

2015

# Bayesian change-point analysis in linear regression model with scale mixtures of normal distributions

Shuaimin Kang  
*Michigan Technological University*

Copyright 2015 Shuaimin Kang

---

## Recommended Citation

Kang, Shuaimin, "Bayesian change-point analysis in linear regression model with scale mixtures of normal distributions", Master's Thesis, Michigan Technological University, 2015.  
<https://digitalcommons.mtu.edu/etds/920>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Statistics and Probability Commons](#)

BAYESIAN CHANGE-POINT ANALYSIS IN LINEAR REGRESSION MODEL WITH  
SCALE MIXTURES OF NORMAL DISTRIBUTIONS

By

Shuaimin Kang

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Mathematical Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2015

© 2015 Shuaimin Kang



This thesis has been approved in partial fulfillment of the requirements for the Degree of  
MASTER OF SCIENCE in Mathematical Sciences.

Department of Mathematical Sciences

Thesis Advisor: *Dr. Min Wang*

Committee Member: *Dr. Howard Qi*

Committee Member: *Dr. Yeonwoo Rho*

Department Chair: *Mark S. Gockenbach*



# Contents

<b>List of Figures</b> . . . . .	<b>vii</b>
<b>List of Tables</b> . . . . .	<b>ix</b>
<b>Acknowledgments</b> . . . . .	<b>xi</b>
<b>Abstract</b> . . . . .	<b>xiii</b>
<b>1 Introduction</b> . . . . .	<b>1</b>
<b>2 The SMN Linear Regression Model with a Variance Change-point</b> . . . . .	<b>7</b>
2.1 The Scale Mixtures of Normal Distributions . . . . .	7
2.2 The SMN Linear Regression Model Setup . . . . .	12
<b>3 Bayesian Analysis of Variance Change-point Problems</b> . . . . .	<b>17</b>
3.1 Bayesian Formulation . . . . .	18
3.2 Inference Procedure Using MCMC . . . . .	21
3.2.1 Gibbs sampling . . . . .	22
3.2.2 Metropolis-Hastings Algorithm . . . . .	27
3.2.3 Acceptance-Rejection Algorithm . . . . .	28

<b>4</b>	<b>Simulations and Real-data Application</b>	<b>31</b>
4.1	Simulation Study	31
4.1.1	Simulation I	32
4.1.2	Simulation II	33
4.2	Coefficient Effect On The Detection Of Variance Change-point	34
4.3	Dow Jones Index With Multiple Variance Change-points	39
<b>5</b>	<b>Concluding Remarks and Future Work</b>	<b>44</b>
	<b>References</b>	<b>47</b>

# List of Figures

2.1	Pdfs of normal and $t$ distributions . . . . .	11
2.2	Pdfs of normal and slash distributions . . . . .	12
2.3	Pdfs of normal and contaminated distributions . . . . .	13
4.1	Gaussian distribution: Detected change points when $\sigma_2 = 5, 6, 7$ . . . . .	36
4.2	Student $t$ distribution: Detected change points when $\sigma_2 = 7, 8, 9$ . . . . .	37
4.3	Contaminated distribution: Detected change points when $\sigma_2 = 5, 6, 7$ . . . . .	38
4.4	Slash distribution: Detected change points when $\sigma_2 = 6, 7, 8$ . . . . .	39
4.5	Time series of $R_t$ . . . . .	43





# List of Tables

4.1	Bayesian posterior estimation of the unknown parameters with standard deviation parameters in Simulation I . . . . .	33
4.2	Bayesian posterior estimation of the unknown parameters with standard deviation parameters in Simulation II . . . . .	34
4.3	Descriptive statistics of $R_t$ . . . . .	42
4.4	Variance change-points of real data . . . . .	42



## Acknowledgments

I would like to thank my advisor Dr. Wang who borrowed me books about Bayesian Analysis, recommended me papers to read, explained to me whenever I had puzzles, and encouraged me when I was lazy. Thanks for your kindness, patience and encouragement. I also thank to Drs. Qi and Rho. Thanks for coming to my thesis dissertation and giving me valuable feedbacks.



## Abstract

In this thesis, we consider Bayesian inference on the detection of variance change-point models with scale mixtures of normal (for short SMN) distributions. This class of distributions is symmetric and thick-tailed and includes as special cases: Gaussian, Student- $t$ , contaminated normal, and slash distributions. The proposed models provide greater flexibility to analyze a lot of practical data, which often show heavy-tail and may not satisfy the normal assumption.

As to the Bayesian analysis, we specify some prior distributions for the unknown parameters in the variance change-point models with the SMN distributions. Due to the complexity of the joint posterior distribution, we propose an efficient Gibbs-type with Metropolis-Hastings sampling algorithm for posterior Bayesian inference. Thereafter, following the idea of [1], we consider the problems of the single and multiple change-point detections. The performance of the proposed procedures is illustrated and analyzed by simulation studies. A real application to the closing price data of U.S. stock market has been analyzed for illustrative purposes.



# Chapter 1

## Introduction

It has been long known that the subject of quick detection of change-points has gained considerable attention in the literature due to its importance in many applications. The change-points analysis can be originally traced to [2], who develops a test for a change in a parameter occurring at an unknown point. Thereafter, the problem of detection of changes has been greatly investigated in a wide range of disciplines, including finance, bioinformatics, climatology, econometrics, network traffic analysis, and so on. For example, to create the safest investment environment, financial investors often pay much attention on the volatility of stock market and develop efficient economic models to monitor the location of a change-point if it exists. Biologists consider the DNA copy number variations, which can be located under some proper change-points detection models for cancer research.



A change point can generally be considered as a location or time point such that the observations follow different distributions before and after that point. To be more specific, we begin with the simplest change-point problem briefly summarized as follows. For a given sample of  $n$  independent observations  $\{Y_1, Y_2, \dots, Y_n\}$ , a change point occurs if there exists a  $k \in [1, n - 1]$  such that the distributions of  $\{Y_1, \dots, Y_k\}$  and  $\{Y_{k+1}, \dots, Y_n\}$  are different with respect to some criteria. Three commonly used criteria are of particular interest:

1. Change in mean: the mean of  $Y_i$  is given by

$$\mu_i = \begin{cases} \mu_1, & \text{if } 1 \leq i \leq k, \\ \mu_2, & \text{if } k + 1 \leq i \leq n, \end{cases}$$

where  $\mu_1 \neq \mu_2$  and the discrete unknown parameter  $k$  indicates the location of the change-point in the sample.

2. Change in regression coefficients: assume that  $X_i$  and  $\varepsilon_i$  are mutually independent and identically distributed (iid) sequences with  $E[\varepsilon_i] = 0$  and  $E[\varepsilon_i^2] = 1$ . Then it follows

$$Y_i = \begin{cases} \beta_0 + \beta_1 X_i + \sigma \varepsilon_i, & \text{if } 1 \leq i \leq k, \\ \gamma_0 + \gamma_1 X_i + \sigma \varepsilon_i, & \text{if } k + 1 \leq i \leq n, \end{cases}$$

where  $\beta_0 \neq \gamma_0$  and  $\beta_1 \neq \gamma_1$ .

3. Change in variance: the variance of  $Y_i$  is given by

$$\sigma_i^2 = \begin{cases} \sigma_1^2, & \text{if } 1 \leq i \leq k, \\ \sigma_2^2, & \text{if } k + 1 \leq i \leq n, \end{cases}$$

where  $\sigma_1^2 \neq \sigma_2^2$ .

Over the years, considerable attention has been devoted to testing and estimation about the change-points problem related to the first two cases. [3] considers the problem of a change in location under several different criteria. [4] studies this problem in the regression coefficients of a linear regression model. [5] and [6] study the change-point problem in the mean of a normal distribution. Later on, [7] extends the change-point problem in generalized linear regression models. In the meantime, the change in variance while the mean or regression coefficient remains common has also been discussed in applied economics and finance. For instance, [8] explores testing and locating multiple variance change points in a sequence of independent Gaussian random variables, assuming known and common mean. It should be noted that many researchers are often interested in studying the three types of change-point problems in regression model under the assumption that the errors follow the normal distribution for mathematical convenience.

As is the case of many real data studies, the normal assumption may be questionable or not be always realistic, because such assumption is very vulnerable to the presence of atypical

observations. Substantial violation of normality assumption could potentially impact the variance change-point detection in these models. To deal with the problem of atypical observations, [9] studies robust statistical regression models with the  $t$  distributed errors. More recently, [10] considers the variance change-point problem in the Student- $t$  linear regression model, which provides heavy-tails, compared with the normal regression model.

It is also well-known that due to some unexpected reasons, real data might have more flexible tails than the Student- $t$  distribution. This observation motivates us to consider the regression models with a more flexible distributed error, which can be either heavier tails compared to the normal distributed errors and includes the normal and Student- $t$  distributions as particular cases. Fortunately, there are various heavy-tailed distributions, such as double-exponential distribution, scale mixtures of uniform distributions studied by [11]. In this thesis, we mainly focus on scale mixtures of normal (for short, SMN) distributions, which are symmetric and thick-tailed and include as special cases: Gaussian, Student- $t$ , contaminated normal, and slash distributions. To the best of our knowledge, there are no published references for the variance change-points problem for the SMN linear regression models, even though detection of change-points is of the utmost importance in statistical literature, especially when data show heavy tails. Moreover, from the Bayesian point of view, the SMN distribution admits a hierarchical representation that allows us to develop an efficient sampling scheme using standard software, such as WinBUGS, JAGS, and/or R programs. The proposed approach can be implemented to obtain Bayesian estimations and standard errors of the change-point and other unknown model parameters. By following

the stepwise and dichotomy method proposed by [1], the proposed approach can also be applied to detect multiple change-points via sampling schemes.

The remainder of this thesis is organized as follows. In Chapter 2, we introduce the SMN class of distribution and then define the SMN linear regression model with a variance change-point. In Chapter 3, we consider the prior specification for the unknown model parameters and propose an efficient sampling algorithm for posterior simulation and estimation. The performance of the proposed approach is illustrated in Chapter 4, by considering the analysis of simulation studies and a real application to the closing price data of U.S. stock market. We summarize our findings and sketch possible extensions in Chapter 5.



## Chapter 2

# The SMN Linear Regression Model with a Variance Change-point

In this chapter, we firstly introduce the class of scale mixtures of normal (for short, SMN) distributions. Then we define the SMN linear regression models with a variance change-point in Section 2.2.

### 2.1 The Scale Mixtures of Normal Distributions

In this section, we recall that a random variable  $Y$  is said to be the SMN class of distributions with location parameter  $\mu \in (-\infty, +\infty)$ , and scale parameter  $\sigma^2 > 0$  if it satisfies

the following stochastic representation

$$Y = \mu + \kappa(u)^{1/2}Z,$$

where  $\kappa(u)$  is a function of  $u$ ,  $u$  is a positive variable, and  $Z$  has a normal distribution with mean 0 and variance  $\sigma^2$ . It is noteworthy that this representation is derived based on Laplace transformation technique studied by [12]. In particular, we call the random variable  $Y$  as the standard SMN distribution if  $\mu = 0$  and  $\sigma^2 = 1$ .

