# DECLARATION

I declare that this thesis represents my own work, except where due acknowledge is made, and that it has not been previously included in a thesis, dissertation or report submitted to this University or to any other institution for a degree, diploma or other qualifications.

*Signed* ________________________________________

Ho Pak Kei

# ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincerely appreciation to my supervisor, Dr. Jennifer S.K. Chan of the Department of Statistics and Actuarial Science at the University of Hong Kong for her continuous guidance. Her ideas, experience, criticisms and encouragement were indispensable Last but not least, I would like to place on record my great gratitude to those who had given me support, and shared some of my hardest time throughout the entire period of study.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1

# BACKGROUND

## 1.1 Development of Geometric Process model

Because of the accumulated wears or ageing effect, data of the successive operating times of a repairable system after repairs or the failure rate of a medical treatment for a patient exhibit a monotone trend. Data of these types which exhibit a monotone trend are common. Lam (1988a, b) first proposed modeling directly the monotone trend by a monotone process called the Geometric Process (GP).

Definition: Let $X_1,\ X_2,\ldots$ be a set of random variables. If there exists a positive real number $a$ such that $\{Y_i = a^{i-1}X_i,\ i = 1,2,\ldots\}$ forms a renewal process (RP), then $\{X_i,\ i = 1,2,\ldots\}$ is called a Geometric Process (GP), and the real number $a$ is called the ratio of the GP. A RP is a sequence of non-negative random variables mutually independent and identically distribution (IID). See Feller (1949) for the definition of a RP. Hence GP is a new class of models for trend data in which the $i$-th outcome can be discounted geometrically $i-1$ times by the ratio $a$ so that the resulting process, the RP, becomes stationary and can be modeled by a parametric distribution. We define

$$E(Y_i) = \mu \quad \text{and} \quad Var(Y_i) = \sigma^2.$$

Then the mean and variance of a GP are given by

$$E(X_i) = \mu/a^{i-1} \quad \text{and} \quad Var(X_i) = \sigma^2/a^{2(i-1)} \tag{1.1}$$

Thus $a$, $\mu$ and $\sigma^2$ are the three parameters in a GP which completely determine the mean and variance of $X$. The mean parameter $\mu$ indicates the initial level of a trend while the ratio parameter $a$ reveals the progression (increasing or decreasing) of the trend while eliminating the random noises.

Previous GP models focus on modeling the inter-arrival times of a series of events which are positive and continuous. In fact a wider class of data including binary, multinominal and Poisson count data are common. Moreover covariate effects and multiple trends may present in data. To extend the scope of the application of GP models, we set the following objectives for this research.

1. To extend the GP models by incorporating covariates in the mean function and adopting multiple GPs to a single series of data with multiple trends for positive continuous data.

2. To generalize the GP models to the class of binary data by assuming an underlying GP such that the observed binary data are indicators of whether the underlying GP are greater than one.

3. To improve the estimation methodologies for the extended models using non-parametric, maximum likelihood and Bayesian methods.

4. To implement the extended GP models in epidemiological studies, business trend studies and industrial reliability studies by using data in Hong Kong and worldwide.

## 1.2 Structure of this thesis

This thesis is divided into three parts. The first part includes Chapter two and Chapter three which give a general introduction and a review of the establishment of the GP model. In the second part, Chapter four and Chapter five focus on further developments of the GP model for positive continuous data. Chapter four introduces a generalized Geometric Process (GGP) model adopting covariates to the mean function of the renewal process. In Chapter five, we consider the multiple Geometric Process (MGP) model for data with multiple trends. A moving window technique is used to detect the turning points in the series and then a GP model is applied to each trend initiated by a turning point. Parameters are estimated by non-parametric method. The third part contains Chapter six which proposes the Binary Geometric Process (BGP) model as an extension of the GP model to binary data. These models are implemented using parametric method. Finally some discussions and suggestions will be made in Chapter seven.

# CHAPTER 2

# INTRODUCTION

## 2.1 Related models

The non-homogeneous Poisson process (NHPP) in which the hazard rate is monotone has been used to model data $\{X_i,\ i = 1, \ldots, n\}$ with trend. A more direct approach is to model the data by a monotone process. Lam (1988a, b) first proposed modeling directly the monotone trend by a monotone process called the Geometric Process (GP). The underlying renewal process (RP) $\{Y_i = a^{i-1}X_i,\ i = 1, \ldots, n\}$ are independent and identically distributed (IID) with mean $\mu$ which indicates the initial level of a GP. However the adoption of a homogenous mean $\mu$ over time is over-simplified. Experience shows that trend data may exhibit covariate effects, say the daily incidence of an epidemic may depend on environmental factors such as temperature and humidity and other external factors such as precaution measures which may change over time. Hence a direct generalization of the GP models to the generalized GP (GGP) models is to use the framework of exponential family of sampling distributions within Generalized Linear Mixed Models (GLMMs), adopting a linear function of covariates log-linked to the mean of some life time distributions for the RP. Such extension is definitely significant in the development of GP models which reduce to GLMMs when $a = 1$. The inclusion of mixture and random effects into the mean function to accommodate the variability of the data due to overdispersion, clustering and population heterogeneity is another area to explore. However the extended RP under the GGP model is no longer IID and hence is called a stochastic process (SP) in general.

Previous GP model and its extension, the GGP models limit to positive continuous data. For example, the inter-arrival times for a series of events. However, a wider class of data including binary, multi-nominal and Poisson count data are more common. Here we assume the presence of an unobserved GP $\{X_i\}$ and the corresponding RP $\{Y_i\}$ such that the observed data is $W_i = I(X_i > 1)$ where $I(\cdot)$ is an indicator function. By adopting some life time distributions to the RP as in the GP model, the model can be implemented using maximum likelihood (ML) and Bayesian methods.

In GP model, equation (1.1) shows that the ratio parameter of a GP affects both its mean and variance and allows them to change over time. Being the ratio of the expected present outcome to the previous outcome, the ratio parameter not only reveals direction and strength of the progression while eliminating the random noises but also accounts for the auto-correlation between observations as in the auto-regressive (AR) model of the time series modeling. However, the way to model such effects is completely different. In particular, we may adopt a linear function of covariates into the ratio function to model such correlation. Moreover, the ratio parameter may also vary stochastically in a way similar to the stochastic volatility model in which the volatilities are allowed vary stochastically. Such adaptive features of GP models open up a wide area for future development and allow broad range of applications. Descriptions of the NHPP models, the GLMM models are given in the coming sections.

### 2.1.1 Nonhomogeneous Poisson Process (NHPP) model

The nonhomogeneous Poisson process (NHPP) has been widely adopted for trend analyses. If the successive inter-arrival times are monotone, the Cox-Lewis model and the Weibull process model are widely used. The Cox-Lewis (CL) model is a NHPP model with hazard function is

$$\rho_1(x) = \exp(\alpha_0 + \alpha_1 x), \qquad -\infty < \alpha_0, \alpha_1 < \infty \quad x > 0$$

The CL model can be applied to data when the successive inter-arrival times are decreasing if $\alpha_1 \geq 0$ and increasing if $\alpha_1 \leq 0$. It is obvious that the CL model is reduced to a HPP model for $\alpha_1 = 0$.

For the Weibull process (WP) model, it is a NHPP model with the hazard function

$$\rho_2(x) = \alpha\theta x^{\theta-1}, \quad \alpha, \theta > 0, \quad x > 0$$

The WP model is applicable to model a series with increasing inter-arrival times if $0 < \theta < 1$ and decreasing inter-arrival times if $\theta \geq 0$. If $\theta = 1$, the WP reduces to a HPP process. In fact, we have a more direct approach to model the data by a monotone process, the Geometric Process (GP)introduced by Lam (1988a, b).

### 2.1.2 Generalized Linear Mixed Model (GLMM)

One important area in statistics is to develop model that can adequately describe the phenomenon of the data. Simple linear model has often been used to investigate the relationship between response and explanatory variables. With $\boldsymbol{Y}$ being a vector of $n$ responses and $\boldsymbol{X}$ being a $n \times p$ matrix of the explanatory variables, the linear model is defined as:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression parameters and $\boldsymbol{\epsilon}$ is a $n \times 1$ vector of error terms which follows a multivariate normal distribution with zero mean and variance $\sigma^2\boldsymbol{I}$ where $\boldsymbol{I}$ is an identity matrix.

However, the simple linear model does not work very well if the observations are serially correlated. This type of data is commonly observed in clinical trials and stock markets. More importantly, the simple linear model cannot be used in the case when the responses are neither normally distributed nor linearly related to the explanatory variables. For instance, the responses may be categorical.

Nelder and Wedderburn (1972) introduced the Generalized Linear Models (GLMs) that allow an exponential family of sampling distributions on the data in order to describe the non-normal responses. These models have unified regression methodology for a wide variety of response (discrete, counted and continuous) that can be assumed to be independent. In GLMs, we first assign a distribution $D$ to any random variable $Y$ in $\boldsymbol{Y}$

within an exponential family is as follows:

$$y_i \overset{indept.}{\sim} f(y_i)$$

where $f(.)$ is the density function of $D$ with mean $\mu = E(\mathrm{Y}) = b'(\boldsymbol{\theta})$ and variance $Var(\mathrm{Y}) = \phi\, b''(\boldsymbol{\theta})/A = \phi v(\mu)/A$ where $b(.)$ is a function of canonical parameters $\boldsymbol{\theta}$, $\phi$ is the dispersion parameter, $A$ is a known prior weight and $v(\mu) = b''[(b')^{-1}(\mu)]$ is called the variance function. There are various kinds of exponential family of distributions including Binomial distribution, Poisson distribution, Gamma distribution and Normal distribution, etc.

After defining the exponential family of distributions for $Y_i$, the next step will be to link a linear function of predictors:

$$\eta_i = \boldsymbol{x_i}\,\boldsymbol{\beta}$$

where $\boldsymbol{x}_i$ is a $1 \times p$ row vector of the matrix of explanatory variables $\boldsymbol{X}$, to the mean $\mu_i$ of the response via a monotone link function $g(.)$ such that $g(\mu_i) = \eta_i$. There are various kinds of link functions. For binary data, the common link functions are logit-link $g(\mu) = \mu/(1-\mu)$ and probit-link $g(\mu) = \Phi^{-1}(\mu)$ for $0 < \mu < 1$. For positive continuous and Poisson count, we have log-link $g(\mu) = ln(\mu)$. For continuous data, we can have an identity-link as $g(\mu) = \mu$. Each member of the exponential family of distributions adopts different canonical link function.

Although the GLMs solve the problem of non-normality in the data, it cannot deal with data which are clustered or longitudinal where overdispersion and auto-correlation are often of concern. Then GLMs were further extended to a more general class of models, the Generalized Linear Mixed Models (GLMMs) by inclusion of random effect terms in the mean function:

$$\eta_i = \boldsymbol{x_i}\,\boldsymbol{\beta} + \boldsymbol{z}_i\,\boldsymbol{u}$$

where $\boldsymbol{z}_i$ is a $1 \times q$ design vector for the unobserved random effects and $\boldsymbol{u}$ is a $q \times 1$ vector of the random effect parameters. GLMMs can be used to account for the serial correlation and clustering effects in the data from longitudinal studies, overcome the problem of over dispersion and at the same time, accommodate the population heterogeneity. Thus, the GLMM model on an independent response given the random effects $\boldsymbol{u}$ becomes:

$$\begin{aligned} f(y_i|\boldsymbol{u}) &\overset{indept.}{\sim} f(y_i) \\ E[y_i|\boldsymbol{u}] &= \mu_i \\ g(\mu_i) &= \boldsymbol{x_i}\,\boldsymbol{\beta} + \boldsymbol{z}_i\,\boldsymbol{u} \end{aligned}$$

Since the responses $\boldsymbol{Y}$ in the GLMMs depend on the random effects $\boldsymbol{u}$ which is unobserved, we usually assign a distribution to the random effects, that is, $\boldsymbol{u} \sim h(\boldsymbol{u})$.

## 2.2 The life time distribution

GP model is first applied to model the inter-arrival times of failures of a system Lam (1988b). Lam and Chan (1998) and Chan *et al.* (2004a) consider parametric inference adopting some life time distributions say the Lognormal and Gamma for the renewal process (RP). The distributions for the RP $\{Y_i = a^{i-1}X_i,\ \ n = 1, 2, \ldots\}$ in the analyses of inter-arrival times data are usually positively skewed, originate from a finite threshold on the left and then tail off to zero on the right since it is rare to have a unusually long life time.

Hazard rate is the instantaneous failure rate at time $t$ given that the system has survived until time $t$. It is an important function associated with a distribution function $F(t)$ and its corresponding density function $f(t)$ defined as

$$h(t) = \frac{f(t)}{1 - F(t)}$$

The characteristics of different hazard functions identify different distributions for positive continuous data.

### 2.2.1 Lognormal distribution

The Lognormal distribution is related to Normal distribution in a way similar to that of Weibull distribution to Extreme-value distribution. If a random variable $X$ is distributed as Normal with mean $\lambda$ and variance $\tau^2$, that is $X \sim N(\lambda, \tau^2)$, then $Y = \exp(X)$ is said to follow a Lognormal distribution, that is $Y \sim LN(\lambda, \tau^2)$. In the reliability and life-time testing, Lognormal distribution is preferred to Normal distribution because it ranges from 0 to $+\infty$ rather than $-\infty$ to $+\infty$. The probability density function (pdf) of the Lognormal distribution is:

$$f(y)=\begin{cases} \frac{1}{\sqrt{2\pi}\tau y} \exp\left[-\frac{1}{2\tau^2}(\ln y - \lambda)^2\right], & y \geq 0, \\ 0, & y < 0, \end{cases} \tag{2.1}$$

and the mean and variance of $Y$ become

$$E(Y) = \exp\left(\lambda + \frac{1}{2}\tau^2\right) \text{ and } Var(Y) = \exp\left(2\lambda + \tau^2\right)\left[\exp(\tau^2) - 1\right] \tag{2.2}$$

respectively. The hazard function $h(y)$ of the Lognormal distribution cannot be expressed in a closed form. However, it was proved that $h(y)$ will be decreasing when $Y$ is large. Therefore, the Lognormal distribution is quite skewed and it is a good model when the failure rate is quite high initially and then decreases as time increases.

### 2.2.2 Exponential distribution

The simplest distribution for modeling life-time data is the one-parameter Exponential distribution with the pdf

$$f(y)=\begin{cases} \lambda e^{\lambda y}, & y \geq 0, \quad \lambda \geq 0, \\ 0, & y < 0, \end{cases} \tag{2.3}$$

where $\lambda$ is called the scale parameter of the distribution, its density curve is reverse J-shaped. The mean and the variance of the distribution are

$$E(Y) = \frac{1}{\lambda} \text{ and } Var(Y) = \frac{1}{\lambda^2}$$

respectively. One characteristic of the Exponential distribution is its constant hazard function.

$$h(y) = \lambda \tag{2.4}$$

Therefore, it is suitable for modeling life time data of a system such that any used system is considered as good as a new one. This is called the 'memory less' property of the Exponential distribution.

### 2.2.3 Gamma distribution

The Gamma distribution arises as the distribution of a sum of $\alpha$ IID random variables each following an Exponential distribution and its corresponding counting process will be Poisson distributed. The pdf of the two parameters Gamma distribution $Ga(\alpha, \lambda)$ is given by

$$f(y)=\begin{cases} \frac{1}{\Gamma(\alpha)}\lambda^{\alpha}y^{\alpha-1}e^{-\lambda y} & y \geq 0, \\ 0, & y < 0, \end{cases} \tag{2.5}$$

where $\alpha(> 0)$ is the shape parameter, $\lambda(> 0)$ is the scale parameter and $\Gamma(\alpha)$ is the Gamma function defined by

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}dt \tag{2.6}$$

The mean and variance of the distribution are

$$E(Y) = \frac{\alpha}{\lambda} \quad \text{and} \quad Var(Y) = \frac{\alpha}{\lambda^2} \tag{2.7}$$

respectively.

Similar to Lognormal distribution, there is no simple closed form expression available for the cumulative distribution function *F(y)* and its hazard rate function $h(y)$. However, some studies show that for $\alpha$ <1, *h(0)* = *0* and $h(y)$ approaches $\lambda$ asymptotically from below as $Y$ tends to $\infty$. For $\alpha > 1$, *h(0)* $=\infty$ and $h(y)$ approaches $\lambda$ asymptotically from above as $Y$ tends to $\infty$. This suggests that Gamma distribution may be useful for modeling the inter-arrival times of failures when the system is under a regular maintenance program. The failure rate may increase initially indicating the "ageing effect" but after certain period the system would become stable as the maintenance program continues. When $\alpha = 1$, the Gamma distribution reduces to an Exponential distribution.

### 2.2.4 Weibull distribution

Weibull distribution provides an alternative generalization of the Exponential distribution. If the random variable $Y = X^\alpha$ where $X$ follows the Exponential distribution and $Y_i$ will follow a Weibull distribution with shape parameter $\alpha$ and scale parameter $\lambda$, its pdf is

$$f(y)=\begin{cases} \alpha\lambda^\alpha y^{\alpha-1}e^{-(\lambda y)^\alpha} & y \geq 0, \\ 0, & y < 0, \end{cases} \tag{2.8}$$

The mean and variance for the distribution are

$$E(Y) = \frac{\Gamma_1}{\lambda}$$
$$Var(Y) = \frac{\Gamma_2 - \Gamma_1^2}{\lambda^2} \tag{2.9}$$

where $\Gamma_k = \Gamma(1 + k\alpha)$ and $\Gamma(.)$ is the Gamma function. Moreover, the hazard function is given by

$$h(y) = \alpha\lambda^\alpha y^{\alpha-1} \tag{2.10}$$

The density curve for this distribution can be either reverse J-shaped or bell-shaped, it depends on the shape parameter $\alpha$. For instance, $h(y)$ is decreasing for $\alpha \leq 1$ and it is increasing for $\alpha \geq 1$. When $\alpha = 1$, $h(y) = \lambda$ and the distribution becomes Exponential distribution. Weibull distribution has been shown experimentally that it provides a good fit for many different type of characteristics and is thus used extensively in life testing and reliability problems. For example, the 'wear-out' or fatigue failure of the vacuum tube failures in Kao (1959). As $\alpha$ becomes large, it gets more peaked and symmetric about 1. As $\alpha$ decreases, it is peaked and asymmetric with a long tail to the right. When $\alpha = 3.6024$, it becomes almost normal in shape.

## 2.3 Methodologies for inference

In statistical inference, Lam (1992b) proposed some non-parametric (NP) methods for the GP models and Lam and Chan (1998) and Chan *et al.* (2004a) considered parametric likelihood method adopting respectively the Lognormal and Gamma distributions to the RP. They derived large sample distributions for the parameter estimates and showed that the parametric methods are more efficient. Lam *et al.* (2003) derived some limit theorems in GP model. Then Lam *et al.* (2004) showed that the GP model performed better than the Cox-Lewis model, the Weibull process model and the homogeneous Poisson process model and was easier to implement using NP methods.

### 2.3.1 Non-parametric inference

In NP inference, we would mainly focus on the method of Least Square Estimate on the observations $X_i$ and $\ln X_i$ with method LSE and log-LSE respectively. In this approach, the parameter estimates are chosen to minimize a quantity called the residual sum of squares.

In Chan *et al.* (2004b), LSE estimates obtained by minimizing the function $LSE = \sum_{i=1}^{n}(X_i - E(X_i))^2$ where $E(X_i) = \frac{\mu}{a^{i-1}}$ is the expected $X_i$ under the GP model. However, Lam (1992b) shows that any set of parameters estimates can also be obtained by minimizing log-LSE function given by log-LSE $= \sum_{i=1}^{n}(X_i - \ln(E(X_i)))^2$. It can be proved in Appendix A that the log-LSE method for parameters estimation is equivalent to linear regression technique. A detail description using LSE and log-LSE is given in Chapter three. In addition, basing on the idea of linear regression technique, a useful way to test the consistency of the GP is to plot $\ln X_i$ against time in order to provide a rough view for the common ratio $a$.

## 2.3.2 Parametric inference

### 2.3.2.1 Maximum Likelihood method

Suppose we have a $n$-dimensional vector of responses $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^T$. We are interested in the parameter estimate $\boldsymbol{\theta}$. We denote the density of $y_i$ as $f(y_i; \boldsymbol{\theta})$. Then the joint density function is expressed as:

$$f(\boldsymbol{y}; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta})$$

In classical Maximum Likelihood (ML) inference, we assign a model $f(\boldsymbol{y}; \boldsymbol{\theta})$ to the observed data $\boldsymbol{y}$ and the information on the data are represented by the likelihood function $L(\boldsymbol{\theta}; \boldsymbol{y}) = f(\boldsymbol{y}; \boldsymbol{\theta})$. The parameter $\boldsymbol{\theta}$ of the model is treated as fixed but unknown. We differentiate the log-likelihood function $\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \ln L(\boldsymbol{\theta}; \boldsymbol{y})$, set $\ell'(\boldsymbol{\theta}; \boldsymbol{y})$ to zero and solve for the parameter estimate $\widehat{\boldsymbol{\theta}}$ which maximizes the likelihood function of the observed data. In the discrete case, the ML estimator of $\boldsymbol{\theta}$ is the value of the parameter $\boldsymbol{\theta}$ that would make the observed sample most likely. Sometimes, the solution of $\ell'(\boldsymbol{\theta}) = \mathbf{0}$ cannot be obtained explicitly. Newton Raphson (NR) method is an iterative method that provides a numerical solution from a system of complicated equations.

Let $\ell'(\boldsymbol{\theta}; \boldsymbol{y})$ and $\ell''(\boldsymbol{\theta}; \boldsymbol{y})$ be the first and second derivatives of $\ell(\boldsymbol{\theta}; \boldsymbol{y})$. In each NR iteration, current parameters estimates $\boldsymbol{\theta}^{(k)}$ in the $k$-th iteration can be updated to $\boldsymbol{\theta}^{(k+1)}$ the $(k+1)$-th iteration by

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - [\ell''(\boldsymbol{\theta}^{(k)}; \boldsymbol{y})]^{-1} \ell'(\boldsymbol{\theta}^{(k)}; \boldsymbol{y}) \tag{2.11}$$

and the procedure continues until $\| \boldsymbol{\theta}^{(k+1)} - \boldsymbol{\theta}^{(k)} \|$ is sufficiently small.

In ML method, no sampling from non-standard distributions is required but we have to write complicated likelihood functions which may involve integration and solves complicated equations. Numerical approximation is required when the likelihood function involves integration.

#### 2.3.2.2 Bayesian method

Bayesian method is an alternative method to ML method. The idea in Bayesian approach can be elaborated as follows.

First, let $\boldsymbol{x}$ denote the observed data, and $\boldsymbol{\theta}$ denote model parameters and missing data. Formal inference requires the setting up of a joint distribution when density function $f(\boldsymbol{x}, \boldsymbol{\theta})$ over all random quantities. The joint distribution consist of two parts, a prior distribution with density function $f(\boldsymbol{\theta})$ and a likelihood function $f(\boldsymbol{X}|\boldsymbol{\theta})$, in which,

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = f(\boldsymbol{x}|\boldsymbol{\theta}) f(\boldsymbol{\theta}) \tag{2.12}$$

where $f(\boldsymbol{X}|\boldsymbol{\theta})$ is the conditional density given $\boldsymbol{\theta}$. Then Bayes theorem is applied to determine the joint posterior density for the parameters $\boldsymbol{\theta}$ conditional on $\boldsymbol{x}$ given by

$$f(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{\theta}) f(\boldsymbol{x}|\boldsymbol{\theta})}{\int f(\boldsymbol{\theta}) f(\boldsymbol{x}|\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{2.13}$$

In Bayesian method, model parameters are not fixed but follow some prior distributions based on prior information. If no prior information is available, non-informative priors with large variance are used. Then a Bayesian estimate of $\boldsymbol{\theta}$ is given by the posterior sample mean or median (some central location measures) of the posterior distribution. Moreover, the posterior samples also enable us to construct confidence intervals for $\boldsymbol{\theta}$. More information on the Bayesian statistics can be found in Lee (1989) and Bernardo and Smith (1994).

In Bayesian inference, no complicated likelihood function and its derivatives are required. However, one has to either derive analytically the posterior densities and their expected values for each parameter which may be complicated or sample the parameters from the posterior distributions. Very often, the posterior distributions are not in standard form. Then one may perform the sampling using some Markov Chain Monte Carlo (MCMC) methods, eg. Gibbs sampler or Metropolis Hasting algorithm. Clayton (1996) and Zeger and Karim (1991) had studied the use of Gibbs sampling approach on the GLMMs. Geman and Geman (1984) and Gelfand and Smith (1990) had discussed the Gibbs sampler in details.

The Gibbs sampling technique is illustrated below. Suppose we have three variables $X$, $Y$, and $Z$. We define the joint distribution as $[X, Y, Z]$ which is a complicated distribution with the conditional distribution of $X$, $Y$ and $Z$ as $[X|Y, Z]$, $[Y|X, Z]$ and $[Z|X, Y]$ respectively which are comparatively simpler.

The algorithm of Gibbs sampler is as follows:

1. Started with an initial value of $X^{(0)}$, $Y^{(0)}$ and $Z^{(0)}$.

2. Draw $X^{(1)}$ from the conditional distribution of $[X|Y^{(0)}, Z^{(0)}]$.

3. Draw $Y^{(1)}$ from the conditional distribution $[Y|X^{(1)}, Z^{(0)}]$ based on $Z^{(0)}$ and the newly simulated $X^{(1)}$.

4. Draw $Z^{(1)}$ from $[Z|X^{(1)}, Y^{(1)}]$ based on the newly simulated $X^{(1)}$ and $Y^{(1)}$.

5. Complete the first iteration. We will repeat this algorithm until $N$ iterations have completed and the simulated values converged to the joint density function.

Suppose totally we have $N$ iterations, the first $K$ iterations in the burn-in period before convergence will be discarded. The remaining $M = N - K$ iterations are used to form posterior samples. To check for convergence, auto-correlation is used by plotting the series of the simulated values and examining their auto-correlation function respectively for each variable. The methodology can be easily implemented using WinBUGS. For ML, it is implemented by writing FORTRAN program with IMSL Library.

## 2.4 Application

Lam (1988a,b) first introduced the geometric process (GP) model as a special monotone process and showed that the GP model can model such kind of data well. During the past fourteen years, increasing attentions have been drawn on the GP models and new area of application has been derived. The GP models have been applied extensively to deteriorating systems modeling and to reliability and maintenance problem in the study of optimal replacement or repairable models. To name a few, Lam (1992a) considered the optimal GP replacement model. Then Lam (1995) and Lam and Zhang (1996b) applied the GP models to analyze a two-component system arranged in parallel and Lam and Zhang (1996a) considered the case when they are arranged in series. Moreover Lam (1997) applied the models to study the rate of occurrence of failures. Furthermore Stanley (1993) used the models in repair replacement problems, Zhang (1999) in a cold standby repairable system, Sheu (1999) in an extended optimal replacement model for deteriorating systems, Neuts *et al.* (2000) in repairable models with operating and repair times governed by phase type distribution, Zhang *et al.* 2001) in an optimal replacement policy with preventive maintenance, Lam *et al.* (2002) in a multi-state degenerative system, Zhang (2002) in a repair model with good-as-new preventive repair, Zhang *et al.* (2002) in an optimal replacement policy for a multi-state repairable system, Lam and Zhang (2003) in a deteriorating system under a random environment and Lam and Tse (2003) in a multistage deteriorating system. Recently the models have received more international attentions and an article entitled Geometrical Process has been written by

Lam, Y. in *Encyclopedia of Statistical Sciences*. See Lam (2004) for a brief review and further reference.

Business cycle and seasonal effect always appear in a economic series such as stock market index and Gross National Income (GNP), etc. GGP model can be applied in these data since GGP can capture the seasonal covariate effects. On the other hand, if the data exhibit more than one trend, MGP model can identify different stages of the series. In Chan (2004b), MGP model is applied to model an epidemic diseases outbreak that contains a growing stage, stabilizing stage and declining stage. For binary data, trend may exist in the underlying probability that contain event occur. The BGP model aims to uncover such trend in a binary data series. To sum up, extending the GP model widens its scope of applications considerably.

# CHAPTER 3

# REVIEW OF GEOMETRIC PROCESS MODEL

## 3.1 Test for Geometric Process

Lam (1992b) provided a detailed description of tests that test whether the data are consistent with a GP and whether the ratio parameter $a$ equals to one. They are reviewed as below.

