy process, the problem caused by misspecification can be solved by using nonparametric approaches as it would only rely on the sample observations.

In the areas of reliability engineering and survival analysis, only limited types of stochastic process with good properties (gamma process and inverse Gaussian process) have been systematically studied. Hence, even if the true underlying process is unknown, gamma and inverse Gaussian process are still assumed for the analysis, which may be incorrect.

Actually, to make the study of degradation processes more general, other processes should be considered, but most of them just lose the convenient property of closure in convolution which causes difficulty in studying the first passage time. For this reason, a more general method is needed.

Equation (2.8) globally generalizes the study to any process and gives the true survival probability if only the Laplace exponents exists and Equation (3.2) further improves it by making no assumption for the certain type of Lévy subordinator. Feuerverger (1989) proved that the empirical Laplace exponent and its derivatives are unbiased estimates of true values and converge to them in probability.

To measure the relative error caused by misspecification, the relative error κ is

defined as

$$\kappa = \frac{|\hat{\theta} - \theta|}{\theta}.$$

Specifically, $\kappa(\tilde{F})$ and $\kappa(F^W)$ represent the relative errors corresponding to the empirical distribution $\tilde{F}$ and incorrectly assumed distribution F^W. The nonparametric estimate of $\kappa(\tilde{F})$ is calculated through the empirical saddlepoint approximations as well as the BS approach.

3.3.1 Bootstrap Estimation

Suppose $Y_1, \cdots, Y_n$ are independent with a common but unknown distribution F, and that the bootstrap method does not sample from the true distribution F but from the empirical distribution $\tilde{F}_n$ defined as

$$\tilde{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i \leq x).$$

The empirical $\tilde{F}_n$ is a consistent estimator of the underlying distribution F. If $\theta = t(F)$ is the quantity of interest, then $\hat{\theta}_n = t(\tilde{F}_n) = \hat{\theta}(Y_1, \cdots, Y_n)$ is a consistent estimator of θ. Conditioned on the observed sample and the hypothetical random sample $Y_1^*, \cdots, Y_n^*$ from the empirical distribution $\tilde{F}$, the bootstrap estimate $\hat{\theta}_n^* = \hat{\theta}(Y_1^*, \cdots, Y_n^*)$ is a consistent estimator of $\hat{\theta}$ as well.

Usually, the sampling distribution of $\hat{\theta}^* - \hat{\theta}$ is not easy to find. No matter whether the underlying distribution is known or unknown, generate a random sample from the parametric distribution F or empirical distribution $\tilde{F}$ and repeat the procedure for a large number of bootstrap runs N. The new estimates obtained in this manner

are denoted by $\hat{\theta}_1^*, \cdots, \hat{\theta}_N^*$, and then the Bootstrap estimate of standard error for $\hat{\theta}$ can be obtained as

$$s.e.(\hat{\theta}) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\hat{\theta}_i^* - \bar{\hat{\theta}}^*)^2}.$$

3.3.2 Simulation Study

Without any assumption for the underlying process but completely following the observations, the empirical Laplace exponent can be calculated and the survival probability can be found through empirical saddlepoint approximations. The Bootstrap samples are acquired by resampling from the original data set with the same size as 1000 and then repeating the procedure 1000 times. Therefore, there are 1000 approximations of survival probability calculated from the 1000 new sample sets and the bootstrap $100p^{th}$ percentile t_p^* is also obtained by the survival probability given by Equation (3.4). The bootstrap relative error is then defined as follows:

$$\kappa^* = \frac{|\hat{t}_p^* - t_p|}{t_p}, \tag{3.9}$$

where $\hat{t}_p^*$ is an estimate of the p^{th} percentile from a bootstrap sample while t_p is the true value of the p^{th} percentile.

If BS approach is applied to approximate the FPT density, the two parameters are calculated from the sample as in Equation (3.4) and μ_1 and $\mu_2 - \mu_1^2$ are replaced by sample mean and sample variance, respectively. Thus, the right-tail cumulative probability $BS(\alpha, \beta)$ is directly taken as the survival probability. This approach will lose much information of the sample since only the mean and variance from the sample

are used.

Two examples from the sample are presented. Mixed gamma process and mixed inverse Gaussian process are taken as the true underlying processes and their survival probabilities can be obtained by Laplace inversion. Without assuming the type of the true process, the relative errors for the empirical saddlepoint approximation, the BS approach as well as when the process is misspecified as a gamma process are determined.

Mixture of two unequally weighted gamma processes

A mixed process can be written as

$$Z_t = pZ_t^{(1)} + (1-p)Z_t^{(2)},$$

where $Z_t^{(1)}$ and $Z_t^{(2)}$ are two independent processes with respective weights p and $1-p$.

Mixed gamma process is considered as a complicated degradation model since the gamma distribution is not closed under convolution when the rate parameters are different and the density function of any point on the degradation path is unavailable in analytic form.

If the two components are assumed to be $Z_t^{(1)} \sim \text{Gamma}(0.05t, 0.06)$ and $Z_t^{(2)} \sim \text{Gamma}(0.04t, 0.05)$, then $\{Z_t\}_{t\geq 0}$ is a mixed gamma process with marginal distribution

$$Z_1 = 0.055 \times \text{Gamma}(0.05, 0.06) + 0.945 \times \text{Gamma}(0.04, 0.05),$$

when the weight on $Z_t^{(1)}$ is specified as $p = 0.055$. Next, 1000 samples are generated from this distribution and taken as observation data.

In Figures 3.5 and 3.6, derived by the original data set, the plots of two nonparametric methods, viz., empirical saddlepoint approximations and BS approach, and one parametric method as the process is misspecified as gamma, are compared with the Laplace inversion based on the true underlying process. Figure 3.5 shows the results of the 90^{th} percentile with different values of threshold ranging from 0.1 to 1. In Figure 3.6, the threshold is fixed as 1 while the percentile p ranges from 0.1 to 0.9.

The red plot represents the estimation if the true underlying process is known; the black dashed line is the empirical saddlepoint approximation if there is no assumption being made regarding the marginal distribution; the green dash-dotted line is the estimation if the process is incorrectly set as gamma process; the blue dotted line is fitted by the BS distribution. It can be seen from Figure 3.5 that empirical saddlepoint approximations (black dashed) is always better than other approaches since the estimate line is more close to the true value (red solid) for any value of threshold for the 90^{th} percentile. From Figure 3.6, the estimation by empirical saddlepoint approximation is clearly more close to the true value than others especially for large values of percentile. When the percentile approaches 0.9, both gamma and BS are divergent while the empirical saddlepoint method is seen to coincide with red plot.

Tables 3.1 and 3.2 show the arithmetic mean of 1000 bootstrap relative errors κ^* and the corresponding bootstrap estimations of standard error using two nonparametric approaches (the empirical saddlepoint approximation and the BS approach) and one parametric approach (assume the underlying process as gamma). These two tables are taken as supplement for the two figures and they provide more details and perspectives. In Table 3.1, the relative errors are investigated for the 90^{th} percentile for different threshold values ranging from 0.1 to 1. All the bootstrap arithmetic

means of percentile estimation based on the empirical saddlepoint approximation are significantly smaller than those for other methods. This phenomenon can also be seen from Figure 3.5. Besides the bootstrap means, the bootstrap estimation of standard error are also very small. Even if the variability of the relative errors is taken into account, the empirical saddlepoint approximation still seems to be more stable and reliable than other methods. Table 3.2 shows the empirical saddlepoint approximation has a smaller relative error and relatively small variability when p is greater than 0.2. For the cases when p is 0.1 or 0.2, direct use of parametric approach seems to be slightly better than the empirical saddlepoint approximation.

Mixture of two equally weighted inverse Gaussian processes

In this section, we alternatively consider an equal-weighted mixed inverse Gaussian process to see whether the empirical saddlepoint approximation can be useful in this case as well.

With the same sample size, there are 1000 samples generated from a mixed inverse Gaussian process $\{Z_t\}_{t\geq 0}$ with marginal distribution as

$$Z_1 = 0.5 \times IG(2, 0.5) + 0.5 \times IG(5, 3).$$

The process includes two independent and equally weighted inverse Gaussian processes with respective parameters (2,0.5) and (5,3).

In Figures 3.7 and 3.8, the plots of two nonparametric methods, viz., empirical saddlepoint approximations and BS approach, and one parametric method as the process is misspecified as gamma, are compared with the Laplace transform based on the true underlying process as a mixed inverse Gaussian process. Figure 3.7

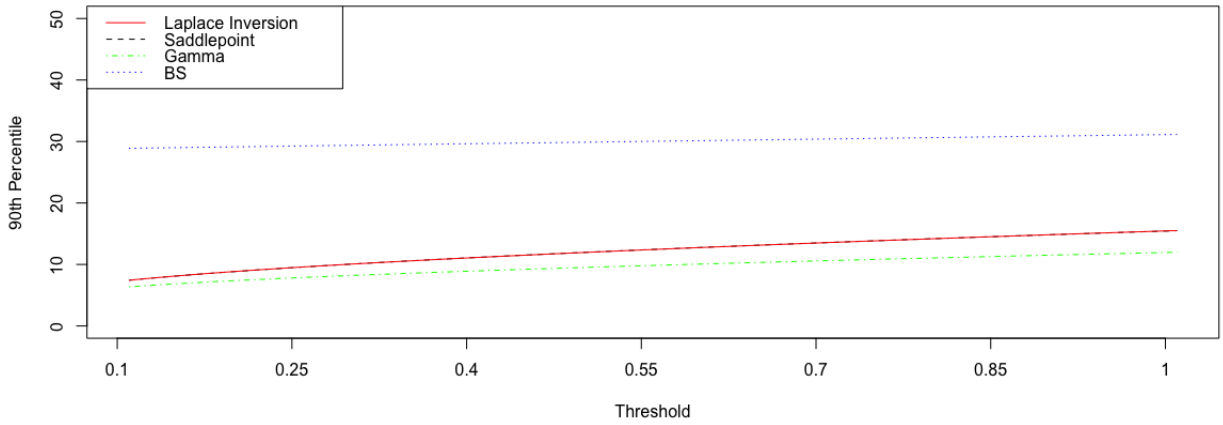

Figure 3.5: Comparisons of the 90^{th} percentiles at different threshold values obtained by Laplace inversion, empirical saddlepoint approximation, the BS approach and the plots of two nonparametric methods, viz., empirical saddlepoint approximations, the BS approach, and the parametric method as the process is misspecified as gamma process when the true underlying process is a Mixed gamma process