The random function  $\kappa(u)$  can be either discrete or continuous and determines the distribution of  $Y$ . In this thesis, we follow the results and comments of [13], [11], and [14] and mainly focus on the case in which  $\kappa(u)$  is set to be  $1/u$ . Consequently, the possibility density function (pdf) of  $Y = \mu + u^{-1/2}Z$  is given by

$$f_Y(y|\mu, u, \sigma) = \int_0^\infty \frac{1}{\sqrt{2\pi}u^{-1/2}\sigma} e^{-\frac{(y-\mu)^2}{2u^{-1}\sigma^2}} \pi(u) du,$$

where  $\pi(\cdot)$  represents the pdf of the random variable  $u$ . The random variable  $Y$  has different distributions with different choices of  $\pi(u)$  summarized as follows.

1. Case I: if  $u = 1$ , then  $Y = \mu + Z$ . In this case,  $Y$  has a normal distribution with mean  $\mu$  and variance  $\sigma^2$ .

2. Case II: if  $uv \sim \chi^2(v)$ ,  $v$  is a positive constant, then we have that

$$(Y \mid \mu, u, v, \sigma) \sim N(\mu, u^{-1}\sigma^2) \quad \text{and} \quad uv \sim \chi^2(v).$$

In this case,  $Y$  has a Student- $t$  distribution with mean  $\mu$ , variance  $\sigma^2$ , and the degrees of freedom  $v$ . It deserves mentioning that Student- $t$  distribution is a heavy-tailed distribution and may thus provide a better fit for a lot of practical data with atypical observations than the one based on the normal assumption.

3. Case III: if  $u \sim \text{Beta}(v, 1)$ , then we have that

$$(Y \mid \mu, u, v, \sigma) \sim N(\mu, u^{-1}\sigma^2) \quad \text{and} \quad u \sim \text{Beta}(v, 1),$$

where  $\text{Beta}(v, 1)$  represents the beta distribution with the pdf given by  $\pi(\lambda \mid v) = \frac{u^{v-1}}{B(v,1)} = vu^{v-1}I_{0 < u < 1}$  with  $B(\cdot, \cdot)$  be a beta function. In this case, the pdf of  $Y$  is

$$f_Y(y \mid \mu, u, v, \sigma) = \int_0^1 \frac{1}{\sqrt{2\pi}u^{-1/2}\sigma} e^{-\frac{(y-\mu)^2}{2u^{-1}\sigma^2}} vu^{v-1} du.$$

We call  $Y$  has a slash distribution, which has a heavier tail than the normal distribution and becomes the normal distribution as  $v$  approaches  $\infty$ .

4. Case IV: if  $u = 1$  with the probability of  $1 - v$  and  $u = \gamma$  with the possibility of  $v$ , then we observe that  $Y$  is a normal mixture, such that  $Y = (1 - v)Y_1 + vY_2$ , where  $Y_1 \sim N(\mu, \sigma^2)$  and  $Y_2 \sim N(\mu, \gamma^{-1}\sigma^2)$ . In this case,  $Y$  has a contaminated

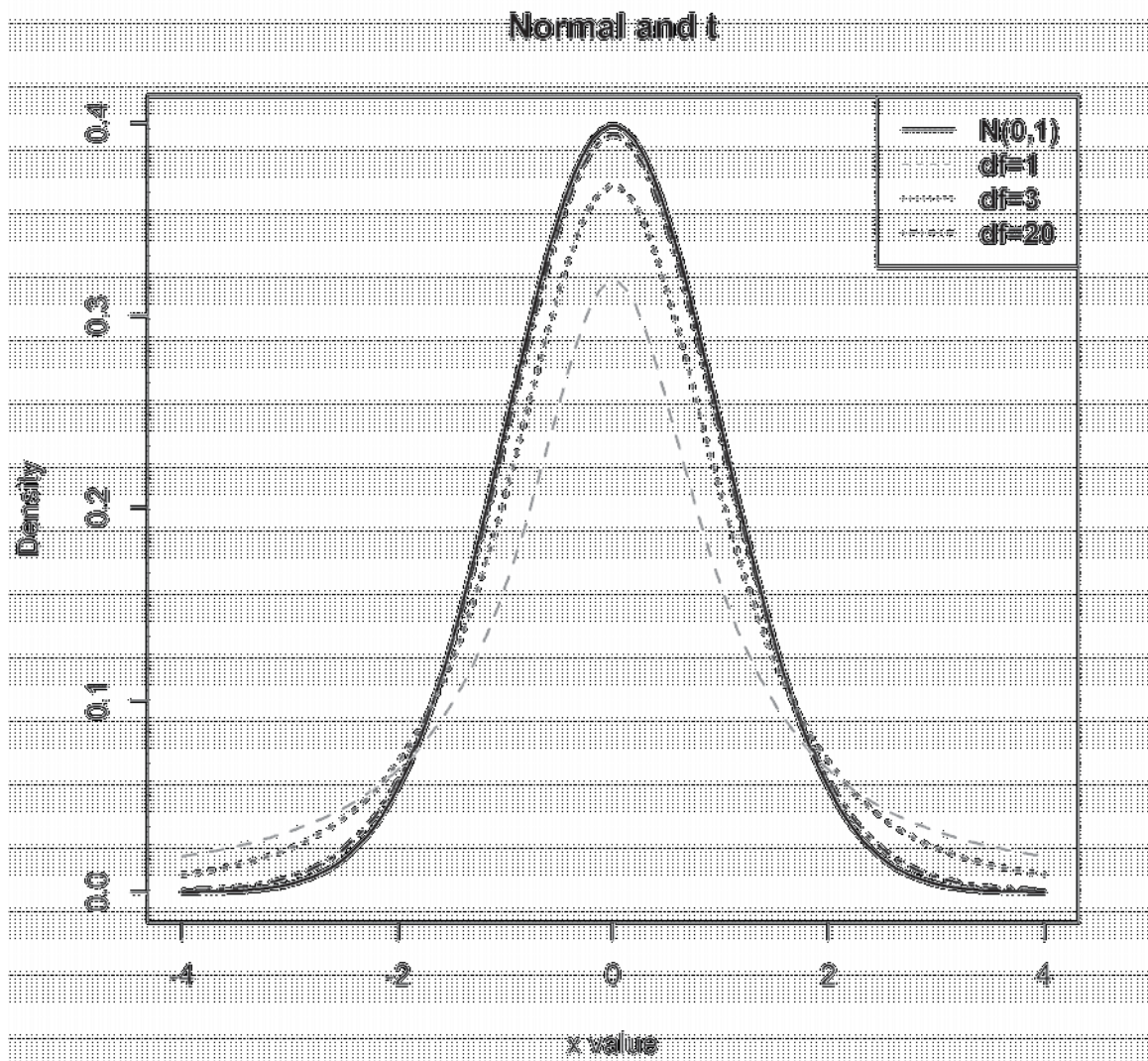


distribution with  $v \in (0, 1)$  and  $\gamma \in (0, 1)$ . Note that the contaminated distribution degenerates to a normal distribution when  $\gamma = 1$ .

Due to the flexible tail of SMN distributions, we consider the variance change-point detection in the context of regression models with SMN distributions, which should be provided to fit the data better. Moreover, from the Bayesian perspective, the above mixture representations of SMN distributions allow us to develop efficient sampling algorithms for posterior analysis.

In order to further compare heavy-tails of the four distributions mentioned above, Figure 2.3 displays their pdfs with various choices of the parameters discussed as follows. We set  $\mu = 0$  and  $\sigma^2 = 1$  for all the distributions under consideration. Then, for Student- $t$  distribution, we choose degrees of freedom  $v = \{1, 3, 20\}$ ; for the slash distribution, we set  $v = \{1, 3, 20\}$ . For the contaminated distribution, we consider three cases which are represented by con1, con2 and con3 in the Figure 2.3. In the case con1,  $v$  and  $\gamma$  have distributions of  $\text{Beta}(1, 0.5)$  and  $\text{Beta}(1, 1)$ , respectively. Similarly, in the case con2,  $v \sim \text{Beta}(2, 3)$ ,  $\gamma \sim \text{Beta}(1, 1)$  and in the case con3,  $v \sim \text{Beta}(6, 4)$ ,  $\gamma \sim \text{Beta}(4, 6)$ . It can be seen from the Figure 2.3 that Student- $t$ , slash, and contaminated distributions all have heavier tails than the normal distribution. What's more, we observe that the tails become thinner with the larger degrees of freedom for Student- $t$  and slash distributions.

In this thesis, we mainly focus on the variance change-point problem for linear regression



**Figure 2.1:** Pdfs of normal and  $t$  distributions

models with the SMN distributions mentioned above. However, it deserves mentioning that besides the above distributions, the SMN distributions include special cases, such as Pearson Type VII distribution [15], variance gamma distribution [16], stable and exponential power distributions [17], among others. The generalization of the proposed algorithm to those distributions is quite straightforward and thus omitted here for simplicity.

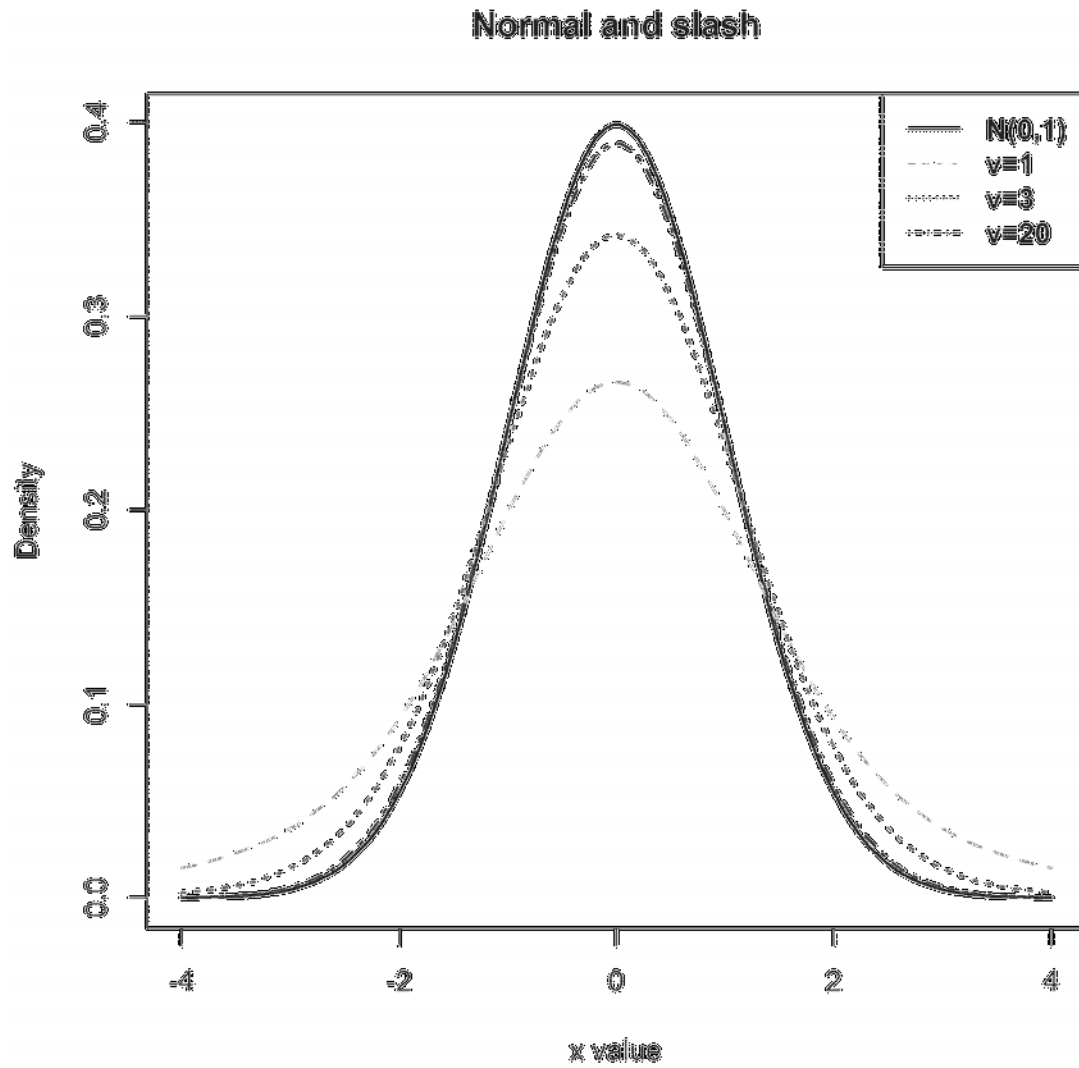


Figure 2.2: Pdfs of normal and slash distributions

## 2.2 The SMN Linear Regression Model Setup

As an illustration, suppose that there exists a change-point for a given sequence of data  $\{Y_1, Y_2, \dots, Y_n\}$ . The variance change-point detection problem can be formulated as a

