Some Contributions to Nonparametric Estimation of Density and Related Functionals for Biased Data

Jun Li

A Thesis
in the Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada
September 2010

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Jun Li**

Entitled: **Some Contributions to Nonparametric Estimation of Density and Related Functionals for Biased Data**

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_______________________ Chair
Dr. D. Kira

_______________________ External Examiner
Dr. R.S. Singh

_______________________ External to Program
Dr. N. Gospodinov

_______________________ Examiner
Dr. A. Sen

_______________________ Examiner
Dr. W. Sun

_______________________ Thesis Supervisor
Dr. Y.P. Chaubey

Approved by _______________________
Dr. J. Garrido, Graduate Program Director

November 29, 2010 _______________________
Dr. B. Lewis, Dean
Faculty of Arts and Science

Abstract

Some Contributions to Nonparametric Estimation of Density and Related Functionals for Biased Data

Jun Li, Ph.D.

Concordia University, 2010

Length biased sampling, as a special case of general biased sampling, occurs naturally in many statistical applications. In problems related with such applications, two different density functions are involved. One of them is the density of interest, which is referred to as the unweighted density, information about which is not observable directly in practice; the other one is referred to as the weighted density, the sample from which could be observed directly. These two densities are connected through a weight function. One aspect regarding data from weighted density is to estimate the unweighted density from the sample obtained using the weighted density. In this thesis we concentrate on the weight function representing length of the sampling unit that results in a sample called length-biased sample. Since most of such data are nonnegative, unweighted density has a non-negative support where common kernel density estimators with symmetric kernel may not be appropriate. Such density estimators usually generate the edge effect, which makes these to have large bias at the lower boundary. One possible reason for this is that symmetric kernels may assign some weights in region of zero probability.

In this thesis, we propose some new smooth density estimators based on Poisson distribution and nonnegative asymmetric kernels for length biased data to take care of

the edge effect. We investigate asymptotic behavior of these proposed density estimators as well as their finite sample performance through extensive simulation studies, that is more meaningful in practice. Also, we compare our new density estimators with other estimators in literature. Further, in addition to density estimators, we also consider smooth estimators of distribution function and some other functionals of the density such as hazard function and mean residual life function.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Yogendra P. Chaubey, for his kindness, constructive suggestions on my thesis and useful help in the completion of this dissertation.

I would also like to accord my heartfelt gratefulness to my teachers for their instruction and direct or indirect help during my studies at Concordia.

Last but not the least I would like to express my deep appreciation to my beloved parents and my dear wife for their understanding, encouragement and consistent support without which I may not be able to complete this difficult task.

Contents

List of Tables

List of Figures

List of Acronyms

AMISE	Asymptotic Mean Integrated Square Error
AIV	Asymptotic Integrated Variance
BCV	Biased Cross-Validation
CDF	Cumulative Distribution Function
CV	Cross-Validation
i.i.d.	independent and identically distributed
ISE	Integrated Squared Error
LB	Length Biased
MISE	Mean Integrated Squared Error
MRL	Mean Residual Life
MSE	Mean Squared Error
NPMLE	Nonparametric Maximum Likelihood Estimator
PDF	Probability Density Function
PWE	Poisson Weights Estimator
UCV	Unbiased Cross-Validation

List of Symbols

χ^2_α	Chi-Square Function with Parameter α
$\xrightarrow{a.s.}$	Convergence Almost Surely
$\xrightarrow{\mathcal{D}}$	Convergence in Distribution
$\exp(\cdot)$	Exponential Function
$\Gamma(\cdot)$	Gamma Function
$I\{\cdot\}$	Indicator Function
$IG(\cdot)$	Inverse Gauss Function
$\log(\cdot)$	Logarithm Function with Base e
$E(\cdot)$	Mathematical Expectation with Respect to the Density of Observed Sample.
$E_*(\cdot)$	Mathematical Expectation with Respect to Density Function $*$
$\max(\cdot)$	Maximum Function
$N(\cdot)$	Normal Distribution
$V(\cdot)$	Variance

Chapter 1

Introduction

1.1 Biased and Length Biased Data

In many statistical applications the observed random variable X_w may have the probability density function (pdf) given by [see Cox (1969) and Rao (1965)]

$$f_w(x) = \frac{w(x)f(x)}{\mu_w} \tag{1.1}$$

where μ_w is the expectation $E_f[w(X)]$, f being a probability density function as well. The distribution of X_w is referred to as the weighted distribution and $w(x)$ is called weight function. The data generated from model (1.1) is called *biased data.* The weight function $w(x)$, usually known, must be non-negative and must have finite expectation. Furthermore, it can be easily seen that for any other weight function $w'(x)$ that is proportional to $w(x)$, $f_w(x)$ and $f_{w'}(x)$ are identical. If $w(x) \neq 1$, $f_w(x)$, the probability law for recording random variable $X_w \sim f_w$, is proportional to $f(x)$ with a weight $w(x)$. However, the main objective concerns the density function $f(x)$. In such a case, the sampling procedure may involve some kind of selection scheme that is related to the weight function $w(x)$. Since the main objective of concern is the probability law $f(x)$, a natural question arises: How can we obtain the information of original random variable $X \sim f$ through the information of recorded random variable X_w? This is the main task of this thesis.

The earliest concept of distribution with weight can be found retrospectively in a classical paper of Fisher (1934). However, a more detailed account of weighted distributions was given by Rao (1965); see also Rao (1977) for a natural example of weighted binomial distribution with $w(x) = x$. Muttlak and Mcdonald (1990) discuss an example of sampling shrubs in the context of *ranked set sampling* where the probability of selection is proportional to the height of shrubs. Though the technique discussed in this thesis can be easily extended to the general weighted case, we concentrate on the special case $w(x) = x$.