Table 3.1: Arithmetic means of bootstrap relative error κ^* of 90^{th} percentile and the corresponding bootstrap estimation of standard error (in brackets) for different values of threshold based on empirical saddlepoint approximations, the the BS approach, and the parametric method as the process is misspecified as gamma process when the true underlying process is a Mixed gamma process

Threshold	Saddlepoint	Gamma	BS
0.1	0.0447(0.0324)	0.1755(0.0296)	2.9847(0.4696)
0.2	0.0440(0.0340)	0.2025(0.0299)	2.3104(0.3646)
0.3	0.0465(0.0348)	0.2189(0.0306)	1.9475(0.3286)
0.4	0.0498(0.0379)	0.2344(0.0310)	1.7018(0.2946)
0.5	0.0502(0.0377)	0.2421(0.0315)	1.5219(0.2757)
0.6	0.0544(0.0382)	0.2492(0.0343)	1.3817(0.2804)
0.7	0.0567(0.0388)	0.2570(0.0338)	1.2560(0.2507)
0.8	0.0567(0.0410)	0.2595(0.0345)	1.1669(0.2483)
0.9	0.0548(0.0397)	0.2641(0.0345)	1.0922(0.2347)
1.0	0.0588(0.0401)	0.2700(0.0347)	1.0151(0.2280)

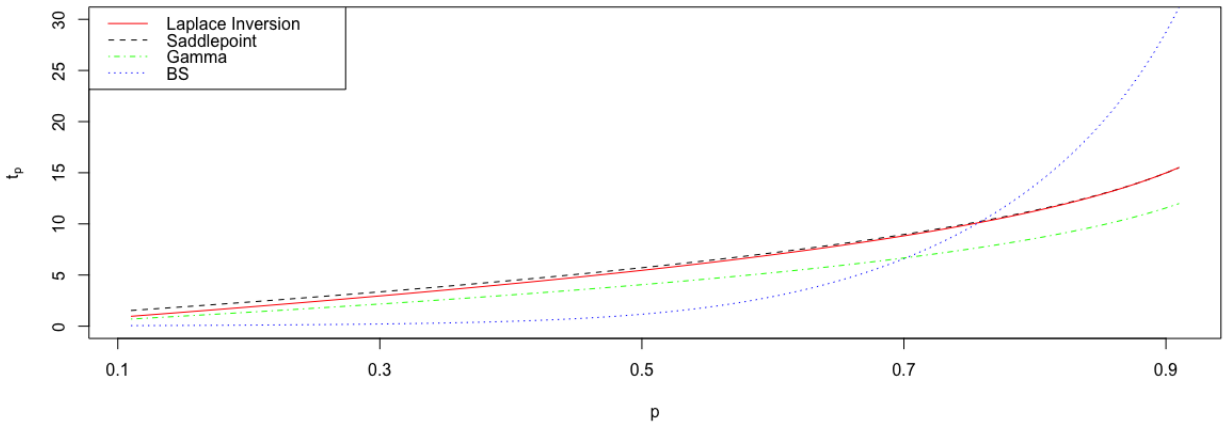

Figure 3.6: Comparisons of the $100p^{th}$ percentiles with fixed threshold as 1 obtained by Laplace inversion, empirical saddlepoint approximation, the BS approach and the parametric method as the process is misspecified as gamma process when the true underlying process is a Mixed gamma process

Table 3.2: Arithmetic means of bootstrap relative error κ^* of $100p^{th}$ percentile and the corresponding bootstrap estimation of standard error (in brackets) for fixed threshold 1 and different values of p based on empirical saddlepoint approximation, the BS approach and the parametric method as the process is misspecified as gamma process when the true underlying process is a Mixed gamma process

p	Saddlepoint	Gamma	BS
0.1	0.5478(0.1295)	0.3247(0.0437)	0.9574(0.0108)
0.2	0.2083(0.0963)	0.3190(0.0419)	0.9564(0.0102)
0.3	0.1065(0.0798)	0.3118(0.0414)	0.9402(0.0132)
0.4	0.0756(0.0583)	0.3055(0.0409)	0.8997(0.0192)
0.5	0.0617(0.0470)	0.3015(0.0395)	0.7999(0.0277)
0.6	0.0573(0.0436)	0.2914(0.0379)	0.5835(0.0438)
0.7	0.0596(0.0421)	0.2875(0.0383)	0.2307(0.0813)
0.8	0.0590(0.0436)	0.2786(0.0385)	0.2855(0.1405)
0.9	0.0551(0.0403)	0.2662(0.0345)	1.0393(0.2304)

investigates the 90^{th} percentile with the value of threshold ranging from 1 to 10 while Figure 3.8 investigates different percentiles when p valuing from 0.1 to 0.9 while the threshold is fixed as 10.

Figure 3.7 demonstrates that the empirical saddlepoint approximation always performs better on estimating the 90^{th} percentile than the other approaches when threshold takes values from 1 to 10. From Figure 3.8, the estimation obtained by the empirical saddlepoint approximation is seen to be the best among the three methods when p is greater than 0.25. When p is smaller than 0.25, black dashed plot and green dash-dotted plot are very close to each other which shows that we can get the same relative error if the gamma process is directly used as the underlying process.

Table 3.3 gives the arithmetic means of bootstrap relative errors for the 90^{th} percentile for different threshold values ranging from 1 to 10. All of bootstrap means based on empirical saddlepoint approximation are significantly smaller than other methods and also the bootstrap estimation of standard error is always the smallest among the three methods. The empirical saddlepoint approximation is seen to be the most stable and reliable compared to others. Table 3.4 shows that the empirical saddlepoint approximation has smallest relative error and its corresponding variability is also smallest when p is greater than 0.2. When p takes the values 0.1 or 0.2, taking the true process as gamma or using the BS approach result in slightly better results than empirical saddlepoint approximation.

3.4 Illustrative Data Analysis

This section is still based on the laser dataset provided by Meeker and Escobar (1998).

Since the survival probability is approximated by the empirical saddlepoint method

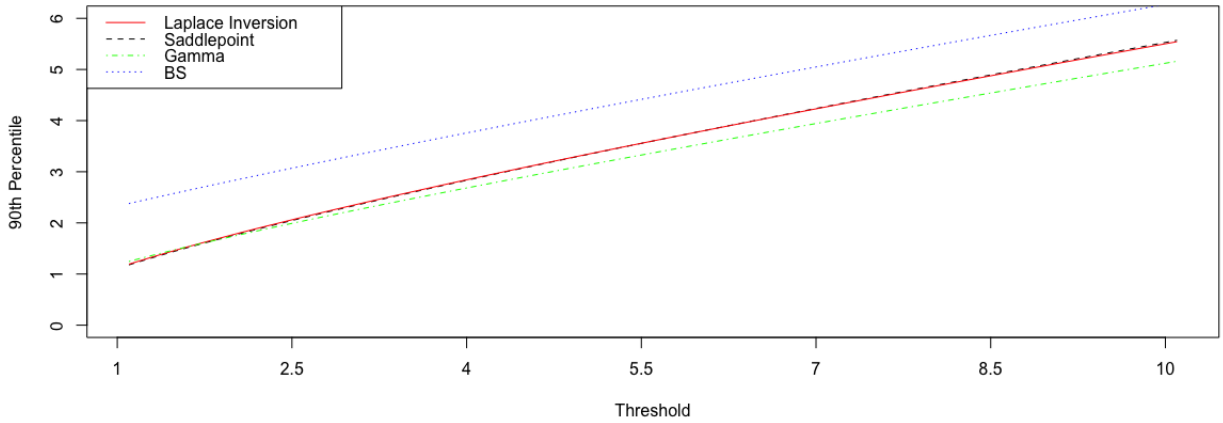

Figure 3.7: Comparison of the 90^{th} percentiles at different threshold values obtained by Laplace inversion, empirical saddlepoint approximation, the BS approach and the parametric method as the process is misspecified as gamma when the true underlying process is a mixed inverse Gaussian process

Table 3.3: Arithmetic means of bootstrap relative error κ^* of 90^{th} percentile and the corresponding bootstrap estimation of standard error (in brackets) for different values of threshold based on empirical saddlepoint approximation, the BS approach, and the parametric method as the process is misspecified as gamma process when the true underlying process is a mixed inverse Gaussian process

Threshold	Saddlepoint	Gamma	BS
1	0.0188(0.0143)	0.0378(0.0182)	0.9397(0.1250)
2	0.0217(0.0150)	0.0313(0.0162)	0.5328(0.0801)
3	0.0217(0.0153)	0.0549(0.0170)	0.3749(0.0645)
4	0.0217(0.0157)	0.0668(0.0174)	0.2837(0.0546)
5	0.0210(0.0152)	0.0725(0.0171)	0.2326(0.0478)
6	0.0207(0.0142)	0.0752(0.0176)	0.1938(0.0422)
7	0.0199(0.0143)	0.0773(0.0177)	0.1678(0.0381)
8	0.0202(0.0151)	0.0786(0.0184)	0.1466(0.0384)
9	0.0197(0.0149)	0.0782(0.0191)	0.1300(0.0353)
10	0.0194(0.0147)	0.0782(0.0187)	0.1173(0.0336)