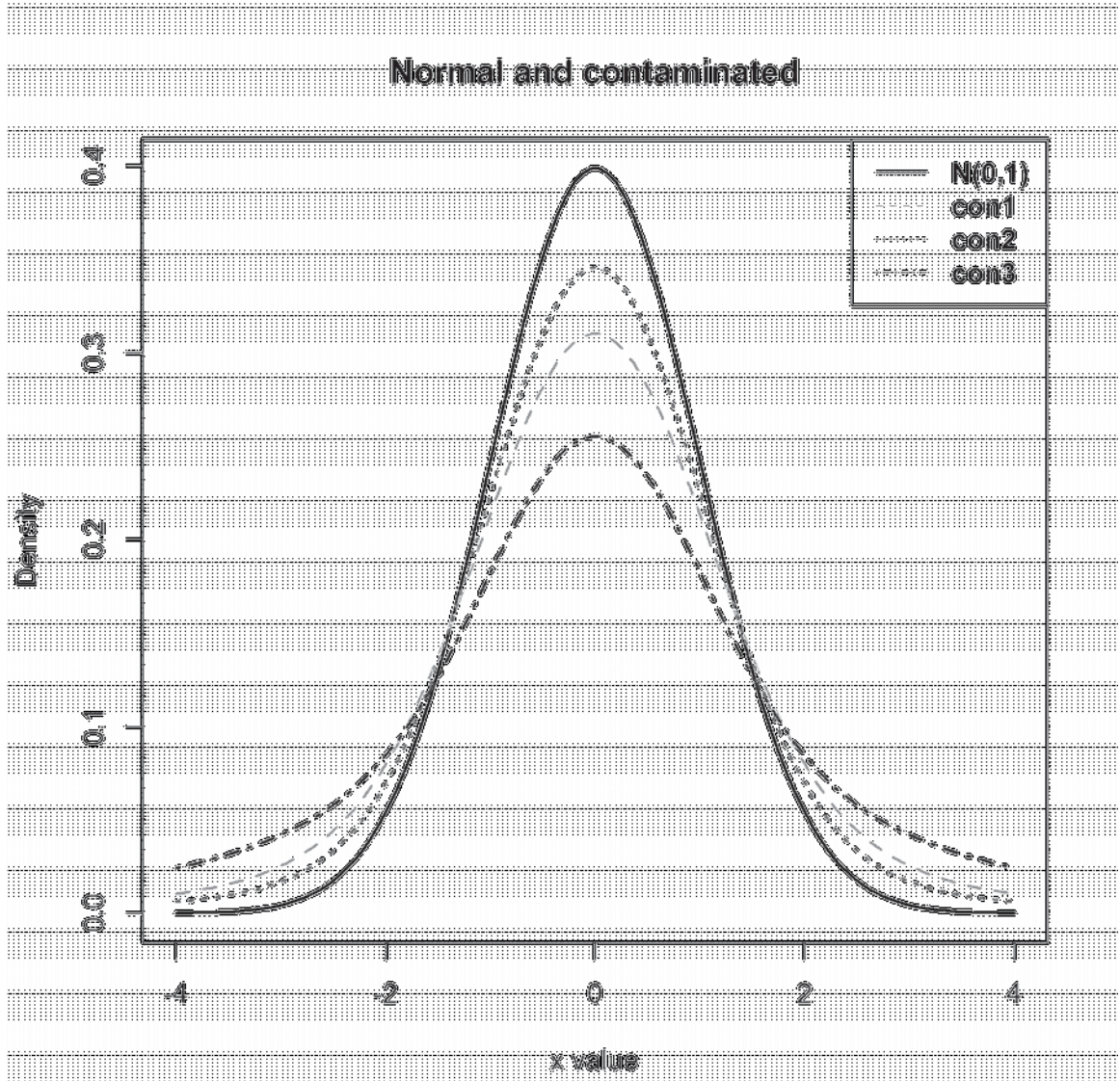


Figure 2.3: Pdfs of normal and contaminated distributions

model selection problem

$$H_0 : Y_i = X_i' \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad (2.1)$$

$$H_1 : Y_i = X_i' \beta + \varepsilon_i^{(1)}, \quad i = 1, \dots, k, \quad (2.2)$$

$$Y_i = X_i' \beta + \varepsilon_i^{(2)}, \quad i = (k + 1), \dots, n,$$

where  $k$  is the unknown position that a variance change exists,  $\beta = (\beta_1, \dots, \beta_p)'$  is a  $p \times 1$  vector of the regression coefficients,  $X_i = (x_{i1}, \dots, x_{ip})'$  denotes the  $p$  covariates, and

$$\varepsilon_i \sim N(0, u_i^{-1}\sigma^2), \quad \varepsilon_i^{(1)} \sim N(0, u_i^{-1}\sigma_1^2), \quad \varepsilon_i^{(2)} \sim N(0, u_i^{-1}\sigma_2^2). \quad (2.3)$$

For notational simplicity, we denote  $Y = (Y_1, \dots, Y_n)'$ ,  $X = (X_1, \dots, X_n)'$ , and  $u = (u_1, \dots, u_n)'$  with all  $u_i$ 's being positive. Then under  $H_1$ , the likelihood function is given by

$$\begin{aligned} L(Y | X, \beta, \sigma_1, \sigma_2, k, u) & \quad (2.4) \\ &= \prod_{i=1}^n f(Y_i | X, \beta, \sigma_1, \sigma_2, k, u) \\ &= \prod_{i=1}^k \frac{1}{\sqrt{2\pi}u_i^{-1/2}\sigma_1} \exp\left\{-\frac{(Y_i - X_i'\beta)^2}{2u_i^{-1}\sigma_1^2}\right\} \prod_{i=(k+1)}^n \frac{1}{\sqrt{2\pi}u_i^{-1/2}\sigma_2} \exp\left\{-\frac{(Y_i - X_i'\beta)^2}{2u_i^{-1}\sigma_2^2}\right\} \\ &\propto \prod_{i=1}^k \frac{u_i^{1/2}}{\sigma_1^k} \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X_i'\beta)^2 u_i}{2\sigma_1^2}\right\} \prod_{i=k+1}^n \frac{u_i^{-1/2}}{\sigma_2^{n-k}} \exp\left\{-\sum_{i=(k+1)}^n \frac{(Y_i - X_i'\beta)^2 u_i}{2\sigma_2^2}\right\}. \end{aligned}$$

We observe that the above likelihood function depends on the distribution of the random variable  $u$ . As the formation of the SMN distributions discussed in Section 2.1, we obtain the four types of linear regression models in terms of the choices of  $u$ : Gaussian model with  $u_i = 1$ , Student- $t$  model with  $u_i v \sim \chi^2(v)$ , slash model with  $u_i \sim \text{Beta}(v, 1)$ , and contaminated model with  $u_i \sim vI_\gamma(u_i) + (1-v)I_1(u_i)$ . Thus, the flexibility of the proposed models indicates that they should be provided to fit the real data better than the one with

either normal or Student- $t$  distributed errors only.



# **Chapter 3**

## **Bayesian Analysis of Variance**

### **Change-point Problems**

In this Chapter, we consider Bayesian analysis for the variance change-point problem in the SMN linear regression models. We discuss the Bayesian formulation of the variance-change problem in Section 3.1. We propose a Gibbs-type with Metropolis-Hastings algorithm for posterior simulation and parameter estimation in Section 3.2.



### 3.1 Bayesian Formulation

Bayesian analysis begins with prior specification for the unknown model parameters. Non-informative prior is often preferred in the lack of the prior knowledge. In this thesis, we specify the priors for the unknown parameters  $\{\beta, \sigma_1, \sigma_2, k\}$  as follows:

$$\pi_1(\beta) \propto \text{constant}, \quad \pi_2(\sigma_1) \propto \frac{1}{\sigma_1}, \quad \pi_3(\sigma_2) \propto \frac{1}{\sigma_2}, \quad \pi_4(k) \propto \frac{1}{n - 2p + 1},$$

With a combination of the likelihood function in (2.4), the joint posterior distribution under the above priors is given by

$$\Pi(\beta, \sigma_1, \sigma_2, k, u \mid Y, X) \propto L(Y \mid X, \beta, \sigma_1, \sigma_2, k, u) \pi_1(\beta) \pi_2(\sigma_1) \pi_3(\sigma_2) \pi_4(k) \pi_5(u),$$

where  $\pi_5(u)$  is the distribution for the random variable  $u$ , which depends on the specific model in the family of the SMN distributions. More specifically, based on the four different treatments of  $u$ , we obtain the four different linear regression models discussed as follows.