Taking $w(x) = x$, (1.1) changes into

$$g(x) = \frac{xf(x)}{\mu} \tag{1.2}$$

where $\mu = E_f(X)$, where $E_f(.)$ refers to expectation with respect to the density f. When there is no ambiguity, $E()$ will refer to expectation with respect to the density g. This weighted distribution is well known as length biased or size biased distribution. The recorded samples generated from the biased distribution (1.2) are called length biased(LB) data. Since $w(x) = x$ is an increasing function of x, the greater the value of X, the better chance of X being observed.

Length biased data is generated naturally in many sampling problems. An interesting example of LB data called *Waiting time paradox* is given in Feller (1966). In this example, buses arrive in accordance with a Poisson process, the expected time between consecutive buses being 1. A passenger arrives at time t, independent of buses. What is the expectation $E(W_t)$ of the passenger's waiting time? Two contradictory answers are given:

(i) The lack of memory of the Poisson process implies that $E(W_t)$ should be independent of t, that is $E(W_t) = E(W_0) = 1$.

(ii) The time of the passenger's arrival is "chosen at random" in the interval between two consecutive buses, so for reasons of symmetry $E(W_t) = 1/2$.

Let us analyze this example precisely. We use X_w to denote the recorded length of time interval between two consecutive buses which covers the waiting passenger. For reasons of symmetry, the conditional expectation $E(W_t|X_w) = X_w/2$. In the solution (ii), it is taken for granted that X_w should have an exponential distribution with mean 1, that is $f_{X_w}(x) = e^{-x}$. Because of this, we have two contradictory answers. Actually, the length of the time interval X_w is recorded with a kind of "choice", that is we require the interval to cover the time t when the passenger arrives at the bus stop. It is obvious that, as it is said in Feller (1966), " a longer interval has a better chance to cover time t than a short one ". In his book, Feller (1966) gave the accurate density function: $f_{X_w} = xe^{-x}$. Then $E(X_w) = 2$, which is doubled, and $E(W_t) = 1$, just same as the solution (i) and paradox gets answered.

From the previous example, we can also see that if we ignore the bias effect, taking biased data as direct data, large mistakes can be made. Technically, the density function of direct data with $f(x) = e^{-x}$ is quite different from the density of LB data with $g(x) = xe^{-x}$ in the shape. So, in some cases, the bias effect can not be ignored. This example also tell us, if not disregarding the bias effect, sometimes we will use the observed samples which are with density $g(x)$ such that $g(0) = 0$ to restore the unobservable density $f(x)$ such that $f(0) \neq 0$. This is a main difficulty in dealing with estimation of density for LB data as well.

Actually, the field of biased or LB data is very wide in scope. The applications of biased data arise in diverse fields that include social sciences, physics, astronomy, market research, reliability, epidemiology, and many other fields. Cook and Martin (1974) took visibility bias into account in studying population density of wild animals. Partil (1984) and Patil *et al.* (1977, 1978) quoted several examples regarding biased data including those generated by PPS (probability proportional to size) sampling scheme, damage-model and sub-sampling. Eberhardt (1978) and Muttlak and McDonald (1990) studied the LB data generated from Line-Intercepts method in studying the density of shrub coverage. Simon (1980) considered the length biased sampling in etiologic studies. Nair and Wang (1989) claimed that size-bias must be considered in the studies of relation between the volume of oil under earth and some related variables. Klein and Sherman (1997) predicted market demand of new product using biased survey data. We can say that if there is sampling, biased data may emerge.

1.2 Nonparametric Functional Estimation for Biased Data

Nonparametric density estimation is a useful method of extracting information directly from data. In other words, a colorful metaphor is used to say that let the data "sing" for themselves. These methods are useful when we can not ascertain a useful parametric family for modeling the data. And the assumed parametric family may not be robust with respect to deviations from the model. As a result the area of nonparametric functional estimation including estimation of density and related functionals is one of the most active fields in statistical research branching in the area of biased data as well. The basic objective of the thesis is to explore various methods for nonparametric density estimation and their application in the area of biased data in general and LB data in particular.

In the area of functional estimation for LB data, the first stone is set by Cox. Cox (1969) suggested

$$F_n(x) = \frac{\sum_{i=1}^{n} \frac{1}{X_i} I\{X_i \leq x\}}{\sum_{i=1}^{n} \frac{1}{X_i}} \tag{1.3}$$

as the counterpart to the empirical distribution function for the LB data where X_i ($i = 1, \ldots, n$) are *i.i.d.* random variables with density $g(x)$ such that $E(X_1^{-1}) < \infty$. This estimator is a nonparametric maximum likelihood estimator (NPMLE) of distribution function under this situation [see Vardi (1982)]. Actually, (1.3) has some beneficial asymptotic properties. Under the condition $E(X_1^{-1}) < \infty$, using the Kolmogorov Strong Law of Large Numbers [see p. 251, Loève (1977)], we have as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} \frac{1}{X_i} I\{X_i \leq x\} \xrightarrow{a.s.} E\big(\frac{1}{X_1} I\{X_1 \leq x\}\big), \tag{1.4}$$

and

$$\frac{1}{n} \sum_{i=1}^{n} \frac{1}{X_i} \xrightarrow{a.s.} E\big(\frac{1}{X_1}\big) = \frac{1}{\mu}. \tag{1.5}$$

The right hand side (1.4) can be seen to be equal to $\frac{1}{\mu} F(x)$ because

$$E(\frac{1}{X_1} I\{X_1 \leq x\} = \int_0^x \frac{1}{t} g(t) dt = \frac{1}{\mu} \int_0^x f(t) dt.$$

Therefore

$$\frac{1}{n} \sum_{i=1}^{n} \frac{1}{X_i} I\{X_i \leq x\} \xrightarrow{a.s.} \frac{1}{\mu} F(x). \tag{1.6}$$

Since we can write $F_n(x) = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\}}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}}$, it follows from (1.6) and (1.5), $F_n(x) \xrightarrow{a.s.} F(x)$. Furthermore, due to the fact that $F_n(x)$ is nondecreasing, we can get the uniform strong consistency of $F_n(x)$, i.e.,

$$\sup_{x \in R^+} |F_n(x) - F(x)| \xrightarrow{a.s.} 0. \tag{1.7}$$

Furthermore, we can obtain the asymptotic normality property of (1.3), namely,

$$\sqrt{n}(F_n(x) - F(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x)), \tag{1.8}$$

where $\delta^2(x) = \mu[\int_0^x \frac{1}{t}f(t)dt - 2F(x)\int_0^x \frac{1}{t}f(t)dt + \bar{\mu}F^2(x)]$ and $\bar{\mu} = \int_0^\infty \frac{f(t)}{t}dt$. We will give the details of proof later.