### 3.1.1 Test for HPP

Laplace test is commonly used to determine whether $\{X_i, \;\; i = 1, 2, ..., n\}$ is a homogeneous Poisson process (HPP) process. Let $T_i$ denote the $i$-th arrival-time of certain events, that is,

$$T_i = \sum_{j=1}^{i} X_j$$

Under the HPP assumption,

$$U = \frac{\frac{\sum_{i=1}^{n} T_i}{n-1} - \frac{T_n}{2}}{T_n\sqrt{\frac{1}{12(n-1)}}} \sim N(0,1)$$

asymptotically. [See Cox and Lewis (1966)]

### 3.1.2 Test for GP

We can further test whether the data come from a GP if a trend probably exists. Lam (1992b) suggested the following method based on two theorems.

Theorem 1. If $\{X_i,\ i = 1, 2, \ldots\}$ is a GP, then $U_i's,\ i = 1, 2, \ldots, m$ are IID random variables where $U_i = \dfrac{X_{2i}}{X_{2i-1}}$ and m is an integer.

Theorem 2. If $\{X_i,\ i = 1, 2, \ldots\}$ is a GP, then $V_i's,\ i = 1, 2, \ldots, m$ are IID random variables, where $V_i = X_i X_{2m+1-i}$ and m is an integer.

Thus, based on theorem 1 and 2, $U_i's$ and $V_i's$ are IID random variables where

$$U_i = \frac{X_{2i}}{X_{2i-1}} \text{ and } V_i = X_i X_{2m+1-i},\ \ i = 1, 2, \ldots, m \ \text{ for } \ n = 2m$$
$$U_i = \frac{X_{2i}}{X_{2i-1}} \text{ and } V_i = X_{i+1} X_{2m+2-i},\ \ i = 1, 2, \ldots, m \ \text{ for } \ n = 2m+1$$

To test whether the variables $U_i's$ and $V_i's$ are IID, the following two tests were introduced by Lam (1992b). Let $\{w_i,\ i = 1, 2, ..., m\}$ be a sequence of IID random variables.

1. Turning point (TP) test

   The test statistics

   $$T_w^* = \frac{T_w - \frac{2(m-2)}{3}}{\sqrt{\frac{16m-29}{90}}} \quad \sim N(0,1)$$

   asymptotically where

   $$T_w = \sum_{i=2}^{m-1} I[(w_i - w_{i-1})(w_{i+1} - w_i) < 0],$$

   where $I(A)$ denotes an indicator function of the event A.

2. Difference sign (DS) test

   The test statistics

$$D_w^* = \frac{D_w - \frac{m-1}{2}}{\sqrt{\frac{m+1}{12}}} \sim N(0,1)$$

   asymptotically where

$$D_w = \sum_{i=2}^{m} I(w_i > w_{i-1})$$

3. Graphical method

   Besides the tests mentioned in above, Lam (1992b) provided another possible graphical test. By taking logarithm of $Y_i = a^{i-1}X_i, \; i = 1, 2, ..., n$ at both sides, it results

$$\ln X_i = \ln Y_i - (i-1)\ln a, \qquad i = 1, 2, ..., n \tag{3.1}$$

   Since the renewal process (RP) are IID variables, Using the Simple Linear Regression (SLR) method, $Y_i$ can be written in the form

$$\ln Y_i = \lambda + \epsilon_i \tag{3.2}$$

   where $E(Y_i) = \lambda$ and $Var(\ln Y_i) = Var(\epsilon_i) = \tau^2$. Combining equations (3.1) and (3.2), we have

$$\ln X_i = (\lambda + \ln a) - i \ln a + \epsilon_i \tag{3.3}$$

   Also, from equation (3.3), we can plot $\ln X_i$ against $i$ and a linear relation indicates that $\{X_i\}$ comes from a GP.

4. Test for $(a = 1)$

Based on equation (3.15), a test is constructed for testing the hypothesis

$$H_0 : a = 1 \qquad \text{vs} \qquad H_1 : a \neq 1.$$

Under $H_0$, the test statistics is

$$T = \frac{n^{\frac{3}{2}}(\hat{a}_{NP1} - 1)}{\sqrt{12}\hat{\tau}} \sim N(0, 1) \tag{3.4}$$

The parameter $\hat{a}_{NP1}$ will be introduced in the next section.

## 3.2 Non-parametric inference

### 3.2.1 Least Square method for $\ln X_i$

From equation (3.3), the sum of squared errors ($SSE$) on $\ln X_i$ is

$$SSE = \sum_{i=1}^{n} [\ln x_i + (i-1)\ln a - \lambda]^2 \tag{3.5}$$

Then the least square estimate (LSE) of $\ln a$ and $\tau^2$ can be obtained by minimizing $SSE$.

The first order derivatives of $SSE$ are

$$\frac{\partial SSE}{\partial \ln a} = 2\sum_{i=1}^{n} [\ln x_i + (i-1)\ln a - \lambda](i-1) \tag{3.6}$$

$$\frac{\partial SSE}{\partial \lambda} = -2\sum_{i=1}^{n} [\ln x_i + (i-1)\ln a - \lambda] \tag{3.7}$$

Setting equations (3.6) and (3.7) to zero and solving for $\ln a$ and $\lambda$, the LSE for $\ln Y_i$ are

$$\widehat{\ln a}_{NP1} = \frac{6}{(n-1)n(n+1)} \sum_{i=1}^{n} (n - 2i + 1) \ln x_i \tag{3.8}$$

$$\widehat{\lambda}_{NP1} = \frac{2}{n(n+1)} \sum_{i=1}^{n} (2n - 3i + 1) \ln x_i \tag{3.9}$$

An estimate for the variance $\hat{\tau}^2$ can be obtained by taking the sample variance of $\hat{\epsilon}_i = (i-1)\widehat{\ln a} + \ln x_i - \hat{\lambda}$ and it can be proved that

$$\hat{\tau}^2_{NP1} = \frac{1}{n-2} \left[ \sum_{i=1}^{n} (\ln x_i)^2 - (\sum_{i=1}^{n} \ln x_i)^2/n - \frac{\widehat{\ln a}}{2} \sum_{i=1}^{n} (n - 2i + 1) \ln x_i \right] \tag{3.10}$$

[See Appendix A for the proof]

As this method considers the LSE for $\ln X_i$, we called this method the log-LSE method. Consequently, the log-LSE estimate of $a$ is given by

$$\widehat{a}_{NP1} = \exp(\widehat{\ln a}_{NP1}) \tag{3.11}$$

Since $\ln Y_i = \lambda + \epsilon_i$ in equation (3.2), we have

$$E(Y_i) = E(e^{\lambda + \epsilon_i}) = e^{\lambda} E(e^{\epsilon_i}) = e^{\lambda} E(1 + \frac{\epsilon_i}{2} + \ldots)$$

Hence,

$$\hat{\mu}_{NP3} = (1 + \frac{\hat{\tau}^2_{NP1}}{2}) \exp(\hat{\lambda}) \tag{3.12}$$

To estimate the mean of $Y_i$, another approach is to use the modified moment method. Let $S_n = \sum_{i=1}^{n} x_i$. Then the GP equation gives $E(S_n) = \frac{\mu(1 - a^{-1})}{1 - a^{-n}}$. Hence, a moment estimate of $\mu$ is given by

$$\hat{\mu}_{NP1} = \frac{S_n(1 - \hat{a}^{-1}_{NP1})}{1 - \hat{a}^{-n}_{NP1}} \tag{3.13}$$

Alternatively, one can consider using $\hat{y}_i = \hat{a}^{i-1}_{NP1} x_i$ to estimate $\mu$, hence another moment estimate is given by

$$\hat{\mu}_{NP4} = \frac{1}{n} \sum_{i=1}^{n} \hat{y}_i \tag{3.14}$$

Among these estimates for $\mu$, simulation studies in Lam (1992b) showed that $\hat{\mu}_{NP1}$ provides the best estimation. Thus, we use $\hat{\mu}_{NP1}$ in the log-LSE method for the NP inference in latter chapters. In Lam *et al.* (2003), large sample theory is derived for $a$, $\lambda$ and $\tau$ given by

$$n^{3/2}(\widehat{a}_{NP1} - a) \xrightarrow{L} N(0, 12a^2\tau^2) \tag{3.15}$$

$$n^{\frac{1}{2}}(\hat{\lambda}_{NP1} - \lambda) \xrightarrow{L} N(0, 4\tau^2) \tag{3.16}$$

$$n^{\frac{1}{2}}(\hat{\tau}^2_{NP1} - \tau^2) \xrightarrow{L} N(0, \frac{2n^2}{(n-2)^2}\tau^4) \tag{3.17}$$

where the notation $\xrightarrow{L}$ means the convergence in distribution as the sample size $n$ increases to infinity. Hence, it is obvious that $\widehat{a}_{NP1}$ is strongly consistent as its variance is very small.

### 3.2.2 Least Square method for $X_i$

Beside the log-LSE method in NP inference, Chan *et al.* (2004b) proposed the LSE for $a$ and $\mu$ by minimizing

$$SSE = \sum_{i=1}^{n}(X_i - \mu a^{-(i-1)})^2$$

directly. Then $\hat{a}_{NP2}$ and $\hat{\mu}_{NP2}$ are the solutions to the equations

$$\begin{cases} \sum_{i=1}^{n}(i-1)(\mu a^{-(2i-1)} - X_i a^{-i}) = 0, \\ \sum_{i=1}^{n}(\mu a^{-(2i-1)} - X_i a^{-i}) = 0, \end{cases}$$

$$\Rightarrow \begin{cases} \left(\sum_{i=1}^{n} \frac{X_i}{a^i}\right)\left(\sum_{i=1}^{n} \frac{i-1}{a^{2i}}\right) = \left(\sum_{i=1}^{n} \frac{(i-1)X_i}{a^i}\right)\left(\sum_{i=1}^{n} \frac{1}{a^{2i}}\right), \\ \mu = \dfrac{\sum_{i=1}^{n} \frac{X_i}{a^i}}{\sum_{i=1}^{n} \frac{1}{a^{2i}}}. \end{cases} \tag{3.18}$$

## 3.3 Parametric inference

In parametric inference, we assume the renewal process $\{Y_i\}$ where $Y_i = a^{i-1}X_i$ follows a distribution, including Lognormal, Exponential, Weibull and Gamma. We consider both the Maximum Likelihood (ML) method and Bayesian method.

### 3.3.1 Lognormal distribution

Lam and Chan (1998) studied the GP model with Lognormal distribution. It can be shown that the parameter estimates using ML method with Lognormal distribution are similar to those from the log-LSE method. If the RP $\{Y_i\}$ follows a Lognormal distribution, that is $\ln Y_i \sim N(\lambda, \tau^2)$ where $E(\ln Y_i) = \lambda$ and $Var(\ln Y_i) = \tau^2$. From equation (2.1), the density function for the observed $X_i$ is given by

$$f_i(x_i) = \frac{1}{\tau\sqrt{2\pi}x_i} \exp\left\{-\frac{[\ln(a^{i-1}x_i) - \lambda]^2}{2\tau^2}\right\}$$

since

$$f_i(x_i) = \frac{dy_i}{dx_i} f(y_i) = a^{i-1} f(y_i)$$

The likelihood function for the observed data $\{X_i\}$ becomes

$$L = \frac{1}{\tau^n (2\pi)^{\frac{n}{2}} \prod_{i=1}^{n} x_i} \exp\left\{-\frac{\sum_{i=1}^{n} [\ln(a^{i-1}x_i) - \lambda]^2}{2\tau^2}\right\}$$

and the log-likelihood function is obtained by taking logarithm on both side and it becomes

$$\ell = -n \ln \tau - \frac{n}{2} \ln(2\pi) - \sum_{i=1}^{n} \ln x_i - \frac{1}{2\tau^2} \sum_{i=1}^{n} (\ln a^{i-1} x_i - \lambda)^2 \tag{3.19}$$

where the last term is similar to the $SSE$ in equation (3.5). Let the ML estimates of $a$ $\lambda$ and $\tau^2$ be $\hat{a}_L$, $\hat{\lambda}_L$ and $\hat{\tau}_L^2$ respectively. Differentiating equation (3.19) with respect to $a$, $\lambda$ and $\tau^2$, $\hat{a}_L$ $\hat{\lambda}_L$ and $\hat{\tau}_L^2$ satisfy the following equations

$$\sum_{i=1}^{n} (n - 2i + 1) \ln(\hat{a}_L^{i-1} x_i) = 0, \tag{3.20}$$

$$\hat{\lambda}_L = \frac{1}{n} \sum_{i=1}^{n} [\ln(\hat{a}_L^{i-1} x_i)], \tag{3.21}$$

$$\hat{\tau}_L^2 = \frac{1}{n} \sum_{i=1}^{n} [\ln(\hat{a}_L^{i-1} x_i) - \hat{\lambda}_L]^2 \tag{3.22}$$

Note that $\hat{\tau}_L^2 = \frac{SSE}{n} = \frac{n-2}{n} \hat{\tau}_{NP1}$ After solving the equations, we get

$$\widehat{\ln a}_L = \frac{6}{(n-1)n(n+1)} \sum_{i=1}^{n} (n - 2i + 1) \ln x_i, \tag{3.23}$$

$$\widehat{\lambda}_L = \frac{2}{n(n+1)} \sum_{i=1}^{n} (2n - 3i + 1) \ln x_i, , \tag{3.24}$$

$$\hat{\tau}_L^2 = \frac{1}{n} \left[ \sum_{i=1}^{n} (\ln x_i)^2 - \frac{1}{n} (\sum_{i=1}^{n} \ln x_i)^2 - \frac{\widehat{\ln a}_L}{2} \sum_{i=1}^{n} (n - 2i + 1) \ln x_i \right] \tag{3.25}$$

Therefore, $\hat{a}_L = \exp(\widehat{\ln a}_L)$. Since the log-likelihood function in equation (3.19) contains the term of $SSE$ in equation (3.5), it is obvious that equation (3.8) is identical to (3.23) and equation (3.9) is identical to (3.24). Thus, we can prove $\hat{a}_L = \widehat{a}_{NP1}$, $\hat{\lambda}_L = \widehat{\lambda}_{NP1}$ and $\hat{\tau}_L^2 = \frac{n-2}{n} \hat{\tau}_{NP}^2$. Lam and Chan (1998) proved that the large sample distributions for $\hat{a}_{NP1}$, $\hat{\lambda}_{NP1}$ and they are given by equations (3.15), (3.16) respectively. Moreover, the large sample distribution for $\hat{\tau}_L^2$, $\hat{\mu}_L^2$ and $\hat{\sigma}_L^2$ are

$$n^{\frac{1}{2}} (\hat{\tau}_L^2 - \tau^2) \xrightarrow{L} N(0, 2\tau^4)$$

$$n^{\frac{1}{2}}(\hat{\mu}_L - \mu) \xrightarrow{L} N(0, \lambda^2\tau^2(4 + \frac{\tau^2}{2}))$$

$$n^{\frac{1}{2}}(\hat{\sigma}_L^2 - \sigma^2) \xrightarrow{L} N(0, 16\sigma^4\tau^2 + 2\tau^4(\lambda^2 + 2\sigma^2)^2)$$

where $\hat{\mu}_L = \exp(\hat{\lambda}_L + \hat{\tau}_L^2/2)$ and $\hat{\sigma}_L^2 = \hat{\mu}_L^2[\exp(\hat{\tau}_L^2) - 1]$. In Lam and Chan (1998), the estimates using the ML method are compared to those using the modified moment (MM) method and results suggested that $\hat{\mu}_{NP1}$ is recommended in data fitting and $\hat{\mu}_L$ is preferred in parameter estimation.

### 3.3.2 Exponential distribution

If $\{Y_i\}$ follows an Exponential distribution with density function given by equation (2.3), the density function for the observed $X_i$ is given by

$$f_i(x_i) = a^{i-1}\lambda \exp\{-a^{i-1}\lambda x_i\} \tag{3.26}$$

since $f_i(x_i) = f(y_i)\dfrac{dy_i}{dx_i} = a^{i-1}f(y_i)$. The likelihood function for the observed data $\{X_i\}$ is

$$L = a^{n(n-1)/2}\lambda^n \exp\{-\lambda \sum_{i=1}^{n} a^{i-1}x_i\}$$

and the log-likelihood function becomes

$$\ell = \frac{n(n-1)}{2}\ln a + n\ln\lambda - \lambda\sum_{i=1}^{n} a^{i-1}x_i \tag{3.27}$$

Differentiating equation (3.27) with respect to $a$ and $\lambda$ and setting them to zero, we obtain the following equations

$$\sum_{i=1}^{n}(n - 2i + 1)\hat{a}_E^{i-1}x_i = 0 \tag{3.28}$$

$$\hat{\lambda}_E = \frac{n}{\sum i = 1^n \hat{a}_E^{i-1} x_i} \quad (3.29)$$

The ML estimates of $\hat{a}_E$ can be solved iteratively and $\hat{\lambda}_E$ can be calculated directly using the closed form in equation (3.29). The GP model with Exponential distribution provides a simple way to demonstrate the parametric inference.

### 3.3.3 Gamma distribution

Chan *et al.* (2004a) investigated the statistical inference for GP with Gamma distribution. If the RP $\{Y_i\}$ follow a Gamma distribution, that is $Y_i \sim Ga(\alpha, \lambda)$ with density function in equation (2.5), the density function for the observed outcome of $X_i$ is given by

$$f_i(x) = \frac{(a^{i-1}\lambda)^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} \exp(-\lambda a^{i-1} x_i) \quad (3.30)$$

where $\Gamma(.)$ is a gamma function. The likelihood function for the observed data $\{X_i\}$ becomes

$$L = \frac{a^{n(n-1)\alpha/2} \lambda^{n\alpha}}{\Gamma(\alpha)^n} (\prod_{i=1}^n x_i)^{\alpha-1} \exp(-\lambda \sum_{i=1}^n a^{i-1} x_i)$$

and the log-likelihood function becomes

$$\ell = \frac{n(n-1)\alpha}{2} \ln a + n\alpha \ln \lambda - n \ln \Gamma(\alpha) + (\alpha - 1) \sum_{i=1}^n \ln x_i - \lambda \sum_{i=1}^n a^{i-1} x_i \quad (3.31)$$

Differentiating equation (3.31) with respect to $a$, $\alpha$ and $\lambda$ and setting tem to zero, we obtain the following equations.

$$\sum_{i=1}^n (n - 2i + 1) \hat{a}_G^{i-1} x_i = 0 \quad (3.32)$$

$$\frac{n(n-1)}{2} \ln \hat{a}_G + n \ln \hat{\alpha}_G - n \ln[\frac{\sum_{i=1}^n \hat{a}_G^{i-1} x_i}{n}] + \sum_{i=1}^n \ln x_i - n\psi(\hat{\alpha}_G) = 0 \quad (3.33)$$

$$\hat{\lambda}_G = \frac{n\hat{\alpha}_G}{\sum\limits_{i=1}^{n} \hat{a}_G^{i-1} x_i} \tag{3.34}$$

where $\psi(.) = \frac{\Gamma'(.)}{\Gamma(.)}$ is a psi function and ML of $\hat{a}_G$ and $\hat{\alpha}_G$ can be solved iteratively using equations (3.32) and (3.33) respectively. The mean and variance of the distribution $Ga(\alpha, \lambda)$ are $\hat{\mu}_G = \dfrac{\hat{\alpha}_G}{\hat{\lambda}_G}$ and $\hat{\sigma}_G^2 = \dfrac{\hat{\alpha}_G}{\hat{\lambda}_G^2}$ respectively. In fact, we can approximate $\hat{\alpha}_G$ in equation (3.33) using this equation,

$$\hat{\alpha}_G = \left(1 - \frac{n}{\bar{y} \sum\limits_{i=1}^{n} \hat{y}_i^{-1}}\right)^{-1}$$

where $\hat{y}_i = \hat{a}_G^{i-1} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} \hat{y}_i$ [See Cohen and Whitten (1988)]. Furthermore, the usual asymptotic properties of ML estimates do not hold unless $\alpha > 2$. Thus, Johnson and Kotz (1970) recommended that ML method should be employed only if $\alpha > 2.5$. Chan *et al.* (2004a) suggested that $\hat{\mu}_{NP1}$ performs better in data fitting but $\hat{\mu}_G$ should be recommended in parameter estimations.

### 3.3.4 Weibull distribution

If the RP $\{Y_i\}$ follows a Weibull distribution, that is $Y_i \sim W(\alpha, \lambda)$ with density function in equation (2.8), the density function of $X_i$ is given by

$$f_i(x_i) = a^{i-1}\alpha\lambda^{\alpha}(a^{i-1}x_i)^{\alpha-1}\exp\{-\lambda^{\alpha}(a^{i-1}x_i)^{\alpha}\}$$

The likelihood function of the observed outcome $\{X_i\}$ becomes

$$L = a^{n(n-1)\alpha/2}(\alpha\lambda^{\alpha})^n(\prod_{i=1}^{n} x_i)^{\alpha-1}\exp\{-\lambda^{\alpha}\sum_{i=1}^{n}(a^{i-1}x_i)^{\alpha}\}$$

and its log-likelihood function is

$$\ell = \frac{n(n-1)\alpha}{2}\ln a + n\ln\alpha + n\alpha\ln\lambda + (\alpha-1)\sum_{i=1}^{n} x_i - \lambda^{\alpha}\sum_{i=1}^{n}(a^{i-1}\lambda)^{\alpha} \tag{3.35}$$

Similarly, we have

$$\sum_{i=1}^{n}(n-2i+1)(\hat{a}_W^{i-1}x_i)^{\hat{\alpha}_W} = 0 \tag{3.36}$$

$$\frac{n-1}{2}\ln\hat{a}_W + \frac{1}{\hat{\alpha}_W} - \frac{1}{n}\sum_{i=1}^{n}\ln x_i - \frac{\sum_{i=1}^{n}(\hat{a}_W^{i-1}x_i)^{\hat{\alpha}_W}\ln(\hat{a}_W^{i-1}x_i)}{\sum_{i=1}^{n}(\hat{a}_W^{i-1}x_i)^{\hat{\alpha}_W}} = 0 \tag{3.37}$$

$$\hat{\lambda}_W = \left[\frac{n}{\sum(\hat{a}_W^{i-1}x_i)^{\hat{\alpha}_W}}\right]^{\frac{1}{n}} \tag{3.38}$$

where $\hat{a}_W$, $\hat{\alpha}_W$ and $\hat{\lambda}_W$ are the ML estimates of $a$, $\alpha$ and $\lambda$ respectively. Similar to Exponential distribution, we can express $\hat{\lambda}_W$ in a close form then $\hat{a}_W$ and $\hat{\alpha}_W$ can be solved iteratively.

# CHAPTER 4

# GENERALIZED GEOMETRIC PROCESS MODEL

## 4.1 The model

Geometric Process (GP) model for continuous data has distinct features over the Linear model because the ratio parameter $a$ of the GP model enables the modeling of trend data as well as reveals important information of the direction and strength of the trend. Moreover, the ratio can be interpreted as the "average" of the ratio of the present observation to the previous observation. After discounting the ratio, the renewal process (RP) $\{Y_i\}$ becomes stationary with constant mean $E(Y_i) = \mu$ and it constitutes a special case of the linear model when $a = 1$. Hence the GP model is in fact a "generalized" linear model in the sense that the ratio parameter is non-one. Therefore, similar to linear model, a direct extension of the GP model is to adopt a linear function of covariates to the mean of the RP.

Data exhibit trend and covariate effects are common. For example, the inter-arrival times between successive failures of a system are subject to certain environmental factors such as humidity, ageing or technological improvement. Therefore, the mean of the RP may evolve over time subject to different external and environment effects, such that covariate and trend effects can be modeled simultaneously by the mean $\mu$ and ratio $a$. In

GLMM, when the data type is not continuous, one should use a *link function* to link a linear function of covariates to the mean of the distribution for the observed data. Here is some common examples.

For continuous data:

$$E(y_i) = \mu_i = \beta_0 + \beta_1 z_i, \qquad y_i \sim N(\mu_i, \sigma^2), t(\mu_i, \sigma^2, \alpha)$$

For positive continuous data:

$$E(y_i) = \mu_i = e^{\beta_0 + \beta_1 z_i}, \qquad y_i \sim exp(\lambda_i), Ga(\alpha_i, \lambda_i)$$

For binary data:

$$E(y_i) = p_i = \frac{e^{\beta_0 + \beta_1 z_i}}{1 + e^{\beta_0 + \beta_1 z_i}}, \qquad y_i \sim Ber(p_i)$$

Adopting the modeling framework of the GLMM, we can extend the GP model to generalized GP (GGP) model to allow the covariate effects. Comparing to the GLMM, the GGP model also shows the direction and strength of trend for trend data.

In this chapter, we set $\mu_i = \beta_0 + \beta_1 z_{i1} + \ldots + \beta_K z_{iK}$ instead of $\mu = \lambda$ , where $\lambda$ is a constant, we set $z_{ij}, j = 1, \ldots, K$ be the $j$-th covariate at time $i$. We will estimate the parameters using non-parametric (NP) and parametric inferences. In parametric inference, both Maximum Likelihood (ML) as well as Bayesian methods will be discussed. The methodologies will be assessed through simulation experiments and demonstrated with real data sets. The definition of the GGP model is given below.

Definition of the GGP model: Let $X_1,\ X_2, \ldots$ be a sequence of random variables. If there exists positive real numbers $a_i$ such that $\{Y_i = a^{i-1} X_i,\ i = 1, 2, \ldots\}$ forms a stochastic process (SP) with means $E(Y_i) = \mu_i$ , then $\{X_i,\ i = 1, 2, \ldots\}$ is called a Generalized Geometric Process (GGP) and $a$ is the ratio of the GGP. A SP is a sequence of non-negative random variables. Here the SP $\{Y_i\}$ is obtained from the GGP $\{X_i\}$ by multiplying the ratio $a$ $(i-1)$ times. Then, the resulting SP $\{Y_i\}$ can be modeled by a parametric distribution such that the mean of $Y_i$ is $\mu_i$.

## 4.2 Non-parametric inference

Assume that $\{X_i\}$ come from a GGP such that the SP is $Y_i = a^{i-1}X_i$ where $a$ is the ratio of the GP. Taking logarithm on both sides, we have

$$\ln y_i = (i-1)\ln a + \ln x_i \tag{4.1}$$

Following the idea in Lam (1992b), we regress $\ln Y_i$ on the covariates $z_i$, such that

$$\ln y_i = \beta_0 + \beta_1 z_i + \epsilon_i \tag{4.2}$$

Therefore, combining equations (4.1) and (4.2), we have

$$\epsilon_i = (i-1)\ln a + \ln x_i - \beta_0 - \beta_1 z_i$$

Then the Least Square Estimates (log-LSE) can be obtained by minimizing the Sum of Squared Error ($SSE$) on $\ln X_i$ given by

$$SSE = \sum_{i=1}^{n}[(i-1)\ln a + \ln x_i - \beta_0 - \beta_1 z_i]^2 \tag{4.3}$$

over parameters $\ln a$, $\beta_0$ and $\beta_1$. The first order derivatives are given by the following equations.

$$\frac{\partial SSE}{\partial \ln a} = \sum_{i=1}^{n} 2[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i](i-1) \tag{4.4}$$

$$\frac{\partial SSE}{\partial \beta_0} = \sum_{i=1}^{n} 2[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i](-1) \tag{4.5}$$

$$\frac{\partial SSE}{\partial \beta_1} = \sum_{i=1}^{n} 2[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i](-z_i) \tag{4.6}$$

However, the above equations cannot be solved easily when set to zero and Newton Raphson (NR) method is applied. The second order derivatives that are required in the NR method are given below.

$$\frac{\partial^2 SSE}{\partial \ln a \partial \beta_0} = \sum_{i=1}^{n} 2(i-1)(-1)$$

$$\frac{\partial^2 SSE}{\partial \ln a \partial \beta_1} = \sum_{i=1}^{n} 2(i-1)(-z_i)$$

$$\frac{\partial^2 SSE}{\partial \beta_0 \partial \beta_1} = \sum_{i=1}^{n} 2z_i$$

$$\frac{\partial^2 SSE}{\partial \ln a^2} = \sum_{i=1}^{n} 2(i-1)^2 = \frac{n(n-1)(2n-1)}{3}$$

$$\frac{\partial^2 SSE}{\partial \beta_0^2} = 2n$$

$$\frac{\partial^2 SSE}{\partial \beta_1^2} = \sum_{i=1}^{n} 2z_i^2$$

## 4.3 Parametric inference

### 4.3.1 Maximum likelihood method

Although NP inference is easier and requires less distribution assumptions, Lam and Chan (1998) and Chan *et al.* (2004a) demonstrated that parametric inference using Maximum Likelihood method provides more accurate estimates than NP inference when the RP $\{Y_i\}$ adopts a Lognormal and Gamma distributions respectively. In this section, we adopt some life time distributions namely the Lognormal, Exponential, Gamma and Weibull distributions for the SP $\{Y_i\}$ with the GGP model framework.