Case I: In Gaussian linear regression model with  $u_i = 1$ , for  $i = 1, \dots, n$ , the posterior

density is

$$\begin{aligned} \Pi_1(\beta, \sigma_1, \sigma_2, k|Y, X) &\propto L(Y|X, \beta, \sigma_1, \sigma_2, k)\pi_1(\beta)\pi_2(\sigma_1)\pi_3(\sigma_2)\pi_4(k) \\ &\propto \sigma_1^{-(k+1)} \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X'_i\beta)^2}{2\sigma_1^2}\right\} \sigma_2^{-(n-k+1)} \exp\left\{-\sum_{i=(k+1)}^n \frac{(Y_i - X'_i\beta)^2}{2\sigma_2^2}\right\}. \end{aligned}$$

Case II: In Student- $t$  linear regression model with  $u_i v \sim \chi^2(v)$ , for  $i = 1, \dots, n$ , we have an additional random variable  $v$ , the degrees of freedom. Due to lack of the prior knowledge, we consider the Jeffrey's prior for  $v$  [18]:

$$\pi_5(v) \propto \left(\frac{v}{v+3}\right)^{\frac{1}{2}} \left[ \psi'\left(\frac{v}{2}\right) - \psi'\left(\frac{v+1}{2}\right) - \frac{2(v+3)}{v(v+1)^2} \right]^{\frac{1}{2}},$$

where  $\psi'(\cdot) = d[\psi(x)]/dx$  and  $\psi(\cdot) = d[\log\Gamma(x)]/dx$  are trigamma and digamma functions, respectively. Thus, the corresponding posterior density is given by

$$\begin{aligned} \Pi_2(\beta, \sigma_1, \sigma_2, u, v, k|Y, X) &\propto L(Y | X, \beta, u, \sigma_1, \sigma_2, k)\pi_1(\beta)\pi(u)\pi_2(\sigma_1)\pi_3(\sigma_2)\pi_5(v)\pi_4(k) \\ &\propto \sigma_1^{-(k+1)} \prod_{i=1}^k u_i^{\frac{v+3}{2}} \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X'_i\beta)^2 \sigma_1^{-2} + v}{2u_i^{-1}}\right\} \\ &\quad \sigma_2^{-(n-k+1)} \prod_{i=k+1}^n u_i^{\frac{v+3}{2}} \exp\left\{-\sum_{i=k+1}^n \frac{(Y_i - X'_i\beta)^2 \sigma_2^{-2} + v}{2u_i^{-1}}\right\} \left(\frac{v}{2}\right)^{\frac{nv}{2}} \\ &\quad \left[\Gamma\left(\frac{v}{2}\right)\right]^{-n} \left(\frac{v}{v+3}\right)^{1/2} \left[ \psi'\left(\frac{v}{2}\right) - \psi'\left(\frac{v+1}{2}\right) - \frac{2(v+3)}{v(v+1)^2} \right]^{1/2}. \end{aligned}$$

Case III: In the slash linear regression model with  $u_i \sim \text{Beta}(v, 1)$ , for  $i = 1, \dots, n$ , we also

have an additional parameter  $v$  to represent the degrees of freedom. We consider the same prior for  $v$  as it is in Student- $t$  linear regression model. Then, the posterior density is

$$\begin{aligned}
& \Pi_3(\beta, \sigma_1, \sigma_2, u, v, k|Y, X) \\
& \propto L(Y|X, \beta, u, \sigma_1, \sigma_2, k)\pi_1(\beta)\pi(u)\pi_2(\sigma_1)\pi_3(\sigma_2)\pi_5(v)\pi_4(k) \\
& \propto \sigma_1^{-(k+1)} \prod_{i=1}^k u_i^{v-\frac{1}{2}} \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X_i'\beta)^2 u_i}{2\sigma_1^2}\right\} \sigma_2^{-(n-k+1)} \\
& \quad \prod_{i=k+1}^n u_i^{v-\frac{1}{2}} \exp\left\{-\sum_{i=k+1}^n \frac{(Y_i - X_i'\beta)^2 u_i}{2\sigma_2^2}\right\} \\
& v^n \left(\frac{v}{v+3}\right)^{1/2} \left[\psi'\left(\frac{v}{2}\right) - \psi'\left(\frac{v+1}{2}\right) - \frac{2(v+3)}{v(v+1)^2}\right]^{1/2}.
\end{aligned}$$

Case IV: In the contaminated linear regression model with  $u_i \sim vI_\gamma(u_i) + (1-v)I_1(u_i)$ ,

we have another two unknown parameters  $\{v, \gamma\}$ . We assign the priors for  $v$  and  $\gamma$

$$v \sim \text{Beta}(v_0, v_1) \quad \text{and} \quad \gamma \sim \text{Beta}(\gamma_0, \gamma_1),$$

respectively, where  $v_0, v_1, \gamma_0$ , and  $\gamma_1$  are the hyperparameters that need to be prespecified.

Then, the posterior density is given by

$$\begin{aligned}
& \Pi_4(\beta, \sigma_1, \sigma_2, u, v, \gamma, k, v_0, v_1, \gamma_0, \gamma_1 | Y, X) \\
& \propto L(Y|X, \beta, u, \sigma_1, \sigma_2, k) \pi_1(\beta) \pi(u) \pi_2(\sigma_1) \pi_3(\sigma_2) \pi_6(v) \pi_7(\gamma) \pi_4(k) \\
& \propto \sigma_1^{-(k+1)} \prod_{i=1}^k u_i^{\frac{1}{2}} \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X'_i \beta)^2 u_i}{2\sigma_1^2}\right\} \sigma_2^{-(n-k+1)} \\
& \quad \prod_{i=k+1}^n u_i^{\frac{1}{2}} \exp\left\{-\sum_{i=k+1}^n \frac{(Y_i - X'_i \beta)^2 u_i}{2\sigma_2^2}\right\} \\
& \quad \prod_{i=1}^n \left[ v I_\gamma(u_i) + (1-v) I_1(u_i) \right] v^{v_0-1} (1-v)^{v_1-1} \gamma^{\gamma_0-1} (1-\gamma)^{\gamma_1-1}.
\end{aligned}$$

It can be seen from the four cases that the conditional posterior density of each parameter is based on other parameters, which allows us to develop an efficient MCMC-type sampling algorithm for posterior Bayesian inference.

### 3.2 Inference Procedure Using MCMC

Due to the complex form of the joint posterior distribution of the unknown parameters, it is prohibitive to directly use numerical techniques to make Bayesian inference. In this Section, we develop an efficient MCMC-type sampling algorithm by generating samples from the posterior distributions. To be more specific, we obtain the full conditional posterior distribution of each parameter under consideration in the model, and then we draw posterior samples from these distributions.

### 3.2.1 Gibbs sampling

In this subsection, we consider the full conditional posterior distributions in Gibbs sampling based on the above four different posterior distributions.

Case I: The Gaussian linear regression model. The full conditional posterior distributions for  $\beta$ ,  $\sigma_1$ ,  $\sigma_2$  and  $k$  are the same as the ones in [10] and are thus omitted for simplicity.

Case II: Student- $t$  linear regression model. Note that the full conditional posterior distributions of  $\beta$ ,  $\sigma_1$ ,  $\sigma_2$ ,  $u$  and  $k$  are also the same as the ones in [10]. However, two main differences between our approach and the one in [10] are the prior specification of  $v$  and the posterior sampling for  $v$ . [10] consider an exponential distribution with a rate  $\lambda$  for  $v$ , where the choice of  $\lambda$  has a large impact on the estimation of the unknown parameters. In addition, their posterior simulation method for  $v$  seems to be computational cumbersome, especially when the value of  $v$  is small. To overcome these two potential difficulties, we consider an objective prior for  $v$  and develop an efficient Metropolis-Hastings algorithm (Section 3.2.2) to draw samples from the full conditional posterior distribution.

Case III: The slash linear regression model. According to the posterior density  $\Pi_3(\cdot)$ , the full conditional posterior distribution can be given as follows.

1. Posterior distribution of  $\beta$  conditional on  $\sigma_1, \sigma_2, u, v, \gamma$ , and  $k$ . Given  $\sigma_1, \sigma_2, u, v, \gamma$ ,

and  $k$ , the posterior distribution of  $\beta$  is proportional to

$$\begin{aligned} & \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X'_i\beta)^2 u_i}{2\sigma_1^2} - \sum_{i=k+1}^n \frac{(Y_i - X'_i\beta)^2 u_i}{2\sigma_2^2}\right\} \\ &= \exp\left\{\frac{(Y - X'\beta)'\Sigma^{-1}(Y - X'\beta)}{2}\right\} \\ &\propto \exp\left\{-\frac{(\beta - \tilde{\beta})'X'\Sigma^{-1}X(\beta - \tilde{\beta})}{2}\right\}, \end{aligned}$$

where  $\Sigma = \text{diag}(\sigma_1^2 u_1^{-1}, \sigma_1^2 u_2^{-1}, \dots, \sigma_1^2 u_k^{-1}, \sigma_2^2 u_{k+1}^{-1}, \dots, \sigma_2^2 u_n^{-1})$  and  $\tilde{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$ . Thus, it follows

$$(\beta \mid u, v, \sigma_1, \sigma_2, k) \sim N(\tilde{\beta}, (X'\Sigma^{-1}X)^{-1}).$$

2. Posterior distribution of  $\sigma_1$  conditional on  $\beta, \sigma_2, u, v, \gamma$ , and  $k$  and posterior distribution of  $\sigma_2$  conditional on  $\beta, \sigma_1, u, v, \gamma$ , and  $k$ . The posterior distribution of  $\sigma_1$  is proportional to

$$\begin{aligned} & \sigma_1^{-(k+1)} \exp\left\{-\sum_{i=1}^k \frac{(Y_i - X'_i\beta)^2 u_i}{2\sigma_1^2}\right\} \\ & \propto \sigma_1^{-2(\frac{k+1}{2})} \exp\left\{-\sigma_1^{-2} \sum_{i=1}^k \frac{(Y_i - X'_i\beta)^2 u_i}{2}\right\}. \end{aligned}$$

Thus, it follows

$$(\sigma_1^{-2} \mid \beta, \sigma_2, u, v, \gamma, k) \sim \text{Gamma}\left(\frac{k}{2}, \sum_{i=1}^k \frac{(Y_i - X'_i\beta)^2 u_i}{2}\right).$$

Similarly, the posterior density of  $\sigma_2$  conditional on  $\beta, \sigma_1, u, v,$  and  $k$  is

$$(\sigma_2^{-2} | \beta, \sigma_1, u, v, \gamma, k) \sim \text{Gamma} \left( \frac{n-k}{2}, \sum_{i=k+1}^n \frac{(Y_i - X'_i \beta)^2 u_i}{2} \right).$$

3. Posterior distribution of  $u$  conditional on  $\sigma_1, \beta, \sigma_2, v, \gamma,$  and  $k$ . The posterior distribution of  $u_i,$  for  $i = 1, \dots, k$  is proportional to

$$\begin{aligned} & u_i^{\frac{1}{2}} \exp \left\{ -\frac{(Y_i - X'_i \beta)^2 u_i}{2\sigma_1^2} \right\} \{v I_\gamma(u_i) + (1-v) I_1(u_i)\} \\ & \propto \exp \left\{ -\frac{(Y_i - X'_i \beta)^2}{2\sigma_1^2} \right\} (1-v) = A(i) \quad \text{when } u_i = 1 \\ & \propto \gamma^{\frac{1}{2}} \exp \left\{ -\frac{(Y_i - X'_i \beta)^2 \gamma}{2\sigma_1^2} \right\} v = B(i) \quad \text{when } u_i = \gamma. \end{aligned}$$

Let  $\psi_1 \sim \text{Binomial} \left( 1, \frac{A(i)}{B(i)} \right),$  then, for  $i = 1, \dots, k,$  it follows

$$u_i = \begin{cases} 1, & \text{if } \psi_1 = 1, \\ \gamma, & \text{if } \psi_1 = 0. \end{cases}$$

Similarly, the posterior distribution of  $u_i$ , for  $i = k + 1, \dots, n$  is proportional to

$$\begin{aligned} & u_i^{\frac{1}{2}} \exp\left\{-\frac{(Y_i - X_i'\beta)u_i}{2\sigma_2}\right\} \left\{vI_\gamma(u_i) + (1-v)I_1(u_i)\right\} \\ & \propto \exp\left\{-\frac{(Y_i - X_i'\beta)}{2\sigma_2}\right\} (1-v) = A(i) \quad \text{when } u_i = 1 \\ & \propto \gamma^{\frac{3}{2}} \exp\left\{-\frac{(Y_i - X_i'\beta)\gamma}{2\sigma_2}\right\} v = B(i) \quad \text{when } u_i = \gamma. \end{aligned}$$

Let  $\psi_2 \sim \text{Binomial}\left(1, \frac{A(i)}{B(i)}\right)$ , then for  $i = (k + 1), \dots, n$ , we have

$$u_i = \begin{cases} 1, & \text{if } \psi_2 = 1, \\ \gamma, & \text{if } \psi_2 = 0. \end{cases}$$

4. Posterior distribution of  $v$  conditional on  $\sigma_1, \beta, \sigma_2, u, \gamma$  and  $k$ . The posterior distribution of  $v$  is proportional to

$$\begin{aligned} & v^{v_0-1} (1-v)^{v_1-1} \prod_{i=1}^n [vI_\gamma(u_i) + (1-v)I_1(u_i)] \\ & \propto v^{v_0-1+n-m} (1-v)^{v_1-1+m}, \end{aligned}$$

where  $m$  is the number of  $u_i = 1$ .