The first kernel density estimator was given by Bhattacharyya *et al.* (1988). In their literature, they proposed the following kernel density estimator for $f(x)$.

$$f_{nB}(x) = (nx)^{-1}\hat{\mu}\sum_{i=1}^{n} k_h(x - X_i) \tag{1.9}$$

where $\hat{\mu} = n(\sum \frac{1}{X_i})^{-1}$ is the consistent estimator of μ proposed by Cox (1969) and $k_h(x) = h^{-1}k(h^{-1}x)$ [$k(\cdot)$ is a kernel function]. The strategy used here is very natural. It can be considered to use two steps to obtain it. First the observed samples are used to build an estimator of weighted density just same as in the procedure of building kernel density estimator with direct data. Then, according to LB model, the estimator obtained in the first step is adjusted to an estimator of unweighted density. However, this strategy is not very satisfactory. Jones (1991) found that, in some situations [f(0)=0], (1.9) will cause large bias near the point $x = 0$ [see Figure 1.1]. This huge bias mainly has two causes. One is that, when the kernel is symmetric as is usually the case in the usual kernel estimation approach, some weights will be assigned below 0 which causes $n^{-1}\sum_{i=1}^{n} k_h(x - X_i)$ [an estimator of $g(x)$] usually does not equal 0 at the boundary when sample is finite; the other is the term x^{-1}, which tends to infinity near the boundary. Combining these facts, Bhattacharyya *et al.* estimator blows up near the boundary under certain circumstances and its graph near the border looks like a vertical line [see Figure 1.1].

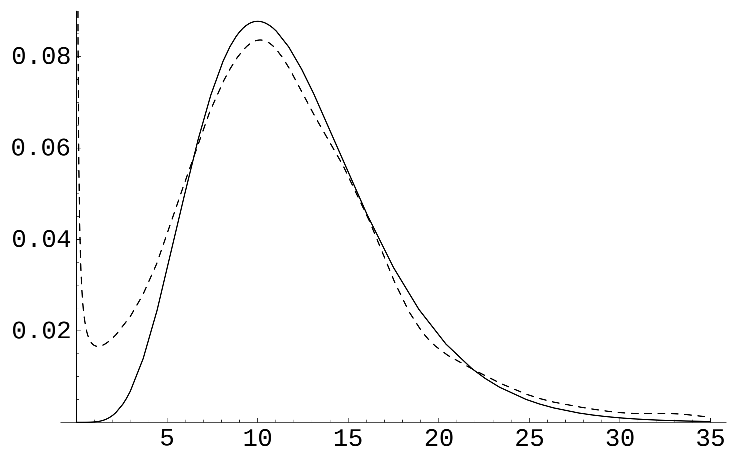


Figure 1.1: Plots of density function of χ^2_{12} and its Bhattacharyya *et al.* kernel estimator. Solid line represents true density and dash line represents the estimator.

Jones (1991) presented an alternative kernel density estimator based on the theory of Cox (1969). This alternative strategy is to smooth the distribution function estimator for $F(x) = \int_{-\infty}^{x} f(t)dt$, as given by Cox (1969) and then use its derivative as the smooth estimator of f. Jones (1991) used this alternative strategy of directly estimating $f(x)$, resulting in the following kernel density estimator:

$$f_{nJ}(x) = n^{-1}\hat{\mu}\sum_{i=1}^{n} X_i^{-1}k_h(x - X_i). \tag{1.10}$$

In his studies, Jones (1991) found that the integrated mean square error (IMSE) of (1.10) is asymptotically less than that of (1.9). Moreover, Wu and Mao (1996) showed that the mean squared error (MSE) of (1.10) is asymptotically lower than that of (1.9) under the minimax criterion.

However, if the kernel function is symmetric, the estimator (1.10) will assign some weights to the undesired region where the value of x is negative [see Figure 1.2]. [This also holds for the Bhattacharyya *et al.* (1988) estimator.] This may cause large bias in the neighborhood of the point $x = 0$.

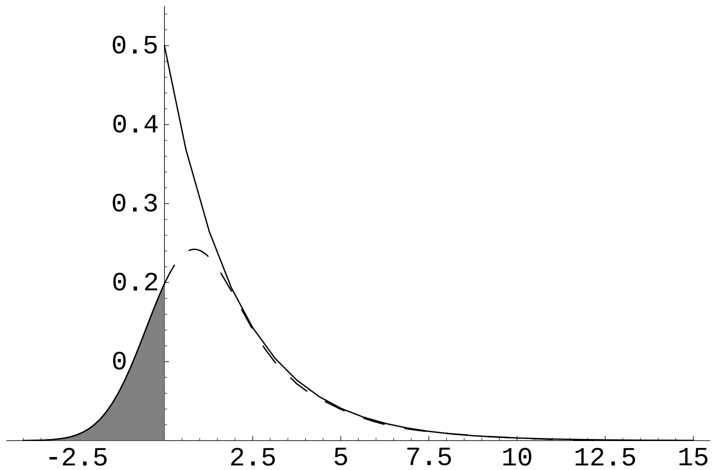


Figure 1.2: Plots of density function of χ_2^2 and its Jones kernel estimator with Normal kernel. Solid line represents true density and dash line represents the estimator.