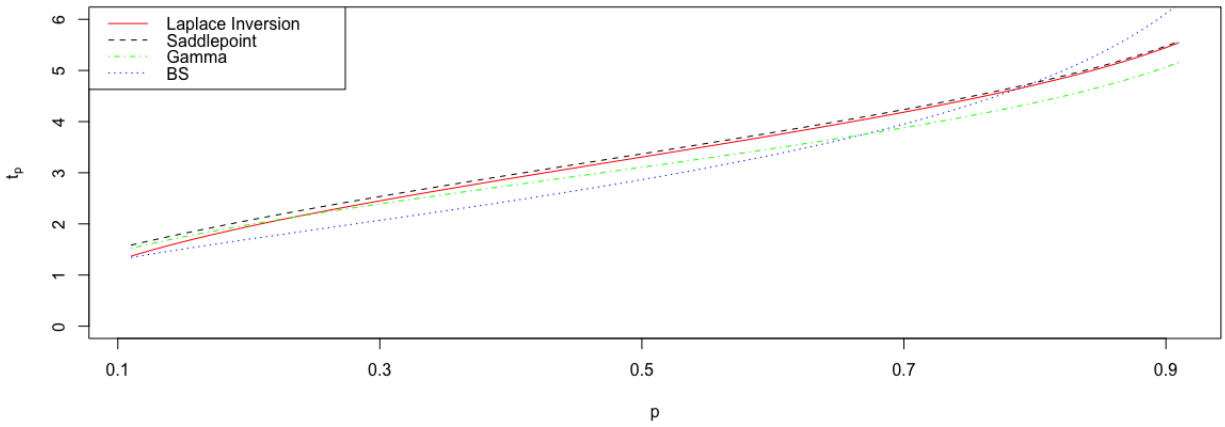

Figure 3.8: Comparisons of the $100p^{th}$ percentiles with fixed threshold as 1 obtained by Laplace inversion, empirical saddlepoint approximation, the BS approach and the parametric method as the process is misspecified as gamma when the true underlying process is a mixed inverse Gaussian process

Table 3.4: Arithmetic means of bootstrap relative error κ^* of $100p^{th}$ percentile and the corresponding bootstrap estimation of standard error (in brackets) for fixed threshold 10 and different values of p based on empirical saddlepoint approximation, the BS approach and the parametric method as the process is misspecified as gamma process when the true underlying process is a mixed inverse Gaussian process

p	Saddlepoint	Gamma	BS
0.1	0.1391(0.0454)	0.0963(0.0519)	0.0511(0.0379)
0.2	0.0445(0.0315)	0.0332(0.0251)	0.1381(0.0455)
0.3	0.0284(0.0208)	0.0435(0.0283)	0.1633(0.0377)
0.4	0.0248(0.0189)	0.0603(0.0294)	0.1598(0.0330)
0.5	0.0216(0.0164)	0.0706(0.0263)	0.1398(0.0276)
0.6	0.0213(0.0160)	0.0795(0.0242)	0.1099(0.0256)
0.7	0.0214(0.0157)	0.0836(0.0234)	0.0650(0.0252)
0.8	0.0200(0.0152)	0.0828(0.0217)	0.0212(0.0163)
0.9	0.0209(0.0151)	0.0797(0.0193)	0.1150(0.0352)

given in Equation (3.2), the CDF of the failure time is equivalent to subtracting the survival probability from 1. To evaluate whether the failure times obtained from the real data follow the approximated CDF, the nonparametric goodness-of-fit test, namely, the Kolmogorov-Smirnov test (K-S test) is used for testing the null hypothesis "the observations follow the approximated distribution $\hat{F}_0$", and the test statistic is $D = \sup_x |\hat{F}_0(x) - \tilde{F}(x)|$, where $\tilde{F}$ is the empirical distribution function of the observed data.

Set the ordered failure times, conditioned on the threshold K, as $\tau^K_{(1)}, \cdots, \tau^K_{(15)}$, where $K = 1, 2, \cdots, 6$. Then K-S test statistic under threshold K is defined as

$$D_K = \max_{1 \le i \le 15} \left(\hat{F}_0\left(\tau^K_{(i)}\right) - \frac{i-1}{15}, \frac{i}{15} - \hat{F}_0\left(\tau^K_{(i)}\right) \right),$$

where $\hat{F}_0$ is obtained by the empirical saddlepoint approximation.

Table 3.5 shows the test statistics D_K, $K = 1, \cdots 6$ as well as the critical value for the K-S test under significance level 0.05 and sample size 15, provided by Kececioglu (1993). The critical value 0.338 is higher than all of the K-S test statistics, and thus the null hypothesis that the distribution of observed failure times is identical to the distribution approximated by the saddlepoint method cannot be rejected.

However, some further evaluation need to be considered, such as the construction of Bootstrap confidence interval for the 90^{th} percentile.

For each test unit, there are 16 observations with the same length of time interval and 240 increments measured for all 15 test units. Since all increments are assumed to independently follow the same distribution, their empirical distribution can be obtained from the observed data. Sampling the increments repeatedly from the data with replacement can make the observation duration to be longer than the original

termination time 4000 hours so that the failure time can be found for any threshold value. For example, if the threshold values are taken from 1 to 10, it is still available to find the failure time when threshold is greater or equal to 7 while this is not possible for the original dataset.

The bootstrap resampling procedure proceeds as follows:

(i) When threshold is K, repeatedly sample increments from the original dataset until the cumulative increment exceeds the given threshold;

(ii) Use linear interpolation to determine the failure time;

(iii) Repeat the Steps (i) and (ii) 10000 times, and get failure times $\tau_1^*, \cdots, \tau_{10000}^*$;

(iv) Find the p^{th} percentile for the failure time t_p^*;

(v) Repeat Steps from (i) to (iv) 1000 times, and get the bootstrap sample set for the p^{th} percentile $t_{p,1}^*, \cdots, t_{p,1000}^*$.

With these bootstrap samples, the bootstrap arithmetic means and bootstrap estimation of standard errors were determined and given in Table 3.6. The 95% bootstrap confidence interval, proposed by Efron (1981), is the so-called percentile method. Given the bootstrap CDF $\hat{G}$ and confidence level α, the bootstrap confidence interval is obtained as $\theta \in \left[\hat{G}^{-1}(\alpha/2), \hat{G}^{-1}(1-\alpha/2)\right]$. For 95% confidence interval, the lower bound $\hat{G}^{-1}(0.025)$ is the 2.5^{th} percentile of the bootstrap sample and upper bound $\hat{G}^{-1}(0.975)$ is the 97.5^{th} percentile. From Table 3.7, it can be seen that neither the misspecified gamma process nor the BS approach has any approximation located inside the 95% bootstrap confidence interval. On the other hand, except when threshold is 1, in all the cases the empirical saddlepoint approximation of

90^{th} percentile falls inside the confidence interval and are quite close to the mean of bootstrap sample according to Table 3.6. This provides that the proposed method performs well for this dataset and also possesses some key advantages as compared to other approaches.

3.5 Concluding Remarks

The nonparametric method based on the empirical saddlepoint approximation gives an effective and accurate approximation for the FPT density. Even when there is no assumption on the exact type of degradation processes, the good approximations for the survival probability and the $100p^{th}$ percentile of lifetime can still be obtained. In most examples of this chapter, the saddlepoint approximation performs significantly better than other methods whether we make assumption or do not make assumption about the underlying processes. The saddlepoint approximation often has smaller bootstrap mean of relative error and bootstrap estimation of standard error compared to other methods.

Of course, we can also include higher-order terms to get the second-order saddlepoint approximation as we consider in Chapter 2, which may further improve the accuracy of the estimation. This may be useful for the case of small p wherein we have observed that the saddlepoint approximation possesses a large relative bias. The development of second-order saddlepoint approximation is expected to find in the future.

When the underlying process is assumed, the true FPT density can be found and these true values are the important benchmarks when comparing the proposed method with other existing approaches. We need to guarantee the existence of Laplace exponent since its close-form expression must be available in terms of Laplace

inversion method. The existence conditions of Laplace transform will be discussed in Appendix.

Table 3.5: K-S test statistics for different thresholds and the critical value under significance level of 0.05 for sample size 15

D_1	D_2	D_3	D_4	D_5	D_6	Critical Value
0.196	0.252	0.201	0.210	0.272	0.306	0.338

Table 3.6: The 90^{th} percentile of the failure time for different thresholds using the empirical saddlepoint approximation, the parametric method as the process is misspecified as gamma process, the BS approach, as well as arithmetic means of bootstrap sample, bootstrap estimation of standard errors and bootstrap confidence intervals when resampling from the original laser dataset

Threshold	Saddlepoint	Gamma	BS	Bootstrap Mean	Bootstrap S.E.	95% Bootstrap CI
1	683.20	678.86	702.34	674.19	2.30	(670.04, 678.57)
2	1248.79	1239.26	1265.61	1248.57	4.01	(1241.07, 1255.66)
3	1796.23	1783.44	1812.31	1793.98	3.69	(1786.59, 1801.03)
4	2334.46	2319.23	2350.34	2330.87	4.52	(2321.90, 2339.49)
5	2866.89	2849.68	2882.80	2863.58	5.31	(2853.33, 2875.00)
6	3395.25	3376.31	3411.29	3392.56	5.79	(3381.84, 3403.80)
7	3920.54	3900.07	3936.78	3918.63	6.46	(3905.41, 3931.34)
8	4443.42	4421.55	4459.88	4442.01	6.58	(4429.59, 4455.61)
9	4964.33	4941.17	4981.03	4964.22	7.27	(4949.98, 4978.21)
10	5483.60	5459.23	5500.55	5485.85	7.66	(5471.44, 5500.00)

Table 3.7: The 90^{th} percentiles derived from the empirical saddlepoint approximation, the parametric method as the process is misspecified as gamma process, the BS approach, are located inside the 95% bootstrap confidence interval (denoted by Y) or not (denoted by N)

Threshold	1	2	3	4	5	6	7	8	9	10
Saddlepoint	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Gamma	N	N	N	N	N	N	N	N	N	N
BS	N	N	N	N	N	N	N	N	N	N

Chapter 4

Optimal Design of Degradation Tests

4.1 Introduction

This chapter discusses the optimal design of the degradation test under both parametric and nonparametric frameworks by assuming the underlying process as the gamma process, the inverse Gaussian process, or the empirical Lévy process. To optimize the degradation experiment, it is necessary to know how the design variables such as the number of test units, the number of measurements, the inspection frequency or the length of time interval can influence the accuracy of the parameter estimation. Conditioned on the experimental cost cannot exceed the pre-specified budget, the values of design variables are determined by minimizing the asymptotic variance of the $100p^{th}$ percentile of the FPT distribution for the parametric scenario or the counterpart bootstrap estimate of variance for the nonparametric scenario.