Thus, we have that

$$(v|\beta, \sigma_1, \sigma_2, u, \gamma, k) \sim \text{Beta}(v_0 + n - m, v_1 + m).$$



5. Posterior distribution of  $\gamma$  conditional on  $\sigma_1, \beta, \sigma_2, u, v$  and  $k$ . The posterior distribution of  $\gamma$  is proportional to

$$\begin{aligned} & \gamma^{\gamma_0-1}(1-\gamma)^{\gamma_1-1} \prod_{i=1}^n [vI_\gamma(u_i) + (1-v)I_1(u_i)] \\ & \propto \gamma^{\gamma_0+n-m-1}(1-\gamma)^{\gamma_1-1}. \end{aligned}$$

Thus, we have that

$$(\gamma|\beta, \sigma_1, \sigma_2, u, v, k) \sim \text{Beta}(\gamma_0 + n - m, \gamma_1).$$

6. Posterior distribution of  $k$  conditional on  $\sigma_1, \beta, \sigma_2, u, v$  and  $\gamma$ . The posterior distribution of  $k$  is proportional to

$$(k|\beta, \sigma_1, \sigma_2, u, v) \sim \frac{L(Y, X|\beta, \sigma_1, \sigma_2, u, v, k)}{\sum_{k=p}^{n-p} L(Y, X|\beta, \sigma_1, \sigma_2, u, v, k)}, k = p, \dots, n - p.$$

Case IV: The contaminated regression model.

The posterior distributions in the contaminated regression model for  $\beta, \sigma_1, \sigma_2$ , and  $k$  are the same as the ones in the slash distribution. We use the Metropolis-Hastings method to calculate posterior distribution of  $v$  in Section 3.2.2.

The posterior distribution of  $u_i$ , for  $i = 1, \dots, k$  conditional on  $\beta, \sigma_1, \sigma_2, v$ , and  $k$  is proportional to

$$u_i^{v-\frac{1}{2}} \exp \left[ -\frac{(Y_i - X_i'\beta)^2 u_i}{2\sigma_1^2} \right].$$

Then, for  $u_i \in (0, 1)$ ,  $i = 1, \dots, k$ , we have that

$$(u_i | \beta, \sigma_1, \sigma_2, v, k) \sim \text{Gamma}\left(v + \frac{1}{2}, \frac{(Y_i - X'_i \beta)^2}{2\sigma_1^2}\right).$$

Similarly, for  $u_i \in (0, 1)$ ,  $i = k + 1, \dots, n$ , we have that

$$(u_i | \beta, \sigma_1, \sigma_2, v, k) \sim \text{Gamma}\left(v + \frac{1}{2}, \frac{(Y_i - X'_i \beta)^2}{2\sigma_2^2}\right).$$

### 3.2.2 Metropolis-Hastings Algorithm

In Student t and slash linear regression models, we use the M-H algorithm to simulate  $v$ .

Given  $\beta, \sigma_1, \sigma_2, u$ , and  $k$ , the posterior density of  $v$  is proportional to

$$\begin{aligned} \pi(v | \beta, \sigma_1, \sigma_2, u, k) & \propto \left(\frac{v}{2}\right)^{\frac{nv}{2}} \left[\Gamma\left(\frac{v}{2}\right)\right]^{-n} \prod_{i=1}^n u_i^{\frac{v+3}{2}} \exp\left(-\sum_{i=1}^n \frac{v}{2u_i}\right) \pi_5(v) \\ & \propto \left(\frac{v}{2}\right)^{\frac{nv}{2}} \left[\Gamma\left(\frac{v}{2}\right)\right]^{-n} \exp(-\eta v), \end{aligned}$$

where  $\pi_5(v)$  is Jeffreys prior of  $v$ , and  $\eta = \frac{1}{2} \sum_{i=1}^n [\log(u_i) + \log(\frac{1}{u_i})] + \log[p(v)]$ .

According to the procedure studied by [18], we use the normal distribution

$\mathcal{N}_{(2 < v < 40)}(\mu_v, \tau_v^2)$  to sample  $v$ , where  $\mu_v = x - \frac{q'(x)}{q''(x)}$ ,  $\tau_v^2 = -\frac{1}{q''(x)}$  and  $q(\cdot) = \log[\pi(v | \beta, \sigma_1, \sigma_2, u, k)]$ .

We assume the four steps for sampling  $v$  as follows:

1. Generate  $v_{\text{new}}$  from  $\mathcal{N}_{(2 < v < 40)}(\mu_v, \tau_v^2)$ , where  $\mu_v = v_j - \frac{q'(v_j)}{q''(v_j)}$ ,  $\tau_v^2 = -\frac{1}{q''(v_j)}$ .
2. Generate  $w$  from the Uniform  $(0, 1)$ .
3. Define  $\alpha$ .

$$\alpha = \min \left( 1, \exp \left\{ \frac{q(v_{\text{new}}|v_j)/\log(p(v_{\text{new}}|v_j))}{q(v_j|v_{\text{new}})/\log(p(v_j|v_{\text{new}}))} \right\} \right),$$

where  $p(\cdot)$  is the pdf of  $\mathcal{N}_{(2 < v < 40)}(\mu_v, \tau_v^2)$ .

4.  $v_{j+1} = v_{\text{new}}I_1(\alpha > w) + v_jI_1(\alpha \leq w)$ .

### 3.2.3 Acceptance-Rejection Algorithm

In the contaminated regression model, the random variable  $u_i$ , for  $i = 1, \dots, n$  has a truncated gamma distribution. It may take too much time by using the truncated gamma function to sample  $u_i$ , for  $i = 1, \dots, n$ . Instead, we use the acceptance-rejection (A-R) algorithm for posterior sampling of  $u_i$ , for  $i = 1, \dots, n$ .

The posterior density of  $u_i$ , for  $i = 1, \dots, k$  conditional on  $\beta, \sigma_1, \sigma_2, v$ , and  $k$  is

$$\pi(u_i|\beta, \sigma_1, \sigma_2, v, k) \propto u_i^{v-\frac{1}{2}} \exp \left\{ -\frac{(Y_i - X_i'\beta)^2 u_i}{2\sigma^2} \right\},$$

where  $\sigma = \sigma_1$  when  $i = 1, \dots, k$ ;  $\sigma = \sigma_2$  when  $i = k + 1, \dots, n$ , and  $0 < u_i < 1$ , for  $i = 1, \dots, n$ .

Since  $0 < u_i < 1$ , for  $i = 1, \dots, n$  and  $v - \frac{1}{2} > 0$ , we have  $\pi(u_i|\beta, \sigma_1, \sigma_2, v, k) < c$ . We set  $c = 1$  in the A-R algorithm which can be explained in the following:

1. Generate  $\tau$  from Uniform  $(0, 1)$ .
2. Calculate  $\alpha = \pi(\tau|\beta, \sigma_1, \sigma_2, v, k)$ .
3.  $u_{ij+1} = \tau I_1(\alpha < c) + u_{ij} I_1(\alpha \geq c)$ .



# Chapter 4

## Simulations and Real-data Application

In this chapter, we conduct the simulation studies for the a variance change-point detection. Then we investigate the effects of the regression coefficients to the detection of variance change-point. Finally, the proposed procedure is applied to the closing price data of U.S. stock market for illustration purposes.

### 4.1 Simulation Study

In this section, we conduct Monte Carlo simulations to detect the change-point in the proposed four models: Gaussian, Student- $t$ , contaminated, and slash linear regression models.

### 4.1.1 Simulation I

Based on the models in Section 2.2, we simulate a sequence of sample  $Y = (Y_1, \dots, Y_n)'$  by setting  $n = 300, \beta = (1, 4, 7, 10)'$ ,  $X = (X_1, \dots, X_{300})'$ ,  $X_i = (x_{i1}, \dots, x_{i4})'$ ,  $x_{ij} \sim \text{Uniform}(0, 1)$ ,  $\sigma_1 = 1, \sigma_2 = 2, u = (u_1, \dots, u_{300})$ , and  $u_i = 1$  for  $i = 1, \dots, 300$ . The location of a change-point is set to be  $k = 200$ . The main goal of this simulation is to evaluate the performance at the proposed procedure discussed in Section 3.2. In the simulation study, we do 100 simulations. In each simulation, we do 5000 iterations. The posterior median of each parameter and its corresponding standard error under the four regression models are listed in Table 4.2.

We set  $u_i = 1$ , for  $i = 1, \dots, 300$ , then the data  $\{Y_1, \dots, Y_{300}\}$  are simulated from the normal distribution based on 2.1. For Table 4.1, the second row is the real value of parameters  $k, \beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$ ,  $\sigma_1, \sigma_2$ . There are two rows corresponding to each model. The upper row are the estimated values of parameters and the lower row are the standard errors of the corresponding parameters. For example, in Student- $t$  regression model, the estimated  $k$  is 200.1 with standard error of 4.383. From Table 4.1, we observe that all the four models locate the variance change-point and estimate the other parameters efficiently. The results from Gaussian and contaminated linear regression models are slightly better than those of Student- $t$  and slash linear regression models. This phenomenon may be caused by the data which are simulated from the normal distribution.

**Table 4.1**  
Bayesian posterior estimation of the unknown parameters with standard deviation  
parameters in Simulation I

Distributions	$k$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\sigma_1$	$\sigma_2$
Data	200	1	4	7	10	1	2
Gaussian	200.2 (5.034)	0.991 (0.234)	4.031 (0.178)	6.987 (0.209)	9.994 (0.201)	0.999 (0.055)	2.027 (0.136)
Student- $t$	200.1 (4.383)	0.995 (0.248)	4.039 (0.187)	6.974 (0.225)	9.991 (0.212)	0.854 (0.057)	1.726 (0.135)
Contaminated	200.2 (5.514)	0.993 (0.256)	4.027 (0.173)	6.990 (0.224)	9.995 (0.229)	1.013 (0.056)	2.051 (0.140)
Slash	200.3 (4.445)	0.993 (0.243)	4.035 (0.182)	6.979 (0.219)	9.994 (0.205)	0.764 (0.046)	1.546 (0.114)

### 4.1.2 Simulation II

In order to have a over look at whether Gaussian and contaminated regression models are always better than Student- $t$  and slash models or not, we conduct another simulation study. In this case, we get the sample based on the formation of Student- $t$  in Section 2.1. We set  $v = 3$  and the other parameters are the same values as they are in the Simulation I study. The detection results are listed in Tables 4.2.