Both of the previous two density estimators have a common defect at the boundary caused by symmetric kernels. This problem is not specific to LB data. It has been recognized in density estimation for nonnegative random variables using direct data [see Silverman (1986)]. In order to overcome this defect, many methods have been proposed particularly in recent years.

Motivated by Hille's approximation lemma [see Lemma 1.1], Chaubey and Sen (1996) proposed a smooth density estimator for nonnegative random variables.

Lemma 1.1 *If $u(x)$ is a bounded, continuous function on $\mathbf{R}^+$, then, as $\lambda \uparrow \infty$,*

$$e^{-\lambda x} \sum_{k \geq 0} u(k/\lambda)(x\lambda)^k/k! \to u(x)$$

uniformly in any finite interval $\mathbf{J}$ contained in $\mathbf{R}^+$.

Chen (2000) obtained two density estimators with asymmetric gamma kernels instead of traditional symmetric kernels. Using gamma kernels

$$K_{\rho_b(x),b}(t) = \frac{t^{\rho_b(x)-1}e^{-t/b}}{b^{\rho_b(x)}\Gamma(\rho_b(x))}, \tag{1.11}$$

the density estimator proposed by Chen (2000) has the form

$$\hat{f}(x) = n^{-1}\sum_{i=1}^{n} K_{\rho_b(x),b}(X_i') \tag{1.12}$$

where $\{X_i'\}_{i=1}^n$ denote *i.i.d.* regular direct data. In his literature, Chen (2000) gave two choices for $\rho_b(x)$. One is

$$\rho_b(x) = x/b + 1 \tag{1.13}$$

which leads to density estimator $\hat{f}_1(x)$; the other is

$$\rho_b(x) = \begin{cases} x/b & \text{if } x \geq 2b; \\ \frac{1}{4}(x/b)^2 + 1 & \text{if } x \in [0, 2b). \end{cases} \tag{1.14}$$

which leads to density estimator $\hat{f}_2(x)$. And he also showed that the $MISE$ of $\hat{f}_2$ is lower than that of $\hat{f}_1$.

Inspired by Chen's idea and using inverse Gaussian density

$$K_{IG(m,\lambda)}(y) = \frac{\sqrt{\lambda}}{\sqrt{2\pi y^3}} \exp\left(-\frac{\lambda}{2m}\left(\frac{y}{m} - 2 + \frac{m}{y}\right)\right), y > 0 \tag{1.15}$$

and reciprocal inverse Gaussian density

$$K_{RIG(m,\lambda)}(z) = \frac{\sqrt{\lambda}}{\sqrt{2\pi z}} \exp\left(-\frac{\lambda}{2m}\left(mz - 2 + \frac{1}{mz}\right)\right), z > 0 \tag{1.16}$$

as kernels, Scaillet (2004) proposed the following two density estimators

$$\hat{f}_{IG}(x) = n^{-1}\sum_{i=1}^{n} K_{IG(x,1/b)}(X_i') \tag{1.17}$$

and

$$\hat{f}_{RIG}(x) = n^{-1} \sum_{i=1}^{n} K_{RIG(1/(x-b),1/b)}(X_i'). \tag{1.18}$$

Using generalized Hille's lemma [see Lemma 1.2], Chaubey, Sen and Sen (2007) suggested a density estimator with asymmetric weights generated from gamma function, extending the estimator in Chaubey and Sen (1996).

Lemma 1.2 *Let $u(t)$ be any continuous and bounded function. $G_{x,n}$, $n = 1, 2, \ldots$ is a family of distributions with mean $\mu_n(x)$ and variance $h_n^2(x)$. Then we have as $\mu_n(x) \to x$ and $h_n(x) \to 0$*

$$\tilde{u}(x) = \int_{-\infty}^{\infty} u(t) dG_{x,n}(t) \to u(x).$$

The convergence is uniform in every subinterval in which $h_n(x) \to 0$ and $\tilde{u}(x)$ is uniformly continuous.

Although Chaubey, Sen and Sen (2007) and Chen (2000) both use asymmetric gamma density function as kernels, the density estimators proposed by them are quite different in form. However, they both can be obtained by using generalized Hille's lemma in two different ways. The density estimators proposed by Chaubey Sen and Sen (2007) are the derivatives of smooth estimators obtained by smoothing empirical function using Hille's lemma; the density estimators in Chen (2000) and Scaillet (2004) can also be obtained by using generalized Hille's lemma to smooth underlying density.

Besides the literature we mentioned above, there are also many other contributions made by statisticians to functional estimation for biased data. Vardi (1982) obtained the nonparametric maximum likelihood estimator for unweighted distribution function based on two sample sets, one from unweighted distribution, the other from weighted distribution. Cox's estimator, as a NPMLE for unweighted distribution function obtained only by weighted sample set, is a special case that considered by him. Vardi (1985) generalized his model to selection bias model. Wu (1996) proposed a nonparametric maximum likelihood smooth estimator for biased data using kernel method. Jones and Kaunamuni (1997) used fourier series method to estimate unweighted density and they found that their estimator perform better than those estimators in Bhattacharyya *et al.* (1988) and Jones (1991). Lloyd and Jones (2000) proposed a nonparametric density estimator for biased data with unknown weight function. In their studies, the weight function is treated as a selection probability. A cross-validation method for selecting smoothing parameter in kernel density estimator with selection biased data was proposed by Wu (1997). Winter and Földes (1988) derived an Kaplan-Meier type estimator for censored biased data. Uña-Álarez (2002) studied its asymptotic properties.