#### 4.3.1.1 Lognormal distribution

We assume that $\ln Y_i$ follows a normal distribution with a mean $\lambda_i = \beta_0 + \beta_1 z_i$ and variance is $\tau^2$. Then $Y_i$ follows a Lognormal distribution $LN(\lambda_i, \tau^2)$ with the probability density function (pdf) of $Y_i$ given by equation (2.1). The corresponding pdf for the observed $X_i$ is

$$f_i(x_i) = \frac{1}{\sqrt{2\pi}\tau x_i} \exp\left\{\frac{-1}{2\tau^2}[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i)]2\right\}$$

The likelihood function for $\{X_i\}$ equal,

$$L = \{(2\pi)^{n/2}\tau^n \prod_{i=1}^{n} x_i\}^{-1} \exp\left\{\frac{-1}{2\tau^2}\sum_{i=1}^{n}[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i]^2\right\}$$

The log-likelihood function is

$$\ell = -\frac{n}{2}\ln(2\pi) - n\ln\tau - \sum_{i=1}^{n}\ln x_i - \frac{1}{2\tau^2}\sum_{i=1}^{n}[\ln(a^{i-1})x_i - \beta_0 - \beta_1 z_i]^2 \qquad (4.7)$$

and its derivatives are

$$\frac{\partial \ell}{\partial \ln a} = \frac{-1}{\tau^2}\sum_{i=1}^{n}[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i](i-1) \tag{4.8}$$

$$\frac{\partial \ell}{\partial \beta_0} = \frac{1}{\tau^2}\sum_{i=1}^{n}[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i] \tag{4.9}$$

$$\frac{\partial \ell}{\partial \beta_1} = \frac{1}{\tau^2}\sum_{i=1}^{n}[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i](z_i) \tag{4.10}$$

$$\frac{\partial \ell}{\partial \tau^2} = \frac{-n}{\tau} + \frac{1}{\tau^3}\sum_{i=1}^{n}[\ln(a^{i-1}x_i) - \beta_0 - \beta_1 z_i]^2 \tag{4.11}$$

The ML estimates $\hat{a}_L$, $\hat{\beta}_{0L}$ and $\hat{\beta}_{1L}$ are solution when equations (4.8) to (4.11) are set to zero. Also, from equation (4.11)

$$\hat{\tau}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln \hat{a}_L^{i-1}x_i - \hat{\beta}_{0L} - \hat{\beta}_{1L}z_i)^2$$

To obtain the ML estimates, Newton Raphson (NR) method is used. The second order derivatives of the log-likelihood function as required in the NR method are given below:

$$\frac{\partial^2 \ell}{\partial \ln a \partial \beta_0} = \frac{\sum_{i=1}^{n}(i-1)}{\tau^2} = \frac{n(n-1)}{2\tau^2}$$

$$\frac{\partial^2 \ell}{\partial \ln a \partial \beta_1} = \frac{\sum_{i=1}^{n}(i-1)z_i}{\tau^2}$$

$$\frac{\partial^2 \ell}{\partial \beta_0 \partial \beta_1} = \frac{-\sum_{i=1}^{n} z_i}{\tau^2}$$

$$\frac{\partial^2 \ell}{\partial \ln a^2} = \frac{-1}{\tau^2}\sum_{i=1}^{n}(i-1)^2$$

$$\frac{\partial^2 \ell}{\partial \beta_0^2} = \frac{-n}{\tau^2}$$

$$\frac{\partial^2 \ell}{\partial \beta_1^2} = \frac{\sum_{i=1}^{n} z_i^2}{\tau^2}$$

Since the $SSE$ in equation (4.3) appears in the log-likelihood function (4.7) and its derivatives from equations (4.8) to (4.10) which govern the estimation of $a$, $\beta_0$ and $\beta_1$ are identical to equations (4.4) to (4.6) in NP inference. Hence parametric method adopting Lognormal distribution for $Y_i$ gives the same set of parameter estimates as in the NP inference using the log-LSE method.

#### 4.3.1.2 Exponential distribution

We assume that $Y_i$ follows an Exponential distribution with a scale parameter $\lambda_i$, mean $\mu_i = \frac{1}{\lambda_i}$. A log link function is used to link the positive mean $\mu_i$ to a linear function of covariate $z_i$ as

$$\ln \mu_i = \beta_0 + \beta_1 z_i$$

such that,

$$\mu_i = e^{\beta_0+\beta_1 z_i} \Rightarrow \lambda_i = \frac{1}{e^{\beta_0+\beta_1 z_i}}$$

and the pdf of $Y_i$ using equation (2.3) is

$$f(y_i) = \frac{1}{e^{\beta_0+\beta_1 z_i}} \exp\left(\frac{-y_i}{e^{\beta_0+\beta_1 z_i}}\right)$$

Refer to equation (3.26), the density function for the observed data $X_i$ is

$$f_i(x_i) = \frac{a^{i-1}}{e^{\beta_0+\beta_1 z_i}} \exp\left(\frac{-a^{i-1} x_i}{e^{\beta_0+\beta_1 z_i}}\right) \tag{4.12}$$

The likelihood function for the observed data $\{X_i\}$ is

$$L = a^{\frac{1}{2}n(n-1)} \left[\prod_{i=1}^{n} e^{-(\beta_0+\beta_1 z_i)}\right] \exp\left[-\sum_{i=1}^{n} \frac{a^{i-1} x_i}{e^{\beta_0+\beta_1 z_i}}\right]$$

and the log-likelihood function is

$$\ell = \frac{n(n-1)}{2} \ln a - \sum_{i=1}^{n} (\beta_0 + \beta_1 z_i) - \sum_{i=1}^{n} \frac{a^{i-1} x_i}{e^{\beta_0+\beta_1 z_i}}$$

Its first order derivatives are

$$\frac{\partial \ell}{\partial a} = \frac{n(n-1)}{2a} - \sum_{i=1}^{n} \frac{(i-1)a^{i-2}x_i}{e^{\beta_0+\beta_1 z_i}}, \tag{4.13}$$

$$\frac{\partial \ell}{\partial \beta_0} = -n + \sum_{i=1}^{n} \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}}, \tag{4.14}$$

$$\frac{\partial \ell}{\partial \beta_1} = -\sum_{i=1}^{n} z_i + \sum_{i=1}^{n} \frac{a^{i-1}x_i z_i}{e^{\beta_0+\beta_1 z_i}} \tag{4.15}$$

The ML estimates are solution to the equations by setting (4.13) to (4.15) to zero. The second order derivatives of the log-likelihood function are

$$\frac{\partial^2 \ell}{\partial a \partial \beta_0} = \sum_{i=1}^{n} \frac{(i-1)a^{i-2}x_i}{e^{\beta_0+\beta_1 z_i}}$$

$$\frac{\partial^2 \ell}{\partial a \partial \beta_1} = \sum_{i=1}^{n} \frac{(i-1)a^{i-2}x_i z_i}{e^{\beta_0+\beta_1 z_i}}$$

$$\frac{\partial^2 \ell}{\partial \beta_0 \partial \beta_1} = \sum_{i=1}^{n} \frac{a^{i-1}x_i z_i}{e^{\beta_0+\beta_1 z_i}}$$

$$\frac{\partial^2 \ell}{\partial a^2} = -\frac{n(n-1)}{2a^2} - \sum_{i=1}^{n} \frac{(i-1)(i-2)a^{i-3}x_i}{e^{\beta_0+\beta_1 z_i}}$$

$$\frac{\partial^2 \ell}{\partial \beta_0^2} = -\sum_{i=1}^{n} \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}}$$

$$\frac{\partial^2 \ell}{\partial \beta_1^2} = -\sum_{i=1}^{n} \frac{a^{i-1}x_i z_i^2}{e^{\beta_0+\beta_1 z_i}}$$

as required in the NR procedures. Setting equation (4.14) to zero and $Y_i = a^{i-1}X_i$, we have

$$e^{\beta_0} = \frac{1}{n}\sum_{i=1}^{n} \frac{a^{i-1}x_i}{e^{\beta_1 z_i}} \Rightarrow \hat{\beta}_0 = \ln\left[\frac{1}{n}\sum_{i=1}^{n} \frac{a^{i-1}x_i}{e^{\beta_1 z_i}}\right] \tag{4.16}$$

implying

$$\hat{\beta}_{0E} = \ln\left[\frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{e^{\beta_{1E} z_i}}\right] \tag{4.17}$$

Hence, $\beta_0$ is independent of $a$ given the SP $\{Y_i\}$. Similarly, substituting equation (4.17) into (4.15) and $Y_i = a^{i-1}X_i$, we have

$$\left(\sum_{i=1}^{n} z_i\right)\left(\sum_{i=1}^{n} \frac{y_i}{e^{\beta_1 z_i}}\right) - n\left(\sum_{i=1}^{n} \frac{y_i z_i}{e^{\beta_1 z_i}}\right) = 0 \tag{4.18}$$

Hence $\beta_1$ is also independent of $a$ given the SP $\{Y_i\}$. From equation (4.13),

$$e^{\beta_0} = \frac{2a}{n(n-1)} \sum_{i=1}^{n} \frac{(i-1)a^{i-2}x_i}{e^{\beta_1 z_i}} \tag{4.19}$$

Equating equations (4.17) and (4.19), we have

$$\frac{1}{n}\sum_{i=1}^{n} \frac{a^{i-1}x_i}{e^{\beta 1 z_i}} = \frac{2a}{n(n-1)} \sum_{i=1}^{n} \frac{(i-1)a^{i-2}x_i}{e^{\beta_1 z_i}}$$

$$\Rightarrow \frac{n(n-1)}{2a}\left(\sum_{i=1}^{n} \frac{a^{i-1}x_i}{e^{\beta 1 z_i}}\right) - n\left(\sum_{i=1}^{n} \frac{(i-1)a^{i-2}x_i}{e^{\beta_1 z_i}}\right) = 0 \tag{4.20}$$

Hence instead of solving the system of equations by setting (4.13) to (4.15) to zero, we can solve two equations by setting (4.16) and (4.19) to zero for $\hat{a}_E$ and $\hat{\beta}_{0E}$, then $\hat{\beta}_{1E}$ can be obtained from equation (4.20).

#### 4.3.1.3 Gamma distribution

When $Y_i$ adopts Gamma distribution, results follow similarly from that of the Exponential distribution. The pdf of a Gamma distribution $Ga(\alpha, \lambda_i)$ with mean $\mu_i = \dfrac{\alpha}{\lambda_i}$ is given by equation (2.5). Again a log link function is needed to link the positive mean $\mu_i$ to a linear function of covariate so that

$$\ln \mu_i = \beta_0 + \beta_1 z_i$$

and we have

$$\lambda_i = \frac{\alpha}{e^{\beta_0 + \beta_1 z_i}}$$

The density function for the observed data $X_i$ is

$$f_i(x_i) = \frac{(\alpha a^{i-1})^\alpha x_i^{\alpha-1} \exp[\frac{-\alpha a^{i-1} x_i}{e^{\beta_0+\beta_1 z_i}}]}{e^{\alpha(\beta_0+\beta_1 z_i)}\Gamma(\alpha)}$$

The likelihood function of the observed data is

$$L = \frac{a^{n(n-1)\alpha/2}\alpha^{n\alpha}(\prod_{i=1}^n x_i)^{\alpha-1}\exp(-\alpha\sum_{i=1}^n \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}})}{(\prod_{i=1}^n e^{\beta_0+\beta_1 z_i})^\alpha \Gamma(\alpha)^n}$$

and the log-likelihood function is

$$\ell = \frac{n(n-1)\alpha}{2}\ln a + n\alpha\ln\alpha + (\alpha-1)\sum_{i=1}^n \ln x_i - n\ln\Gamma(\alpha) - \alpha\sum_{i=1}^n(\beta_0+\beta_1 z_i) - \alpha\sum_{i=1}^n \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}}$$

Moreover, we set

$$\ln\alpha = \gamma \qquad \Rightarrow \qquad \alpha = e^\gamma$$

so that $\alpha$ is always positive. Hence,

$$\ell = e^\gamma[\frac{n(n-1)}{2}\ln a + n\gamma - \sum_{i=1}^n(\beta_0+\beta_1 z_i) - \sum_{i=1}^n \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}}] + (e^\gamma - 1)\sum_{i=1}^n \ln x_i - n\ln\Gamma(e^\gamma)$$

The first and second order derivatives of the log-likelihood function are

$$\frac{\partial\ell}{\partial a} = e^\gamma\left[\frac{n(n-1)}{2a} - \sum_{i=1}^n \frac{(i-1)a^{i-2}x_i}{e^{\beta_0+\beta_1 z_i}}\right] \tag{4.21}$$

$$\frac{\partial\ell}{\partial\beta_0} = e^\gamma\left(-n + \sum_{i=1}^n \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}}\right) \tag{4.22}$$

$$\frac{\partial\ell}{\partial\beta_1} = e^\gamma\left(-\sum_{i=1}^n z_i + \sum_{i=1}^n \frac{a^{i-1}x_i z_i}{e^{\beta_0+\beta_1 z_i}}\right) \tag{4.23}$$

$$\frac{\partial\ell}{\partial\gamma} = e^\gamma[\frac{n(n-1)}{2}\ln a + n\gamma + n + \sum_{i=1}^n \ln x_i - n\psi(e^\gamma) - \sum_{i=1}^n(\beta_0+\beta_1 z_i) - \sum_{i=1}^n \frac{a^{i-1}x_i}{e^{\beta_0+\beta_1 z_i}}] \tag{4.24}$$

where $\psi(.) = \dfrac{\Gamma'(.)}{\Gamma(.)}$ is the psi function. Since equations (4.21) to (4.23) are similar to equations (4.13) to (4.15) apart from the multiple $e^{\gamma}$, the ML estimates $\hat{a}_G$, $\hat{\beta}_{0G}$ and $\hat{\beta}_{1G}$ are the same as $\hat{a}_E$, $\hat{\beta}_{0E}$ and $\beta_{1E}$ given the observed data $\{X_i\}$ and covariate $\boldsymbol{z}$. Hence, $\hat{\alpha}_G$ affect only the shape of the Gamma distribution assigned to $\{Y_i\}$.

The difference in the estimation when $Y_i$ adopts a Gamma distribution instead of an Exponential distribution lies in the presence of shape parameter $\alpha$ in the distribution, Gamma function in the density function and psi function in its first and second order derivatives. Numerical approximations for the Gamma and psi functions are required and they are done by calling subroutines in the International Mathematical and Statistical Library (IMSL) library. These approximations lower the accuracies and lead to a slower convergent rate. In this case, the minimum number of iterations is set to be 100 in the NR updating procedures as compared with 20 iterations when $Y_i$ adopts an Exponential distribution.

As our interest is to determine the trend effect as measure by the ratio $a$ and the covariate effect as measured by $\beta_0$ and $\beta_1$, we do not consider Gamma distribution in ML method since both distributions give the same set of these parameters but the parameters when $Y_i$ follows Gamma distribution are subject to more errors.

#### 4.3.1.4 Weibull distribution

If $Y_i$ follows Weibull distribution with shape parameter $\alpha$ and scale parameter $\lambda_i$, the pdf of for $Y_i$ is given by equation (2.8). Then we use a log function to link mean $\mu$ to a linear function of covariate $z_i$. It results

$$\mu_i = \frac{1}{\lambda_i}\Gamma(1+\alpha) = e^{\beta_0+\beta_1 z_i} \Rightarrow \lambda_i = \frac{\Gamma(1+\alpha)}{e^{\beta_0+\beta_1 z_i}}$$

Since the scale parameter $\lambda_i$ and hence the log-likelihood function contain Gamma function $\Gamma(1+\alpha)$ which involve the scale parameter $\alpha$, ML estimates especially for $\alpha$ are difficult to determine. Numerical approximation to the Gamma and psi function in the first and second order derivatives lower the accuracy and make the convergence of estimates in NR procedures in ruin. Moreover the parameter $\alpha$ lies in the indices of certain terms in the log-likelihood function and its derivatives, making the estimates very sensitive to the accuracy of $\alpha$. As a result, the Weibull distribution will not be considered in the ML method of inference.

### 4.3.2 Bayesian method

The Bayesian method of inference is implemented using WinBUGS. We assign Exponential distribution to SP $\{Y_i\}$. Gamma and Weibull distributions are not considered because of the difficulties in evaluating the gamma function which is not supported in the WinBUGS program. Moreover estimates using ML method when $Y_i$ adopts a Lognormal distribution are equivalent to those of NP method. Hence, we do not consider the SP using Lognormal distribution in the Bayesian method of inference.

To estimate the unknown parameters, we draw every 10-th sample start from iteration 5000 till iteration 10000 resulting in 500 posterior samples of $a$, $\beta_0$ and $\beta_1$. Parameter estimates are given by the sample mean of the posterior samples. A user defined likelihood function for the observed $X_i$ when $Y_i$ follows an Exponential distribution is given by (4.12) The priors for the parameters are given by

$$a \sim U[0.95, 1.05]$$

$$\beta_0 \text{ and } \beta_1 \sim N(0, 1000)$$

WinBUGS commands for the implemented program using WinBUGS are given in Appendix B.

## 4.4 Simulation

In the simulation experiment, we assign Exponential distribution to the stochastic process (SP) $\{Y_i\}$. For reasons stated in the ML and Bayesian method of inference, we do not consider Gamma, Weibull and Lognormal distributions.

For each set of parameter values, we perform simulation experiment 200 times resulting in 200 ($N$=200) realizations. In each realization, we simulate 240 data ($n$=240) for the SP $\{Y_i, i = 1, 2, \ldots, 240\}$ which follows an Exponential distribution, with $E(Y_i) = \mu_i = e^{\beta_0 + \beta_1 z_i}$ where $z_1 = z_2 = z_3 = 1$, $z_4 = z_5 = z_6 = 2$, $z_7 = z_8 = z_9 = 3$, $z_{10} = z_{11} = z_{12} = 4$ and cycle repeats again. The values of $z_i$ are set to resemble a time series of monthly data such that $z_i$ is a constant within a season but is increasing across seasons. Such trend is periodic with a period of one year.

IMSL subroutine DRNEXP for Exponential is used to generate sequences of IID random variables denoted by $\{V_i, i = 1, 2, \ldots, 240\}$ following a given distribution with a mean $\mu_i$ and the scale parameter $\lambda_i$ being set to one. Then we can perform the following transformations:

$$Y_i = \frac{V_i}{\lambda_i}$$

and

$$X_i = \frac{Y_i}{a^{i-1}}$$

where $\lambda_i = \frac{1}{e^{\beta_0+\beta_1 z_i}}$. Thus, the simulated sequence of IID random variables $\{V_i\}$ is transformed to a realization denoted by $\{X_i\}$ which is a GGP with a ratio parameter $a$ and a scale parameter $\lambda_i$.

By varying the values of the parameters $a$, $\beta_0$ and $\beta_1$, we obtain different sets of simulated data. The true values for the parameter true values set in the simulation experiment are given below:

$$(\beta_0, \beta_1) = (1.5, 0.8), (2.5, 0.5), (2.5, 0.8)$$

Then each realization $\{X_i\}$ is fitted to a GGP model using NP and ML methods of inference. Bayesian method is not considered in this simulation experiment because the implementation using WinBUGS is time-consuming and complicated when it has to be run with a large number of data. Different number of iterations for the NR procedures are required for different estimation methods because their speeds of convergence are not the same. For NP inference, around 10 iterations are required. However, for parametric inference using ML method approximately 50 iterations are required.

Then parameter estimates $\hat{\beta}_0$ and $\hat{\beta}_0$ are given by the sample mean of $\hat{\beta}_{0j}$ and $\hat{\beta}_{1j}$ respectively. That is $\hat{\beta}_0 = \frac{1}{N}\sum_{j=1}^{N}\hat{\beta}_{0j}$ and $\hat{\beta}_1 = \frac{1}{N}\sum_{j=1}^{N}\hat{\beta}_{1j}$. Moreover the standard deviation ($SD$) for $\hat{\beta}_0$ and $\hat{\beta}_1$ are given by the sample $SD$ of $\hat{\beta}_{0j}$. That is $SD(\hat{\beta}_0) = \left[\frac{\sum_{j=1}^{N}(\hat{\beta}_{0j}-\hat{\beta}_0)^2}{N-1}\right]^{\frac{1}{2}}$ and $SD(\hat{\beta}_1) = \left[\frac{\sum_{j=1}^{N}(\hat{\beta}_{1j}-\hat{\beta}_1)^2}{N-1}\right]^{\frac{1}{2}}$.

In order to compare the performances between the NP, ML and Bayesian methods of inference, we define the following two criteria. The first criterion evaluates the deviation ($DM$) of the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ from their true values defined as

$$DM = \frac{|\hat{\beta}_0 - \beta|}{\beta_0} + \frac{|\hat{\beta}_1 - \beta|}{\beta_1} \tag{4.25}$$

Alternatively we compare the $i$-th simulated value $X_{ij}$ in the $j$-th realization, with the fitted value $\hat{X}_{ij}$, $i = 1, \dot{,} n \;\; j = 1, \ldots, N$ and sum the squared deviation over data $i$ and realization $j$. That is

$$MSE = \frac{1}{nN}\sum_{j=1}^{N}\sum_{i=1}^{n}(\hat{x}_{ij} - x_{ij})^2 \tag{4.26}$$

The first criterion which accesses the accuracy of parameter estimates is useful when parameter estimation is the main focus. The second criterion which measures the accuracy of data fitting and it is useful when our objective is prediction. Both criteria are ranked across methods of inference and method with the lowest rank, say rank 1, is considered to be the best method.

Note that for each set of parameters, we simulate a SP $\{Y_i\}$ before the GP $\{X_i\}$. As the SP $\{Y_i\}$ depends on the parameters $\beta_0$ and $\beta_1$ only through $\lambda_i = \frac{1}{\mu_i} = \frac{1}{e^{\beta_0+\beta_1 z_i}}$. Therefore, $\{Y_i\}$ is independent of $a$ given $\beta_0$, $\beta_1$ and a set of IID random variables $\{V_i\}$. Moreover equations (4.17) and (4.18) show that $\hat{\beta}_0$ and $\hat{\beta}_1$ are independent of $a$ and hence constant across $a$ given $\{Y_i\}$ as shown in Table 4.1. In real situation, only $\{X_i\}$ is observed and hence the parameter estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ are not independent of $a$.

The simulation results given in Table 4.1, shows that ML method gives better performance than the NP method in terms of both parameter estimates and data fitting when SP is adopted Exponential distribution. The poor performance of the estimates $\beta_0$ using NP method leads to a poor performance in data fitting resulting in a larger $MSE$. The better performance of the ML may be due to the fact that the data sets used in this simulation study are simulated from Exponential distribution. If we have no prior knowledge about which distribution the data follow, we may attempt all methods of inference assigning different distributions and select the method which has the best goodness-of-fit.

Moreover, it should be noted that the $SD$ for $\hat{a}$ is consistently lower but the $SD$s for $\hat{\beta}_0$ and $\hat{\beta}_1$ using the NP method are consistently higher.

**************************************************

Table 4.1 are shown here.

**************************************************

## 4.5 Real data analysis

The methodologies developed for the GGP model is used to analyze two data sets: "Severe Acute Respiratory Syndrome (SARS) in Hong Kong" and the "Imports of goods index" data. There is an downward trend in the number of daily infected person of SARS as the control polices were gradually implemented. Thus the ratio of the GGP should be greater than one. For the second data set, the imports of goods indices should be increasing due to increasing demand of consumption overtime and hence the ratio of the GGP should be smaller than one.

Before we fit the data using GGP models, we test whether the data follow a GP. Among the many methods suggested in Lam (1992b), the simplest method is to plot the logarithm of the observed data $\{X_i\}$ against time. A linear relationship (either increasing or decreasing) in the plot indicates that the data follow a GP. If the data follow a GP, the NP, ML and Bayesian methods are applied to estimate the parameters of the GGP model. To compare the performance of GGP models with GP models, parameter estimates for the GP model using NP method are also reported.

### 4.5.1 Hong Kong SARS data

The Severe Acute Respiratory Syndrome (SARS) pose one of the most serious global health threats because of its high infectiousness and potential to spread through air travel. The disease was first found in a southern city in Mainland China in late November, 2002.

Afterward, the epidemic was widely spread throughout the world in a few months from February 2003. Basically, SARS cases in Hong Kong were reported to World Health Organization (WHO). Instead the SARS cases were posted on the website of the Department of Health in Hong Kong during that period. Further information about the SARS cases in Hong Kong can be obtained from

$$\texttt{http}://\texttt{www.hku.hk/sars} \tag{4.27}$$

The data were collected is from March 12 and June 11 2003, we have 92 data and 1755 reported cases totally. During this period, the number of cases per day is 19 and the average temperature is 24.4 $^oC$. The data set is given in Appendix E. It contains number of daily infected cases, daily recovery cases as well as daily death cases. In this thesis, we will focus on the modeling of the number of daily infected rate. We set temperature as a covariate in the mean function because many epidemics are affected by temperature and SARS virus is proved to become less active and even die in a high temperature environment.

******************************************************

Table 4.2, Figure 4.1 & 4.2 are shown here.

******************************************************

Table 4.2 gives parameter estimates using NP, ML and Bayesian methods. Note that the first set of parameter estimates refers to the GP model with constant mean (i.e. $\beta_1 = 0$). Models are selected according to $MSE$ and the model with the lowest rank is the best model. As shown in Figure 4.2, there seems to be two trends inside the series which indicates different stage of an epidemic outbreak. To better model the data, multiple GP model will be applied in Chapter five.

### 4.5.2 Imports of goods index

Imports of goods are defined as goods which have been produced or manufactured in places outside Hong Kong and brought into Hong Kong for domestic use or for subsequent re-export. They also include Hong Kong products re-imported back into Hong Kong. The data contain monthly indices of goods imported in Hong Kong from January 1982 to April 2004 resulting in 268 data as shown in Appendix E. The data is collected from the web site of Census and Statistics department in the section of "Key economic and social indicator" and its corresponding address is:

http://www.info.gov.hk/censtatd/eng/hkstat/fas/ex-trade/trade2/im_nsa_std_index.html

In merchandising industries, the amount of goods imported usually exhibits certain seasonal effects. Figure 4.3 shows that the amount of goods imported in January and February are significantly less than those in other months throughout the period. Therefore, we add a seasonal covariate $z_i$ which is an indicator, say $z_i = 1$ for the month January and February otherwise $z_i = 0$, to the mean function $\mu_i$ of the GGP model to account for such seasonal effect.

****************************************************

Table 4.3, Figure 4.3 & 4.4 are shown here.

****************************************************

A plot of the $\ln X_i$ against time in Figure 4.3 shows an obvious increasing trend indicating that the data come from a GP. Table 4.3 shows that NP, ML and Bayesian methods give almost the same set of parameter estimates except that $\beta_0$ when using the Bayesian method is relatively large. Appendix C contains history plots of samples extracted after the burn-in period that indicates the parameter estimates using Bayesian method are stable. Moreover, the autocorrelation functions are reasonably low suggesting that the samples are independent. According to $MSE$, parameter estimates using the NP method are chosen. Figure 4.4 plots the fitted indices using the NP method with the observed indices and the plot reveals that there are more than one GPs in the data possibly caused by the economic turmoil in Asian financial crisis that made the imports of goods lower than expected. Thus, a multiple GP (MGP) model is fitted to the data in Chapter five.

## 4.6 Discussion

Previous GP models limits to the modeling of inter-arrival times which are positive and continuous. The RP is assumed to have a constant mean $\mu$ which indicates the initial level of a GP. This assumption is too strong and fails to respond to the diverse effects exerted on a process. In this chapter, we extend the GP model to the GGP model by accommodating a linear function of covariates log-linked to the mean of a lifetime distribution, including Exponential, Weibull, Gamma and Lognormal distributions for the SP $\{Y_i\}$. The extended RP is no longer IID and hence is called a stochastic process (SP) in general. The extension is straight forward and the methods of inference follow similarly from those of the GP model, NP, ML and Bayesian methods. Again, simulation studies and real data analyses are performed to demonstrate the models.

Similar to the GP models, ML and Bayesian inferences adopting a Weibull distribution for the SP are problematic. This is probably due to the complicated density function of the Weibull distribution when its shape parameter $\alpha$ appears in the indices, making the log-likelihood function and its derivatives functions extremely sensitive to the current parameter values. On the other hand, the ML method adopting a Lognormal distribution for the SP gives the same set of parameter estimates as the NP method. Moreover, the ML method adopting Gamma distribution also gives the same set of parameter estimates as the Exponential distribution. As a result, simulation studies comparing the NP and ML methods are only conducted when the SP follow Exponential.