From Table 4.2, we observe that the results based on Student- $t$  and slash distributions outperform the ones from Gaussian and contaminated regression models. This result is quite reasonable, because data is simulated from Student- $t$  distribution. The result also shows that it is beneficial to use different models for different data to detect the most efficient variance change-point.



**Table 4.2**  
Bayesian posterior estimation of the unknown parameters with standard deviation  
parameters in Simulation II

Distributions	$k$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\sigma_1$	$\sigma_2$
Data	200	1	4	7	10	1	2
Gaussian	197.9 (26.728)	0.962 (0.333)	4.006 (0.301)	6.983 (0.302)	10.055 (0.345)	1.559 (0.137)	3.259 (0.627)
Student- $t$	199.9 (5.498)	0.969 (0.261)	4.020 (0.251)	6.995 (0.232)	10.017 (0.275)	0.961 (0.071)	1.870 (0.175)
Contaminated	200.6 (11.544)	0.978 (0.286)	4.002 (0.295)	6.994 (0.273)	10.021 (0.328)	1.749 (0.151)	3.663 (0.761)
Slash	200.0 (6.270)	0.970 (0.265)	4.006 (0.258)	6.994 (0.247)	10.033 (0.302)	1.033 (0.068)	2.007 (0.162)

## 4.2 Coefficient Effect On The Detection Of Variance

### Change-point

In the previous settings, we assume that coefficients of the data are the same, whereas it may not be reasonable for the real life data. In this section, we introduce the coefficient change-point of data  $Y = \{Y_1, \dots, Y_n\}$ . Suppose that the coefficient change-point is  $h$  and the variance change-point is  $k$ . Then the model can be written as:

$$Y_i = X_i' \beta^{(1)} + \varepsilon_i^{(1)}, \quad i = 1, \dots, h;$$

$$Y_i = X_i' \beta^{(2)} + \varepsilon_i^{(1)}, \quad i = (h + 1), \dots, k;$$

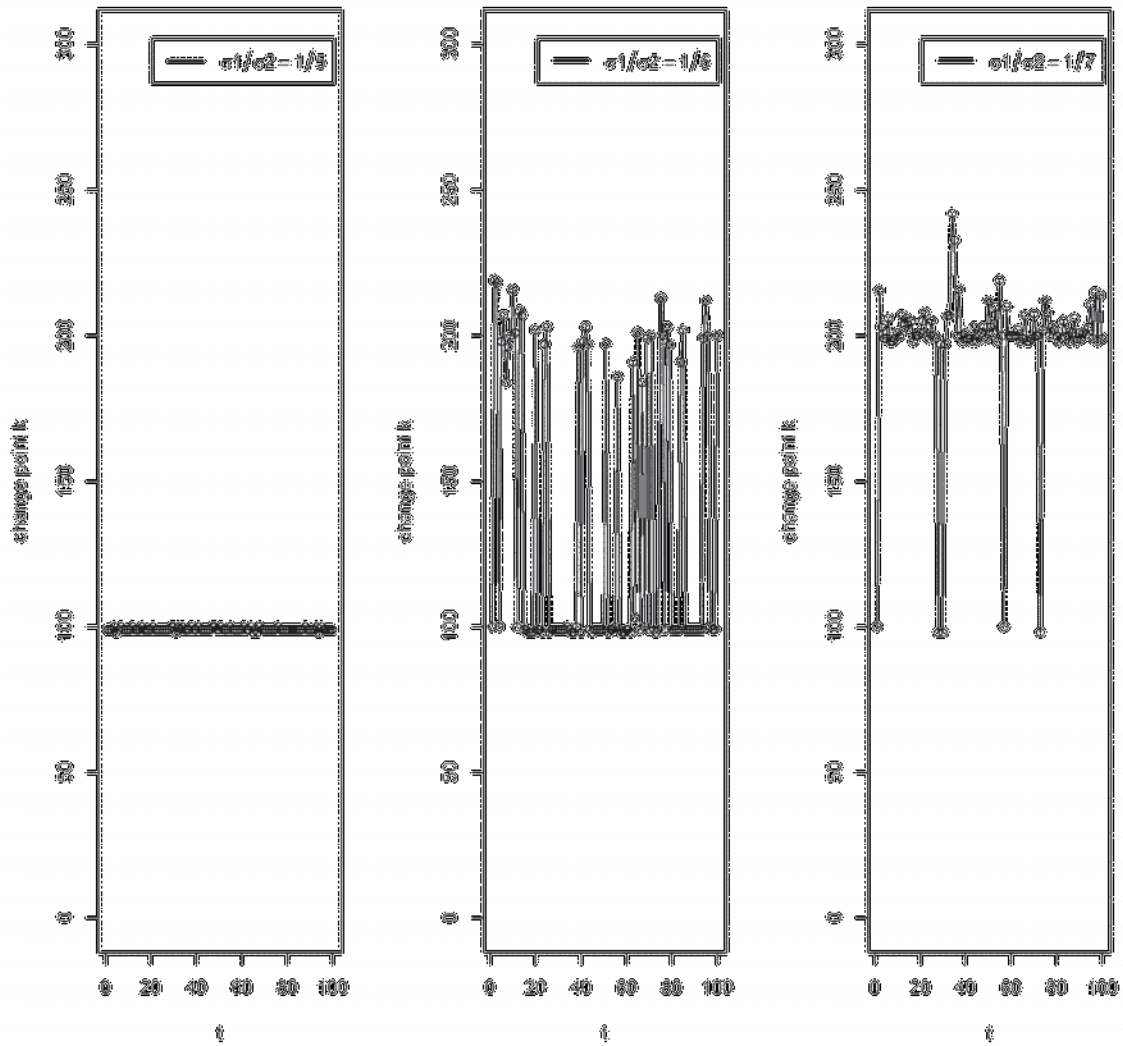
$$Y_i = X_i' \beta^{(2)} + \varepsilon_i^{(2)}, \quad i = (k + 1), \dots, n.$$

We conduct Monte Carlo simulations to investigate the performance of the proposed procedure in the case of existing both coefficient change-point and variance change-point. We set  $h = 100$ ,  $\beta^{(1)} = (1, 4, 7, 10)'$ , and  $\beta^{(2)} = (1, 1, 1, 1)'$ . Values of  $Y, X, n$ , and  $k$  are the same as the ones in the Section 4.1.1. We choose  $\sigma_1 = 1$ , and in order to have a look at the coefficient effects on the variance change-point, a sequence of  $\sigma_2$  are selected, i.e.,  $\sigma_2 = \{4, 5, \dots, 9\}$ .

We do 100 simulations. Each simulation has 5000 iterations. Figures 4.1,4.2,4.3,4.4 are the detected change-points under different values of  $\sigma_2$ .

Based on Figures 4.1, 4.2, 4.3, 4.4, several conclusions can be made as follows. Firstly, we detect the changes in terms of both coefficient and variance of the data. Secondly, the detected change-point is 100, which is a coefficient change-point, when the value of  $\sigma_2$  is small and it changes to 200, which is a variance change-point, when the value of  $\sigma_2$  is large. This phenomenon shows that the detected change-point may be a coefficient change-point. It happens when the variance change is not obvious. Thirdly, in the Student- $t$  linear regression model, the variance change-point 200 is not detected until  $\sigma_2$  is 8 which means that Student- $t$  regression model may be easily affected by the coefficient change point in the process of variance change-points detection.

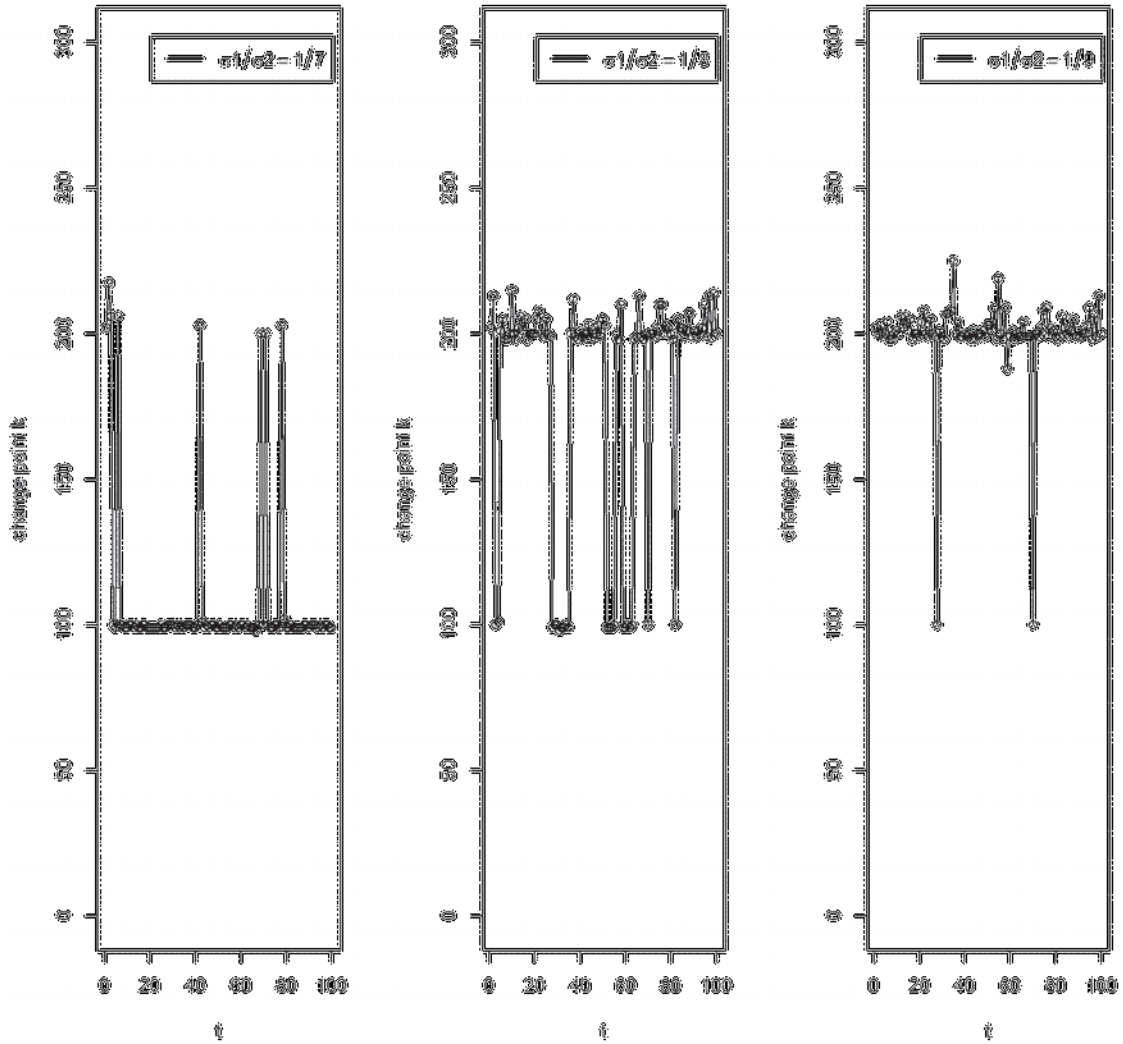
In order to check if 100 is a variance change-point, we detect the variance change-point of data  $Y_1, \dots, Y_{200}$ . At this time, 100 is not detected which means that 100 is not a variance change-point. The finding is consistent with the model of coefficient change-point of 100.