1.3 Motivation of the Estimators

The examples of biased data present themselves mostly as non-negative data where the traditional kernel methods of density estimators may not be appropriate. Recently, as mentioned previously, there have been significant advances in the area of density estimation for non-negative data. We would like to incorporate the new estimators for biased data in this thesis that is mainly motivated by the use of Hille's lemma and Cox's proposal for estimating the distribution function for the biased data. Chaubey and Sen (1996) proposed a smooth estimator of the distribution function for the i.i.d. case using the Hille's lemma that incorporates Poisson weights for functional smoothing of non-negative functions. The empirical distribution function used for the i.i.d. case may be replaced by Cox's (1969) estimator of the distribution function for the LB data. The recent generalization [Chaubey, Sen and Sen (2007)] of Chaubey and Sen (1996), using weights generated by non-negative asymmetric kernels such as gamma kernels, may be adapted to the case of LB data as well.

1.4 Objectives

Since the LB data are commonly non-negative, the use of traditional kernel estimator may not be appropriate; it may cause large bias at the boundary. It is expected that the methods developed in Chaubey and Sen (1996) and in Chaubey, Sen and Sen (2007) can be satisfactorily adapted for the LB case and thus we have chosen to study these in the present thesis. Actually, for LB data, there are mainly two strategies to estimate unweighted density. One is, starting from Cox's estimator, to directly estimate unweighted density [as in Jones (1991)]; the other is to estimate weighted density first and adjust it to estimate the original density [as in Bhattacharyya *et al.* (1988)]. Is there a relatively better strategy or do the two strategies produce similar results? We plan to find a answer to this question. In order to compare the proposed estimators, we will simulate for some standard distributions and use the mean integrated squared error (MISE) as a global measure of estimator's behavior and mean square error (MSE) as a local indicator of estimator's performance. Comparison between our proposed density estimators and other density estimators with asymmetric kernels will be carried out as well. Our plan includes investigating estimators of other functions, such as, distribution function estimator, hazard function estimator and mean residual life function estimator also.

1.5 Organization of the Thesis

This thesis is organized as follows. In Chapter 2, based on Cox's estimator for distribution function, we propose some distribution and density estimators with Poisson weights or asymmetric weights and study their asymptotic properties. Motivated by Chen (2000) and Scaillet (2004), we also obtain some density estimators with asymmetric kernels for LB data which are different from our proposed estimators in form. An alternative method starting from the usual empirical distribution function based on observed samples is used in Chapter 3 to find some new density and distribution function estimators with Poisson weights or asymmetric weights. Asymptotic properties of these estimators are investigated as well. Through extensive simulation for some standard distributions, Chapter 4 will show how the smoothing parameters in density estimators are selected and how each density estimator performs globally and locally. We dedicate Chapter 5 to the estimators of some functionals related to density and distribution functions and their asymptotic properties. These functionals include hazard function and mean residual life function. Dependency or censoring, as some situations frequently happening in statistical applications, may emerge with biased data at the same time. In future, we are planning to consider these situations as well. The details are contained in Chapter 6.

Chapter 2

Smooth Estimators of Density and Distribution Functions Based on Cox's Estimator

2.1 Introduction

In this chapter, we will use Cox's estimator (F_n) of the distribution function proposed for the LB data to obtain some smooth estimators of the underlying true density and the corresponding distribution function. Motivated through Hille's lemma and Cox's proposal, it is easy to obtain smooth estimator of a distribution function in the length bias case similar to that obtained by Chaubey and Sen (1996) for the *i.i.d.* direct data. Since the smooth estimator is differentiable, it is reasonable to use its derivative as an estimator of the underlying density. We will consider Hille's lemma that uses Poisson weights as well as its generalized version that uses weights generated by asymmetric kernels. Thus, based on F_n, we get two kinds of density estimators, the first using Poisson weights and the other using weights from asymmetric kernels. In Section 2.2, we will study theoretical properties of smooth estimators with Poisson weights, such as strong consistency and asymptotic normality. The smooth estimators include distribution and density estimator. Similar theoretical properties of estimators with asymmetric kernels are investigated in Section 2.3. In this section, a perturbation and boundary correction are applied to density estimator. They will effectively enhance the accuracy of density estimator under certain circumstances. In Chaubey *et al.* (2010) extensive simulation studies have been carried out to compare the density estimators using Poisson weights with kernel estimators proposed by Bhattacharyya *et al.* (1988) and Jones (1991). The study in the above paper demonstrates that the kernel estimators with symmetric kernels do not perform very well for LB data. In order to make a fair comparison between our proposed estimators and other estimators [see Chapter 4], we only consider density estimators with asymmetric kernels in this thesis. Therefore, besides our proposed estimators, we will apply the idea of Chen (2000) and Scailltet (2004) also to obtain some other density estimators with asymmetric kernels in Section 2.4.

2.2 Estimators of Distribution and Density Functions with Poisson Weights

2.2.1 Smooth Estimator of Cumulative Distribution Function

The raw estimator (1.3) [Cox's estimator for distribution function] is a step function and not differentiable. In order to obtain a smooth estimator with differentiable property, we apply Lemma 1.1 by replacing $u(\cdot)$ with $F_n(\cdot)$. Since $F_n(.)$ is not continuous function, this lemma is not directly applicable, but may be considered as a motivation for the suggested estimator. As we investigate the convergence properties of the proposed estimator, it becomes clear that it provides an stochastic approximation to the integral in Lemma 1.1 that replaces $u(x)$ by $F(x)$, which is a continuous function. The combination of Cox's estimator and Lemma 1.1 results in the following smooth estimator of distribution function, namely,

$$\widetilde{F}_n(x) = \sum_{k \geq 0} p_k(x\lambda_n) F_n(k/\lambda_n) \tag{2.1}$$

where $p_k(u) = \frac{u^k}{k!} e^{-u}$ and λ_n such that, as $n \to \infty$, $\lambda_n \to \infty$. Actually, λ_n controls the smoothness of the smooth estimator. A stochastic choice of λ_n is proposed by Chaubey and Sen (1996, 1998) as follows.

$$\lambda_n = \begin{cases} \dfrac{n}{\max\{X_1, \ldots, X_n\}} & \text{if } X_1 \text{ has an infinite support} \\ \dfrac{n}{X_{n-r_n+1:n} \log\log n} & \text{if } X_1 \text{ has a finite support} \end{cases}$$

where $r_n = o(\log\log n)$, provided that $E(X_1) < \infty$. Chaubey and Sen (2009) provide a more comprehensive numerical study for the choice of λ_n in the context of density estimation for the *i.i.d.* data. We use their approach for the LB data while discussing the smooth density estimation later in this section.