Simulation studies show that the ML method adopting an Exponential distribution for the SP is preferred to the NP method. Since the implementation of Bayesian method through WinBUGS is complicated, involving a large number of parameter sets for the posterior samples, Bayesian method is only applied in real data analyses. Results from real data analyses show that NP estimates are better with a lower $MSE$ indicating that the underlying SP for the data may not follow an Exponential. As the NP method persistently gives better estimates in data analyses, we suggest the NP method should be recommended in data fitting and the ML method adopting an Exponential distribution for the SP should be recommended in parameter estimation in the GGP model. Also, one may attempt all methods and select the set of parameter estimates with the lowest $MSE$.

Table 4.1: Simulation studies for GGP model

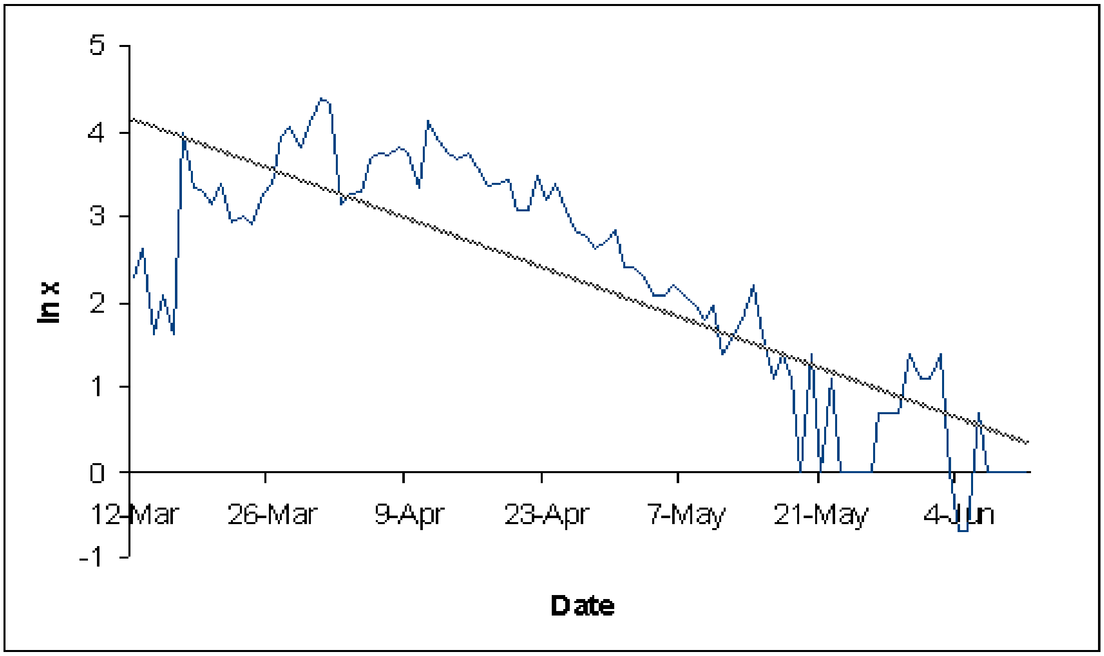
| *Parameter* | $a$ | | | | | $\beta_0$ | $\beta_1$ | rank | rank |
|---|---|---|---|---|---|---|---|---|---|
| *true value* | *0.96* | *0.98* | *1.0* | *1.02* | *1.04* | *1.5* | *0.5* | *DM* | *MSE* |
| ML | | | | | | | | | |
| Estimate | 0.96006 | 0.98006 | 1.00006 | 1.02006 | 1.04006 | 1.52088 | 0.79105 | 1 | 1 |
| SD | 5.74E-06 | 6.23E-06 | 6.75E-06 | 7.31E-06 | 7.91E-06 | 4.23E-02 | 3.40E-03 | | |
| NP | | | | | | | | | |
| Estimate | 0.95995 | 0.97994 | 0.99995 | 1.01995 | 1.03994 | 0.93770 | 0.79274 | 2 | 2 |
| SD | 1.53E-06 | 1.59E-06 | 1.66E-06 | 1.73E-06 | 1.80E-06 | 6.17E-02 | 4.83E-03 | | |
| *true value* | *0.96* | *0.98* | *1.0* | *1.02* | *1.04* | *2.5* | *0.5* | *DM* | *MSE* |
| ML | | | | | | | | | |
| Estimate | 0.96006 | 0.98006 | 1.00006 | 1.02006 | 1.04006 | 2.02087 | 0.49105 | 1 | 1 |
| SD | 5.74E-06 | 6.23E-06 | 6.75E-06 | 7.31E-06 | 7.91E-06 | 4.23E-02 | 3.40E-03 | | |
| NP | | | | | | | | | |
| Estimate | 0.95995 | 0.97994 | 0.99995 | 1.01995 | 1.03994 | 1.43770 | 0.49274 | 2 | 2 |
| SD | 1.53E-06 | 1.59E-06 | 1.66E-06 | 1.73E-06 | 1.80E-06 | 6.17E-02 | 4.83E-03 | | |
| *true value* | *0.96* | *0.98* | *1.0* | *1.02* | *1.04* | *2.5* | *0.8* | *DM* | *MSE* |
| ML | | | | | | | | | |
| Estimate | 0.96006 | 0.98006 | 1.00006 | 1.02006 | 1.04006 | 2.52088 | 0.79105 | 1 | 1 |
| SD | 5.74E-06 | 6.23E-06 | 6.75E-06 | 7.31E-06 | 7.91E-06 | 4.23E-02 | 3.40E-03 | | |
| NP | | | | | | | | | |
| Estimate | 0.95995 | 0.97994 | 0.99995 | 1.01995 | 1.03994 | 1.93769 | 0.79274 | 2 | 2 |
| SD | 1.53E-06 | 1.59E-06 | 1.66E-06 | 1.73E-06 | 1.80E-06 | 6.17E-02 | 4.83E-03 | | |

Table 4.2: Summary of parameter estimates for SARS in HK using GGP model

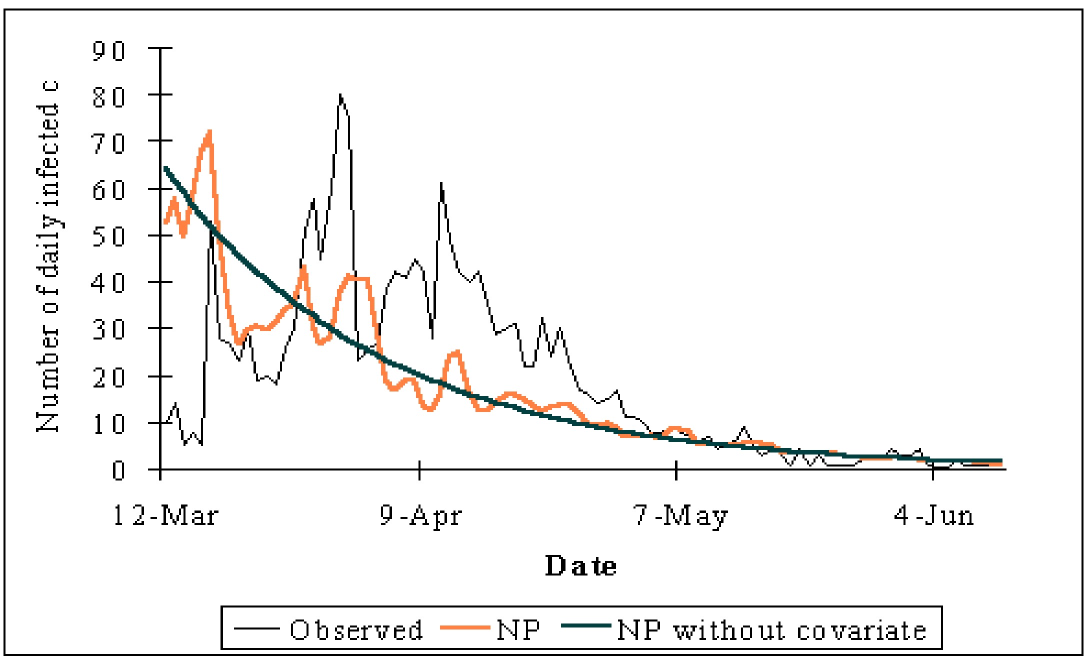
| *method* | *Distribution* | $a$ | $\beta_0$ | $\beta_1$ | *MSE* |
|---|---|---|---|---|---|
| NP | - | 1.0429 | 4.1645 | - | 356.6986 |
| NP | - | 1.0545 | 2.0610 | 0.1070 | 316.9569 |
| ML | EXP | 1.0578 | 2.8031 | 0.0918 | 562.7144 |
| ML | GAMMA | 1.0570 | 3.0064 | 0.0819 | 573.3660 |
| Bayesian | EXP | 1.046 | 3.823 | 0.03033 | 432.8422 |

Table 4.3: Summary of parameter estimates for imports of goods index in HK using GGP model

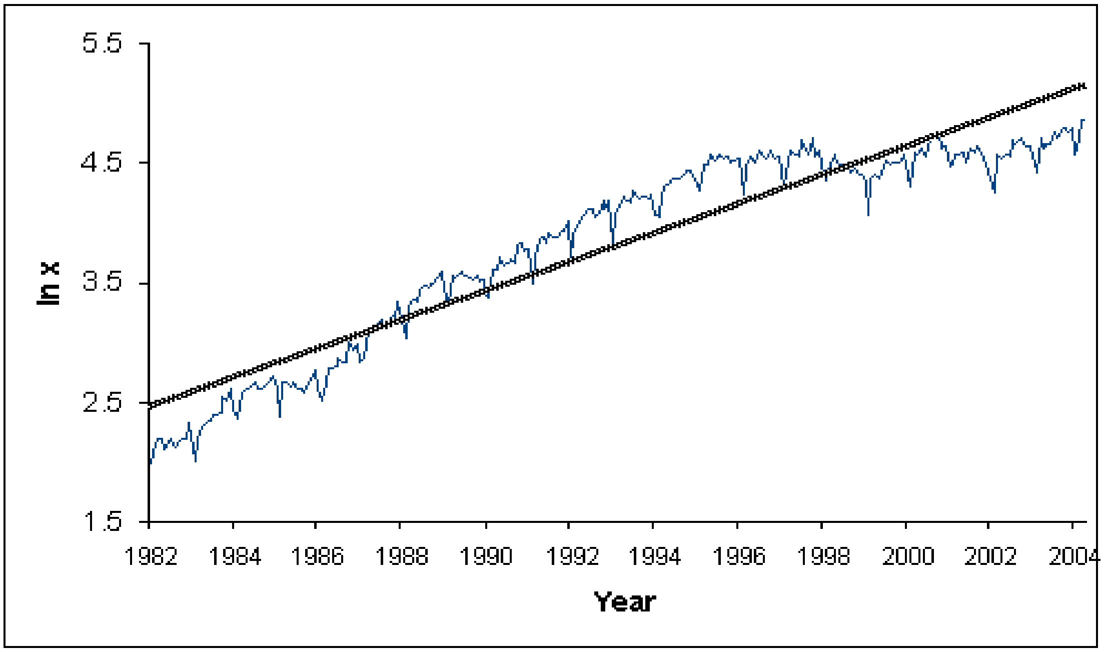
| *Method* | *Distribution* | $a$ | $\beta_0$ | $\beta_1$ | *MSE* |
|---|---|---|---|---|---|
| NP | - | 0.9897 | 2.471 | - | 456.57 |
| NP | - | 0.9899 | 2.489 | -0.1886 | 443.16 |
| ML | EXP | 0.9898 | 2.486 | -0.1882 | 451.41 |
| Bayesian | EXP | 0.9898 | 2.518 | -0.1828 | 618.21 |



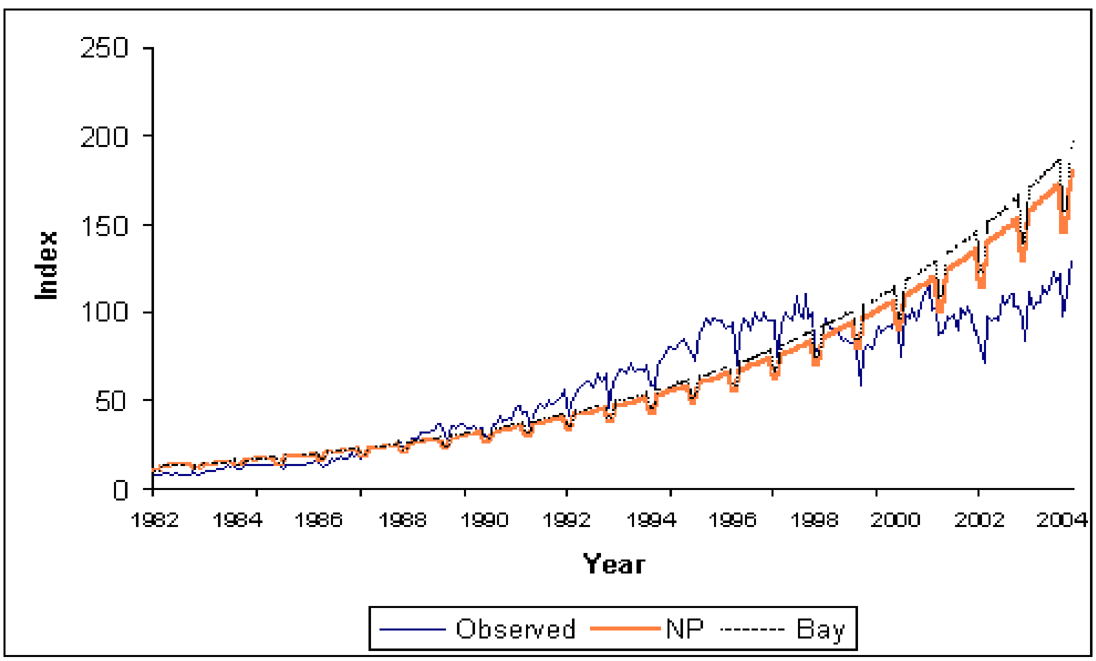
5
4
3
2
1
0
-1
ln x
12-Mar
26-Mar
9-Apr
23-Apr
7-May
21-May
4-Jun
Date

90
80
70
60
50
40
30
20
10
0
Number of daily infected c
12-Mar
9-Apr
7-May
4-Jun
Date
Observed
NP
NP without covariate



5.5
4.5
3.5
2.5
1.5
ln x
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
Year

250
200
150
100
50
0
Index
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
Year
Observed
NP
Bay

# CHAPTER 5

# MULTIPLE GEOMETRIC PROCESS MODEL

## 5.1 Non-parametric inference

Previous Geometric Process (GP) models limit to trend data with a single trend, but in real daily life, epidemic diseases or sales of certain commodity usually experience a growing stage, followed by a stabilized stage and a declining stage.

Chan *et al.* (2004b) developed the Multiple Geometric Process (MGP) model as an extension from the single GP model using non-parametric (NP) inference. NP inference for each GP model is given by equations (3.8) to (3.10) in Section 3.2.1. We choose the ratio estimate $\hat{a}_{NP1}$ in log-LSE method and it is given by

$$\widehat{a}_{NP1} = \exp(\widehat{\ln a}_{NP1}).$$

Since equation (3.8) requires the logarithm of the observed data $\{X_i\}$, therefore, any zero $\{X_i\}$ will be averaged with all previous observations up to the first non-zero observation. For example, if $X_{i-2} = 1$, $X_{i-1} = 0$, $X_i = 0$ and $X_{i+1} = 1$ , they are adjusted to $X_{i-2} = X_{i-1} = X_t = 0.33$ and $X_{i+1} = 1$. In this way, any zero $\{X_i\}$ is removed and the sum of $\{X_i\}$, that is $S_n$, is unchanged.

Given that $a$ is estimated by $\widehat{a}_{NP1}$, different NP estimators were proposed for the mean parameter $\mu$ in Lam (1992b). Among them, the estimate

$$\hat{\mu}_{NP1} = \frac{S_n(1-\widehat{a}_{NP1}^{-1})}{1-\widehat{a}_{NP1}^{-n}} \quad \text{where} \quad S_n = \sum_{i=1}^{n} X_i$$

is preferred according to mean squared error ($MSE$) defined as

$$MSE = \sum_{m=1}^{N} \sum_{i=1}^{n} (\widehat{X}_{mi} - X_{mi})^2 / (nN) \quad \text{where} \quad \widehat{X}_{mi} = \widehat{\mu}_m \widehat{a}_m^{-(i-1)},$$

where $X_{mi}$ is the $i$-th observation in the $m$-th simulated data and $N$ is the number of realization in simulation.

Beside the log-LSE method in NP inference, Chan *et al.* (2004b) proposed the LSE method for the estimation of $a$ and $\mu$ in equation (3.18) and the estimates are denoted as $\hat{a}_{NP2}$ and $\hat{\mu}_{NP2}$ respectively.

## 5.2 Trend identification

In real life data, multiple trends are commonly exist in a series. Chan *et al.* (2004b) extend the GP models by fitting a separate GP to each stage of the series that identified by a turning point. In the MGP models, one or more turning points can be detected.
Let $X_i,\ i = 1, \ldots, n$ be the observed data, $T_m,\ m = 1, \ldots, M$ be the turning points of the GPs and $n_m$ be the number of observations for the $m$-th GP defined as

$$GP_m = \{X_i,\ T_m \le i < T_{m+1}\}, \quad m = 1, \ldots M$$

where $T_1 = 1$, $T_m = 1 + \sum_{j=1}^{m-1} n_j,\ m = 2, \ldots, M$ and $\sum_{m=1}^{M} n_m = n$.
Then we define the corresponding renewal process (RP) as

$$RP_m = \{a_m^{i-T_m} X_i,\ T_m \le i < T_{m+1}\}$$

where $a_m > 0$. The ratio $a_m$ is called the ratio parameter of the $m$-th GP. Also, the mean and the variance for the $m$-th GP is,

$$E(X_{T_m+t}) = \mu_m / a_m^t \quad \text{and} \quad Var(X_{T_m+t}) = \sigma_m^2 / a_m^{2t},\ t = 0, \ldots, n_m - 1$$

where $E(X_{T_m}) = \mu_m$ and $Var(X_{T_m}) = \sigma_m^2$. Thus $a_m$, $\mu_m$ and $\sigma_m^2$ are the three parameters which completely determine the mean and variance of $X_{T_m+t}$ in the $m$-th GP.

For series with multiple trends, there exists a problem of detecting turning point(s), that is the moment when the series changes its direction. Among the methodologies that available to estimate the turning point(s), Chan *et al.* (2004b) proposed a moving window technique in which GP models are applied successively to each subset of data of fixed length $L$ starting from time $i = 1$, $i = 2$ and so on up to $i = n - L + 1$. Since the ratio $a$ of a GP changes according to the moving windows, turning point(s) $T_m$ can be located by a change of $a$ from "less than 1" to "greater than 1" and vice versa. Depending on the nature of data, we may detect more than one turning points and the resulting GPs are $GP_m = \{X_i,\ T_m \leq i < T_{m+1}\}$, $m = 1, \ldots, M$ where $T_1 = 1$ and $T_{M+1} = n + 1$.

Studies show that $a_m$ may fluctuate around 1 frequently so that the resulting GPs may be too short and the parameter estimates for these GPs will not be stable. Moreover, the change of $a$ around 1 may be small, indicating a "noisy" rather than a real change of trend. To avoid capturing the noise, we add two criteria in the detection of turning points. The first criterion is to limit the length of a new GP to be at least $h = 10$ say. The desirable level for $h$ depends on the length of the series in real situation. The second criterion is to constraint the change of $a$ at the turning point to be reasonably large, for example, $|a_{T_k} - a_{T_k - 1}| > d = 0.01$. These two criteria prevent the moving window technique detecting too many trends resulting in over-fitted and complicated model. In fact, we can try different set of $h$ and $d$ until they produce reasonable result.

Moreover, we can attempt different value for the window width $L$ in order to find the best $L$ for the data set. Obviously, different window widths produce different set of turning points. A short $L$ identifies changes of $a$ more precisely resulting in shorter GPs. In this case, noise may also be captured. However a large $L$ may average the changes of $a$ too much leading to a loss of information. For an extreme case, only one GP will be detected when $L$ is equal to $n$. In the real data analyses, $L$ is set to be within the range 20 to 50 and for each $L$, the corresponding set of GPs and parameters are identified. To select the best set of turning points and model parameters, we use the Adjusted Mean Square Error ($AMSE$) as a goodness-of-fit measure. $AMSE$ consist of the sum of Mean Square Error ($MSE$) and a penalized term, which is a scalar multiple of the number of parameters $2M$. It is defined as

$$AMSE{=}k(2M) + \frac{1}{n}\sum_{i=1}^{n}(X_i - \widehat{\mu}_m \widehat{a}_m^{-(i-T_m)})^2, \qquad T_m \leq i < T_{m+1}.$$

The number of GPs $M$ will be over-estimated if we do not penalized the model by its number of model parameters ($k = 0$) since more parameters always give a lower $MSE$. On the other hand, if the model is heavily penalized say ($k = 1$) by its number of parameters, the number of GPs will possibly be under-estimated. To strike a balance, we set $k = \frac{1}{2}\ln(\bar{X})$ where $\bar{X}$ is the average number of cases $X_i$, so that higher average leads to higher penalty.

## 5.3 Real data analysis

### 5.3.1 Hong Kong SARS data

Chan *et al.* (2004b) demonstrated MGP model using the data of Severe Atypical Respiratory Syndromes (SARS) in Hong Kong. Some general descriptions for the SARS in

Hong Kong data are given in Chapter four. In Chapter four, we model the daily number of infections of SARS in Hong Kong by a single GP model with the mean $\mu$ set to be a linear function of temperature. In this chapter, we focus on MGP model which enables the identification of trends in the whole data. To begin with, we test the existence of multiple trends. The plot of $\ln X_i$ against time provides an objective way to detect any potential turning points and trends inside the data. It also provides us a quick view of the characteristics in the data. In Figure 5.1, it is obvious that there exists a turning point between April 12 and April 15, even though we cannot locate the point exactly.

*******************************************************

Figure 5.1 is shown at here.

*******************************************************

MGP model is fitted to the Hong Kong SARS data using two NP methods, namely the log-LSE (NP1) method and the LSE (NP2) method and the results are given in Table 5.1. Since the observations $X_{74}, X_{86}, X_{89}$ are zero, the data are modified so that $(X_{73}, X_{74}, X_{85}, X_{86}, X_{88}, X_{89})$ changes from (2,0,1,0,2,0) to (1,1,0.5,0.5,1,1) respectively when log-LSE method is used. Moreover, since zero $X_i$ are not common in the data, such modifications have minimal effect on the results. From Table 5.1, LSE method is preferred according to $AMSE$. Both log-LSE and LSE methods identify one turning point and give similar estimates of the turning point. The log-LSE method locates the turning point on April 14 whereas the LSE method locates that on April 13. Moreover, their estimates are chosen from the same window width $L = 20$. The two ratio estimates $a$ from either method are increasing from "less than one" to "greater than one" indicating a growing stage followed by declining stage in the data.

**************************************************

Table 5.1, Figure 5.2 & 5.3 are shown here.

**************************************************

During the growing stage, the trend using log-LSE method starts at a lower level ($\hat{\mu}_{NP1} = 14.85$) but increases at a faster rate of around 5% per day ($\hat{a}_{NP1} = 0.95$) whereas the trend using LSE method starts at a higher level ($\hat{\mu}_{NP2} = 21.30$) but increases at a slower rate of around 3% per day ($\hat{a}_{NP2} = 0.97$). Starting from Apr 14 and Apr 13 respectively using the two methods, the trend changes direction and drop from 41.22 and 44.02 cases respectively at 7% per day. Figure 5.2 and 5.3 show that the fitted values using the two methods are similar during the declining stage.

**************************************************

Figure 5.4 is shown here.

**************************************************

Prediction is important in modeling and we will demonstrate the power of prediction of the MGP model using this data. During May 2003, the government is concerned with the removal of travel advice that released by WHO in order to recover the economy. In May 20, the Department of Health in Hong Kong predicted the travel advice would be removed before late June. For the removal of travel advice, one requirement was the 3-day average of infected cases was less than 5. We refit the MGP model using data up to April 30. The window width of $L = 20$ is again chosen and it is obvious that the turning point $T$ remains unchanged. The predicted daily infections $X_i$ after April 30 for the two methods are given in Figure 5.4 and their 3-day averages cut the line $X = 5$ on May

19 and 20 respectively which are very close to the observed date of May 17. The travel advice was removed on May 23.

### 5.3.2 Imports of goods index

The "Imports of goods index" data is described in Chapter four and Figure 5.5 shows that there exists more than one trend inside the data. Although a trend line has been added to Figure 4.3 to show the presence of a GP in the data, it can be seen in Figure 5.6 that two trend lines model the data $\ln X_i$ better. From Table 5.2, the two methods of inference, namely the log-LSE and LSE methods, give very different turning point estimates that is $T_2 = 194$ and $T_2 = 255$ respectively and very different set of parameters estimates. As both ratio estimates using either method are less than one, the two trends in the data are both increasing but from different levels and at different rates. As mentioned in Chapter four, the change in trend may due to the economic turmoil in Asian financial crisis during late 1997, which change the level and the rate of the trend but not the direction. The turning point using the log-LSE method is Feb 1998 which agrees reasonably well with the actual date of the Asian financial crisis.

As shown in Figure 5.6, the trend using the log-LSE method starts at a level of 8.15 and increases at a rate of 1.5% per month until Feb 1998. Thereafter, the indices drop from 131.06 to 81.59 and increases at a much lower rate of 0.4% per month. These two trends using the log-LSE method agree reasonably well with the observed data as indicated by much lower $AMSE$ (90.22 versus 461.23) than of the GP model when only a single GP is fitted to the data. However, results from using the LSE method are much less satisfactory. The turning point is estimated to be close to the end of the data ($T_2 = 255$ lies on Mar

2003). Hence the two trends differ not much from that of the GP model. Moreover the $AMSE$ also shows not much improvement (415.19 versus 461.23). The reason for the much less satisfactory result sing log-LSE method remains to be investigated. Needless to say, parameters using the log-LSE method are preferred.

******************************************************

Table 5.2, Figure 5.5 & 5.6 are shown here.

*******************************************************

### 5.3.3 No. 3 data

The data set contains the arrival times of unscheduled maintenance actions for the U.S.S. Halfbeak No.3 main propulsion diesel engine. Arrival times of the scheduled engine overhauls are discarded because our focus is on the failures which cause the unscheduled maintenance actions but not the overhaul. Totally there are 71 data and the data set is given in Appendix E and in Andrews and Herzberg (1985). The data were first studied by Ascher and Feirgold (1969 and 1981) and then by Lam (1992b), Lam and Chan (1988) and Lam *et al.* (2004) using GP models.

Lam (1992b) and Lam *et al.* (2004) applied the GP model to model the inter-arrival times and parameters were estimated using the log-LSE method. Lam and Chan (1998) used a parameter method with Lognormal distribution for $\{Y_i\}$ to estimate the parameters. We apply the MGP model to model the inter-arrival times.

Lam (1992b) has tested that the data were consistent with a GP. The plotting of $\ln X_i$ against number of unscheduled maintenance action in Figure 5.7 shows a declining trend inside the data. However, regarding the 50-th unscheduled maintenance action as a turning point, we can divide the whole data into 2 trends, a downward trend at the beginning followed by a slightly upward trend. The identification of 2 separate GPs provides better model fit especially after 50-th unscheduled maintenance action because the MGP model fits an increasing trend of inter-arrival times against the number of unscheduled maintenance actions instead of a decreasing trend when the data is modeled by a single GP.

**********************************************************

Table 5.3 and Figure 5.7, 5.8 & 5.9 are shown here.

***********************************************************

Table 5.3 lists the estimates of the MGP models as well as those of the GP model. Figure 5.8 and 5.9 show the plot of inter-arrival times and arrival times for unscheduled maintenance action respectively. It can be observed that the MGP model captures the shape for the observed trend very well. Both the GP and MGP give a declining trend at the beginning but the trend from the MGP starts at a higher level of 2108.6 and 1619.3 respectively using log-LSE and LSE methods as compared to 1079.9 for the GP model. Also, trends estimated from MGP decreases at faster rates of 10.3 % and 7.2 % respectively versus 1.5% decreasing rate from the GP model. Near the end of the data, log-LSE method identifies a turning point at the 50-th unscheduled maintenance action and the trend increases thereafter at a rate of 3.0% from the level of 2108.6. However the LSE method produces a completely different set of parameter estimates. The method identifies a turning point at the 64-th unscheduled maintenance action which is much closer to the end of the data. Then the trend decreases again at a much faster rate of

33.4% from a level of 423.7. Since inter-arrival times $X_{64} = 280$ and $X_{65} = 714$ are much higher than the neighbourhood values, it is possible that the estimation of turning point using the LSE method is seriously affected by the outlying effects around $i = 64$ and 65. Particularly, when the LSE method considers the $SSE$ of $X_i$ instead of $SSE$ of $\ln X_i$ and hence the LSE method is more sensitive to outliers. According to $AMSE$, parameters using the LSE method are preferred. However, fitted values after the 50-th unscheduled maintenance action provide a better description of the level and direction of trend apart from the outliers at the 64-th and 65-th unscheduled maintenance actions.