The problem of optimal design has been studied by a series of literatures from

various aspects. Yu and Tseng (1999) addressed the degradation path by a quasilinear model and optimally determined the design variables including sample size, number of measurements and termination time by minimizing the variance of the estimated $100p^{th}$ percentile of the failure time distribution. Tseng *et al.* (2009) assumed the underlying process as a gamma process and optimize the degradation test by minimizing the asymptotic variance of the mean-time-to-failure. Tsai *et al.* (2012) further discussed the heterogeneity problem by assuming the underlying process as a gamma process incorporated with random effects and conducted the degradation test by minimizing the asymptotic variance of the estimated $100p^{th}$ percentile.

This chapter is organized as follows: In Section 4.2, the design variables, influential factors and procedures of the optimal design are described and the gamma process, the inverse Gaussian process or the empirical Lévy process are introduced as the underlying processes. Then each underlying process is discussed regarding the parameter estimation, objective function and how to proceed the corresponding optimization. In Section 4.3, an practical example regarding laser data is discussed to illustrate the proposed approaches.

4.2 Model Description

When conducting a degradation test, the experimental cost and the precision of reliability estimation are the two major concerned factors. Under different assumptions, the design values such as the sample size of test units, the measurement times, and the inspection frequency can influence both of these two factors and inappropriate decision for the design variables will waste experimental resources and reduce the estimation efficiency. To deal with such issue of optimal design, we need to determine

the values of design variables under the constraint that the total experimental cost does not exceed the predetermined budget.

Assume n as the number of test units , m as the number of measurements made every ft_u where t_u is length of each time interval, then the termination time is given as $t_m = fmt_u$. For each test unit, the sample paths can be observed as $Z_{t_1}, Z_{t_2}, \ldots, Z_{t_m}$ and the time interval between two consecutive observations is $ft_u = t_j - t_{j-1}$ where $j = 1, \ldots, m$, with the initial time $t_0 = 0$.

Considering the cost from the time to conduct the experiment, purchasing the tested units and the cost of measurements, the total cost of the degradation test, $TC(n, f, m)$, is given as a linear model to include three cost sources

$$TC(n, f, m) = C_{op} fm + C_{mea} nm + C_{it} n,$$

where C_{op} denotes the cost of operation, C_{mea} denotes the cost of each measurement, and C_{it} is the cost purchase each test unit. If C_b denotes the pre-specified budget of the experiment, then in any case we should make $TC \leq C_b$ satisfied.

The degradation processes are is given by the three different models including the gamma process, the inverse Gaussian process and the empirical Lévy process. All of these models are investigated conditioned on all the degradation paths are independent and the increments are identically distributed.

The optimization criterion for each of the parametric and nonparametric models is defined differently. The criterion for the parametric model is the asymptotic variance of the $100p^{th}$ percentile of the FPT, then the optimization problem can be states as

$$\text{Minimize } AVar(t_p(\theta|\xi) \text{ subject to } TC(\xi) \leq C_b,$$

where the vector $\xi = (n, f, m)$ represents the design vector.

Unlike the parametric models, the optimization criterion for the nonparametric model is alternatively defined as the bootstrap estimate of variance and optimization statement changes to

$$\text{Minimize } Var(t_p(\tilde{F})|\xi) \text{ subject to } TC(\xi) \leq C_b.$$

In the rest of this section, we will individually discuss each of the degradation models.

4.2.1 Gamma Process

Start with the gamma process Z_t with the shape parameter α and the rate parameter β, it can be represented as

$$Z_t \sim \text{Gamma}(\alpha t, \beta).$$

Suppose $Z_t^{(i)}$, $1 \leq i \leq n$ represents the degradation path of i^{th} test unit, $Y_{ij} = Z_{t_j}^{(i)} - Z_{t_{j-1}}^{(i)}$, $1 \leq j \leq m$ is the increment from t_{j-1} to t_j. In actual dataset, and the time interval of two consecutive measurements is ft_u. If α and β are given as the parameters of the increment within each time unit t_u, then the observed increments should follow $Gamma(\alpha f, \beta)$. As all of the test units are independent and the increments are assumed to be identically distributed, the likelihood function for the model can be given in terms of design variables n, f, m as well as the parameters

$\theta = (\alpha, \beta)$

$$L(\theta) = \prod_{i=1}^{n}\prod_{j=1}^{m} \frac{\beta^{\alpha f}}{\Gamma(\alpha f)} y_{ij}^{\alpha f - 1} e^{-\beta y_{ij}}.$$

Then, the log-likelihood function is the natural logarithm of $L(\theta)$

$$l(\theta) = \ln(L(\theta)) = nm \ln\left(\frac{\beta^{\alpha f}}{\Gamma(\alpha f)}\right) + \sum_{i=1}^{n}\sum_{j=1}^{m} \ln\left(y_{ij}^{\alpha f - 1} e^{-\beta y_{ij}}\right).$$

According to the expression of $l(\theta)$, we can see that n and m are not distinguished in terms of positions.

Based on the assumption that the test units are independent and the increments observed for each test unit are i.i.d., then we can say the increments for all the test units can be collected to obtain a dataset containing $n \times m$ i.i.d. random variables.

We can take derivative of the log-likelihood function with respect to each parameter and find the MLE of θ possess the relations specified as below

$$nm\left(\ln(\alpha f) - \ln\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij}}{nm}\right) - \psi(\alpha f) + \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} \ln(y_{ij})}{nm}\right),$$

and

$$\beta = \frac{\alpha f}{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij}}{nm}},$$

where the derivative of the logarithm of the gamma function given as

$$\psi(\alpha) = \frac{d}{d\alpha} \ln(\Gamma(\alpha))$$

is known as the digamma function.

Using Delta method, the asymptotic distribution for $t_p(\theta)$ can be found as

$$\sqrt{n}\left(t_p(\hat{\theta}) - t_p(\theta)\right) \xrightarrow{d} N\left(0, \frac{h' I^{-1}(\theta) h}{f^2(t_p(\theta))}\right),$$

where the vector h is defined as

$$h = \left(\frac{\partial F(t)}{\partial \alpha}, \frac{\partial F(t)}{\partial \beta}\right)\Bigg|_{t=t_p(\theta)}.$$

$I(\theta)$ is the Fisher information matrix with each entry as the negative second-order derivatives of log-likelihood function given by

$$\begin{aligned} I(\theta)_{11} &= nmf^2\psi_1(\alpha f), \\ I(\theta)_{22} &= \frac{\alpha f}{\beta^2}, \\ I(\theta)_{12} &= I(\theta)_{21} = -\frac{f}{\beta}. \end{aligned}$$

and $f(\cdot)$ is the pdf of the FPT.

Since by the cdf of the FPT given in (2.6)

$$F(t) = 1 - P(\tau_x > t) = 1 - \mathcal{L}^{-1}\left(\frac{e^{-\phi(\omega)t}}{\omega}\right),$$

where the Laplace exponent $\phi(\omega)$ is

$$\phi(\omega) = \alpha \ln\left(1 + \frac{\omega}{\beta}\right),$$

we can calculate the pdf $f(t)$ by taking the derivative with respect to t as

$$\begin{aligned} f(t) &= \frac{dF(t)}{dt} \\ &= -\frac{d\mathcal{L}^{-1}\left(\frac{e^{-\phi(\omega)t}}{\omega}\right)}{dt} \\ &= -\mathcal{L}^{-1}(\frac{d\frac{e^{-\phi(\omega)t}}{\omega}}{dt}) \\ &= \mathcal{L}^{-1}\left(\frac{\phi(\omega)}{\omega}e^{-\phi(\omega)t}\right). \end{aligned} \tag{4.1}$$

We can interchange the sequence of inverse Laplace transform $\mathcal{L}^{-1}$ and the derivative because by the definition of inverse Laplace transform, we have

$$\mathcal{L}^{-1}\left(\frac{e^{-\phi(\omega)t}}{\omega}\right) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{\omega x}\frac{e^{-\phi(\omega)t}}{\omega}d\omega,$$

where the integrand is continuous and t is irrelevant to the integrated variable.