Similar asymptotic results as given in Chaubey and Sen (1996) for the smooth estimator $\tilde{F}_n(x)$ in the non-weighted case can be established. These are given in the following theorems. First we establish the uniform strong consistency.

Thoerem 2.1 *If* $0 < E(X_1^{-1}) < \infty$, *$F(x)$ is continuous (a.e.) and* $\lambda_n \to \infty$ *,then, as* $n \to \infty$*,*

$$\|\widetilde{F}_n(x) - F(x)\| = \sup_{x \in R^+} \{|\widetilde{F}_n(x) - F(x)|\} \xrightarrow{a.s.} 0$$

Remark 2.1: In Theorem 3.1 of Chaubey and Sen (1996), additional condition on λ_n, namely that $n^{-1}\lambda_n \to 0$ is assumed that is not required for the above theorem to hold. It may be noted that the estimator in Chaubey and Sen (1996) uses truncated Poisson weights, where such a condition may be necessary.

Next, we discuss the closeness of (2.1) to the raw estimator $F_n(x)$. This also helps in establishing the the asymptotic distribution of the smooth estimator. Along the lines of the proof of Theorem 3.2 in Chaubey and Sen (1996) using Lemma 2.1 with $b_n = n^{-\frac{1}{2}}(\log n)^{\frac{1+\theta}{2}}$ [see also the treatment in Sen (1984)], we establish the following theorem.

Thoerem 2.2 *If $E(X_1^{-2}) < \infty$, $\lambda_n \to \infty$, and $n^{-1}\lambda_n \to 0$, $f(x)$ is absolutely continuous with bounded derivative $f'(x)$ on $\mathbf{R}^+$, then for some $\delta > 0$, as $n \to \infty$,*

$$||\widetilde{F}_n(x) - F_n(x)|| = O(n^{-3/4}(\log n)^{1+\delta}) \text{ a.s. } \forall x \in \mathbb{R}^+. \tag{2.2}$$

Note that

$$\sqrt{n}(\tilde{F}_n(x) - F(x)) = \sqrt{n}(F_n(x) - F(x)) + \sqrt{n}(\tilde{F}_n(x) - F_n(x))$$

and from Theorem 2.2,

$$\sqrt{n}(\tilde{F}_n(x) - F_n(x)) = O(n^{-1/4}(\log n)^{1+\delta}, \text{ a.s..}$$

Then we can see that the asymptotic law for $\tilde{F}_n(x)$ is same as that of $F_n(x)$ under the condition of Theorem 2.2. Therefore to study the asymptotic distribution of $\tilde{F}_n(x)$, we just need to find out the asymptotic distribution of $F_n(x)$. We can write $F_n(x)$ as

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

By the strong law of large numbers, we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\} \stackrel{a.s.}{\to} F(x)/\mu$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} \stackrel{a.s.}{\to} 1/\mu.$$

So we can expand $F_n(x)$ as

$$\begin{aligned}
F_n(x) &= F(x) + \mu\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\} - F(x)/\mu\right] - \mu F(x)\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}\right] \\
&\quad + O\Big((\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\} - F(x)/\mu)^2 \\
&\quad + (\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\} - F(x)/\mu)(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}) \\
&\quad + (\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu})^2\Big) \\
&= F(x) + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mu}{X_i}I\{X_i \le x\} - \frac{\mu F(x)}{X_i}\right) \\
&\quad + O\Big((\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\} - F(x)/\mu)^2 \\
&\quad + (\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}I\{X_i \le x\} - F(x)/\mu)(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}) + (\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu})^2\Big)(a.s.).
\end{aligned}$$

Note that since the last term in above equation has an order $o_p(\frac{1}{\sqrt{n}})$, the asymptotic distribution of $\sqrt{n}F_n(x)$ is same as that of

$$\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mu}{X_i}I\{X_i \le x\} - \frac{\mu F(x)}{X_i}\right)\right]. \tag{2.3}$$

Therefore, to obtain the asymptotic distribution of $F_n(x)$, it is sufficient to consider the asymptotic distribution of (2.3). For the term (2.3), we have

$$E\left(\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mu}{X_i}I\{X_i \le x\} - \frac{\mu F(x)}{X_i}\right)\right]\right) = 0$$

and

$$V\left(\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mu}{X_i}I\{X_i \le x\} - \frac{\mu F(x)}{X_i}\right)\right]\right) = \delta^2(x)$$

where $\delta^2(x) = \mu[\int_0^x \frac{1}{t} f(t)dt - 2F(x) \int_0^x \frac{1}{t} f(t)dt + \bar{\mu}F^2(x)]$ and $\bar{\mu} = \int_0^\infty \frac{f(t)}{t} dt$. Then we have

$$\sqrt{n}(F_n(x) - F(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x)),$$

Therefore we have following theorem.

Thoerem 2.3 *If $E(X_1^{-2}) < \infty$, $\lambda_n \to \infty$, and $n^{-1}\lambda_n \to 0$, $f(x)$ is absolutely continuous with bounded derivative $f'(x)$ on $\mathbf{R}^+$, then, as $n \to \infty$,*

$$\sqrt{n}(\widetilde{F}_n(x) - F(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x)),$$

specifically

$$\delta^2(x) = \mu[\int_0^x \frac{1}{t} f(t)dt - 2F(x) \int_0^x \frac{1}{t} f(t)dt + \bar{\mu}F^2(x)]$$

where $\bar{\mu} = E_f(\frac{1}{X})$.