## 5.4 Discussion

In the MGP model, each turning point is estimated to be the beginning of a GP when the ratio parameter $a$ "change direction", that is, from "$a < 1$" to "$a > 1$" and vice versa. Specifically, when the ratio parameter $a_i$ in the $i$-th GP, that is $GP_i = \{X_i, \ldots, X_{i+L-1}\}$, enclosed by the $i$-th window changes direction, we set the turning point $T_{m+1} = i$ where $m$ indicates the number of such changes. In general, the moving window technique may detect more than one turning points. However, other choices of turning point estimates are possible. For instance, the turning point may be estimated to lie at the middle of the $GP_i$, that is,

$$T_{m+1} = \begin{cases} i + \dfrac{L-1}{2}, & \text{if } L \text{ is odd} \\ i + \dfrac{L}{2}, & \text{if } L \text{ is even} \end{cases}$$

or at the end of the $GP_i$, that is

$$T_{m+1} = i + L - 1$$

We prefer the choice of setting the turning point to lie at the beginning of the $GP_i$, that is $T_{m+1} = i$ because it has more appealing interpretation. When the ratio parameter $a_i$ changes direction, the whole $GP_i$ changes from increasing to decreasing and vice versa. Hence the whole $GP_i$ initiated by $X_i$ should be interpreted as a new trend.

In the MGP model, ad-hoc 'fine-tuning' procedures such as setting the minimum length of each trend to be $h = 10$, the minimal change of $a$ at the turning point to be $d = 0.005$ and the penalty for additional parameters to be $c(2K) = \frac{1}{2}\ln \bar{X}(2K)$ is used to detect significant trends. While it works reasonably well in most cases, more formal procedures for detecting significant trends and for hypothesis testing using large sample theory should be derived. For example, whether "$a = 1$" can be tested using the test statistics (3.4) in Section 3.1.2. Further research should also be done on improving the $AMSE$ for model selection on one side and adopting other model selection criteria such as the cross validation ($CV$) on the other.

To choose between the log-LSE and LSE methods of inference, we apply both methods and choose the set of parameter estimates with a lower $AMSE$. The LSE method is preferred in the SARS in HK data whereas the log-LSE method is preferred in the imports of goods index data and the No. 3 data. The LSE method does not require the averaging of zero $X_i$, therefore, it is preferred when the data contain some zeros. However the method also requires the solving of equation (3.18) which may be problematic when $X_i$ are mostly zero or linear, giving infinite or no solution for $a$. In these cases, it is reasonable to set $\widehat{a}_{NP2} = 1$ and hence $\widehat{\mu}_{NP2} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ which is a moment estimate.

Another problem facing the LSE method is its high sensitivity of the $SSE$ to outlying observations as demonstrated in the No.3 data. From the results of the three data analyses, the MGP model obviously provides much better model fits than that of the GP and GGP models. Hence we can conclude that when both effects of covariates and multiple trends are present in a data, the effect from multiple trends tends to be more influential. At present, we have derived models to account for the two effects separately. To improve the models and extend their scope of application, we may adopt a linear function of covariates to the mean parameter $\mu$ for each GP, resulting in the multiple GGP (MGGP) model. The set of equations in (3.6) and (3.7) using the log-LSE method and (3.18) using the LSE method can be similarly derived. Again the moving window technique can be applied to detect the turning points and the width $L$ of the moving windows can be selected according to $AMSE$.

At present, the method of inference for the MGP model and possibly the MGGP model is limited to NP inference. In fact, the ML estimate for $a$ can be obtained by solving equation (3.28) and then $\mu$ be evaluated in equation (3.29) when the RP $\{Y_i\}$ adopts an Exponential distribution. However for other life time distributions with scale parameter $\lambda$ and shape parameter $\alpha$, the parameters $a$, $\lambda$ and $\alpha$ have to be solved from a system of equations, such as the equations in (3.32) to (3.34) for Gamma distribution, using the NR method for each moving window. As iterations are required for each moving window, the methods become too computational intensive and infeasible. On the other hand, as the ML method adopting an Exponential distribution for the RP is similar to that of NP method, we do not consider the ML method in this research. Moreover, the implementation of the moving window technique using Bayesian method of inference is again not automatic in WinBUGS and hence the Bayesian method is not considered too.

A more straight forward method for estimating the $M$ turning points $T_m$, $m =$

$1, \ldots, M$, will be to set them as model parameters and estimate them using the ML or Bayesian methods. In Bayesian inference, the turning points may be set as $T_m = [t_{(m)}]$ where $t_m$ follows a uniform prior distribution $U(0, n)$, $t_{(m)}$ denotes the $m$-th order statistics of $\{t_m\}$ and $[\cdot]$ denotes a round-off function. Then the number of turning points $M$ can again be selected according to some goodness-of-fit measures say $AMSE$ or $CV$.

The weakness of the MGP model in prediction lies in the uncertainty about the turning point(s) when the events are still on-going. We have to obtain sufficient observations before we can locate the turning point approximately and hence predict the remaining trend fairly accurately. If the data show no sign of a "turn", we may model the data by a single GP. Otherwise, we may use the MGP model. Prediction using MGP model is demonstrated using the SARS in Hong Kong data. Results show that the prediction of the date for the removal of travel advice is very accurate even though the data are observed for only 18 more days after the turn from a growing to declining stage.

Table 5.1: Summary of parameter estimates for SARS in HK using MGP model

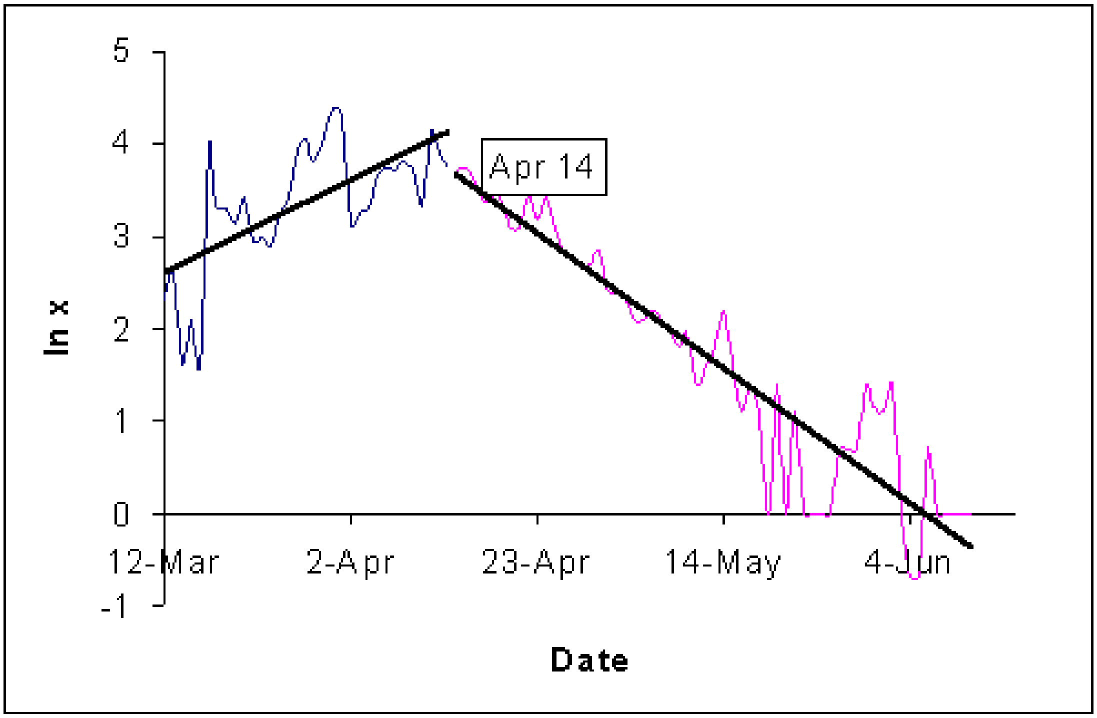
| *Method* | $M$ | $T_m$ | $L$ | $AMSE$ | $a$ | | $\mu$ | |
|---|---|---|---|---|---|---|---|---|
| log-LSE | 1 | N.A. | N.A. | 359.65 | 1.0429 | N.A. | 64.36 | N.A. |
| log-LSE | 2 | 34 | 20 | 118.98 | 0.9540 | 1.0718 | 14.854 | 41.222 |
| LSE | 2 | 33 | 20 | 101.09 | 0.9711 | 1.0713 | 21.298 | 44.020 |

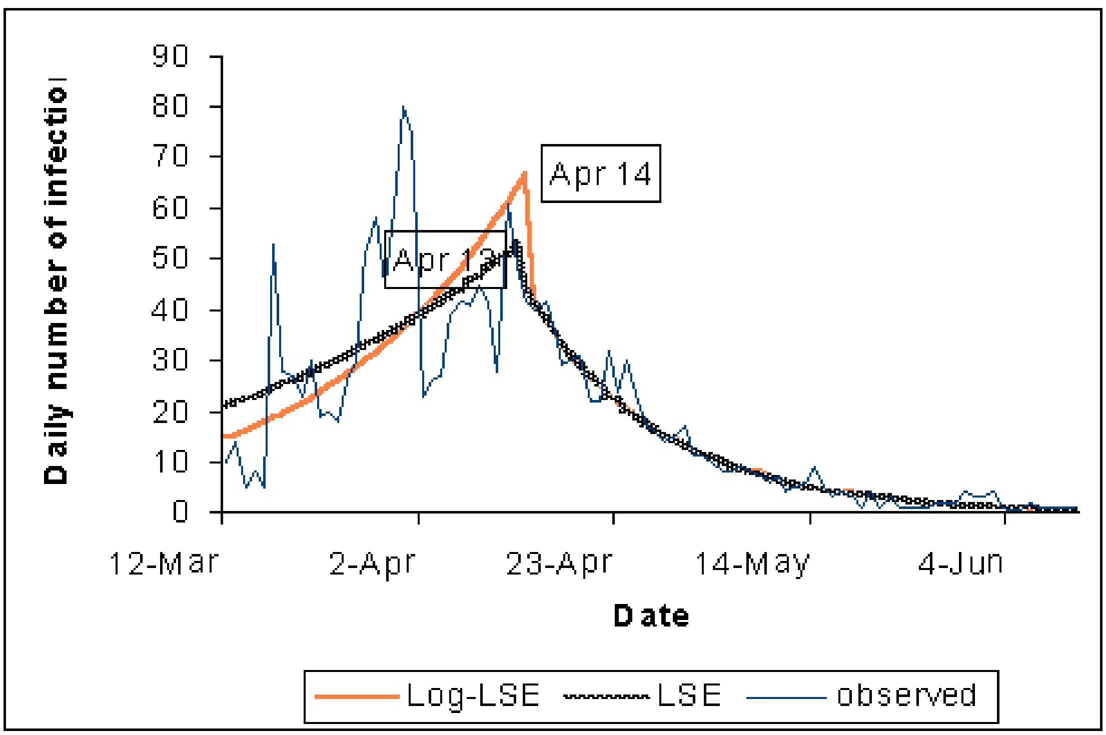
Table 5.2: Summary of parameter estimates for imports of goods index in HK using MGP model

| *Method* | $M$ | $T_m$ | $L$ | $AMSE$ | $a$ | | $\mu$ | |
|---|---|---|---|---|---|---|---|---|
| log-LSE | 1 | N.A. | N.A. | 461.23 | 0.9897 | N.A. | 11.0849 | N.A. |
| log-LSE | 2 | 194 | 24 | 90.22 | 0.9856 | 0.9956 | 8.1482 | 81.5923 |
| LSE | 2 | 255 | 20 | 415.192 | 0.9934 | 0.9893 | 21.2503 | 105.7379 |

Table 5.3: Summary of parameter estimates for No.3 using MGP model

| *Method* | $M$ | $T_m$ | $L$ | $AMSE$ | $a$ | | $\mu$ | |
|---|---|---|---|---|---|---|---|---|
| log-LSE | 1 | N.A. | N.A. | 195990.2 | 1.0146 | N.A. | 1079.9 | N.A. |
| log-LSE | 2 | 50 | 21 | 184178.7 | 1.1034 | 0.9705 | 2108.6 | 104.5 |
| LSE | 2 | 64 | 39 | 170508.3 | 1.0718 | 1.3338 | 1619.3 | 423.7 |


5
4
3
2
1
0
-1
ln x
Apr 14
12-Mar
2-Apr
23-Apr
14-May
4-Jun
Date


90
80
70
60
50
40
30
20
10
0
Daily number of infectio
Apr 14
Apr 13
12-Mar
2-Apr
23-Apr
14-May
4-Jun
Date
Log-LSE
LSE
observed

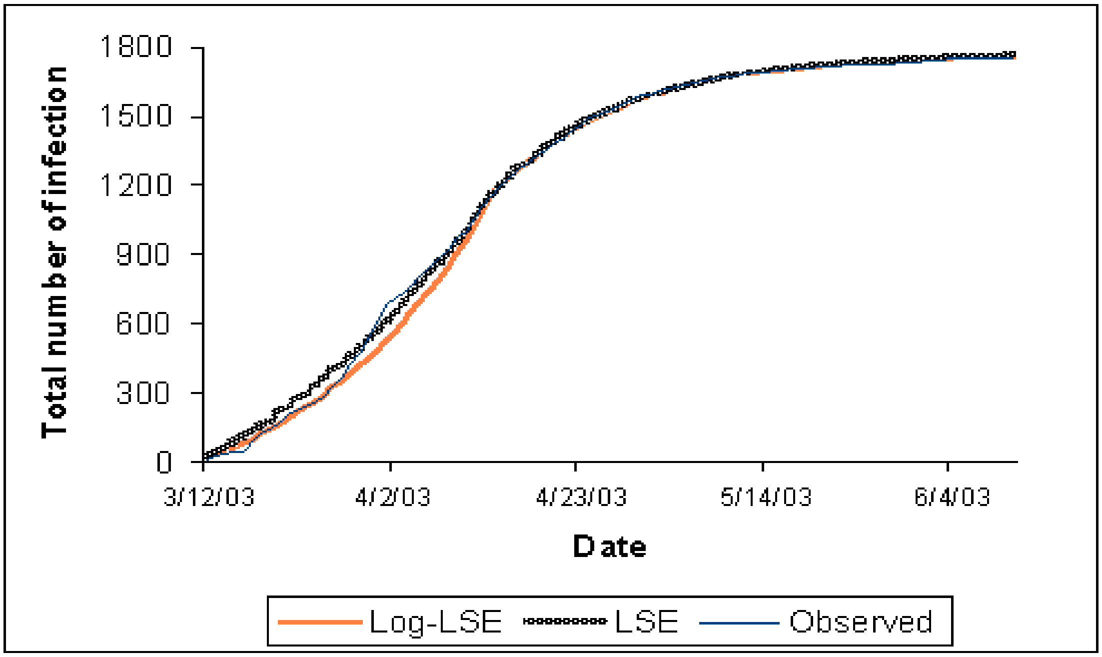
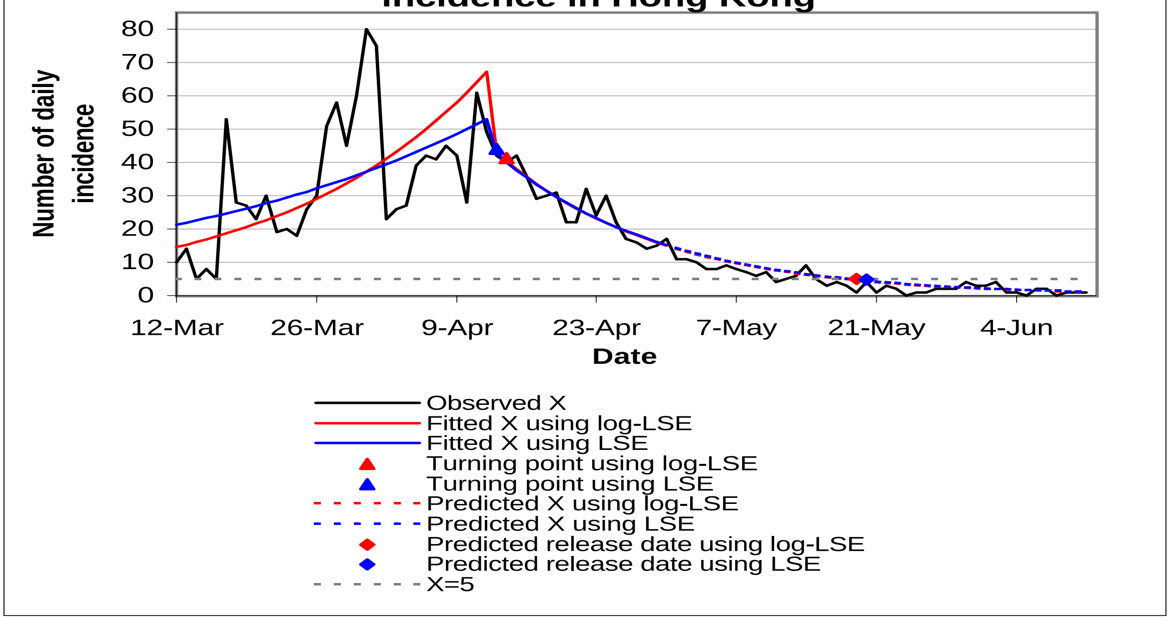
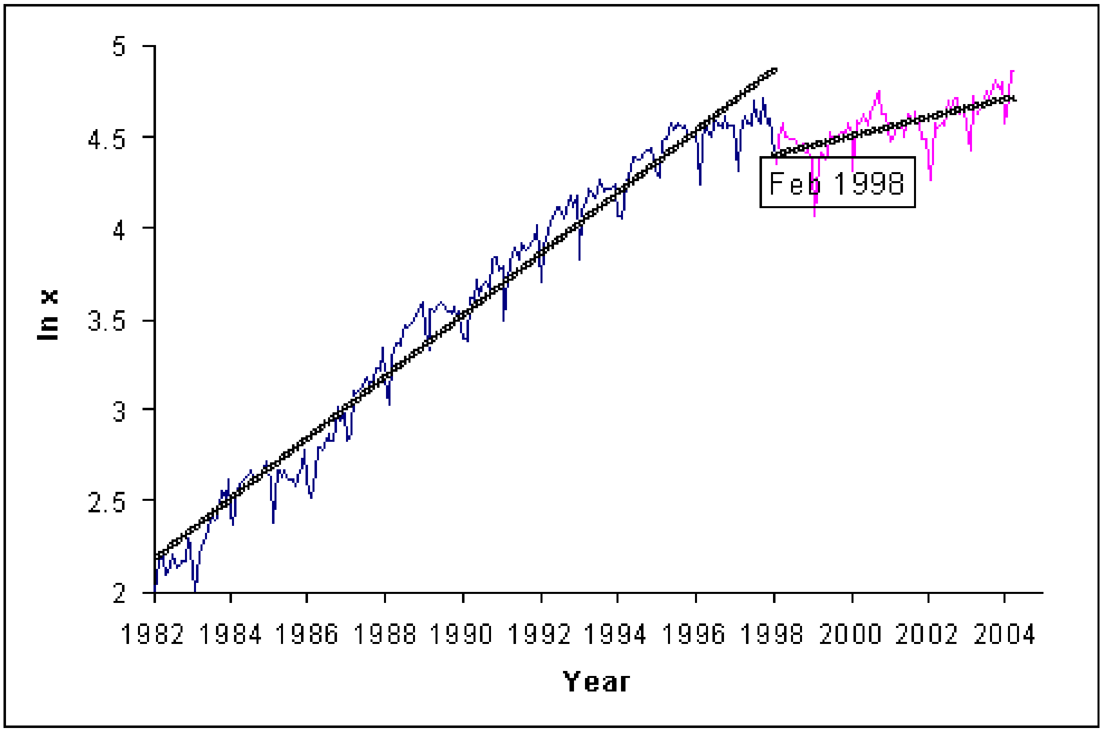
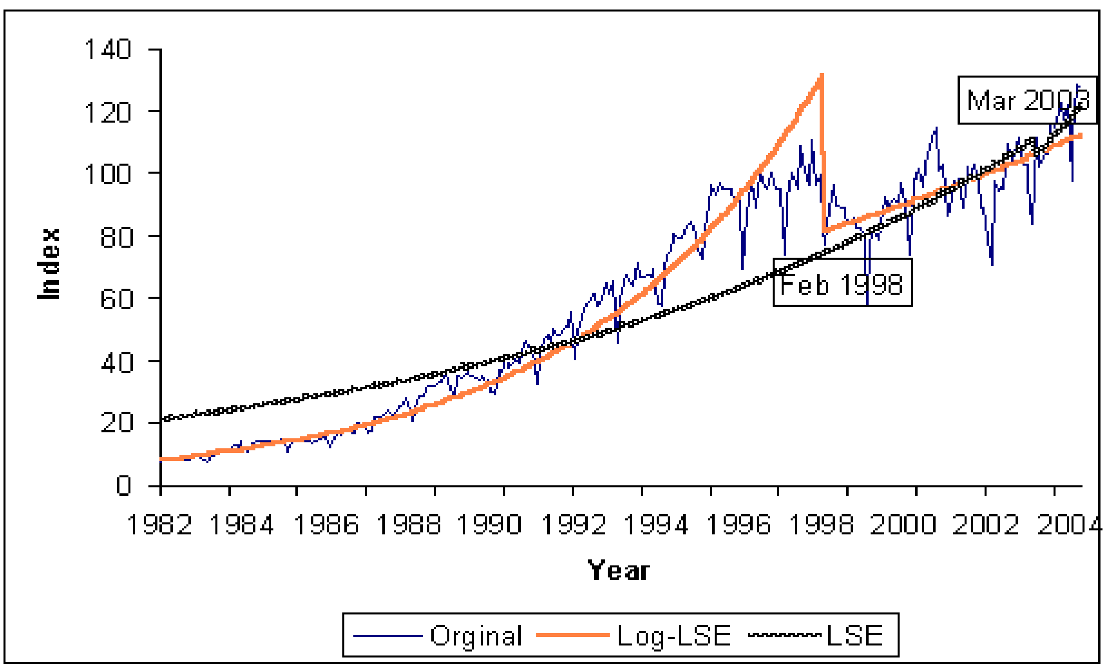
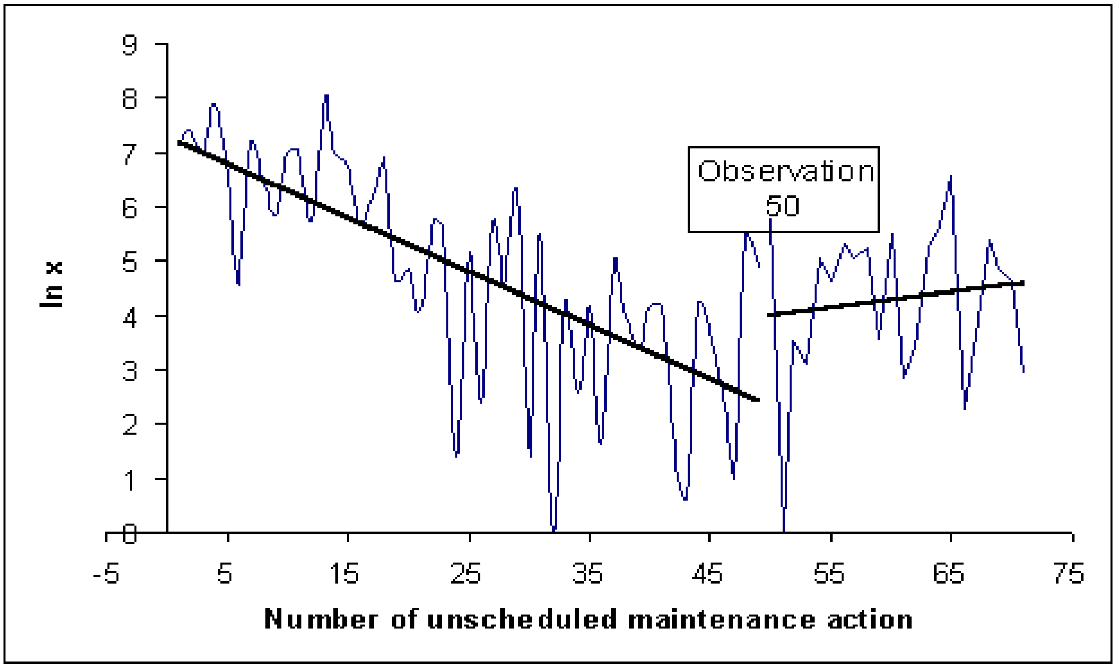



Figure 5: Predicted number of daily SARS incidence in Hong Kong

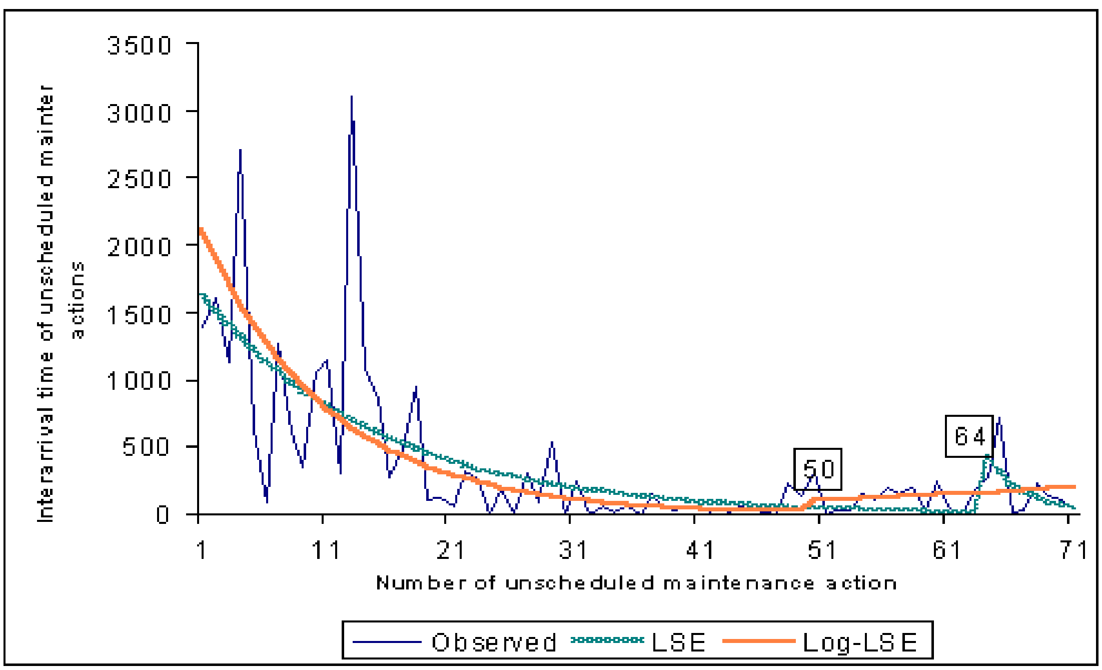



5
4.5
4
3.5
3
2.5
2
ln x
1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004
Year
Feb 1998

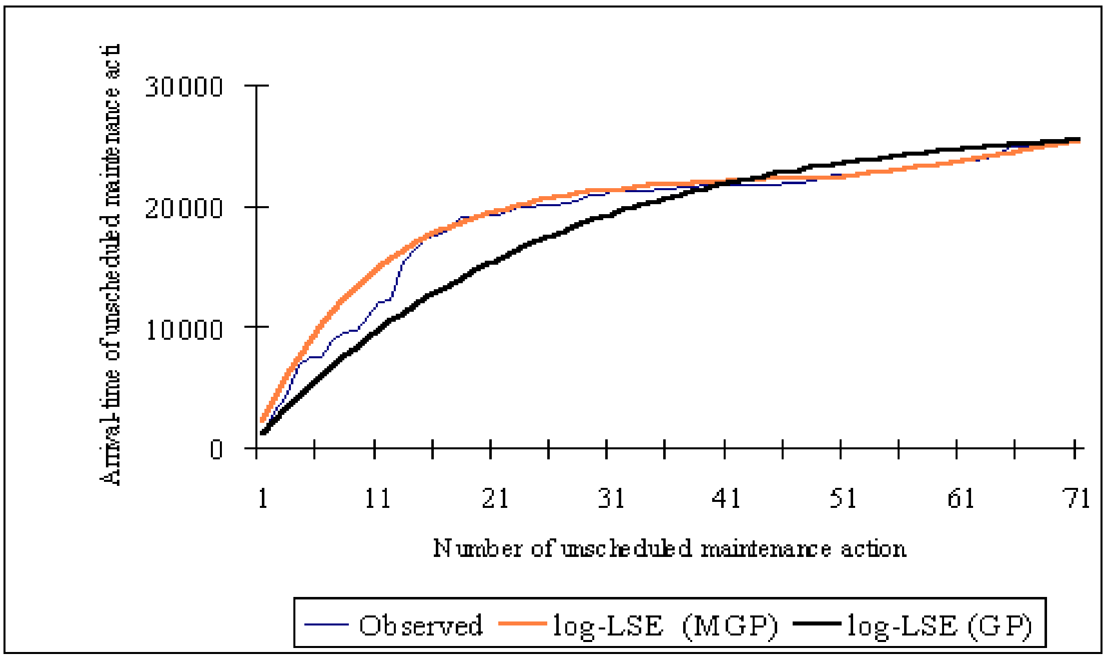

140
120
100
80
60
40
20
0
Index
1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004
Year
Mar 2003
Feb 1998
Orginal
Log-LSE
LSE



9
8
7
6
5
4
3
2
1
0
ln x
Observation 50
-5
5
15
25
35
45
55
65
75
Number of unscheduled maintenance action

3500
3000
2500
2000
1500
1000
500
0
Interarrival time of unscheduled mainter actions
50
64
1
11
21
31
41
51
61
71
Number of unscheduled maintenance action
Observed
LSE
Log-LSE


Arrival time of unscheduled maintenance acti
30000
20000
10000
0
1
11
21
31
41
51
61
71
Number of unscheduled maintenance action
Observed
log-LSE (MGP)
log-LSE (GP)

# CHAPTER 6

# BINARY GEOMETRIC PROCESS MODEL

## 6.1 Introduction

Binary data are often encountered in real situation. Geometric process (GP) model has been developed for positive continuous data. In this chapter, we will extend the GP model to binary data. Borrowing the idea in GLMMs, we assume that there is an underlying unobserved GP and we observed indicators of whether $X_i$ are greater than certain level. For simplity, we may set $S = 1$. The two methods of inference that using parametric approach will be considered: the Maximum Likelihood (ML) method and the Bayesian method. We adopt different distributions, for example, Exponential and Weibull distribution for the underlying renewal process (RP) $\{Y_i\}$.