**Figure 4.1:** Gaussian distribution: Detected change points when  $\sigma_2 = 5, 6, 7$ .

Consequently, it is beneficial to check if the detected change-point is caused by variance change, when we are locating the variance change-point.

Suppose that we detect change-points at locations  $k_1, k_2, k_3$  and that there are no other



**Figure 4.2:** Student  $t$  distribution: Detected change points when  $\sigma_2 = 7, 8, 9$ .

change-points between  $k_1$  and  $k_3$  except  $k_2$ , in order to decide if  $k_2$  is a variance change-point, we need to detect the variance change-point between  $Y_{k_1}$  and  $Y_{k_3}$ . If we cannot locate  $k_2$ ,  $k_2$  is not a variance change-point. If we can still locate  $k_2$ , then  $k_2$  is a variance change-point.

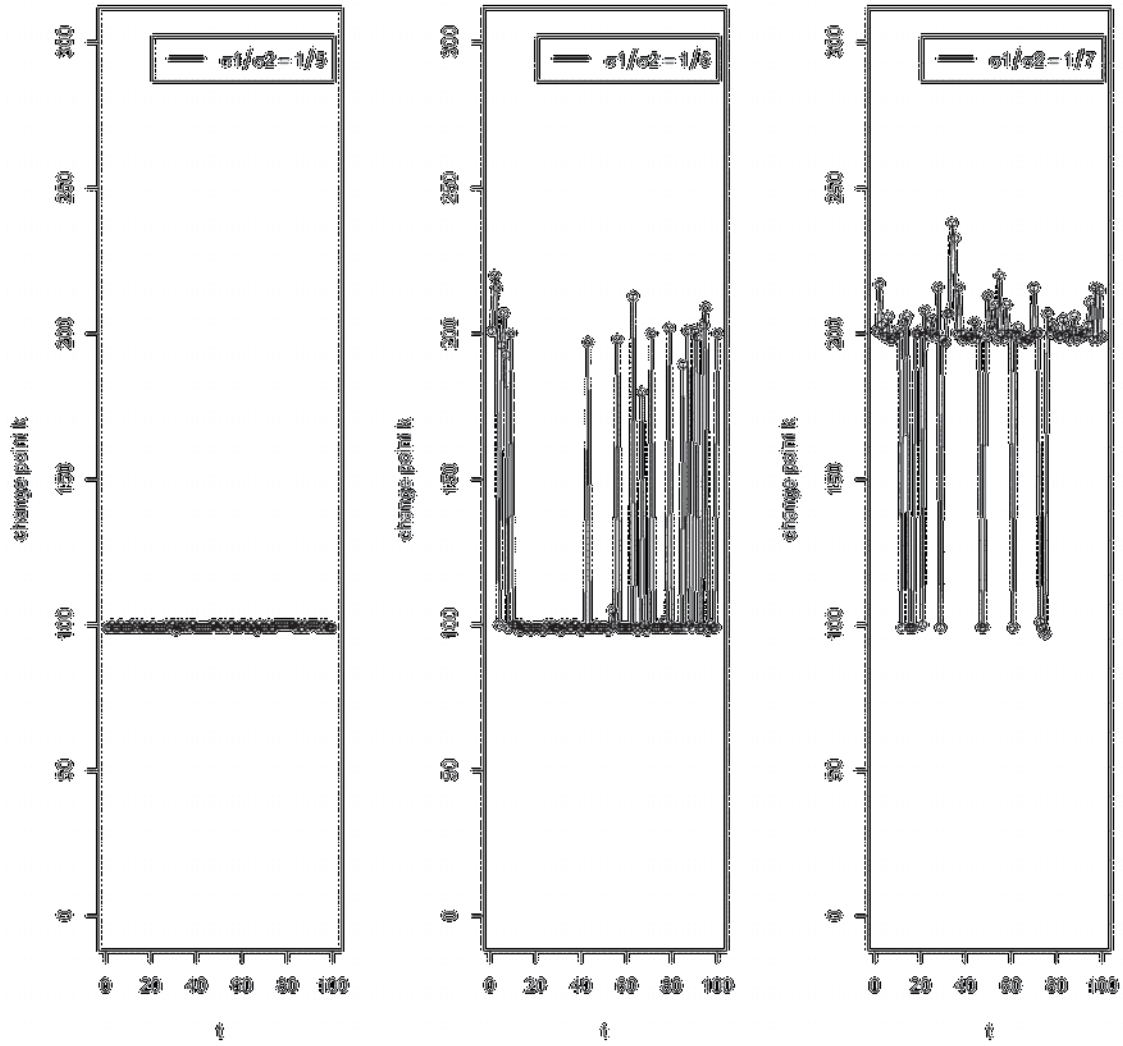


Figure 4.3: Contaminated distribution: Detected change points when  $\sigma_2 = 5, 6, 7$ .

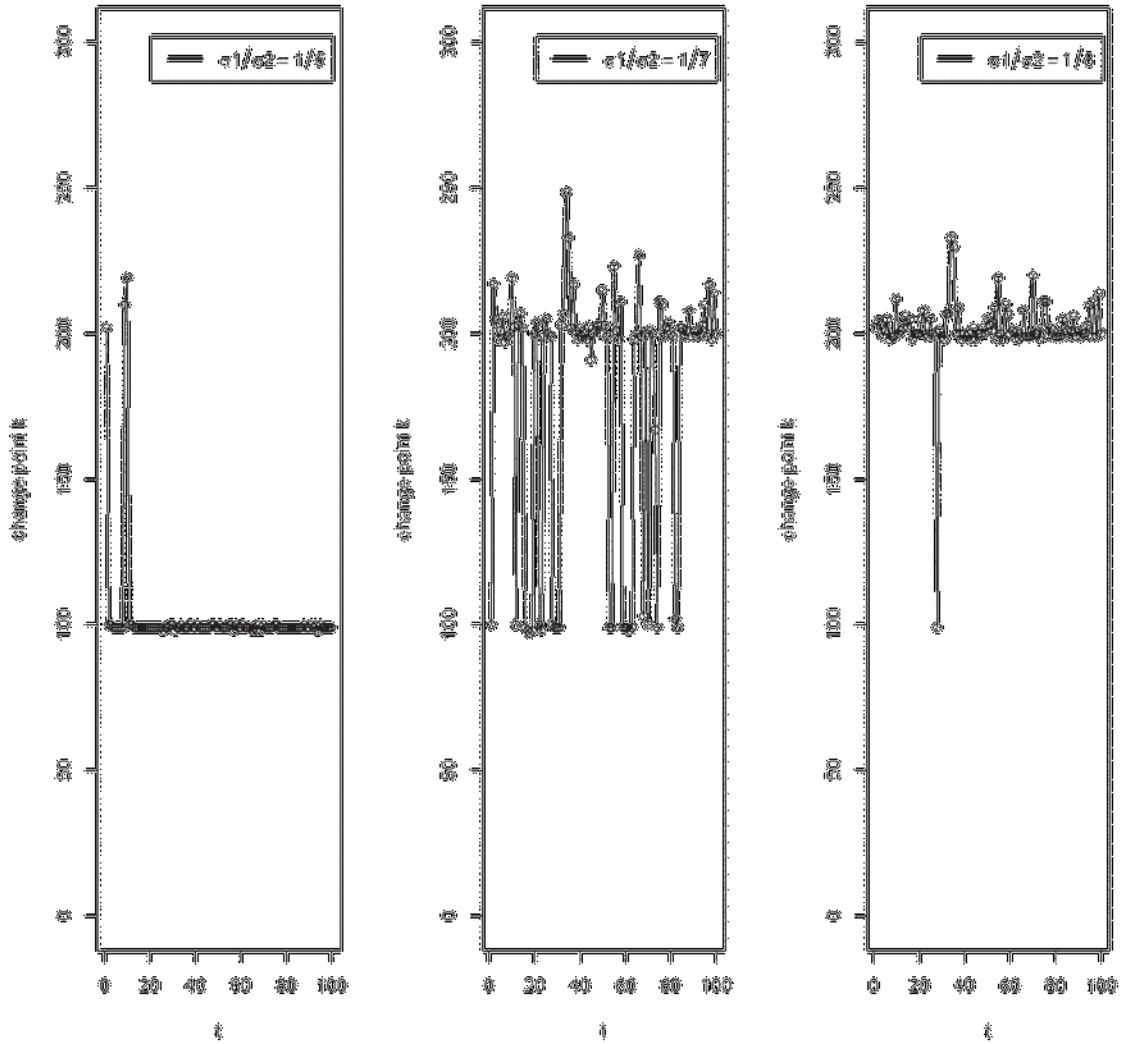


Figure 4.4: Slash distribution: Detected change points when  $\sigma_2 = 6, 7, 8$ .

### 4.3 Dow Jones Index With Multiple Variance Change-points

In practical applications, there are usually more than one variance change-points for the real data, especially for a long time data. In this section, we discuss how to locate all the variance change-points of the data. The method we consider is based on the binary segmentation procedure developed by .

Steps of detecting multiple variance change-points can be summarized as follows:

1. Finding a change point  $k_1$  which results in two separate data  $(1 : k_1)$  and  $(k_1 : n)$ .
2. Find the second change point  $k_2$  from data  $(1 : k_1)$  and the third change point  $k_3$  from data  $(k_1 : n)$ . If the change-points  $k_2$  and  $k_3$  exist, go to Step 3.
3. Do the same procedures as Step 2 until no more change-points are detected.

To illustrate how to apply our models into the real data and check if they are effective.

We use the Dow Jones Index closing price data  $Y_1, \dots, Y_{252}$  from Oct 22, 2008 to Oct 22, 2009 at <https://finance.yahoo.com/q/hp?s=%5EDJI&a=09&b=22&c=2008&d=09&e=22&f=2009&g=d>. Before we detect the variance change-points, we transform the closing price into yearly yield rate by using  $R_t = 252 \frac{Y_{t+1} - Y_t}{Y_t}$  to reduce the non-stationarity of the original data. Table 4.3 summaries the statistics of data  $R_1, \dots, R_{252}$  and Figure 4.5 is a time series plot of data  $R_1, \dots, R_{252}$ .

To detect the multiple variance change-points of yearly yield rate data, we firstly construct

the model:

$$\begin{aligned}
 R_t &= \beta + \varepsilon_t^{(1)}, & t = 1, \dots, k_1; \\
 R_t &= \beta + \varepsilon_t^{(2)}, & t = (k_1 + 1), \dots, k_2; \\
 &\dots\dots \\
 R_t &= \beta + \varepsilon_t^{(k_s)}, & t = (k_{s-1} + 1), \dots, k_s,
 \end{aligned}$$

where  $k_s$  is the number of variance change points and  $\varepsilon_t^{(j)}$ , for  $t = 1, \dots, n$  and  $j = 1, \dots, k_s$ , have distributions corresponding to Gaussian, Student- $t$ , contaminated, and slash linear regression models. Combing the binary segmentation procedure and the sampling algorithm in Section 3.2, numerical results are listed in Table 4.4. We draw conclusions as follows.