The elements of the vector h can be calculated by taking the partial derivatives of the cdf with respect to each parameter

$$\begin{aligned} \frac{\partial F(t)}{\partial \alpha} &= \mathcal{L}^{-1}\left(\frac{te^{-\phi(\omega)t}}{\omega}\cdot\frac{\partial\phi(\omega)}{\partial\alpha}\right) \\ &= \mathcal{L}^{-1}\left(\frac{te^{-\phi(\omega)t}}{\omega}\cdot\ln\left(1+\frac{\omega}{\beta}\right)\right), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial F(t)}{\partial \beta} &= \mathcal{L}^{-1}\left(\frac{te^{-\phi(\omega)t}}{\omega}\cdot\frac{\partial\phi(\omega)}{\partial\beta}\right) \\ &= \mathcal{L}^{-1}\left(-\frac{\alpha te^{-\phi(\omega)t}}{\beta^2+\beta\omega}\right). \end{aligned}$$

The optimization criterion is defined as the asymptotic variance of t_p given as

$$AVar(t_p(\theta)) = \frac{h' I^{-1}(\theta) h}{n f^2(t_p(\theta))}. \tag{4.2}$$

4.2.2 Inverse Gaussian Process

Next, we consider another commonly used degradation process, the inverse Gaussian process. Suppose μ is the mean and λ is the shape parameter, the inverse Gaussian process can be represented as

$$Z_t \sim IG(\mu t, \lambda t^2).$$

Since the increments are actually observed every $f t_u$ and μ, λ are defined as the parameters of the increment within each time interval t_u, the observed increments follow $IG(\mu f, \lambda f^2)$.

With the observations of increment $Y_{ij} = Z^{(i)}_{t_j} - Z^{(i)}_{t_{j-1}}$, the likelihood function can be written as

$$L(\theta) = \prod_{i=1}^{n} \prod_{j=1}^{m} \left(\frac{\lambda f^2}{2\pi y_{ij}^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\lambda f^2 (y_{ij} - \mu f)^2}{2\mu^2 f^2 y_{ij}} \right),$$

where $\theta = (\mu, \lambda)$. And its corresponding log-likelihood function is

$$l(\theta) = \ln\left(L(\theta)\right) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} \left(\ln(\lambda) + \ln(f^2) - \ln(2\pi y_{ij}^3) \right)}{2} - \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\lambda (y_{ij} - \mu f)^2}{2\mu^2 y_{ij}}.$$

Take the derivatives of $l(\theta)$ with respect to μ and λ, the MLEs of the parameter are

given as

$$\hat{\mu} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij}}{nmf},$$
$$\hat{\lambda} = \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij} - \hat{\mu}f)^2}{nm\hat{\mu}^2 y_{ij}} \right)^{-1}.$$

Then, the Fisher information matrix can derived as

$$I(\theta) = nm \cdot \begin{bmatrix} \frac{\lambda f}{\mu^3} & 0 \\ 0 & \frac{1}{2\lambda^2} \end{bmatrix}.$$

To find the vector $h = (\frac{\partial F}{\partial \mu}, \frac{\partial F}{\partial \lambda})$, the first-order derivatives of the cdf with respect to μ and λ can be calculated by inverting the Laplace transforms

$$\frac{\partial F}{\partial \mu} = \mathcal{L}^{-1}\left(\frac{te^{-\phi(\omega)t}}{\omega} \cdot \frac{\partial \phi(\omega)}{\partial \mu} \right),$$
$$\frac{\partial F}{\partial \lambda} = \mathcal{L}^{-1}\left(\frac{te^{-\phi(\omega)t}}{\omega} \cdot \frac{\partial \phi(\omega)}{\partial \lambda} \right),$$

where $\phi(\omega)$ is the Laplace exponent of inverse Gaussian distribution given as

$$\phi(\omega) = \frac{\lambda}{\mu}\left(\sqrt{1 + \frac{2\mu^2\omega}{\lambda}} - 1 \right),$$

and its derivatives with respect to each parameter are

$$\frac{\partial \phi(\omega)}{\partial \mu} = \frac{\lambda}{\mu^2} - \frac{\lambda^2}{\mu^3}\left(\frac{\lambda^2}{\mu^2} + 2\lambda\omega \right)^{-\frac{1}{2}}$$
$$\frac{\partial \phi(\omega)}{\partial \lambda} = \left(\frac{\lambda^2}{\mu^2} + 2\lambda\omega \right)^{-\frac{1}{2}}\left(\frac{\lambda}{\mu^2} + \omega \right) - \frac{1}{\mu}.$$

With the pdf of the FPT $f(t)$ given in (4.1), we can conduct the optimal design based on the optimization criterion defined in (4.2).

4.2.3 Empirical Lévy Process

The definition of empirical Lévy process can be referred to Chapter 3. Without giving any process-type assumption, the misspecification problem can be avoid by using the empirical Lévy process. Same as the parametric models, we assume the degradation paths are independent and their increments $Y_j = Z_{t_j} - Z_{t_{j-1}}$ follow the identical distribution. The sample size of the increments is still $n \times m$. Therefore, with the i.i.d. increments $Y_{11}, Y_{12}, \ldots, Y_{nm} \sim F$, the empirical distribution $\tilde{F}(y)$ is defined as

$$\tilde{F}(y) = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \mathbb{1}_{Y_{ij} \leq y}. \tag{4.3}$$

The empirical distribution $\tilde{F}$ is a consistent estimator of the underlying distribution F. Since the $100p^{th}$ percentile $t_p(F)$ is the quantity of interest, then $t_p(\tilde{F})$ is also the consistent estimator of $t_p(F)$. Conditioned on the original sample $Y_1, \ldots, Y_{nm}$ and the hypothetical random sample $Y_1^*, \ldots, Y_{nm}^*$ from the empirical distribution $\tilde{F}$ obtained by the original sample set, the bootstrap estimate $t_p(\tilde{F}^*)$ is a consistent estimator of $t_p(\tilde{F})$.

Although the exact distribution of $t_p(\tilde{F})$ is unavailable to find, the bootstrap estimate of variance can be obtained as follows. First generate random samples from the empirical distribution $\tilde{F}$ and estimate $100p^{th}$ percentile by the new empirical distribution $\tilde{F}^*$. Then repeat the procedure for a large number of bootstrap runs N and the new estimates obtained in this manner are denoted by $t_p(\tilde{F}_1^*), \ldots, t_p(\tilde{F}_N^*)$.

Finally, the Bootstrap estimate of variance for $t_p(\tilde{F})$ can be obtained as

$$Var(t_p(\tilde{F})) = \frac{1}{N-1}\sum_{i=1}^{N}\left(t_p(\tilde{F}_i^*) - \bar{t}_p(\tilde{F}^*)\right)^2, \tag{4.4}$$

where $\bar{t}_p(\tilde{F})^* = \frac{\sum_{i=1}^{N} t_p(\tilde{F}_i^*)}{N}$.

In this nonparametric setting, although it is obvious that the empirical distribution is directly related to the number of test units n and the number of measurements for each test unit m, the relation between t_p and f is not very straightforward to be figured out. For example, for a data set with design (n, f, m) as well as the length of each time unit is t_u, the new sample set can always be constructed by resampling from the original sample set for any values of n and m. However, if f changes to $\frac{f}{2}$ which is the half of the original inspection frequency, it is unavailable to get any new observation beside the observations from the original sample set.

Another thing to be noticed is that in equation (2.6), the time t is not the actual testing time but is the number of time intervals ft_u. The actual testing time should be $ft_u \cdot t$. Then for degradation process $Z_{f_1 t_u t}$ with inspection frequency $f_1 t_u$, its Laplace exponent $\phi_1(\omega)$ satisfies

$$\mathbb{E}(e^{-\omega Z_{f_1 t_u t}}) = e^{-t f_1 t_u \phi_0(\omega)} = e^{-t\phi_1(\omega)},$$

where $\phi_0(\omega)$ is the Laplace transform of the increment in the single time unit and $\phi_1(\omega) = f_1 t_u \phi_0(\omega)$. Similarly, if $\phi_2(\omega)$ is the Laplace exponent with measurement time interval $f_2 t_u$, then it satisfies $\phi_2(\omega) = f_2 t_u \phi_0(\omega)$. Therefore, the ratio between

$\phi_1(\omega)$ and $\phi_2(\omega)$ is

$$\frac{\phi_1(\omega)}{\phi_2(\omega)} = \frac{f_1}{f_2}.$$

To find the first passage time in terms of the actual testing time, the equation (2.6) should be rewritten as

$$P(\tau_x > f t_u t) = \mathcal{L}^{-1}\{\frac{\exp\left(-\phi(\omega) f t_u t\right)}{\omega}\}(x). \tag{4.5}$$

Suppose $\hat{\phi}(\omega)$ is the empirical Laplace exponent defined in (3.1) based on the original sample set with inspection frequency f, then the new empirical Laplace exponent $\hat{\phi}^*(\omega) = \hat{\phi}(\omega) \cdot \frac{f^*}{f}$ for the new frequency f^*. Hence, the survival probability $P(\tau_x > t)$ under f^* is

$$P(\tau_x > f^* t_u t) = \mathcal{L}^{-1}\{\frac{\exp\left(-\hat{\phi}^*(\omega) \cdot t\right)}{\omega}\}(x) = \mathcal{L}^{-1}\{\frac{\exp\left(-\hat{\phi}(\omega) \cdot \frac{f^*}{f} \cdot t\right)}{\omega}\}(x). \tag{4.6}$$

Replace $f^* t_u t$ by t^*, the equation (4.6) changes to

$$P(\tau_x > t^*) = \mathcal{L}^{-1}\{\frac{\exp\left(-\frac{\hat{\phi}(\omega)}{f t_u} \cdot t^*\right)}{\omega}\}(x). \tag{4.7}$$

Therefore, the estimate of t_p does not depend on f^* and the optimal value of it f^* is always fixed as 1. The optimal solutions of the other two design variables n^* and m^* can be obtained by the complete enumeration method described in Yu and Tseng (1999).

4.3 Data Analysis

In the section, we will still keep analysing the laser data and the optimal designs are compared under the three different settings of degradation processes. The optimization criterion of the degradation test is the asymptotic variance given in (4.2) for the gamma process and inverse Gaussian process or the bootstrap estimate of variance given in (4.5) for the empirical Lévy process. The bootstrap runs N is specified as large as 100,000.