We consider the following model. Suppose $\{X_i\}$ is the unobserved GP and $\{Y_i\}$ is a RP with $Y_i = a^{i-1}X_i$. Now, we define $W_i = I(X_i > 1) = I(Y_i > a^{i-1})$ where $I$ is an indicator function and $\{W_i\}$ are the observed binary data. Then, we have

$$P_i = P(W_i = 1) = P(X_i > 1) = I(Y_i > a^{i-1}) = 1 - P(Y_i < a^{i-1}) = 1 - F(a^{i-1}),$$

where $F(Y_i)$ is the cumulative distribution function (cdf) of $Y_i$. The likelihood function for the observed data $\{W_i\}$ is

$$L(\boldsymbol{\theta}) = \prod_{I=1}^{n} [1 - F(a^{i-1})]^{W_i} [F(a^{i-1})]^{1-W_i} \tag{6.1}$$

where $\boldsymbol{\theta}$ is a vector of model parameters and the log-likelihood function for the observed data $\{W_i\}$ becomes

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} w_i \ln[1 - F(a^{i-1})] + (1 - w_i) \ln[F(a^{i-1})] \tag{6.2}$$

Different methods of inference adopting different distribution will be evaluated through simulation and finally the Binary Geometric Process (BGP) model is fitted a Coal mining disaster data.

## 6.2 Methodology

### 6.2.1 Maximum Likelihood method

In this section, we assume that the RP $\{Y_i\}$ follows some life time distributions including Exponential, Weibull, Lognormal and Gamma. The likelihood and log-likelihood functions are given by (6.1) and (6.2) respectively substituting the probability density function (pdf) and cumulative distribution function (cdf) of each distribution. Newton Raphson (NR) method is used to evaluate the ML estimates that maximize the log-likelihood function. In the method, parameters are updaed according to the following equation.

$$\boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^{k} - 0.5\left(\frac{\partial^2 \ell}{\partial \boldsymbol{\theta}^2}\right)^{-1}\left(\frac{\partial \ell}{\partial \boldsymbol{\theta}}\right)\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^k}$$

Note that the updating procedures are slightly different from equation (2.11) in Section 2.3.2.1 by halving and hence lowering the speed of updating steps to avoid problems encountered if the convergence rate is too fast. Theoretically, these updating procedures will not affect the final parameter estimates but more iterations are required.

Moreover, the initial values in the updating procedures should not be too far away from the true values as the updating steps are very sensitive to current parameter estimates. Our suggestion is to use the Bayesian estimates to be the initial values.

#### 6.2.1.1 Lognormal distribution

Suppose the underlying RP $\{Y_i\}$ follows a Lognormal distribution with local parameter $\lambda > 0$ and scale parameter $\tau^2 > 0$, and the pdf of $Y_i$ is given by equation (2.1) and the cdf is

$$F(y) = \int_0^y \frac{1}{t\sqrt{2\pi\tau}} \exp\left[\frac{-(\ln t-\lambda)^2}{2\tau^2}\right] dt, \quad y \geq 0. \tag{6.3}$$

So the likelihood function for the observed data $\{W_i\}$ becomes

$$L(a,\lambda,\tau^2)=\prod_{i=1}^{n}\left\{\int_{a^{i-1}}^{\infty} \frac{1}{t\sqrt{2\pi\tau}} \exp\left[\frac{-(\ln t-\lambda)^2}{2\tau^2}\right] dt\right\}^{w_i} \cdot \left\{\int_0^{a^{i-1}} \frac{1}{t\sqrt{2\pi\tau}} \exp\left[\frac{-(\ln t-\lambda)^2}{2\tau^2}\right] dt\right\}^{1-w_i}$$

and the log-likelihood function is

$$\ell(\boldsymbol{\theta})=\sum_{i=1}^{n} w_i \ln\left\{\int_{a^{i-1}}^{\infty} \frac{1}{t\sqrt{2\pi\tau}} \exp\left[\frac{-(\ln t-\lambda)^2}{2\tau^2}\right] dt\right\} + (1-w_i)\ln\left\{\int_0^{a^{i-1}} \frac{1}{t\sqrt{2\pi\tau}} \exp\left[\frac{-(\ln t-\lambda)^2}{2\tau^2}\right] dt\right\}.$$

where $\boldsymbol{\theta} = (a, \lambda, \tau^2)$ is a vecttor of model parameters. We can derive the first and second order derivatives of the log-likelihood function as required in the NR method. However, since the cdf given in (6.3) involves integration, the evaluations of the log-likelihood functions and its derivatives are subject to error in approximating the integrals. Hence, the updating procedures in the NR method are more problematic. As a result, we do not consider the Lognormal distribution for the RP $\{Y_i\}$.

#### 6.2.1.2 Exponential distribution

Suppose that the underlying RP $\{Y_i\}$ follows an Exponential distribution with scale parameter $\lambda > 0$, and The pdf is given by equation (2.3) and the cdf is

$$F(y) = 1 - e^{-\lambda y}, \quad y \geq 0.$$

So the likelihood function for the observed data $\{W_i\}$ becomes

$$L(a,\lambda) = \prod_{i=1}^{n} \left(e^{-\lambda a^{i-1}}\right)^{w_i} \left(1 - e^{-\lambda a^{i-1}}\right)^{1-w_i}, \tag{6.4}$$

and the log-likelihood function is

$$\ell(\boldsymbol{\theta}) = -\sum_{i=1}^{n} w_i \lambda a^{i-1} + \sum_{i=1}^{n} (1-w_i) \ln\left(1 - e^{-\lambda a^{i-1}}\right).$$

where $\boldsymbol{\theta} = (a, \lambda)$ is a vector of model parameters On differentiation, we obtain the following first order derivatives

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial a} = -\sum_{i=1}^{n} w_i \frac{1}{a}(i-1)(\lambda a^{i-1}) + \sum_{i=1}^{n}(1-w_i)\frac{1}{a}(i-1)(\lambda a^{i-1}) e^{-(\lambda a^{i-1})}[1 - e^{-(\lambda a^{i-1})}]^{-1}$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} = -\sum_{i=1}^{n} w_i a^{i-1} + \sum_{i=1}^{n}(1-w_i)\frac{1}{\lambda}(\lambda a^{i-1}) e^{-(\lambda a^{i-1})}[1 - e^{-(\lambda a^{i-1})}]^{-1}$$

The second order derivatives when $\{Y_i\}$ follows an Exponential distribution as well as other distributions used in BGP model are given in Appendix D.

#### 6.2.1.3 Gamma distribution

Suppose that the RP $\{Y_i\}$ follows a Gamma distribution with shape parameter $r > 0$ and scale parameter $\lambda > 0$. The pdf of $Y_i$ is given by equation (2.5) and the cdf is

$$F(y) = \int_0^y \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt, \quad y \geq 0.$$

where $\Gamma(.)$ is a Gamma function. Then the likelihood function for the observed data $\{W_i\}$ becomes

$$L(a, \alpha, \lambda) = \prod_{i=1}^{n} \left\{ \int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt \right\}^{w_i} \cdot \left\{ \int_0^{a^{i-1}} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt \right\}^{1-w_i},$$

and the log-likelihood function is

$$\ell(\boldsymbol{\theta}) = -\sum_{i=1}^{n} w_i \ln\left\{ \int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt \right\} + (1-w_i) \ln\left\{ \int_0^{a^{i-1}} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt \right\}.$$

where $\boldsymbol{\theta} = (a, \alpha, \lambda)$ is a vector of model parameters. On differentiation, we obtain the following first order derivative equations for $a$, $\alpha$ and $\lambda$.

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial a} = \sum_{i=1}^{n} w_i \left[ \frac{\frac{\lambda^\alpha}{\Gamma(\alpha)} (i-1) a^{i\alpha-\alpha-1} e^{-\lambda a^{i-1}}}{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt} \right] + \sum_{i=1}^{n} (1-w_i) \left[ \frac{\frac{\lambda^\alpha}{\Gamma(\alpha)} (i-1) a^{i\alpha-\alpha-1} e^{-\lambda a^{i-1}}}{\int_0^{a^{i-1}} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\, dt} \right]$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = \sum_{i=1}^{n} w_i \left[ \frac{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \left( \ln \lambda + \ln t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \right) \, dt}{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \, dt} \right] +$$

$$\sum_{i=1}^{n} (1 - w_i) \left[ \frac{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \left( \ln \lambda + \ln t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \right) \, dt}{\int_{0}^{a^{i-1}} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \, dt} \right]$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} = \sum_{i=1}^{n} w_i \left[ \frac{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \left( \frac{\alpha}{\lambda} - t \right) \, dt}{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \, dt} \right] +$$

$$\sum_{i=1}^{n} (1 - w_i) \left[ \frac{\int_{a^{i-1}}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \left( \frac{\alpha}{\lambda} - t \right) \, dt}{\int_{0}^{a^{i-1}} \frac{\lambda^\alpha}{\Gamma(\alpha)} \, t^{\alpha-1} e^{-\lambda t} \, dt} \right]$$

To evaluate these integrals, we use the numerical approximation function namely QDAGI and QDAGS in the IMSL Library implemented in FORTRAN programs.

#### 6.2.1.4 Weibull distribution

Suppose that the RP $\{Y_i\}$ follows a Weibull distribution with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$. The pdf of $Y_i$ is given by equation (2.8) the its cdf is

$$F(y) = 1 - e^{-(\lambda y)^\alpha}, \quad y \geq 0.$$

So the likelihood function for the observed data $\{w_i\}$ is

$$L(a, \alpha, \lambda) = \prod_{i=1}^{n} \{ e^{[-(\lambda a^{i-1})^\alpha]} \}^{w_i} \cdot [1 - e^{-(\lambda a^{i-1})^\alpha}]^{1-w_i}, \tag{6.5}$$

and the log-likelihood function is

$$\ell(\boldsymbol{\theta}) = \ln L(a, \alpha, \lambda) = -\sum_{i=1}^{n} w_i (\lambda a^{i-1})^\alpha + \sum_{i=1}^{n} (1 - w_i) \ln\{1 - e^{[-(\lambda a^{i-1})^\alpha]}\}.$$

where $\boldsymbol{\theta} = (a, \lambda, \tau^2)$ is a vecttor of model parameters. On differentiation, we obtain the following first order derivatives.

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial a} = -\sum_{i=1}^{n} w_i \frac{1}{a}[(i-1)\alpha](\lambda a^{i-1})^{\alpha} + \sum_{i=1}^{n}(1-w_i)\frac{1}{a}[(i-1)\alpha](\lambda a^{i-1})^{\alpha} e^{-(\lambda a^{i-1})^{\alpha}}[1-e^{-(\lambda a^{i-1})^{\alpha}}]^{-1}$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = -\sum_{i=1}^{n} w_i (\lambda a^{i-1})^{\alpha} \ln(\lambda a^{i-1}) + \sum_{i=1}^{n}(1-w_i)\ln(\lambda a^{i-1})(\lambda a^{i-1})^{\alpha} e^{-(\lambda a^{i-1})^{\alpha}}[1-e^{-(\lambda a^{i-1})^{\alpha}}]^{-1}$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} = -\sum_{i=1}^{n} w_i \frac{\alpha}{\lambda}(\lambda a^{i-1})^{\alpha} + \sum_{i=1}^{n}(1-w_i)\frac{\alpha}{\lambda}(\lambda a^{i-1})^{\alpha} e^{-(\lambda a^{i-1})^{\alpha}}[1-e^{-(\lambda a^{i-1})^{\alpha}}]^{-1}$$

### 6.2.2 Bayesian method

Similar to Section 4.3.2, the implementatin of Bayesian method using WinBUGS is infeasible when $\{Y_i\}$ follows Lognormal and Gamma distributions because integration and Gamma function are involved in the likelihood function respectively and they are not supported in WinBUGS. In Bayesian inference, we set the RP $\{Y_i\}$ follows an Exponential distribution. As before, Bayesian method is implemented using a user defined likelihood function for the observed data $\{W_i\}$ given by equations (6.4) and (6.5) when $Y_i$ follows Exponential and Weibull distribution respectively. Moreover the priors for the model parameters are given below:

$$a \sim \; U[0.95, 1.05], \;\; \alpha \sim Ga(c, d), \;\; \lambda \sim Ga(c, d)$$

where $c = 1 \times 10^{-9}$ and $d = 1 \times 10^{-6}$ so that the priors for $\lambda$ and $\alpha$ are non-informative since the mean of the Gamma distribution is $\frac{c}{d} = 0.001$ and the variance is $\frac{c}{d^2} = 1000$. The WinBUGS commands for the BGP model adopting with an Exponential distribution for the RP $\{Y_i\}$ are given in Appendix B.

To obtain the parameter estimates, we discard the first 1000 iterations in the burn-in period, and take the parrameter estimates in every 10-th iteration from 1001 to 7000 resulting in samples of size 600. The auto-correlation functions, densities and history plots of all posterior samples of parameters are given in Appendix C and they show that the samples are stable and independent.

## 6.3 Simulation

In simulation studies, we investigate the accuracy of parameter estimates using ML method only for reasons described in Section 4.4. Moreover, we assume that the RP $\{Y_i\}$ follows an Exponential distribution only because of the reasons stated in Section 6.2.1. Among the many factors that affect the accuracy of parameter estimates, we study the effects of the length $n$ of the data. To do this, we set the scale parameter of the Exponential distribution to be $\lambda = 2.0$, the sample size $n = 100$, 200 and 300 and the number of realizations $N = 200$.

***************************************************

Table 6.1 is shown here.

***************************************************

Table 6.1 shows that different length $n$ do not have a large effect on the accuracy of parameter estimation unless the $n$ is too small. The ratio parameter $a$ is consistently underestimated when $a$ is less than one and overestimated when $a$ is larger than one. Accuracy in $\hat{a}$ decreases when $a$ is further away from one from above and $n$ is smaller. Similarly accuracy in $\hat{a}$ decreases when $a$ is further away from one from below but to a smaller magnitude. Moreover, the accuracy of $\hat{\lambda}$ also decreases when $a$ is further away from one and the decrease is more when $a$ is greater than one. Note that the standard deviation of $\lambda$ is high especially when $a$ is further away from one from above. These results may be due to the fact that binary data is less informative and hence parameter estimations are highly sensitive to small change in $a$ when the trend is increasing at a faster rate, that is, $a$ is further away from one from above.

*************************************************

Table 6.2 is shown here.

*************************************************

Table 6.2 gives parameter estimates with different scale parameters $\lambda$. It can be seen that if $\lambda$ is large, say $\lambda > 3.0$, the accuracy of $\hat{\lambda}$ becomes worse. On the whole, simulation results show that the BGP model using ML method is restricted to data when the underlying GP has ratio $a$ close to one and the data is longer say $n \geq 100$

## 6.4 Real data analysis

We use the data set coal mining disasters data that is reported in Andrews and Herzberg (1985). The data consist of 190 inter-arrival times between successive disaster in Great Britain. It was originally given by Maguire *et al.* (1952) and was studied by Cox and Lewis (1966) in the analysis of trend. Afterward, the data were extended so that it cover the period from $15^{th}$ May 1851 to $22^{nd}$ March 1962, total to 40550 days. Lam (1992b), Lam and Chan (1998), Lam *et al.* (2004) and Chan *et al.* (2004) fitted the data to GP models.

We transform the inter-arrival times to binary data by first dividing a year into four quarters and then defining an indicator function $W_i$ to indicate whether a disaster occurs to these quarters. Since the data cover a period of 112 years, we have totally 448 quarters and 448 data in the transformed data set. In reality, indicators of hazardous or failure levels are required to monitor the operation of certain systems. If the likelihood of a failure to the system is high regardless of the intensity of the disaster, certain control policy should be implemented. Hence the informations from these binary indicators of disaster are important.

Both ML and Bayesian give similar parameter estimates which are 0.5889 and 0.5843 for $\lambda$ using the ML and Bayesian methods respectively, and 1.003 for $a$ using both methods. In order to assess the goodness-of-fit, we simulate 200 sets of data using each set of parameters and then take the average of the simulated data to obtain a set of expected values $E(W_i)$. Plot of cumulated $W_i$ and $E(W_i)$ using the ML and Bayesian method is shown in Figure 6.1. The plot shows that ML method provdies better model fit.

## 6.5 Discussion

In this chapter, we extend the GP model to binary data by modeling the underlying RP to follow some lifetime distributions, including Exponential, Weibull, Gamma and Lognormal. Parameters are estimated using ML and Bayesian methods. Non-parametric method of inference is generally not considered for binary data because the Sum of Squared Error ($SSE$) defined as

$$SSE = \sum_{i=1}^{n}[w_i - E(W_i)]^2$$

as required in the LSE method will be defined in terms of the observed data $W_i$ and its expected value $E(W_i)$. Such definition is not proper for binary data.

Theoretically the RP can follow one of the four life time distributions mentioned above. However the log-likelihood functions for the Gamma and Lognormal distributions involve integration and hence numerical approximations to these functions and their first and second order derivatives produce additional source of errors in the NR procedures. When the ratio parameter $a$ is relatively far away from one, these functions become very sensitive to the current parameter estimates making the NR procedures highly problematic. Moreover the scale parameter $\lambda$ of the Weibull distribution appears in the indices of some terms, say $(\lambda a^{i-1})^\alpha$ and $e^{-(\lambda a^{i-1})^\alpha}$ in the log-likelihood function and its derivatives, again making these functions highly sensitive to the current parameter estimate of $\alpha$. As a result, simulation studies as well as real data analyses for the BGP model when the RP adopts Lognormal, Gamma and Weibull distributions are not successful.

Similar problems arise when the Bayesian method is used and is implemented by WinBUGS. It is not possible to run WinBUGS when the RP follows Gamma and Lognormal distributions because the user defined likelihood functions for the observed data $W_i$ involve integrations which are not supported in WinBUGS. Moreover the parameter estimates when the RP adopts a Weibull distribution are again problematic. Parameter estimates of $\mu$ and $a$ deviate a lot from the true values in the simulation studies.

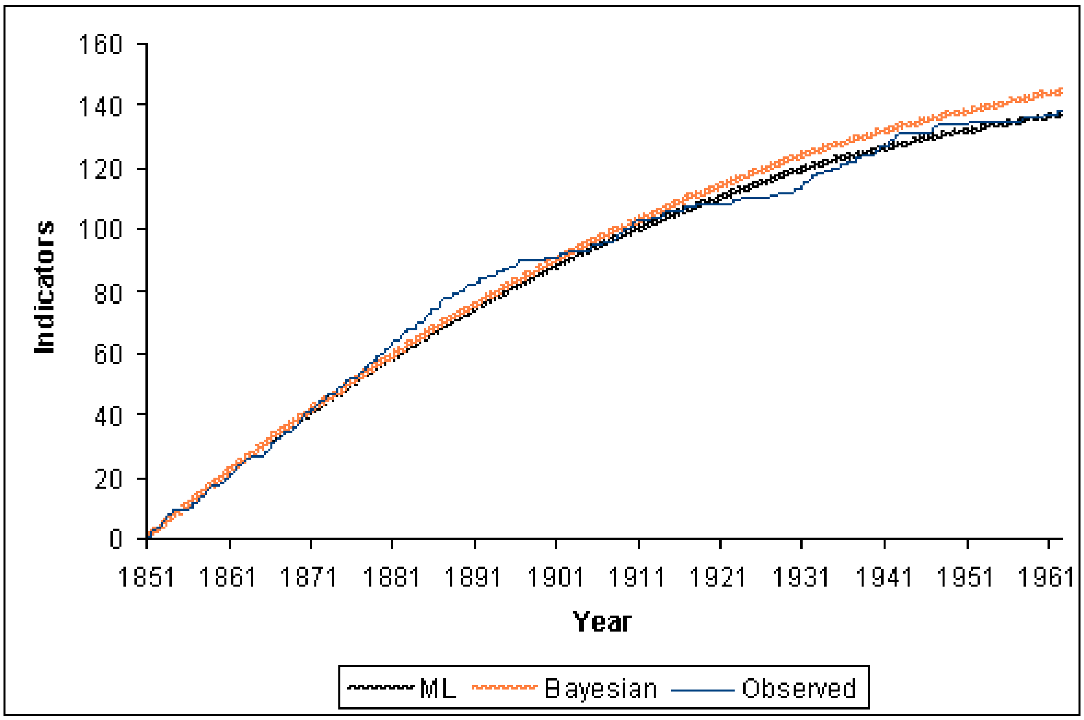
In fact binary data are much less informative and hence there may not be sufficient information in the data to estimate precisely the shape parameter $\alpha$ of more advanced life-time distributions. In other words, the adoption of different lifetime distributions may have little effect on the model parameters $a$ and $\lambda$. Moreover, there are some limitations in the ML method. If the starting values for the NR procedures are far away from the true values, the updating parameters cannot converge. Simulation studies also show that estimates using ML method may not be accurate if the ratio $a$ is not close to one, say $a < 0.95$ or $a > 1.05$. Hence to fit the BGP model to binary data, we recommend the adoption of Exponential distribution to the underlying RP and the use of Bayesian method for inference if the ML method fails as the two methods give similar estimates.

Table 6.1: Simulation study for BGP model ($\lambda = 2.0$)

| | $n = 100$ | | $n = 150$ | | $n = 300$ | |
|---|---|---|---|---|---|---|
| *true a* | $\hat{a}$ | $\hat{\lambda}$ | $\hat{a}$ | $\hat{\lambda}$ | $\hat{a}$ | $\hat{\lambda}$ |
| *0.96* | | | | | | |
| *Estimate* | 0.9575 | 2.5030 | 0.9575 | 2.5218 | 0.9574 | 2.5137 |
| *SD* | 0.0100 | 1.5866 | 0.0090 | 1.7115 | 0.0086 | 1.6978 |
| *0.98* | | | | | | |
| *Estimate* | 0.9801 | 2.1805 | 0.9796 | 2.1677 | 0.9789 | 2.2348 |
| *SD* | 0.0057 | 0.7848 | 0.0036 | 0.6072 | 0.0029 | 0.6547 |
| *1.0* | | | | | | |
| *Estimate* | 0.9993 | 2.1585 | 1.0002 | 2.0234 | 1.0079 | 1.9709 |
| *SD* | 0.0050 | 0.8023 | 0.0027 | 0.4965 | 0.0495 | 0.2683 |
| *1.02* | | | | | | |
| *Estimate* | 1.0651 | 2.1709 | 1.0577 | 2.1554 | 1.0428 | 2.1788 |
| *SD* | 0.1222 | 1.3436 | 0.0940 | 1.3097 | 0.0518 | 1.2686 |
| *1.04* | | | | | | |
| *Estimate* | 1.1328 | 2.9937 | 1.1061 | 2.9994 | 1.0647 | 3.0750 |
| *SD* | 0.1688 | 2.4824 | 0.1179 | 2.4343 | 0.0586 | 2.3297 |

Table 6.2: Simulation study for Exponential distribution ($n = 150$)

| | $\lambda = 1$ | | $\lambda = 3$ | | $\lambda = 4$ | |
|---|---|---|---|---|---|---|
| $a$ | $\hat{a}$ | $\hat{\lambda}$ | $\hat{a}$ | $\hat{\lambda}$ | $\hat{a}$ | $\hat{\lambda}$ |
| *0.96* | | | | | | |
| *Estimate* | 0.9559 | 1.4301 | 0.9552 | 3.0384 | 0.9583 | 4.5725 |
| *SD* | 0.0152 | 0.9305 | 0.0021 | 2.8432 | 0.0082 | 1.6979 |
| *0.98* | | | | | | |
| *Estimate* | 0.9787 | 1.1626 | 0.9801 | 3.1805 | 0.9796 | 4.2637 |
| *SD* | 0.0083 | 0.3470 | 0.0050 | 0.9949 | 0.0046 | 1.2762 |
| *1.0* | | | | | | |
| *Estimate* | 0.9848 | 1.0766 | 0.9999 | 3.2311 | 1.0056 | 4.8683 |
| *SD* | 0.0917 | 0.3670 | 0.0057 | 1.1369 | 0.0186 | 2.8948 |
| *1.02* | | | | | | |
| *Estimate* | 1.0200 | 1.0930 | 1.0788 | 4.1234 | 1.0573 | 6.0061 |
| *SD* | 0.0074 | 0.4591 | 0.1242 | 2.5259 | 0.1042 | 1.8868 |
| *1.04* | | | | | | |
| *Estimate* | 1.0624 | 1.0324 | 1.1036 | 5.0736 | 1.0694 | 6.1044 |
| *SD* | 0.0590 | 0.6070 | 0.1551 | 2.6345 | 0.1398 | 1.8223 |



160
140
120
100
80
60
40
20
0
Indicators
1851
1861
1871
1881
1891
1901
1911
1921
1931
1941
1951
1961
Year
ML
Bayesian
Observed

# CHAPTER 7

# DISCUSSION

The non-homogeneous Poisson process (NHPP) in which the hazard rate is monotone has been used to model data with trend. Lam (1988a, b) first proposed modeling directly the monotone trend by a monotone process called the Geometric Process (GP). Previous GP model limits to the modeling of inter-arrival times of a series of events generated from a process or system such that the underlying renewal process (RP) has a constant mean $\mu$ which indicates the initial level of a trend and a constant ratio $a$ which indicates the direction and strength of the trend. Previous researches on GP model concentrate mostly on the inferences and applications of GP models to reliability and maintenance problems. In this research, we aim to generalize the model in terms of the modeling methodologies, methods of inference and fields of application.

Equation (1.1) shows that the ratio parameter $a$ of a GP affects both its mean and variance and allows them to change over time. Hence the ratio parameter plays a significant role in the whole modeling methodologies and it reduces the GP model to the Generalized Linear Mixed Models (GLMMs) when $a = 1$. Following the framework of the exponential family of distributions within the GLMMs, GP models are extended to the generalized GP (GGP) models in Chapter four by including a linear function of covariates log-linked to the mean of some (lifetime) sampling distributions for the underlying stochastic process (SP) $\{Y_i\}$, in order to respond to the diverse effects exerted from the environment onto the GP. In general, GGP models may also accommodate clustering effects, heterogeneity and over-dispersion by inclusion of mixture and random effects in the mean function. Such extension is important and worths further investigation.

By extending the mean of a RP to a mean function of covariates and possibly mixture and random effect terms, we implicitly assume that the effects on a GP can be separated into two types: effects which affect the evolving but latent baseline levels of the process and effects which affect the progression of the process (the rate of change), geometric in nature, across time. These two types of effects are separately measured by two sets of model parameters, the parameters in the mean function $\mu_i$ (say $\beta_0$ and $\beta_1$) and the parameter(s) in the ratio $a$. The ratio parameter $a$ actually reveals the direction and strength of the progression which may change over time. A more generalized GP model can be developed by log-linking the ratio parameter $a$ to a linear function of covariates in a way similar to that of the mean parameter $\mu$, that is

$$\ln a_i = \beta_{a0} + \beta_{a1} z_i + \ldots + \beta_{aK}, \quad i = 1, \ldots, n$$

where $z_{ik}$ are some covariates such as the time $i$. In this way, we have $Y_i = {a_i}^{i-1} X_i$, $E(X_i) = \mu_i / {a_i}^{i-1}$ and the ratio $a_i$ changes according to different stages of the process. For example, $a_i$ may change from "less than one" to "greater than one" indicating the gradual transition from a growing stage to a declining stage across $i$. Such adaptive feature of a GGP model opens up a wide area of future development and allows a broad range of applications.