We observe that the results in the four models are consistent and show that the locations of the variance change-points are about 31, 122, 185. By using the SIC method introduced by Chen and Gupta in 1997, the detected variance change-points are 28, 122 and 188 which are quite similar with the ones from our four regression models. This means that our four linear regression models are effective. What's more, the running time of our method is much less than Lin and Chen's proposed sampling. This highlights that our procedure is more efficient than the one in [10].

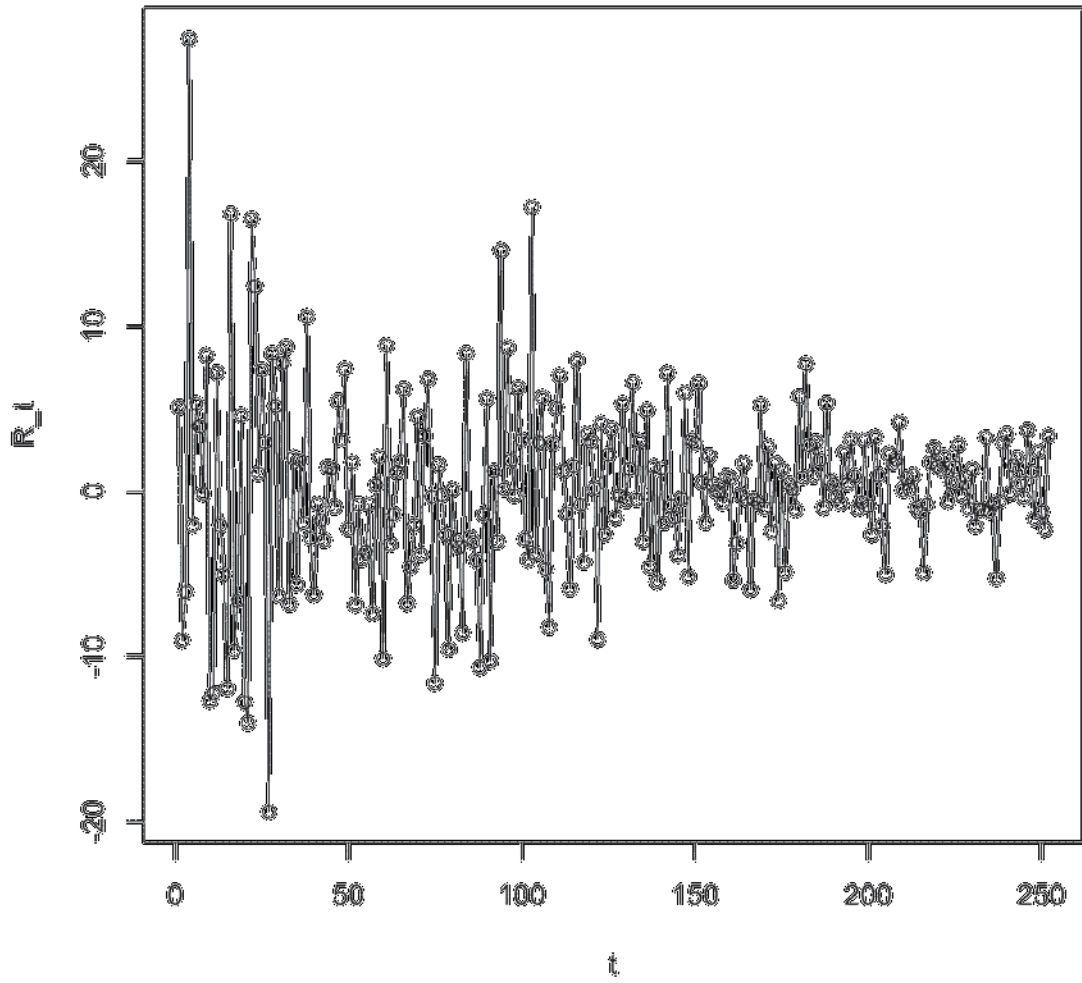


**Table 4.3**  
Descriptive statistics of  $R_t$

Sample size	Mean	Standard error	Median
252	0.225	5.374	0.144
Minimum	Maximum	Skewness	Kurtosis
-19.407	27.412	0.395	6.499

**Table 4.4**  
Variance change-points of real data

Distributions	Gaussian	Student- $t$	Contaminated	Slash
k1	122 (7.339)	123 (15.125)	122 (7.318)	123 (13.953)
k2	31 (4.466)	33 (10.582)	31 (4.780)	32 (4.016)
k3	186 (13.156)	185 (23.572)	186 (13.040)	184 (16.576)



**Figure 4.5:** Time series of  $R_t$

# Chapter 5

## Concluding Remarks and Future Work

In this thesis, we not only discuss Bayesian inferences of variance change-points detection in the family of the SMN distributions, which includes special cases such as Gaussian, Student- $t$ , contaminated and slash distributions, but also do we use simulation studies to prove the effectiveness of these four linear regression models on the detection of variance change-points. In the Student- $t$  regression model, we consider an objective prior for the degrees of freedom, which makes our approach be more persuasive. Moreover, by studying the effects of coefficient change-points on the detection of variance change-points, we note that the detected change-point belongs to the variance change-points when the variance change is more obvious than the mean change in the four models under consideration. Finally, we apply the four models to detect the change-points of Dow Jones Index closing price data and obtain the similar results as the one based on the SIC method.

It deserves mentioning that different models should be used based on the properties of the data, such as the skewness and the kurtosis, etc.



# References

- [1] Vostrikova, L. *Soviet Mathematics* **1981**, 24, 55–59.
- [2] Page, E. S. *Biometrika* **1954**, 41, 100–115.
- [3] Talwar, P. P. *Journal of Econometrics* **1983**, 23(3), 353 – 367.
- [4] Chen, J. *Communications in Statistics. Theory and Methods* **1998**, 27(10), 2481–2493.
- [5] Horváth, L.; Steinebach, J. *Journal of Statistical Planning and Inference* **2000**, 91(2), 365 – 376.
- [6] Moreno, E.; Casella, G.; Garcia-Ferrer, A. *Stochastic Environmental Research and Risk Assessment* **2005**, 19(3), 191–204.
- [7] Zhou, H.; Liang, K.-Y. In *Beyond parametrics in interdisciplinary research: Festschrift in honor of Professor Pranab K. Sen*, Vol. 1 of *Inst. Math. Stat. Collect.*; Inst. Math. Statist., Beachwood, OH, 2008; pages 305–320.

- [8] Chen, J.; Gupta, A. K. *Journal of the American Statistical Association* **1997**, 92(438), 739–747.
- [9] Lange, K. L.; Little, R. J. A.; Taylor, J. M. G. *Journal of the American Statistical Association* **1989**, 84(408), pp. 881–896.
- [10] Lin, J.-G.; Chen, J.; Li, Y. *Computational Economics* **2012**, 40(3), 203–217.
- [11] Choy, S. T. B.; Chan, J. S. K. *Australian & New Zealand Journal of Statistics* **2008**, 50(2), 135–146.
- [12] Andrews, D. F.; Mallows, C. L. *Journal of the Royal Statistical Society. Series B. Methodological* **1974**, 36, 99–102.
- [13] Lange, K.; Sinsheimer, J. S. *Journal of Computational and Graphical Statistics* **1993**, 2(2), 175–198.
- [14] Abanto-Valle, C. A.; Migon, H. S.; Lachos, V. H. *Brazilian Journal of Probability and Statistics* **2012**, 26(4), 402–422.
- [15] Johnson, N. L.; Kotz, S.; Balakrishnan, N. *Continuous univariate distributions. Vol. 1*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics; John Wiley & Sons, Inc., New York, second ed., 1994.
- [16] Madan, D. B.; Seneta, E. *The Journal of Business* **1990**, 63(4), 511–24.
- [17] Choy, S.; Smith, A. *Test* **1997**, 6(1), 205–221.

- [18] Abanto-Valle, C. A.; Dey, D. K. *Computational Statistics & Data Analysis* **2014**, *71*, 274–287.
- [19] Abanto-Valle, C. A.; Lachos, V. H.; Ghosh, P. *Applied Stochastic Models in Business and Industry* **2012**, *28*(5), 430–447.
- [20] Abraham, B.; Wei, W. W. S. *Metrika. International Journal for Theoretical and Applied Statistics* **1984**, *31*(3-4), 183–194.
- [21] Bhattacharya, P. K.; Zhou, H. In *Change-point problems (South Hadley, MA, 1992)*, Vol. 23 of *IMS Lecture Notes Monogr. Ser.*; Inst. Math. Statist., Hayward, CA, 1994; pages 57–65.
- [22] Bhattacharyya, G. K.; Johnson, R. A. *Annals of Mathematical Statistics* **1968**, *39*, 1731–1743.
- [23] Brodsky, B. E.; Darkhovsky, B. S. *Nonparametric methods in change-point problems*, Vol. 243 of *Mathematics and its Applications*; Kluwer Academic Publishers Group, Dordrecht, 1993.
- [24] Car, R.; Parrinello, M. *Physical Review Letters* **1985**, *55*, 2471.
- [25] Chen, J. *Inference about the change points in a sequence of Gaussian random vectors using information criterion*; ProQuest LLC, Ann Arbor, MI, 1995.
- [26] Chen, C. W. S.; Chan, J. S. K.; Gerlach, R.; Hsieh, W. Y. L. *Statistics and Computing* **2011**, *21*(3), 395–414.



- [27] Chen, J.; Gupta, A. K. *Mathematical Methods of Statistics* **2003**, 12(1), 116–130.
- [28] Chen, J.; Gupta, A. K. *Advances and Applications in Statistics* **2007**, 7(1), 115–125.
- [29] Chin Choy, J. H.; Broemeling, L. D. *Technometrics. A Journal of Statistics for the Physical, Chemical and Engineering Sciences* **1980**, 22(1), 71–78.
- [30] Chin Choy, J. H.-L. T. *A Bayesian analysis of a change linear model*; ProQuest LLC, Ann Arbor, MI, 1977.
- [31] De la Cruz, R. *Journal of Pharmaceutical Statistics* **2014-02-01 00:00:00.001**, 13(5), 81–93.
- [32] Gupta, A. K.; Ramanayake, A. *Journal of Statistical Planning and Inference* **2001**, 93(1-2), 181–195.
- [33] Hawkins, D. L. *Journal of Statistical Planning and Inference* **1992**, 33(2), 233–244.
- [34] Horvath, L. *The Annals of Statistics* **1993**, 21(2), 671–680.
- [35] Hsu, D. A. *Journal of the American Statistical Association* **1979**, 74(365), 31–40.
- [36] Liu, C. *Journal of the American Statistical Association* **1996**, 91(435), 1219–1227.
- [37] Page, E. S. *Biometrika* **1955**, 42, 523–527.
- [38] Sen, A.; Srivastava, M. S. *The Annals of Statistics* **1975**, 3, 98–108.
- [39] Sen, A.; Srivastava, M. S. *Annals of the Institute of Statistical Mathematics* **1975**, 27(3), 479–486.

- [40] Tang, J.; Gupta, A. K. *Journal of Statistical Computation and Simulation* **1987**, 27(2), 155–173.
- [41] Wang, Y. *Biometrika* **1995**, 82(2), 385–397.
- [42] Worsley, K. J. *Journal of the American Statistical Association* **1979**, 74(366, part 1), 365–367.
- [43] Worsley, K. J. *Biometrika* **1983**, 70(2), 455–464.
- [44] Worsley, K. J. *Biometrika* **1986**, 73(1), 91–104.