The cost configurations are supposed to be $C_{op} = \$13$/unit time, $C_{mea} = \$0.05$/measurement, $C_{it} = \$51$/unit and $t_u = 24$ hours. If the 90^{th} percentile of the FPT $t_{0.9}$ under threshold 10 is of our interest, the optimal designs with various budgets C_b under different underlying procecsses are shown in Table 4.1-4.3. For example, when the budget $C_b = \$1250$, the optimal designs under the gamma process are given as $\xi^* = (n^*, f^*, m^*) = (16, 2, 16)$ which means 16 laser devices are selected for the experiment and each device is measured for 16 times with time interval 48 hours. The termination time for the experiment is $16 \times 48 = 768$ hours and total cost is \$1244.8. The asymptotic variance for 90^{th} percentile is 4767.339.

The values of the design variables ξ^* showed in Table 4.1 and 4.2 are completely identical and it is understandable that ξ^* in Table 4.3 are different for all the budget levels since its frequency variable f of the nonparametric model is always equal to one.

We can see that the bootstrap estimates of variance $BVar$ for all the budget level C_b in Table 4.3 are significantly greater than the counterpart asymptotic variance $AVar$ in Table 4.1 and 4.2. Such phenomenon means the parametric models may be more suitable for the degradation test based on the laser dataset when the optimization

Table 4.1: Optimal Test Plans Under Gamma Process

C_b	n*	f*	m*	AVar$(t_{0.9}\|\xi^*)$	Cost
1000	13	1	24	9592.011	990.6
1250	16	2	16	4767.339	1244.8
1500	20	3	12	2722.985	1500

Table 4.2: Optimal Test Plans Under Inverse Gaussian Process

C_b	n*	f*	m*	AVar$(t_{0.9}\|\xi^*)$	Cost
1000	13	1	24	9236.610	990.6
1250	16	2	16	4593.874	1244.8
1500	20	3	12	2625.163	1500

criterion is set as the asymptotic variance. As mentioned in Chapter 3, the parametric model can have good performance if this model can correctly capture enough information from the dataset but once misspecification happens, the nonparametric model is an effective approach to avoid such problem.

Substitute the MLE of the parameters, we investigate the goodness-of-fit of the gamma distribution $Gamma(7.19, 14.11)$ and the inverse Gaussian distribution $IG(0.51, 3.41)$ for the laser dataset. To verify it, the Kolmogorov-Smirnov test is conducted. From the results in Table 4.4, the p-values for the hypothesis test are 0.1508 and 0.6216 for the gamma and inverse Gaussian distribution respectively and therefore there is no significant evidence to claim the data are inappropriately fitted by the two parametric models. Then, it is reasonable that the gamma and inverse Gaussian models have comparatively smaller variances since they are seemed as the 'correct' models and can capture enough information to 'beat' nonparametric model when the criterion is the asymptotic variance. In the comparisons between the two parametric models, the test

Table 4.3: Optimal Test Plans Under Empirical Lévy Process

C_b	n*	f*	m*	BVar($t_{0.9}\|\xi^*$)	Cost
1000	10	1	36	10834.54	996.00
1250	13	1	43	6981.344	1249.95
1500	15	1	53	4917.395	1493.75

statistic of the inverse Gaussian model (0.048626) is smaller than the test statistic of the gamma model (0.073379) which means the inverse Gaussian process can fit the degradation data better. Intuitively, we can conjecture it may be the reason why the inverse Gaussian model can get even smaller asymptotic variance than the gamma model but this issue will not be further discussed in this thesis.

The smaller variance only reflects the estimated $100p^{th}$ percentile $t_p(\hat{\theta})$ can be found more stable regard to the true parameter-based $t_p(\theta)$. It does not mean $t_p(\hat{\theta})$ can be more close to the true value of the percentile of the observed FPT. The prediction accuracy is intrinsically limited by the pre-assumed model since no parametric model can capture all the information revealed from the data. We know that $t_p(\theta)$ will change with the selection of the underlying process since the information volume contained in θ are different. However, the true percentile of the observations will not change with the underlying model. In Chapter 3, we compared the parametric and the nonparametric models with the criterion defined as the prediction accuracy. Such criterion puts the different approaches on the same platform. Therefore, the nonparametric approach still possesses the advantages as it includes more information from the data than the parametric models in most occasions.

Table 4.4: The K-S test regarding the gamma and inverse Gaussian processes

	Gamma	IG
D-statistics	0.073379	0.048626
p-value	0.1508	0.6216

4.4 Concluding Remarks

This chapter discusses the optimal design for degradation test when the underlying process is specifically assumed or unspecified. Although the gamma or inverse Gaussian process is usually used as the degradation process when study the degradation test, it is difficult to find an appropriate parametric process to accurately fit the observed degradation data. For the parametric models, the asymptotic variance of $100p^{th}$ percentile is taken as the optimization criterion which is a function of the design variables including number of test units n, the number of measurements m and inspection frequency f. The optimal designs can be determined when the asymptotic variance is minimized. Unlike the parametric evaluation of the optimal design, the inspection frequency f is irrelevant to the value of $100p^{th}$ percentile under the empirical Lévy process and only the n and m are considered subject to the cost not exceeding the budget. From the results of the illustrative example, the parametric models can get the smaller variances for the special optimization criterion although their prediction accuracy is worse than nonparametric approach according to the results of Chapter 3. Therefore, it is appropriate to try different optimization criterion on the degradation tests and it will be the extension of our works in the future.

Chapter 5

Summary and Conclusions

The thesis investigates the FPT of degradation processes from the scope of Lévy jump processes. Our approach releases the constraints that the close-form pdf of any point along the degradation path should be existed and provides a more convenient and general way to study the system reliability. It can be also developed to the distribution-free method which can accurately capture the reliability information and efficiently estimate the FPT distribution percentile. In this chapter, we will summarize each chapter and discuss the potential future works.

5.1 Summary of Results

In Chapter 2, the proposed parametric approach based on Laplace transforms expand the selection of degradation processes to more general one-side Lévy processes which were previously limited to the gamma process or the inverse Gaussian process in most cases. The FPT of the underlying process can be numerically obtained by inverse Laplace transform. With such techniques, we can implement more complicated

models, for example, the multi-sensors model. If a component or a part is very crucial for the whole system, it will probably be monitored by multiple sensors. Since each sensor may report a different degradation path, a weight-convolution of these degradation paths is taken as the underlying degradation process reflecting the overall reliability. One of difficulties of the complicated models is that its likelihood function may be intractable. Then we can solve this estimation problem from other aspects. If the Laplace transform can be in analytic form, a distance-based estimation method such as GMM can optimally estimate the parameters by minimizing the distance between the parametric and empirical Laplace transforms.

When the underlying process is specifically assumed, the numerical FPT density can be found only if its analytic Laplace exponent exists so that the numerical Laplace inversion can be conducted. In Chapter 3, these numerical results are taken as true values and taken as standards to evaluate the nonparametric methods. Rather than exactly giving any process-type assumption and looking for the close-form Laplace transform, we directly obey the information of the sample set itself by replacing the parametric Laplace transform to the empirical Laplace transform. Then the FPT density can be effectively approximated by inverting the Laplace transform using the empirical saddlepoint method. With the FPT density, we can find the $100p^{th}$ percentile of lifetime distribution which is the important index of reliability. As it is always very hard to find a perfect model for a dataset, an incorrectly assumed model may bring large bias to the estimates and make statistical inference inefficiently. The nonparametric method can also help us to get rid of the misspecification problem. From the examples of Chapter 3, it can be seen that our proposed nonparametric method performs significantly better than arbitrarily modelling the data with the

gamma process.

As a practical application, the optimal design for degradation test is studied in Chapter 4 when the underlying processes are assumed to be the gamma process, the inverse Gaussian process or the empirical Lévy process. In this experiment, we look for the optimal values for the design factors which minimize the variation of the reliability index subject to the cost not exceeding the budget. The objective function is defined as the asymptotic variance of the 100^pth percentile of the lifetime distribution for the parametric models and as the bootstrap estimate of variance for the nonparametric model. Since both the gamma process and the inverse Gaussian process can fit the degradation data well, the values of the optimal design variables are completely the same with each other while the empirical Lévy process gives a different result. Compared with the parametric evaluations, the inspection frequency factor under the nonparametric model is irrelevant to the value of $100p^{th}$ percentile. Hence, ignoring the influence of this factor, only the number of test units and the number of measurements are concerned under this scenario.

5.2 Possible Future Work

We can observe that the saddlepoint approximation possesses a large relative bias for the left tail of the FPT distribution and it can be improved by developing the second-order saddlepoint approximation or even including higher-order terms. This future work is also mentioned in the concluding remarks in Chapter 2 and 3.

Another potential direction is developing the single-dimensional degradation process to the multidimensional model since many systems are subjected to multidimensional degradation processes in practice. Therefore, to capture the reliability information of

the system, it is necessary to propose an effective way to measure the competing risks and examine the dependence structures. The multidimensional degradation process can be constructed using the Lévy copula proposed by Kallsen and Tankov (2006). Here each marginal degradation process is assumed to be an one-sided Lévy process.

In terms of the multidimensional degradation process, Wang and Pham (2012) gave a typical example regarding the human body system to illustrate the multivariate competing risks model. The longevity risk of human relies on the health condition of each unit in our body from cells, tissues to organs. These biological units experience gradual degradation of its function as well as the sudden shocks. For a human cardiovascular function, its efficiency of delivering blood to all over the body will begin to reduce due to the emerging risks of vascular sclerosis which causes the arteries harden or blocked. Therefore, the health conditions of organs are multiply correlated and the human life is exposed to the catastrophic risks when the accumulated damage of any organ reaches the certain failure level. Consequently, proposing a method to consider the FPT of a complex system of degradations with various of dependence structure is important in practical significance.