Then GP models are further extended to multiple GP (MGP) models in Chapter five to capture multiple trend data that exhibit a growing, developing and declining stages of some processes. We investigate methods of detecting turning points and apply GP models repeatedly to each trend. We also remark that multiple GGP (MGGP) models are similar defined when a GGP instead of GP is fitted to each trend. We discuss future area of research in this new model. In fact, the application of GP models in series to a process offers one approach to the generalization of GP models. Sure one may consider the application of GP models to a parallel system.

Another approach to the generalization of GP models will be to consider a wider class of data including binary, multi-nominal and Poisson count data. Again the modeling methodologies are borrowed from that of the exponential family of sampling distributions within the GLMMs framework. We assume that there is an underlying GP $\{X_i\}$ such that the observed data $W_i = I(X_i > 1)$ and that the underlying RP $\{Y_i\}$ follows some life-time distributions. Results show that the simple life-time distribution, the Exponential distribution is the best choice. Following the idea, extension to other type of data, say the Poisson counts, is straight forward. For instance, the problem of zero daily infection in the SARS in Hong Kong data can be easily solved by adopting a Poisson distribution to $X_i$ with the means forming a GP, that is $X_i \sim Poisson(\mu_i)$ and $\mu_i = \mu/a^{i-1}$ where the parameters $\mu$ and $a$ indicate the initial level and ratio respectively of a GP. The resulting model is in fact a nonhomogeneous Poisson process (NHPP) adopting a "deterministic" GP to the means.

Apart from model developments, the other objective of this research is to improve the methods of inference for these extended models. One important contribution is the use of Bayesian method implemented by WinBUGS. As the software can be downloaded free and is easy to apply in a lot of sampling distributions, it saves us from deriving complicated log-likelihood functions and their derivatives and writing a FORTRAN program to implement the models using ML method. Lastly a wide range of applications of the extended GP models have been demonstrated using data in epidemiology, business and industry.

# Appendix A

## Estimate for the variance in the log-LSE method

Considering the regression model:

$$\ln(a^{i-1}X_i) = \lambda + \epsilon_i \Leftrightarrow \ln X_i = \lambda - (i-1)\ln a + \epsilon_i$$

It can be written in the form of

$$Y_i = \beta_0 + \beta_1 z_i$$

where $Y_i = \ln X_i$, $\beta_0 = \lambda$, $\beta_1 = -\ln a$ and $z_i = i - 1$. Thus, an estimate for the variance $\tau^2 = Var(\epsilon_i)$ is given by

$$\hat{\tau}^2 = \frac{SSE}{n-2}$$

where

$$SSE = \sum_{i=1}^{n}(y_i - \bar{y})^2 - \hat{\beta}\sum_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z}),$$

and

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}\left[\ln x_i - \frac{1}{n}\sum_{i=1}^{n}\ln x_i\right]^2 = \sum_{i=1}^{n}(\ln x_i)^2 - \frac{1}{n}\left(\sum_{i=1}^{n}\ln x_i\right)^2$$

$$\begin{aligned}
\sum_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z}) &= \sum_{i=1}^{n}\left(\ln x_i - \frac{1}{n}\sum_{i=1}^{n}\ln x_i\right)\left(i - 1 - \frac{n-1}{2}\right) \\
&= -\frac{1}{2}\sum_{i=1}^{n}\left(\ln x_i - \frac{1}{n}\sum_{i=1}^{n}\ln x_i\right)(n - 2i + 1) \\
&= -\frac{1}{2}\sum_{i=1}^{n}(n - 2i + 1)\ln x_i + \frac{1}{2n}\left(\sum_{i=1}^{n}\ln x_i\right)\left[\sum_{i=1}^{n}(n - 2i + 1)\right] \\
&= -\frac{1}{2}\sum_{i=1}^{n}(n - 2i + 1)\ln x_i
\end{aligned}$$

since $\sum_{i=1}^{n}(n - 2i + 1) = 0$. Hence,

$$\hat{\tau}^2 = \frac{1}{n-2}\left[\sum_{i=1}^{n}(\ln x_i)^2 - \frac{1}{n}\left(\sum_{i=1}^{n}\ln x_i\right)^2 - \frac{\widehat{\ln a}}{2}\sum_{i=1}^{n}(n - 2i + 1)\ln x_i\right]$$

# Appendix B

# WinBUGS program

## GGP model with Exponential distribution

```
model
{ c <- 10000
for (i in 1:N) {
ones[i] <- 1
ones[i] ~ dbern( p[i] )
p[i] <- pow(a,i-1)*(1/(beta0+beta1*quarter[i]))*
exp(-1*pow(a,i-1)*(1/(beta0+beta1*quarter[i]))*x[i]) / c
}
beta0 ~ dgamma(0.001,0.001)
beta1 ~ dgamma(0.001,0.001)
a ~ dunif(0.95,1.05)
}


   list(beta0=0.1, beta1=0.1, a=1.01)
list(N=267)
time[] x[] quarter[]
```

## BGP model with Exponential distribution

```
model
{ c <- 10000
for (i in 1:N){
ai1[i] <- pow(a, (i-1))
e[i] <- exp(-1*lam*ai1[i])
ones[i] <- 1
ones[i] ~ dbern( p[i] )
p[i] <- pow(e[i],x[i])*pow(1-e[i],1-x[i]) / c
}
lam ~ dgamma(0.001,0.001)
a ~ dunif(0.95,1.05)
}


list(lam=0.6, a=1.01)
list(N=448)
x[]
```

BGP model with Weibull distribution

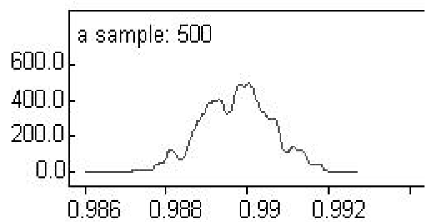
```
model
{ c < − 10000
for (i in 1:N){
ai1r[i] < − pow(a, (i-1)*r)
lamr[i] < − pow(lam,r)
e[i] < − exp(-1*lamr[i]*ai1r[i])
ones[i] < − 1
ones[i] ∼ dbern( p[i] )
p[i] < − pow(e[i],x[i])*pow(1-e[i],1-x[i]) / c
}
lam ∼ dgamma(0.001,0.001)
r ∼ dgamma(0.001,0.001)
a ∼ dunif(0.95,1.05)
}
```

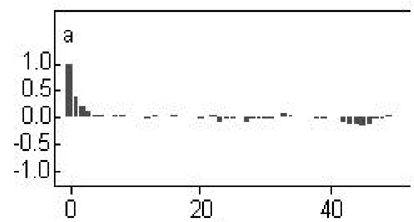
```
list(lam=1.0, r=1.0, a=1.01)
list(N=189)
x[]
```

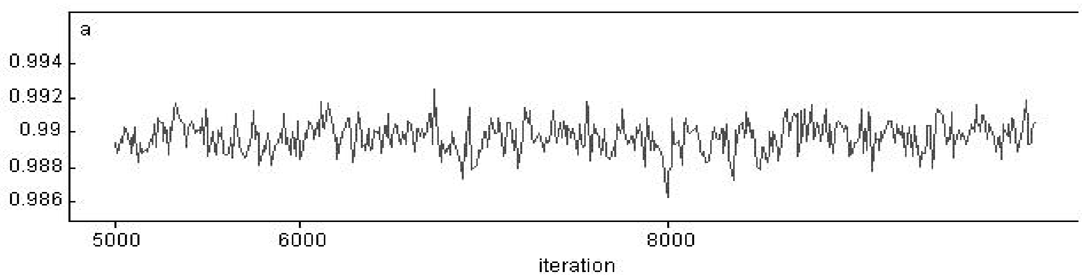
# Appendix C

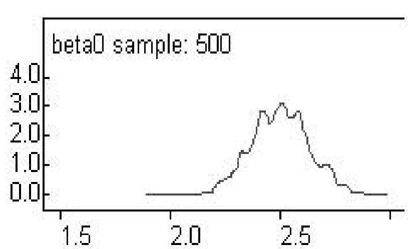
## Auto-correlation functions, densities and history plots of all posterior samples in WinBUGS program

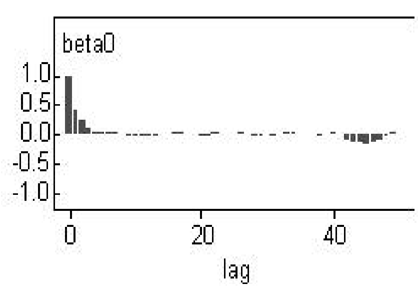
The GGP model with Exponential distribution

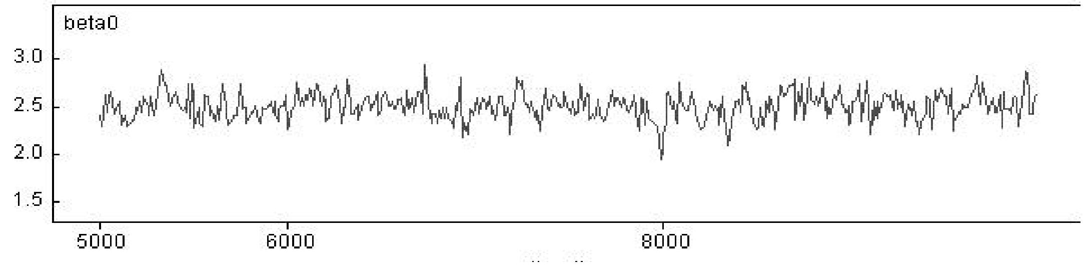


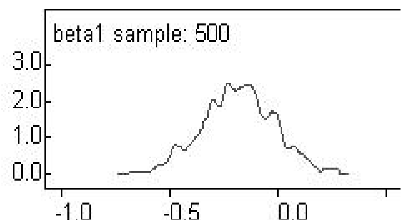



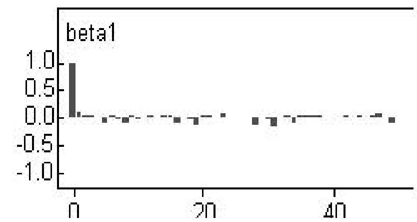



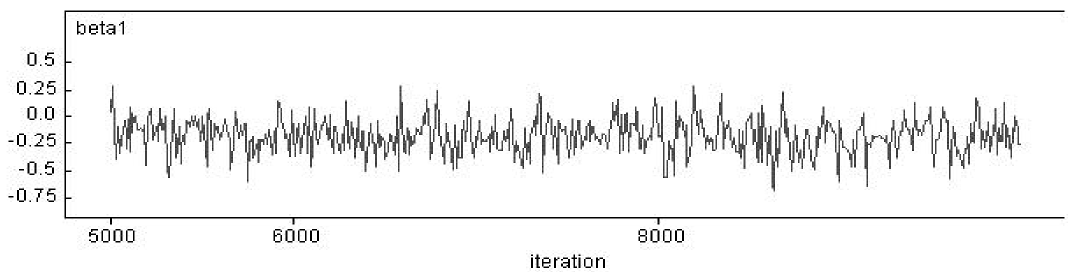



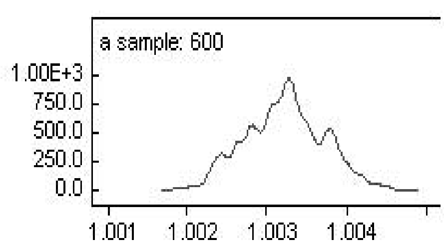


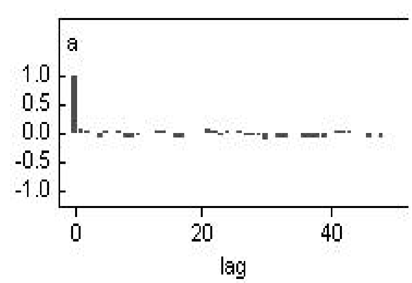



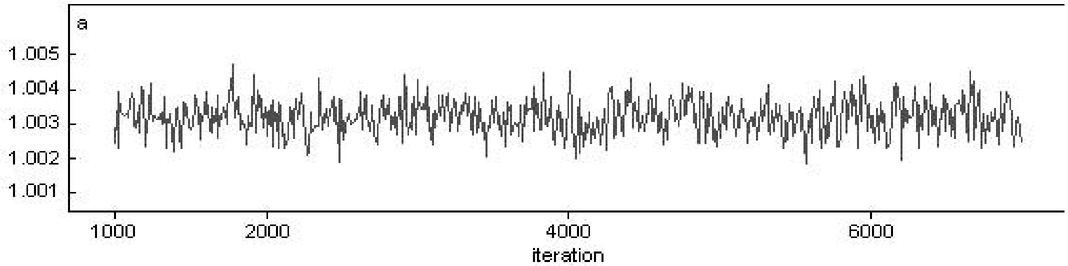



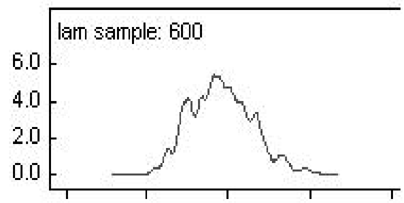



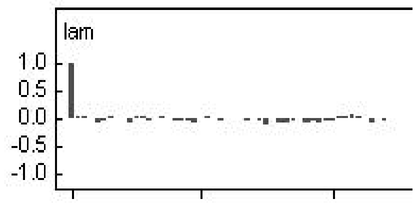

## The BGP model with Exponential distribution

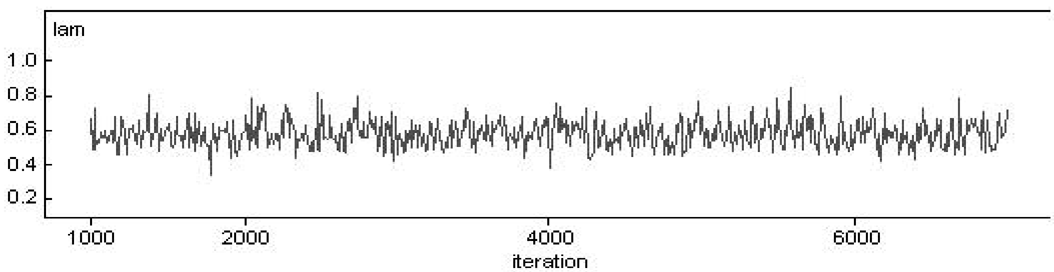



a
1.005
1.004
1.003
1.002
1.001
1000
2000
4000
6000
iteration


lam sample: 600
6.0
4.0
2.0
0.0


lam
1.0
0.5
0.0
-0.5
-1.0


lam
1.0
0.8
0.6
0.4
0.2
1000
2000
4000
6000
iteration

# Appendix D

# Second order derivatives required in NR method using BGP model

1. Exponential distribution

$$\frac{\partial\ell^2(\boldsymbol{\theta})}{\partial a^2}=-\sum_{i=1}^{n}w_i\frac{1}{a^2}(i-1)(i-2)(\lambda a^{i-1})+\sum_{i=1}^{n}(1-w_i)(i-1)$$
$$\Big\{(1-e^{-(\lambda a^{i-1})})\frac{1}{a}(\lambda a^{i-1})e^{-(\lambda a^{i-1})}\{\frac{1}{a}(i-2)-\frac{1}{a}(i-1)(\lambda a^{i-1})\}$$
$$-e^{-(\lambda a^{i-1})}e^{-(\lambda a^{i-1})}\frac{1}{a}(i-1)(\lambda a^{i-1})\Big\}[1-e^{-(\lambda a^{i-1})}]^{-2}$$
$$=-\sum_{i=1}^{n}w_i\frac{1}{a^2}(i-1)(i-2)(\lambda a^{i-1})+\sum_{i=1}^{n}(1-w_i)\frac{1}{a^2}(i-1)(\lambda a^{i-1})e^{-(\lambda a^{i-1})}$$
$$\Big\{(i-2)-(i-1)(\lambda a^{i-1})-(i-2)e^{-(\lambda a^{i-1})}\Big\}[1-e^{-(\lambda a^{i-1})}]^{-2}$$

$$\frac{\partial\ell^2(\boldsymbol{\theta})}{\partial\lambda^2}=\sum_{i=1}^{n}(1-w_i)\{[1-e^{-(\lambda a^{i-1})}]\frac{1}{\lambda}(\lambda a^{i-1})e^{-(\lambda a^{i-1})}$$
$$[-\frac{1}{\lambda}(\lambda a^{i-1})]-\frac{1}{\lambda}(\lambda a^{i-1})e^{-(\lambda a^{i-1})}e^{-(\lambda a^{i-1})}\frac{1}{\lambda}(\lambda a^{i-1})\}[1-e^{-(\lambda a^{i-1})}]^{-2}$$
$$=\sum_{i=1}^{n}(1-w_i)\frac{1}{\lambda^2}(\lambda a^{i-1})e^{-(\lambda a^{i-1})}-(\lambda a^{i-1})[1-e^{-(\lambda a^{i-1})}]^{-2}$$

$$\frac{\partial\ell^2(\boldsymbol{\theta})}{\partial a\partial\lambda}=-\sum_{i=1}^{n}w_i\frac{1}{\lambda a}(i-1)(\lambda a^{i-1})+\sum_{i=1}^{n}(1-w_i)\frac{1}{a}(i-1)$$
$$\{[1-e^{-(\lambda a^{i-1})}](\lambda a^{i-1})e^{-(\lambda a^{i-1})}[\frac{1}{\lambda}-\frac{1}{\lambda}(\lambda a^{i-1})]-(\lambda a^{i-1})e^{-(\lambda a^{i-1})}e^{-(\lambda a^{i-1})}\frac{1}{\lambda}(\lambda a^{i-1})\}$$
$$[1-e^{-(\lambda a^{i-1})}]^{-2}$$
$$=-\sum_{i=1}^{n}w_i\frac{1}{\lambda a}(i-1)(\lambda a^{i-1})+\sum_{i=1}^{n}(1-w_i)\frac{1}{a\lambda}(i-1)(\lambda a^{i-1})e^{-(\lambda a^{i-1})}$$
$$[1-(\lambda a^{i-1})-e^{-(\lambda a^{i-1})}][1-e^{-(\lambda a^{i-1})}]^{-2}$$

## 2. Weibull distribution

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial a^2}=-\sum_{i=1}^{n} w_i \frac{1}{a^2}[(i-1)\alpha][(i-1)\alpha-1](\lambda a^{i-1})^\alpha+\sum_{i=1}^{n}(1-w_i)\frac{1}{a^2}[(i-1)\alpha](\lambda a^{i-1})^\alpha e^{-(\lambda a^{i-1})^\alpha}$$
$$\Big\{(i-1)\alpha-[(i-1)\alpha](\lambda a^{i-1})^\alpha-(i-1)\alpha e^{-(\lambda a^{i-1})^\alpha}-K(a,\alpha,\lambda)\Big\}[K(a,\alpha,\lambda)]^{-2}$$
$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial \alpha^2}=-\sum_{i=1}^{n} w_i(\lambda a^{i-1})^\alpha[\ln(\lambda a^{i-1})]^2+\sum_{i=1}^{n}(1-w_i)(\lambda a^{i-1})^\alpha e^{-(\lambda a^{i-1})^\alpha}[\ln(\lambda a^{i-1})]^2$$
$$[K(a,\alpha,\lambda)-(\lambda a^{i-1})^\alpha][K(a,\alpha,\lambda)]^{-2}$$
$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial \lambda^2}=-\sum_{i=1}^{n} w_i \frac{1}{\lambda^2}\alpha(\alpha-1)(\lambda a^{i-1})^\alpha+\sum_{i=1}^{n}(1-w_i)\frac{\alpha}{\lambda^2}(\lambda a^{i-1})^\alpha e^{-(\lambda a^{i-1})^\alpha}$$
$$[\alpha-\alpha(\lambda a^{i-1})^\alpha-\alpha e^{-(\lambda a^{i-1})^\alpha}-K(a,\alpha,\lambda)][K(a,\alpha,\lambda)]^{-2}$$
$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial a\partial \alpha}=-\sum_{i=1}^{n} w_i \frac{1}{a}(i-1)[(\lambda a^{i-1})^\alpha+\alpha(\lambda a^{i-1})^\alpha\ln(\lambda a^{i-1})]+\sum_{i=1}^{n}(1-w_i)\frac{1}{a}(i-1)e^{-(\lambda a^{i-1})^\alpha}(\lambda a^{i-1})^\alpha$$
$$[K(a,\alpha,\lambda)+\alpha\ln(\lambda a^{i-1})-\alpha(\lambda a^{i-1})^\alpha\ln(\lambda a^{i-1})-\alpha\ln(\lambda a^{i-1})e^{-(\lambda a^{i-1})^\alpha}][K(a,\alpha,\lambda)]^{-2}$$
$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial a\partial \lambda}=-\sum_{i=1}^{n} w_i \frac{\alpha^2}{\lambda a}(i-1)(\lambda a^{i-1})^\alpha+\sum_{i=1}^{n}(1-w_i)\frac{\alpha^2}{a\lambda}(i-1)(\lambda a^{i-1})^\alpha e^{-(\lambda a^{i-1})^\alpha}$$
$$[K(a,\alpha,\lambda)-(\lambda a^{i-1})^\alpha][K(a,\alpha,\lambda)]^{-2}$$
$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial \alpha\partial \lambda}=-\sum_{i=1}^{n} w_i \frac{1}{\lambda}(\lambda a^{i-1})^\alpha[1+\alpha\ln(\lambda a^{i-1})]+\sum_{i=1}^{n}(1-w_i)\frac{1}{\lambda}e^{-(\lambda a^{i-1})^\alpha}(\lambda a^{i-1})^\alpha$$
$$[K(a,\alpha,\lambda)+\alpha\ln(\lambda a^{i-1})-\alpha\ln(\lambda a^{i-1})(\lambda a^{i-1})-\alpha\ln(\lambda a^{i-1})e^{-(\lambda a^{i-1})^\alpha}][K(a,\alpha,\lambda)]^{-2}$$

where $K(a,\alpha,\lambda)=1-e^{-(\lambda a^{i-1})^\alpha}$

## 3. Gamma distribution

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial a^2}=\sum_{i=1}^{n} w_i\left\{\frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}\left[\frac{i\alpha-\alpha-1}{a}-\lambda(i-1)a^{i-2}\right]}{\int_{a^{i-1}}^{\infty}\ V(\alpha,\lambda,t)\ dt}-\left[\frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}}{\int_{a^{i-1}}^{\infty}\ V(\alpha,\lambda,t)\ dt}\right]^2\right\}$$
$$+\sum_{i=1}^{n}(1-w_i)\left\{\frac{\frac{\lambda^r}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}\left[\frac{i\alpha-\alpha-1}{a}-\lambda(i-1)a^{i-2}\right]}{\int_{0}^{a^{i-1}}\ V(\alpha,\lambda,t)\ dt}-\right.$$
$$\left.\left[\frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}}{\int_{0}^{a^{i-1}}\ V(\alpha,\lambda,t)\ dt}\right]^2\right\}$$

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial \alpha^2}=\sum_{i=1}^{n} w_i \left\{ \frac{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left[\left(\ln\lambda+\ln t-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)^2-\frac{\Gamma(\alpha)\Gamma''(\alpha)-\Gamma'(\alpha)^2}{\Gamma(\alpha)^2}\right] dt}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt} - \right.$$

$$\left. \left[\frac{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\ln\lambda+\ln t-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right) dt}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt}\right]^2 \right\} +$$

$$\sum_{i=1}^{n}(1-w_i) \left\{ \frac{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left[\left(\ln\lambda+\ln t-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)^2-\frac{\Gamma(\alpha)\Gamma''(\alpha)-\Gamma'(\alpha)^2}{\Gamma(\alpha)^2}\right] dt}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt} - \right.$$

$$\left. \left[\frac{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left(\ln\lambda+\ln t-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right) dt}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt}\right]^2 \right\}$$

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial \lambda^2}=\sum_{i=1}^{n} w_i \left\{ \frac{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left[\left(\frac{\alpha}{\lambda}-t\right)^2-\frac{\alpha}{\lambda}\right] dt}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt} - \left[\frac{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\frac{\alpha}{\lambda}-t\right) dt}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt}\right]^2 \right\} +$$

$$\sum_{i=1}^{n}(1-w_i) \left\{ \frac{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left[\left(\frac{\alpha}{\lambda}-t\right)^2-\frac{\alpha}{\lambda}\right] dt}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt} - \left[\frac{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left(\frac{\alpha}{\lambda}-t\right) dt}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt}\right]^2 \right\}$$

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial a \partial \alpha}=\sum_{i=1}^{n} w_i \left\{ \frac{\displaystyle\frac{\lambda^\alpha}{\Gamma(\alpha)}\;(i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}\lambda^\alpha\left(\ln\lambda+(i-1)\ln a-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt} - \right.$$

$$\left. \left[\frac{\displaystyle\frac{\lambda^\alpha}{\Gamma(\alpha)}\;(i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt}\right]\left[\frac{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\ln\lambda+\ln t-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right) dt}{\displaystyle\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\; dt}\right] \right\} +$$

$$\sum_{i=1}^{n}(1-w_i) \left\{ \frac{\displaystyle\frac{\lambda^\alpha}{\Gamma(\alpha)}\;(i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}\lambda^\alpha\left(\ln\lambda+(i-1)\ln a-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt} - \right.$$

$$\left. \left[\frac{\displaystyle\frac{\lambda^\alpha}{\Gamma(\alpha)}\;(i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt}\right]\left[\frac{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left(\ln\lambda+\ln t-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right) dt}{\displaystyle\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\; dt}\right] \right\}$$

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial a \partial \lambda} = \sum_{i=1}^{n} w_i \left\{ \frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}\lambda^\alpha \left(\frac{\alpha}{\lambda} - a^{i-1}\right)}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt} - \right.$$

$$\left. \left[\frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt}\right] \left[\frac{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\frac{\alpha}{\lambda} - t\right)\ dt}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt}\right] \right\} +$$

$$\sum_{i=1}^{n}(1-w_i) \left\{ \frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}\lambda^\alpha \left(\frac{\alpha}{\lambda} - a^{i-1}\right)}{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\ dt} - \right.$$

$$\left. \left[\frac{\frac{\lambda^\alpha}{\Gamma(\alpha)}\ (i-1)a^{i\alpha-\alpha-1}e^{-\lambda a^{i-1}}}{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\ dt}\right] \left[\frac{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left(\frac{\alpha}{\lambda} - t\right)\ dt}{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\ dt}\right] \right\}$$

$$\frac{\partial \ell^2(\boldsymbol{\theta})}{\partial \alpha \partial \lambda} = \sum_{i=1}^{n} w_i \left\{ \frac{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left[\frac{1}{\lambda} + \left(\frac{\alpha}{\lambda} - t\right)\left(\ln\lambda + \ln t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)\right]\ dt}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt} - \right.$$

$$\left. \left[\frac{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\ln\lambda + \ln t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)\ dt}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt}\right] \left[\frac{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\frac{\alpha}{\lambda} - t\right)\ dt}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt}\right] \right\} +$$

$$\sum_{i=1}^{n}(1-w_i) \left\{ \frac{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left[\frac{1}{\lambda} + \left(\frac{\alpha}{\lambda} - t\right)\left(\ln\lambda + \ln t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)\right]\ dt}{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\ dt} - \right.$$

$$\left. \left[\frac{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\left(\ln\lambda + \ln t - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)\ dt}{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\ dt}\right] \left[\frac{\int_{a^{i-1}}^{\infty} V(\alpha,\lambda,t)\left(\frac{\alpha}{\lambda} - t\right)\ dt}{\int_{0}^{a^{i-1}} V(\alpha,\lambda,t)\ dt}\right] \right\}$$

where $V(\alpha,\lambda,t) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\ t^{\alpha-1}e^{-\lambda t}$