From Theorem 2.1 and 2.3, we see that the some of the key asymptotic properties of the raw estimators (1.3) may be exhibited also for the smooth estimator (2.1).

2.2.2 Smooth Density Estimator

Since $\widetilde{F}_n(x)$ converges strongly to $F(x)$, it is reasonable to believe that their derivatives should be close. Since we have

$$\frac{dp_k(\lambda x)}{dx} = -\lambda[p_k(\lambda x) - p_{k-1}(\lambda x)],$$

for $k \geq 0$, where we interpret $p_{-1}(.) = 0$, the derivative of $\tilde{F}_n(x)$ is given by

$$\frac{d\tilde{F}_n(x)}{dx} = -\lambda_n \left[\sum_{k \geq 0} p_k(\lambda_n x) F_n\left(\frac{k}{\lambda_n}\right) - \sum_{k \geq 1} p_{k-1}(\lambda_n x) F_n\left(\frac{k}{\lambda_n}\right) \right].$$

This simplifies to

$$\frac{d\tilde{F}_n(x)}{dx} = \lambda_n \sum_{k \geq 0} p_k(\lambda_n x) \left[F_n\left(\frac{k+1}{\lambda_n}\right) - F_n\left(\frac{k}{\lambda_n}\right) \right].$$

Hence, our proposed smooth density estimator is

$$\tilde{f}_n(x) = \lambda_n \sum_{k\geq 0} p_k(\lambda_n x)\left[F_n\left(\frac{k+1}{\lambda_n}\right) - F_n\left(\frac{k}{\lambda_n}\right)\right]. \tag{2.4}$$

as the smooth estimator of density $f(x)$. We also obtain the asymptotic properties of (2.4) as follows.

2.2.2.1 Asymptotic Properties of $\tilde{f}_n(x)$

The strong consistency of $\tilde{f}_n(.)$ is provided in the following theorem. Note that the moment condition used in this theorem implies the boundedness of the density $f(x)$.

Thoerem 2.4 *If $E(X_1^{-2}) < \infty$, $f'(x)$ is bounded on $\mathbf{R}^+$ and $\lambda_n = O(n^\alpha)$ for some $0 < \alpha < 1$, then, as $n \to \infty$,*

$$\|\tilde{f}_n(x) - f(x)\| \xrightarrow{a.s.} 0$$

In order to obtain the weak convergence of $\tilde{f}_n$, we need $f'(x)$ to satisfy a Lipschitz order α condition. That is, for some $\alpha > 0$, there exits a finite positive K, such that

$$|f'(s) - f'(t)| \leq K|s-t|^\alpha, \text{ for every } t,\ s \in \mathbf{R}^+. \tag{2.5}$$

If $\lambda_n = O(n^{2/5})$, $MSE(\tilde{f}_n(x))$ achieve the lowest order [see Remark 2.2]. We establish the following representation theorem.

Thoerem 2.5 *If $E(X_1^{-2}) < \infty$, $\lambda_n = O(n^{2/5})$(nonstochastic) and (2.5) holds, then, for a compact set $\mathcal{C} \subset \mathbf{R}^+$,*

$$\left\{\left(n^{2/5}[\tilde{f}_n(x) - f(x)] - \frac{1}{2\delta^2} f'(x)\right), x \in \mathcal{C}\right\} \xrightarrow{\mathcal{D}} \text{ Gaussian process}$$

with mean zero and covariance function $\gamma_x^2 \delta_{xt}$ where $\gamma_x^2 = \dfrac{\mu}{2}(\pi x^3)^{-1/2} f(x)\delta$, $\delta_{xt} = 0$ for $x \neq t$ and 1 for $x = t$ and $\delta = \lim_{n\to\infty}(n^{-1/5}\lambda_n^{1/2})$

Remark 2.2: In order to understand the order of bias and MSE of the density estimator, we see that under condition (2.5) for $\lambda_n = cn^h$ using the steps in proofs of Theorems 2.4 and 2.5, we have

$$Bias^2(\tilde{f}_n(x)) \approx c^{-2}(f'(x)/2)^2 n^{-2h} \tag{2.6}$$

and

$$V(\tilde{f}_n(x)) \approx \frac{\mu}{2}\sqrt{\frac{c}{\pi x^3}} f(x) n^{\frac{h}{2}-1}, \tag{2.7}$$

then we have

$$MSE(\tilde{f}_n(x)) \approx c^{-2}(f'(x)/2)^2 n^{-2h} + \frac{\mu}{2}\sqrt{\frac{c}{\pi x^3}} f(x) n^{\frac{h}{2}-1} \tag{2.8}$$

When $\lambda_n = cn^{2/5}$, (2.8) achieve the order $O(n^{-4/5})$, which is same as classical kennel estimators. In order to achieve the same order $O(n^{-4/5})$, Poisson weights estimator just need the information of first derivative of density. However, kennel estimators require the existence of second derivative [see Jones (1996)].

2.2.2.2 Proof of Theorems

First, we will introduce an important lemma, which plays a critical role in the proof of strong consistency of $\tilde{f}_n(x)$.

Lemma 2.1 *If $E(X_1^{-2}) < \infty$, $f'(t)$ is bounded on $\mathbf{R}^+$ and $b_n \to 0$, then for a sequence $\{b_n\}_{n\geq 1}$ such that $0 < b_n^{-1} < O(n^{1-\gamma})(0 < \gamma < 1)$,*

$$\sup_{t\in R^+} \sup_{|\beta|\leq b_n} \{|F_n(t+\beta) - F_n(t) - F(t+\beta) + F(t)|\} = O(b_n^{\frac{1}{2}} n^{-\frac{1}{2}} (\log n)^{1+\theta}) \ \ a.s.$$

where $\theta(> 0)$ is arbitrary.