In reliability area, Pan *et al.* (2013), Wang *et al.* (2015b) proposed multidimensional degradation model driven by the Wiener process to estimate the residual life of the products and the gamma-process driven model was discussed by Pan and Balakrishnan (2011) and Wang *et al.* (2015a). Sacerdote *et al.* (2016) gave a numerical method to study the FPT of two-dimensional correlated diffusion processes and Cai *et al.* (2017) constructed the multivariate insurance risk model to study the joint-ruin problems of two risk undertakers by defining the marginal process as Markovian arrival process.

With the proposed methods in the thesis, we can develop the multidimensional

degradation process to the nonparametric frameworks. Rather than giving specific process-type, the marginal process can be empirically modelled with the sample set. Then the competing risks of the system can be investigated by both the copula dependence structure and each marginal process.

Appendix A

Itô Lemma

As an important part composing the foundations of the stochastic calculus, Itô lemma was proved by Kiyoshi Itô in 1944. It extends the methods of calculus to stochastic processes. This lemma is widely used in mathematical finance specifically in derivation of the Black-Scholes equation for option pricing.

Assume X_t is a Itô drift-diffusion process given by the following SDE

$$dX_t = \mu dt + \sigma dW_t,$$

where W_t is a Wiener process. For a twice differentiable scalar function $f(t, x)$ with two entries t and x, it satisfies the following SDE

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma \frac{\partial f}{\partial x} dW_t.$$

Itô lemma is derived by expansion in the regular Taylor series

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dx^2 + \cdots .$$

If x is replaced by X_t and except $dW_t dW_t = dt$, all other products including $dW_t dt$, $dtdt$ and higher order terms are equal to zero.

For example, if X_t follows the geometric Brownian motion satisfies $\frac{dX_t}{X_t} = \mu dt + \sigma dW_t$, the differentiation of $f(X_t) = \ln(X_t)$ can be obtained with Itô lemma

$$d\ln(X_t) = \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dW_t,$$

and solving the equation gives the expression of X_t

$$X_t = X_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right).$$

The underlying process of the Black-Scholes formula is given the geometric Brownian motion and Itô lemma is the essential theoretical support for this option pricing formula.

Appendix B

Existence of the Laplace Transform

We will present the sufficient conditions for the existence of Laplace transform and prove the theorem.

Definition B.0.1 (Piecewise continuous function)

If the function f is a piecewise continuous function on the interval $[a, b]$, then

1. *f is continuous in each subinterval (t_i, t_{i+1}), $i = 0, 1, \ldots, n-1$ as the interval $[a, b]$ is split into finite subintervals with separation points $a = t_0 < t_1 < \cdots < t_n = b$.*

2. *for the discontinuity point t_i we have*

$$\left| \lim_{t \to t_i^{\pm}} f(t) \right| < \infty, \ i = 0, 1, \ldots, n-1.$$

Definition B.0.2 (Exponential Order)

f is of exponential order if there exists constants $c, M > 0$ as well as $A > 0$ such that

$$\left| f(t) \right| \leq M e^{ct}, \ t > A.$$

Theorem B.0.1 (Existence of Laplace Transform)

The Laplace transform $\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)dt$ *exists if*

1. *f is piecewise continuous,*

2. *f is of exponential order.*

Proof. With a positive number A, we can separate the Laplace transform into two parts

$$\int_0^\infty e^{-st} f(t)dt = \int_A^\infty e^{-st} f(t)dt + \int_0^A e^{-st} f(t)dt,$$

and denote $I_1 = \int_A^\infty e^{-st} f(t)dt$ and $I_2 = \int_0^A e^{-st} f(t)dt$. And I_2 exists since f is piecewise continuous.

By the definition of exponential order, we can find constants A, M, c such that $\Big|f(t)\Big| \leq Me^{ct}$. Then for I_1, we have

$$\Big|e^{-st} f(t)\Big| \leq Me^{-(s-c)t}.$$

I_1 also exists for $s > c$ since

$$\int_A^\infty \Big|e^{-st} f(t)\Big|dt \leq M \int_A^\infty e^{-(s-c)t}dt \leq M \int_0^\infty e^{-(s-c)t}dt = \frac{M}{s-c}.$$

Hence, the Laplace transform exists for $s > c$. □

Appendix C

Glivenko-Cantelli Theorem

The Glivenko-Cantelli theorem was proposed by Valery Glivenko and Francesco Cantelli in 1933. It determines the asymptotic behaviour of the empirical distribution function. The proof of theorem refers to the personal website of David Stephens.

Theorem C.0.1 (Glivenko–Cantelli Theorem)

$X_1, \ldots, X_n$ are i.i.d. random variables with cdf F and corresponding empirical distribution function is denoted as F_n. Then,

$$P\left(\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0\right) = 1.$$

Proof. Let $\epsilon > 0$ and fix $k > \frac{1}{\epsilon}$, then we can split $\mathbb{R}$ with the partition points $t_0, t_1, \ldots, t_k$ such that

$$-\infty = t_0 < t_1 \leq \cdots \leq x_{k-1} = \infty.$$

If for $j = 1, \ldots, k-1$

$$F(t_j^-) = P(X_j < t_j) \text{ and } F(t_j) = P(X_j \leq t_j),$$

then

$$F(t_j^-) \leq \frac{j}{k} \leq F(t^j).$$

With $t_{j-1} < t_j$, we have

$$F(t_j^-) - F(t_{j-1}) \leq \frac{j}{k} - \frac{j-1}{k} = \frac{1}{k} < \epsilon.$$

Since with the pointwise convergence, we have

$$\lim_{n\to\infty} F_n(t_j) = F(t_j) \text{ and } \lim_{n\to\infty} F_n(t_j^-) = F(t_j^-).$$

Denote Δ_n as

$$\Delta_n = \max_{j=1,\ldots,k-1} \{|F_n(t_j) - F(t_j)|, |F_n(t_j^-) - F(t_j^-)|\},$$

and it can be seen that

$$\lim_{n\to\infty} \Delta_n = 0.$$

If x lies between t_{j-1} and t_j such that

$$t_{j-1} \leq x < t_j,$$

we have

$$F_n(t_{j-1}) - F(t_{j-1}) - \epsilon \leq F_n(t_{j-1}) - F(t_j^-) \leq F_n(x) - F(x)$$
$$\leq F_n(t_j^-) - F(t_{j-1}) \leq F_n(t_j^-) - F(t_j^-) + \epsilon.$$

Therefore,

$$|F_n(x) - F(x)| \leq \Delta_n + \epsilon,$$

and as n goes to infinity

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \to \infty.$$

If A_ϵ denotes the set of events that the convergence is observed such that $P(A_\epsilon) = 1$, then

$$A = \cap_{\epsilon > 0} A_\epsilon = \lim_{\epsilon \to 0} A_\epsilon \text{ which leads to}$$
$$P(A) = P(\lim_{\epsilon \to 0} A_\epsilon) = \lim_{\epsilon \to 0} P(A_\epsilon) = 1.$$

Therefore,

$$P\left(\lim_{n \to \infty} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| = 0 \right) = 1.$$

□

Appendix D

Derivation of the Saddlepoint Approximation

This proof of Equation (2.7) is basedon Edgeworth expansion proposed by Daniels (1954).

If $\phi(\cdot)$ is the pdf of standard normal distribution, the pdf $g(z)$ of a random variable Z with mean 0 and variance 1 can be expanded into the Gram-Charlier type A series

$$g(z) = \sum_{n=0}^{\infty} \frac{c_n \phi^{(n)}(z)}{n!},$$

where

$$c_n = (-1)^n \int_{-\infty}^{\infty} H_n(z) g(z) dz,$$

and $H_z(z)$ is the n^{th} Hermite polynomial.

Denote μ_i as i^{th} moment of Z, then

$$g(z) = \phi(z)\left(1 + \frac{1}{6}\mu_3(z^3 - 3z) + \frac{1}{24}(\mu_4 - 3)(z^4 - 6z^2 + 3) + \cdots\right).$$

Let Y be a random variable with pdf $f(y)$ which has mean μ and variance σ^2, then $Z = \frac{Y-\mu}{\sigma}$ and $f(y)$ can be derived in the form of Edgeworth expansion

$$f(y) = \frac{\phi(z)}{\sigma}\left(1 + \frac{\rho_3}{6}(z^3 - 3z) + \frac{\rho_4}{24}(z^4 - 6z^2 + 3) + \cdots\right),$$

where

$$\rho_i = \frac{K^{(i)}(0)}{K''(0)^{\frac{i}{2}}}.$$

$K(\cdot)$ is CGF defined in Chapter 2.

Denote

$$f(y; s) = e^{sy - K(s)} f(y)$$

which is a density function since $\int_{-\infty}^{\infty} e^{sy-K(s)} f(y)dy = 1$. And for Y_s follows $f(y; s)$, we can compute

$$\mathbb{E}(Y_s) = K'(s) \text{ and}$$
$$Var(Y_s) = K''(s).$$

Applying Edgeworth expansion to $f(y;\hat{s})$ where $K'(\hat{s}) = y$, we have

$$f(y;\hat{s}) = \frac{1}{\sqrt{2\pi K''(\hat{s})}}\left(1 + \frac{1}{8}\rho^4 + \cdots\right).$$

Therefore, the first-order saddlepoint approximation is given by

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi K''(\hat{s})}}\exp(K(\hat{s}) - \hat{s}y).$$

We can also include the higher-order terms. For example, the second-order saddlepoint approximation can be obtained as

$$\hat{f}(y) = \frac{1}{2\pi K''(\hat{s})}\exp(K(\hat{s}) - \hat{s}y)\left(1 + \frac{K^{(4)}(\hat{s})}{8(K''(\hat{s}))^2} - \frac{5(K^{(3)}(\hat{s}))^2}{24(K''(\hat{s}))^3}\right).$$

Bibliography

Balakrishnan, N., Leiva, V., Sanhueza, A., and Cabrera, E. (2009). Mixture inverse Gaussian distribution and its transformations, moments and applications. *Statistics*, **43**, 91–104.