# Appendix E

# Data set

Table E.1: Number of daily SARS cases in Hong Kong

| *Date* | $^oC$ | *Cases* | *Date* | $^oC$ | *Cases* | *Date* | $^oC$ | *Cases* | *Date* | $^oC$ | *Cases* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 12-Mar | 17.8 | 10 | 4-Apr | 23.7 | 27 | 27-Apr | 25 | 16 | 20-May | 26.9 | 4 |
| 13-Mar | 19.1 | 14 | 5-Apr | 20.2 | 39 | 28-Apr | 25.3 | 14 | 21-May | 27.2 | 1 |
| 14-Mar | 18.2 | 5 | 6-Apr | 19.4 | 42 | 29-Apr | 26.1 | 15 | 22-May | 27.5 | 3 |
| 15-Mar | 20.2 | 8 | 7-Apr | 21.1 | 41 | 30-Apr | 25.5 | 17 | 23-May | 28 | 1 |
| 16-Mar | 22.1 | 5 | 8-Apr | 21.6 | 45 | 1-May | 23.5 | 11 | 24-May | 28.5 | 1 |
| 17-Mar | 23.2 | 53 | 9-Apr | 19.1 | 42 | 2-May | 24.1 | 11 | 25-May | 27.5 | 1 |
| 18-Mar | 20.3 | 28 | 10-Apr | 18.8 | 28 | 3-May | 24.6 | 10 | 26-May | 27.1 | 1 |
| 19-Mar | 16.9 | 27 | 11-Apr | 21.8 | 61 | 4-May | 25.5 | 8 | 27-May | 26.4 | 2 |
| 20-Mar | 15.5 | 23 | 12-Apr | 25.8 | 49 | 5-May | 24.9 | 8 | 28-May | 25.8 | 2 |
| 21-Mar | 16.8 | 30 | 13-Apr | 26.6 | 42 | 6-May | 28.1 | 9 | 29-May | 26.4 | 2 |
| 22-Mar | 17.7 | 19 | 14-Apr | 23.3 | 40 | 7-May | 28.5 | 8 | 30-May | 28.1 | 4 |
| 23-Mar | 17.9 | 20 | 15-Apr | 21.4 | 42 | 8-May | 28.7 | 7 | 31-May | 28.1 | 3 |
| 24-Mar | 18.9 | 18 | 16-Apr | 21.6 | 36 | 9-May | 25.5 | 6 | 1-Jun | 27.3 | 3 |
| 25-Mar | 20.2 | 26 | 17-Apr | 23.5 | 29 | 10-May | 25.4 | 7 | 2-Jun | 27.5 | 4 |
| 26-Mar | 21 | 30 | 18-Apr | 24.9 | 30 | 11-May | 25.9 | 4 | 3-Jun | 27.4 | 1 |
| 27-Mar | 23.3 | 51 | 19-Apr | 25.4 | 31 | 12-May | 26.3 | 5 | 4-Jun | 27.6 | 0.5 |
| 28-Mar | 20.6 | 58 | 20-Apr | 25.4 | 22 | 13-May | 27.2 | 6 | 5-Jun | 27.7 | 0.5 |
| 29-Mar | 19.9 | 45 | 21-Apr | 25 | 22 | 14-May | 28 | 9 | 6-Jun | 27.8 | 2 |
| 30-Mar | 20.9 | 60 | 22-Apr | 24.5 | 32 | 15-May | 28.6 | 5 | 7-Jun | 27.2 | 1 |
| 31-Mar | 24 | 80 | 23-Apr | 25.8 | 24 | 16-May | 29.1 | 3 | 8-Jun | 27 | 1 |
| 1-Apr | 25.5 | 75 | 24-Apr | 26.5 | 30 | 17-May | 29.2 | 4 | 9-Jun | 25.7 | 1 |
| 2-Apr | 25.8 | 23 | 25-Apr | 27 | 22 | 18-May | 26.7 | 3 | 10-Jun | 25.9 | 1 |
| 3-Apr | 26.2 | 26 | 26-Apr | 26.3 | 17 | 19-May | 25.4 | 1 | 11-Jun | 25.7 | 1 |

Table E.2: Inter-arrival time of unscheduled maintanence actions

| *Case number* | *Inter-arrival time* | *Case number* | *Inter-arrival time* | *Case number* | *Inter-arrival time* |
|---|---|---|---|---|---|
| 1 | 1382 | 25 | 177 | 49 | 130 |
| 2 | 1608 | 26 | 11 | 50 | 323 |
| 3 | 1134 | 27 | 299 | 51 | 1 |
| 4 | 2703 | 28 | 94 | 52 | 34 |
| 5 | 645 | 29 | 532 | 53 | 22 |
| 6 | 95 | 30 | 4 | 54 | 155 |
| 7 | 1278 | 31 | 248 | 55 | 101 |
| 8 | 605 | 32 | 1 | 56 | 202 |
| 9 | 344 | 33 | 68 | 57 | 156 |
| 10 | 1054 | 34 | 13 | 58 | 186 |
| 11 | 1145 | 35 | 65 | 59 | 35 |
| 12 | 307 | 36 | 5 | 60 | 248 |
| 13 | 3113 | 37 | 142 | 61 | 17 |
| 14 | 1084 | 38 | 55 | 62 | 31 |
| 15 | 855 | 39 | 30 | 63 | 184 |
| 16 | 280 | 40 | 62 | 64 | 280 |
| 17 | 490 | 41 | 65 | 65 | 714 |
| 18 | 945 | 42 | 5 | 66 | 10 |
| 19 | 105 | 43 | 2 | 67 | 38 |
| 20 | 127 | 44 | 66 | 68 | 220 |
| 21 | 61 | 45 | 42 | 69 | 132 |
| 22 | 326 | 46 | 13 | 70 | 100 |
| 23 | 254 | 47 | 3 | 71 | 18 |
| 24 | 4 | 48 | 235 | | |

Table E.3: Imports of goods index in Hong Kong

| *Year* | *month* | *Index* | *Year* | *month* | *Index* | *Year* | *month* | *Index* | *Year* | *month* | *Index* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1982 | Jan | 7.3 | | May | 16.2 | | Sep | 39 | 1995 | Jan | 76 |
| | Feb | 8.1 | | Jun | 16.5 | | Oct | 45.7 | | Feb | 72.2 |
| | Mar | 8.8 | | Jul | 17.5 | | Nov | 46.4 | | Mar | 88 |
| | Apr | 8.9 | | Aug | 17 | | Dec | 43 | | Apr | 88.8 |
| | May | 8.2 | | Sep | 17 | 1991 | Jan | 43.9 | | May | 96.6 |
| | Jun | 8.5 | | Oct | 20.3 | | Feb | 32.6 | | Jun | 93.3 |
| | Jul | 9 | | Nov | 18.9 | | Mar | 41.3 | | Jul | 96.9 |
| | Aug | 8.4 | | Dec | 19.9 | | Apr | 47.6 | | Aug | 95 |
| | Sep | 8.6 | 1987 | Jan | 17 | | May | 48.6 | | Sep | 94.8 |
| | Oct | 8.8 | | Feb | 17.5 | | Jun | 46 | | Oct | 95 |
| | Nov | 8.8 | | Mar | 22.2 | | Jul | 50 | | Nov | 90.4 |
| | Dec | 10.2 | | Apr | 22.1 | | Aug | 48.3 | | Dec | 92.4 |
| 1983 | Jan | 8.4 | | May | 22.6 | | Sep | 48.5 | 1996 | Jan | 94.3 |
| | Feb | 7.4 | | Jun | 23.3 | | Oct | 50.1 | | Feb | 69.5 |
| | Mar | 9 | | Jul | 24.1 | | Nov | 51.2 | | Mar | 87.4 |
| | Apr | 9.8 | | Aug | 23.1 | | Dec | 55.7 | | Apr | 96.3 |
| | May | 10.1 | | Sep | 23.7 | 1992 | Jan | 40.5 | | May | 94 |
| | Jun | 10.5 | | Oct | 25.1 | | Feb | 49.3 | | Jun | 89.6 |
| | Jul | 11 | | Nov | 24.8 | | Mar | 53.5 | | Jul | 100.3 |
| | Aug | 10.9 | | Dec | 28.1 | | Apr | 56.9 | | Aug | 95.6 |
| | Sep | 11.2 | 1988 | Jan | 24 | | May | 59 | | Sep | 94.6 |
| | Oct | 12.8 | | Feb | 20.6 | | Jun | 61.2 | | Oct | 100 |
| | Nov | 12.3 | | Mar | 27 | | Jul | 61.5 | | Nov | 94.9 |
| | Dec | 13.7 | | Apr | 29 | | Aug | 57.4 | | Dec | 95.1 |
| 1984 | Jan | 11.5 | | May | 28.6 | | Sep | 60 | 1997 | Jan | 95.4 |
| | Feb | 10.7 | | Jun | 31.3 | | Oct | 65.2 | | Feb | 74 |
| | Mar | 13.1 | | Jul | 32.1 | | Nov | 61.6 | | Mar | 92.7 |
| | Apr | 13.4 | | Aug | 31.7 | | Dec | 65.3 | | Apr | 100.7 |
| | May | 13.7 | | Sep | 32.4 | 1993 | Jan | 45.9 | | May | 97.4 |
| | Jun | 14 | | Oct | 33.1 | | Feb | 60 | | Jun | 95.3 |
| | Jul | 14.3 | | Nov | 34.9 | | Mar | 62.9 | | Jul | 109.3 |
| | Aug | 13.8 | | Dec | 36.3 | | Apr | 67.8 | | Aug | 101.2 |
| | Sep | 13.5 | 1989 | Jan | 31.2 | | May | 65.7 | | Sep | 96.4 |
| | Oct | 14.2 | | Feb | 27.8 | | Jun | 64.6 | | Oct | 110.7 |
| | Nov | 14.3 | | Mar | 35 | | Jul | 71.4 | | Nov | 96.3 |
| | Dec | 15.1 | | Apr | 34.7 | | Aug | 67.3 | | Dec | 99.8 |
| 1985 | Jan | 13.8 | | May | 35.2 | | Sep | 66.6 | 1998 | Jan | 81.9 |
| | Feb | 10.8 | | Jun | 36.2 | | Oct | 67.7 | | Feb | 77.3 |
| | Mar | 14.4 | | Jul | 35.3 | | Nov | 66.9 | | Mar | 89.5 |
| | Apr | 14.3 | | Aug | 34.7 | | Dec | 69.5 | | Apr | 96.4 |
| | May | 13.9 | | Sep | 34 | 1994 | Jan | 58.2 | | May | 90.1 |
| | Jun | 14.3 | | Oct | 35.3 | | Feb | 57.4 | | Jun | 89.2 |
| | Jul | 13.7 | | Nov | 33.5 | | Mar | 69.7 | | Jul | 89.1 |
| | Aug | 13.8 | | Dec | 34.3 | | Apr | 73.8 | | Aug | 86.8 |
| | Sep | 13.2 | 1990 | Jan | 29.7 | | May | 75.5 | | Sep | 83.7 |
| | Oct | 14.5 | | Feb | 29.3 | | Jun | 80.2 | | Oct | 85.9 |
| | Nov | 14.7 | | Mar | 37 | | Jul | 79.1 | | Nov | 82.4 |
| | Dec | 16.1 | | Apr | 36.7 | | Aug | 79.4 | | Dec | 82.7 |
| 1986 | Jan | 13.6 | | May | 40.9 | | Sep | 82.2 | 1999 | Jan | 77.8 |
| | Feb | 12.3 | | Jun | 37.7 | | Oct | 84.1 | | Feb | 58.1 |
| | Mar | 14.4 | | Jul | 39.4 | | Nov | 84.7 | | Mar | 78.4 |
| | Apr | 16.3 | | Aug | 40.3 | | Dec | 80.8 | | Apr | 82 |

Table E.4: Imports of goods index in Hong Kong (con't)

| *Year* | *month* | *Index* | *Year* | *month* | *Index* |
|---|---|---|---|---|---|
| | May | 78.8 | | Sep | 117.4 |
| | Jun | 83.7 | | Oct | 122.3 |
| | Jul | 92.4 | | Nov | 117.1 |
| | Aug | 88.2 | | Dec | 121.7 |
| | Sep | 90.3 | 2004 | Jan | 97.3 |
| | Oct | 91.8 | | Feb | 111.9 |
| | Nov | 90.5 | | Mar | 128.8 |
| | Dec | 96.7 | | Apr | 127.4 |
| 2000 | Jan | 90.6 | | | |
| | Feb | 73.9 | | | |
| | Mar | 98.7 | | | |
| | Apr | 96.7 | | | |
| | May | 101.8 | | | |
| | Jun | 95.9 | | | |
| | Jul | 103.2 | | | |
| | Aug | 108.9 | | | |
| | Sep | 113.2 | | | |
| | Oct | 115 | | | |
| | Nov | 101 | | | |
| | Dec | 102.7 | | | |
| 2001 | Jan | 87 | | | |
| | Feb | 88.6 | | | |
| | Mar | 97.2 | | | |
| | Apr | 96.4 | | | |
| | May | 98.6 | | | |
| | Jun | 89.6 | | | |
| | Jul | 102.2 | | | |
| | Aug | 98.7 | | | |
| | Sep | 103.1 | | | |
| | Oct | 97.9 | | | |
| | Nov | 89.3 | | | |
| | Dec | 88.8 | | | |
| 2002 | Jan | 79.5 | | | |
| | Feb | 71 | | | |
| | Mar | 97.4 | | | |
| | Apr | 94.4 | | | |
| | May | 96.5 | | | |
| | Jun | 95 | | | |
| | Jul | 109.6 | | | |
| | Aug | 104.4 | | | |
| | Sep | 109.4 | | | |
| | Oct | 111 | | | |
| | Nov | 102.9 | | | |
| | Dec | 103.2 | | | |
| 2003 | Jan | 96.4 | | | |
| | Feb | 83.9 | | | |
| | Mar | 111.6 | | | |
| | Apr | 102.3 | | | |
| | May | 105.5 | | | |
| | Jun | 105.9 | | | |
| | Jul | 115.7 | | | |
| | Aug | 110.7 | | | |

Table E.5: Coal mining disasters

| Year | Quarter | Indicator | Year | Quarter | Indicator | Year | Quarter | Indicator | Year | Quarter | Indicator |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1851 | 1 | 1 | 1864 | 1 | 0 | 1877 | 1 | 1 | 1890 | 1 | 1 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 1 | | 3 | 0 | | 3 | 1 | | 3 | 0 |
| | 4 | 1 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1852 | 1 | 0 | 1865 | 1 | 0 | 1878 | 1 | 1 | 1891 | 1 | 0 |
| | 2 | 1 | | 2 | 1 | | 2 | 1 | | 2 | 1 |
| | 3 | 0 | | 3 | 0 | | 3 | 1 | | 3 | 1 |
| | 4 | 1 | | 4 | 1 | | 4 | 0 | | 4 | 0 |
| 1853 | 1 | 1 | 1866 | 1 | 1 | 1879 | 1 | 1 | 1892 | 1 | 0 |
| | 2 | 1 | | 2 | 1 | | 2 | 0 | | 2 | 0 |
| | 3 | 1 | | 3 | 0 | | 3 | 1 | | 3 | 1 |
| | 4 | 0 | | 4 | 1 | | 4 | 0 | | 4 | 0 |
| 1854 | 1 | 1 | 1867 | 1 | 0 | 1880 | 1 | 1 | 1893 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 1 | | 3 | 1 | | 3 | 1 |
| | 4 | 0 | | 4 | 1 | | 4 | 1 | | 4 | 0 |
| 1855 | 1 | 0 | 1868 | 1 | 0 | 1881 | 1 | 1 | 1894 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 1 | | 3 | 0 | | 3 | 1 |
| | 4 | 0 | | 4 | 1 | | 4 | 1 | | 4 | 0 |
| 1856 | 1 | 0 | 1869 | 1 | 0 | 1882 | 1 | 1 | 1895 | 1 | 0 |
| | 2 | 1 | | 2 | 1 | | 2 | 1 | | 2 | 1 |
| | 3 | 1 | | 3 | 1 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 1 | | 4 | 1 | | 4 | 0 |
| 1857 | 1 | 1 | 1870 | 1 | 1 | 1883 | 1 | 0 | 1896 | 1 | 1 |
| | 2 | 1 | | 2 | 0 | | 2 | 0 | | 2 | 1 |
| | 3 | 1 | | 3 | 1 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 1 | | 4 | 0 |
| 1858 | 1 | 1 | 1871 | 1 | 1 | 1884 | 1 | 1 | 1897 | 1 | 0 |
| | 2 | 1 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 1 | | 3 | 0 | | 3 | 0 |
| | 4 | 1 | | 4 | 1 | | 4 | 1 | | 4 | 0 |
| 1859 | 1 | 0 | 1872 | 1 | 1 | 1885 | 1 | 1 | 1898 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 1 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 1 | | 4 | 1 | | 4 | 0 |
| 1860 | 1 | 1 | 1873 | 1 | 1 | 1886 | 1 | 0 | 1899 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 1 | | 3 | 0 | | 3 | 1 | | 3 | 1 |
| | 4 | 1 | | 4 | 0 | | 4 | 1 | | 4 | 0 |
| 1861 | 1 | 1 | 1874 | 1 | 0 | 1887 | 1 | 1 | 1900 | 1 | 0 |
| | 2 | 0 | | 2 | 1 | | 2 | 1 | | 2 | 0 |
| | 3 | 1 | | 3 | 1 | | 3 | 0 | | 3 | 0 |
| | 4 | 1 | | 4 | 1 | | 4 | 0 | | 4 | 0 |
| 1862 | 1 | 1 | 1875 | 1 | 0 | 1888 | 1 | 0 | 1901 | 1 | 0 |
| | 2 | 0 | | 2 | 1 | | 2 | 1 | | 2 | 1 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 1 | | 4 | 1 | | 4 | 0 | | 4 | 0 |
| 1863 | 1 | 1 | 1876 | 1 | 0 | 1889 | 1 | 1 | 1902 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 1 |
| | 4 | 1 | | 4 | 1 | | 4 | 1 | | 4 | 0 |

Table E.6: Coal mining disasters (cont'd)

| Year | Quarter | Indicator | Year | Quarter | Indicator | Year | Quarter | Indicator | Year | Quarter | Indicator |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1903 | 1 | 0 | 1916 | 1 | 0 | 1929 | 1 | 0 | 1942 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 1 |
| | 3 | 0 | | 3 | 1 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 0 | | 4 | 1 |
| 1904 | 1 | 0 | 1917 | 1 | 0 | 1930 | 1 | 1 | 1943 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 1 | | 4 | 0 |
| 1905 | 1 | 1 | 1918 | 1 | 1 | 1931 | 1 | 1 | 1944 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 1 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 1 | | 4 | 0 |
| 1906 | 1 | 0 | 1919 | 1 | 0 | 1932 | 1 | 1 | 1945 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 1 | | 4 | 0 | | 4 | 1 | | 4 | 0 |
| 1907 | 1 | 0 | 1920 | 1 | 0 | 1933 | 1 | 0 | 1946 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 1 | | 4 | 1 |
| 1908 | 1 | 1 | 1921 | 1 | 0 | 1934 | 1 | 0 | 1947 | 1 | 1 |
| | 2 | 1 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 1 | | 3 | 0 | | 3 | 1 | | 3 | 1 |
| | 4 | 0 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1909 | 1 | 1 | 1922 | 1 | 0 | 1935 | 1 | 0 | 1948 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 1 | | 3 | 1 | | 3 | 0 |
| | 4 | 1 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1910 | 1 | 0 | 1923 | 1 | 0 | 1936 | 1 | 0 | 1949 | 1 | 0 |
| | 2 | 1 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 1 | | 3 | 1 | | 3 | 0 |
| | 4 | 1 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1911 | 1 | 0 | 1924 | 1 | 0 | 1937 | 1 | 0 | 1950 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 1 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1912 | 1 | 0 | 1925 | 1 | 0 | 1938 | 1 | 0 | 1951 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 1 | | 2 | 1 |
| | 3 | 1 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1913 | 1 | 0 | 1926 | 1 | 0 | 1939 | 1 | 0 | 1952 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 0 | | 3 | 0 |
| | 4 | 1 | | 4 | 0 | | 4 | 1 | | 4 | 0 |
| 1914 | 1 | 0 | 1927 | 1 | 1 | 1940 | 1 | 0 | 1953 | 1 | 0 |
| | 2 | 1 | | 2 | 0 | | 2 | 1 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 1 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 0 | | 4 | 0 |
| 1915 | 1 | 0 | 1928 | 1 | 1 | 1941 | 1 | 0 | 1954 | 1 | 0 |
| | 2 | 0 | | 2 | 0 | | 2 | 0 | | 2 | 0 |
| | 3 | 0 | | 3 | 0 | | 3 | 1 | | 3 | 0 |
| | 4 | 0 | | 4 | 0 | | 4 | 1 | | 4 | 0 |

Table E.7: Coal mining disasters (cont'd)

| Year | Quarter | Indicator |
|---|---|---|
| 1955 | 1 | 0 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 0 |
| 1956 | 1 | 0 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 0 |
| 1957 | 1 | 0 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 1 |
| 1958 | 1 | 0 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 0 |
| 1959 | 1 | 0 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 0 |
| 1960 | 1 | 0 |
| | 2 | 0 |
| | 3 | 1 |
| | 4 | 0 |
| 1961 | 1 | 0 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 0 |
| 1962 | 1 | 1 |
| | 2 | 0 |
| | 3 | 0 |
| | 4 | 0 |

# References

[1] Andrews, D.F., Herzberg, A.M., (1985). *Data.* Springer, New York, 1985.

[2] Ascher, H., Feingold, H. 'Bad-as-Old' analysis of system failure data. *Annals of Assurance Sciences* , Gordon and Breach, New York, (1969).

[3] Ascher, H., Feingold, H., (1981). *Repairable Systems Reliability.* Marcel Dekker, New York.

[4] Barndorff-Nielsen, O.e. and Cox, D.R. (1994) *Inference and asmptotics*, Chapman & Hall, London.

[5] Bernardo, J. M. and Smith, A. F. M. (1994), *Bayesian Statistics*, Wiley.

[6] Chan, S.K., (1992). Parametric statistical inference for geometric processes. M. Phil. *Thesis.* Department of Statistics, The Chinese University of Hong Kong.

[7] Chan, S.K., Lam, Y. and Leung, D.Y.P (2004a) Statistical Inference for Geometric Processes with Gamma distributions. *Comp. Stat. & Data analysis*, 47, 565-581.

[8] Chan, S.K., Yu, L.H., Lam, Y and Ho, P.K. (2004b) Modelling SARS data using threshold Geometric Process. Submitted for publication.

[9] Clayton, D. G. (1996), "Generalized linear mixed models," in *Markov chain Monte Carlo in practice*, eds. W.R. Gilks, S. Richardson and D.J. Spiegelhalter, London: Chapman & Hall, 275-302.

[10] Cohen, A.C., and Whitten, B.j. (1988), *Parameter Estimation in Reliability and Life Span Models*, Marcel dekker, New York.

[11] Cox, D.R., Lewis, P.A., (1966). *The statistical Analysis of Series of Events.* Mathuen, London.

[12] Engle, R. F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation," *Econometrica*, **50**, 987-1006.

[13] Feller, A. (1949) *Trans. Amer. Math. Soc.*, **67**, 98-119.

[14] Gelfand, A. E. and Smith, A. F. M. (1990), "Sampling-based approaches to calculating marginal densities," *Journal of American Statistical Association*, **85**, 398-409.

[15] Geman, S., and Geman, D. (1984) "Stochastic relaxation, Gibbs Distributions and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721-741.

[16] Johnson, N.L. and Kotz, S. (1970), *Continuous Univariate Distributions*, I. Houghton Mifflin, Boston.

[17] Kao, J.H.K. (1959), *A Graphical Estimation of Mixed Weibull Parameters in Life Testing Election Tubes*, Technometrics 1, 389-407

[18] Jarrett, R. G. (1979) "A Note on the Intervals Between Coal-mining Disasters", *Biometrika*, 66, 191-193.

[19] Lam, Y., (1988a). A note on the optimal replacement problem. *Adv. Appl. Prob.* 20.

[20] Lam, Y., (1988b). Geometric process and replacement problem. *Acta Math. Appl. Sinica* 4, 366-377.

[21] Lam, Y., (1992a). Optimal geometric process replacement model. *Acta Math Appl. Sinica* 8, 73-81.

[22] Lam, Y., (1992b). Nonparametric inference for geometric processes. *Commun. Statist. Theory Meth.* 21, 2083-2105.

[23] Lam, Y. (1995) "Calculating the Rate of Occurrence of Failures for Continuous-time

Markov Chains with Application to a Two-component Parallel System", *J. Oper. Res. Soc.*, 45, 528-536.

[24] Lam, Y. and Zhang, Y.L. (1996a) Analysis of a two-component series system with a geometric process model, *Naval Research Logistics*, **43**, 491-502.

[25] Lam, Y. and Zhang, Y.L. (1996b) Analysis of a parallel system with two different units, *Acta Mathematicae Applicatae Sinica*, **12**, 408-417.

[26] Lam, Y. (1997) The rate of occurrence of failures, *J. Appl. Prob.*, 34, 234-247.

[27] Lam, Y., Chan, J.S.K., (1998). Statistical inference for geometric processes with lognormal distribution. Computional Statistics & Data Analysis 27, 99-112

[28] Lam, Y. Zhang, Y.L. and Zheng, Y.H. (2002) A geometric process equivalent model for a multistate degenerative system, *European Journal of Operational Research*, **142**, 21-29.

[29] Lam, Y. and Tse, Y.K. (2003) Optimal maintanance model for a multistate deteriorating system: a geometric process approach, *International Journal of Systems Science*, **34**, No. 5: 303-308.

[30] Lam, Y. and Zhang, Y.L. (2003) A geometric-process maintenance model for a deteriorating system under a random environment, *IEEE Trans. Reliab.*, **52** No. 1: 83-89.

[31] Lam, Y., Zheng, Y.H. and Zhang, Y.L. (2003) Some Limit Theorems in Geometric Processes, *Acta Math. Appl. Sinica*, **19** No. 3: 405-416.

[32] Lam, Y. (2004) Geometric process. In *The Encyclopedia of Statistical Sciences, 2nd edition*, Balakrishnan, N., Read, C., Kotz, S. and Vidakovic, B. ed. John Wiley & Sons, Inc., New York. 2004, to appear.

[33] Lam, Y., Zhu, L.X., Chan, J.S.K. and Liu, Q. (2004) Analysis of data from a series of events by a Geometric Process model, *Acta Mathematica Applicatae Sinca, English Series*, 20, No.2, 263-282.

[34] Lee, P. M. (1989), *Bayesian Statistics: An Introduction* , Edward Arnold.

[35] Lehmann, E.L., 1983. Theory of Point Estimation. Wiley, New York.

[36] Lewis, T. and the M345 Course Team (1986) M345 *Statistical Methods*, *Unit* 2: *Basic Methods: Testing and Estimation*, *Milton Keynes:* The Open University, 16.

[37] Lieblein, J. and Zelen, M.(1956) , *Statistical Investigation of the Fatigue Life of Deep Groove Ball Bearing,* J. of Res. Natl. Bur, Std., 57, 273-315

[38] Maguire, B. A., Pearson, E. S. and Wynn, A. H. A. (1952). "The Time Intervals Between Industrial Accidents", *Biometrika*, 39, 168-180.

[39] Nelder, J. A. and Wedderburn, R.W.M. (1972), "Generalized linear models", *Journal of the Royal Statistical Society, A, 135, 370-384*

[40] Neuts, M.F., Perez-Ocon and Torres-castro, I. (2000) Repairable models with operating and repair times governed by phase type distributions, *Adv. Appl. Prob.*, **32**, 468-479.

[41] Proschan, F. (1963) "Theoretical Explanation of Observed Decreasing Failure Rate", *Technometrics*, 5, 375-383.

[42] Sheu, S.H. (1999) Extended optimal replacement model for deteriorating systems, *European Journal of Operational Research*, **112**, 503-516.

[43] Sinha, S.K. (1986), *Reliability and Life Testing*, John Wiley & Sons, New York.

[44] Stanley, A.D.J. (1993) On Geometric Processes and Repair Replacement Problems,

*Microelectron Reliab.*, **33** No. 4: 489-49.

[45] Zeger, S. L. and Karim, M. R. (1991), "Generalized linear models with random effecrs; A Gibbs sampling approach," *Journal of American Statistical Association*, **86**, 79-86.

[46] Zhang, Y.L. (1999) An optimal geometric process model for a cold standby repairable system, *Reliab. Eng. Syst. Safe*, **63** No. 1: 107-110.

[47] Zhang, Y.L., Yam, R.C.M. and Zuo, M.J. (2001) Optimal replacement policy for a deteriorating production system with preventive maintenance, *Int. J. Syst. Sci.*, **32** No. 10: 1193-1198.

[48] Zhang, Y.L. (2002) A geometric-process repair model with good-as-new preventive repair, *IEEE Trans. Reliab.*, **51**, No. 2: 223-228.

[49] Zhang, Y.L., Yam, R.C.M. and Zuo, M.J. (2002) Optimal replacement policy for a multistate repairable system, *J. Oper.Res. Soc.*, **53**, No. 3: 336-341.