Berman, M. (1981). Inhomogeneous and modulated gamma process. *Biometrika*, **68**, 143–152.

Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press, Cambridge.

Bertoin, J., van Harn, K., and Steutel, F. (1999). Renewal theory and level passage by subordinators. *Statistics and Probability Letters*, **45**, 65–69.

Birnbaum, Z. and Saunders, S. (1969a). A new family of life distribution. *Journal of Applied Probability*, **6**, 319–327.

Birnbaum, Z. and Saunders, S. (1969b). Estimation for a family of life distribution with application to fatigue. *Journal of Applied Probability*, **6**, 327–347.

Braun, M., Meintanis, S., and Melas, V. (2008). Optimal design approach to GMM estimation of parameters based on empirical transforms. *International Statistical Review*, **76**, 387–400.

Cai, J., Landriault, D., Shi, T., and Wei, W. (2017). Joint insolvency analysis of a shared MAP risk process: A capital allocation application. *North American Actuarial Journal*, **21**, 178–192.

Carrasco, M. and Florens, J. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory*, **16**, 797–834.

Carrasco, M. and Florens, J. (2002). Efficient GMM estimation using the empirical characteristic function. *Working Paper, Department of Economics, University of Rochester*.

Chen, P. and Ye, Z. (2016). Random effects models for aggregate lifetime data. *IEEE Transactions on Reliability*, **66**, 76–83.

Daniels, H. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, **25**, 631–650.

Davison, A. and Hinkley, D. (1988). Saddlepoint approximation in resampling methods. *Biometrika*, **75**, 417–431.

Doksum, K. and Hóyland, A. (1992). Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. *Technometrics*, **34**, 74–82.

Doksum, K. and Normand, S. (1995). Gaussian models for degradation processes-part I: Methods for the analysis of biomaker data. *Lifetime Data Analysis*, **1**, 131–144.

Dufresne, F., Gerber, H., and Shiu, E. (1991). Risk theory with the gamma process. *ASTIN Bulletin*, **21**, 177–192.

Durham, S. and Padgett, W. (1997). A cumulative damage model for system failure with application to carbon fibers and composites. *Technometrics*, **39**, 34–44.

Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Canadian Journal of Statistics*, **9**, 139–172.

Eliazar, I. and Klafter, J. (2004). On the first passage of one-sided Lévy motions. *Physica A*, **336**, 219–244.

Feuerverger, A. (1989). On the empirical saddlepoint approximation. *Biometrika*, **76**, 457–464.

Feuerverger, A. and McDunnough, P. (1981a). On some fourier methods for inference. *Journal of the American Statistical Association*, **76**, 379–387.

Feuerverger, A. and McDunnough, P. (1981b). On the efficiency of empirical characteristic function procedures. *Journal of the Royal Statistical Society, Series B*, **43**, 20–27.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, **50**, 1029–1054.

Hua, D., Al-Khalifa, K., Hamouda, A., and Elsayed, E. (2013). Multi-sensor degradation data analysis. *Chemical Engineering Transactions*, **33**, 31–36.

Jorgensen, B., Seshadri, V., and Whitmore, G. (1991). On the mixture of the inverse Gaussian distribution with its complementary reciprocal. *Scandinavian Journal of Statistics*, **18**, 77–89.

Kalbfleisch, J. and Prentice, R. (2002). *The Statistical Analysis of Failure Time Data, 2nd Edition*. John Wiley & Sons, New York.

Kallsen, J. and Tankov, P. (2006). Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis*, **97**, 1551–1572.

Lageras, A. (2005). A renewal-process type expression for the moments of inverse subordinators. *Journal of Applied Probability*, **42**, 1134–1144.

Lawless, J. and Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, **10**, 213–227.

Levuliene, R. (2002). Semiparametric estimates and goodness-of-fit tests for tire wear and failure time data. *Nonlinear Analysis: Modeling and Control*, **7**, 61–95.

Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons, New York.

Padgett, W. and Tomlinson, M. (2004). Inference from accelerated degradation and failure data based on Gaussian process models. *Lifetime Data Analysis*, **10**, 191–206.

Pan, Z. and Balakrishnan, N. (2011). Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. *Reliability Engineering and System Safety*, **96**, 949–957.

Pan, Z., Balakrishnan, N., Sun, Q., and Zhou, J. (2013). Bivariate degradation

analysis of products based on Wiener processes and copulas. *Journal of Statistical Computation and Simulation*, **83**, 1316–1329.

Park, C. and Padgett, W. (2005). Accelerated degradation models for failure based on geometric brownian motion and gamma processes. *Lifetime Data Analysis*, **11**, 511–527.

Peng, C. (2015). Inverse Gaussian processes with random effects and explanatory variables for degradation data. *Technometrics*, **57**, 100–111.

Prentice, R. and Kalbfleisch, J. (1979). Hazard rate models with covariates. *Biometrics*, **35**, 25–39.

Sacerdote, L., Tamborrino, M., and Zucca, C. (2016). First passage times of two-dimensional correlated processes: Analytical results for the Wiener process and a numerical method for diffusion processes. *Journal of Computational and Applied Mathematics*, **296**, 275–292.

Schmidt, P. (1982). An improved version of the Quandt-Ramsey MGF estimator for mixtures of normal distributions and switching regressions. *Econometrica*, **50**, 501–524.

Shu, Y., Feng, Q., and Coit, D. (2015). Life distribution analysis based on Lévy subordinators for degradation with random jumps. *Naval Research Logistics*, **62**, 483–492.

Singleton, K. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, **102**, 111–141.

Slater, L. (1960). *Confluent Hypergeometric Functions*. Cambridge University Press, Cambridge, England.

Smith, W. (1959). On the cumulants of renewal processes. *Biometrika*, **46**, 1–29.

Storn, R. and Price, K. (1997). Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, **11**, 341–359.

Tankov, P. and Cont, R. (2004). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Financial Mathematics Series, Boca Raton, Florida.

Todorov, V. and Tauchen, G. (2012). Inverse realized Laplace transforms for non-parametric volatility density estimation in jump-diffusions. *Journal of the American Statistical Association*, **107**, 622–635.

Tsai, C.C., Tseng, S.T., and Balakrishnan, N. (2011). Mis-specification analyses of gamma and Wiener degradation processes. *Journal of Statistical Planning and Inference*, **141**, 3725–3735.

Tsai, C.C., Tseng, S.T., and Balakrishnan, N. (2012). Optimal design for degradation tests based on gamma processes with random effects. *IEEE Transactions on Reliability*, **61**, 604–613.

Tseng, S.T., Balakrishnan, N., and Tsai, C.C. (2009). Optimal step-stress accelerated degradation test plan for gamma degradation processes. *IEEE Transactions on Reliability*, **58**, 611–618.

Valsa, J. and Brancik, L. (1998). Approximate formulae for numerical inversion of Laplace transforms. *International Journal of Numerical Modelling*, **11**, 153–166.

van Noortwijk, J. M. (2009). A survey of application of gamma processes in maintenance. *Reliability Engineering and Systems Safety*, **94**, 2–21.

Veillette, M. and Taqqu, M. (2010). Numerical computation of first-passage times of increasing Lévy processes. *Methodology and Computing in Applied Probability*, **12**, 695–729.

Wang, X. (2010). Wiener processes with random effects for degradation data. *Journal of Multivariate Analysis*, **101**, 340–351.

Wang, X. and Xu, D. (2010). An inverse Gaussian process model for degradation data. *Technometrics*, **52**, 188–197.

Wang, X., Balakrishnan, N., Guo, B., and Jiang, P. (2015a). Residual life estimation based on bivariate non-stationary gamma degradation process. *Journal of Statistical Computation and Simulation*, **85**, 405–421.

Wang, X., Balakrishnan, N., and Guo, B. (2015b). Residual life estimation based on nonlinear multivariate Wiener processes. *Journal of Statistical Computation and Simulation*, **85**, 1742–1764.

Wang, Y. and Pham, H. (2012). Modeling the depdent competing risks with multiple degradation processes and random shock using time-varying copulas. *IEEE Transactions on Reliability*, **61**, 13–22.

Wasan, M.T. (1968). On an inverse Gaussian process. *Scandinavian Actuarial Journal*, **51**, 69–96.

Whitmore, G. (1995). Estimating degradation by a Wiener diffusion process subject to measurement error. *Lifetime Data Analysis*, **1**, 307–319.

Whitmore, G. and Schenkelberg, F. (1997). Modeling accelerated degradation data using Wiener diffusion with a scale transformation. *Lifetime Data Analysis*, **3**, 27–45.

Yang, Y. and Klutke, G. (2000). Lifetime-characteristic and inspection-schemes for Lévy degradation processes. *IEEE Transactions on Reliability*, **49**, 377–382.

Yao, Q. and Morgan, B. (1999). Empirical transform estimation for indexed stochastic models. *Journal of the Royal Statistical Society, Series B*, **61**, 127–141.

Ye, Z. and Chen, N. (2014). The inverse Gaussian processes as a degradation model. *Technometrics*, **56**, 302–311.

Ye, Z., Chen, L., Tang, L., and Xie, M. (2014). Accelerated degradation test planning using the inverse Gaussian process. *IEEE Transactions on Reliability*, **63**, 750–763.

Yin, G., Ma, Y., Liang, F., and Yuan, Y. (2011). Stochastic generalized method of moments. *Journal of Computational and Graphical Statistics*, **20**, 714–727.

Yu, H.F. and Tseng, S.T. (1999). Designing a degradation experiment. *Naval Research Logistics*, **46**, 698–706.

Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Reviews*, **23**, 93–123.