In order to prove Lemma 2.1, we need the following two lemmas. For convenience, we denote

$$U_i(t,\beta) = \frac{\mu}{X_i} I\big\{ \min(t,t+\beta) < X_i \le \max(t,t+\beta)\big\} - |F(t+\beta) - F(t)| \quad (i = 1,\ldots,n) \tag{2.9}$$

Lemma 2.2 *If $E(X_1^{-2}) < \infty$, then, for any $t \ge 0$ and $t+\beta \ge 0$,*

$$\frac{1}{n}\sum_{i=1}^{n} U_i(t,\beta) = o(n^{-1/2}(\log n)^{(1+\theta)/2}) \ \ a.s. \tag{2.10}$$

Proof of Lemma 2.2: In order to prove the lemma, we need the Kolmogorov's Proposition A in M. Loève (p. 250). We state the proposition here.

Proposition: *If the integrable r.v.'s X_n are independent, then $\sum \frac{\sigma^2(X_n)}{a_n^2} < \infty$, $a_n \uparrow \infty$, entails $\frac{S_n - ES_n}{a_n} \xrightarrow{a.s.} 0$. where $S_n = \sum_{i=1}^n X_i$ and $\sigma^2(X_i)$ means the variance of X_i.*

Under the assumption $E(X_1^{-2}) < \infty$, for any $t \ge 0$ and $t+\beta \ge 0$, we have that

$$\sum_{n=1}^{\infty} \frac{\sigma^2(U_n(t,\beta))}{\big(n^{1/2}(\log n)^{(1+\theta)/2}\big)^2} \le \sum_{n=1}^{\infty} \frac{E(X_n^{-2})}{\big(n^{1/2}(\log n)^{(1+\theta)/2}\big)^2} < \infty.$$

By the Proposition of Kolmogorov , we have

$$\frac{\sum_{i=1}^n U_i(t,\beta)}{n^{1/2}(\log n)^{(1+\theta)/2}} \xrightarrow{a.s.} 0. \tag{2.11}$$

It is obvious that

$$\frac{1}{n}\sum_{i=1}^{n} U_i(t,\beta) = n^{-1/2}(\log n)^{(1+\theta)/2} \frac{\sum_{i=1}^n U_i(t,\beta)}{n^{1/2}(\log n)^{(1+\theta)/2}}. \tag{2.12}$$

By (2.11) and (2.12), we obtain the desired result.

Lemma 2.3 *If $E(X_1^{-2}) < \infty$ and $f'(t)$ is bounded on $\mathbf{R}^+$, then there exits $d > 0$ such that, for any $t \geq 0$, $0 < b_n^{-1} < O(n^{1-\gamma})(0 < \gamma < 1)$, $-b_n < \beta < b_n$, $D = b_n^{\frac{1}{2}} n^{\frac{1}{2}} (\log n)^{1+\theta}$, we have*

$$P\big\{|\sum_{i=1}^{n} U_i(t,\beta)| > 2dD\big\} \leq O(n^{-4}). \tag{2.13}$$

The order $O(n^{-4})$ does not depend on t and β.

Proof of Lemma 2.3: First we should verify several facts. For any $\delta > 2$, we have

$$E\big|\sum_{i=1}^{p} U_i(t,\beta)\big|^{\delta} \leq (p(\log n)^{1+\theta})^{\delta/2}, \tag{2.14}$$

since

$$E\big|\sum_{i=1}^{p} U_i(t,\beta)\big|^{\delta} = p^{\delta} E\big|\frac{1}{p}\sum_{i=1}^{p} U_i(t,\beta)\big|^{\delta}$$

and, by Lemma 2.2,

$$\frac{1}{p}\sum_{i=1}^{p} U_i(t,\beta) = o(p^{-1/2}(\log p)^{(1+\theta)/2}) \ \ a.s.$$

At the same time, we have

$$\begin{aligned}
E(x_1(t,\beta))^2 &= E\big(\frac{\mu^2}{X_1^2} I\big\{\min(t,t+\beta) < X_i \leq \max(t,t+\beta)\big\}\big) \\
&\quad -|F(t+\beta) - F(t)|^2 \\
&= \big|\int_t^{t+\beta} \frac{\mu f(x)}{x} dx\big| - |F(t+\beta) - F(t)|^2 \qquad (2.15)\\
&= O(|\beta|). \qquad (2.16)
\end{aligned}$$

The conclusion of the last step follows because $E(X_1^{-2}) < \infty$ and that $f'(x)$ is bounded. Since, $|f(x)/x| = |f'(\eta)| < M$, ($\eta \in (0,x)$ and M is finite), the first term of (2.15) has an order $O(|\beta|)$. And since $f(x)$ is bounded, the second term of (2.15) has an order $O(\beta^2)$.

So, using (2.16) and the independence of $U_i(t,\beta)(i=1,\ldots,n)$, we can also establish (2.7) in Lemma 2.1 of Babu and Singh (1978), that is

$$E(\xi_1^2) \le O(pb_n). \tag{2.17}$$

Substituting (2.4) in Lemma 2.1 of Babu and Singh (1978) with (2.14), taking $\delta = 60/\gamma$ and $p = [n^{\gamma/2}]$, and following the proof of Lemma 2.1 of Babu and Singh (1978), we can obtain the result.

Remark 2.3: The second term $\exp(-8D^2n^{-1}b_n^{-1})$ in (2.1) of Babu and Singh(1978) disappears in our inequality, because under our choice of D, this term is much smaller than $O(n^{-4})$.

Proof of Lemma 2.1 : Let

$$H_n(t,\beta) = F_n(t+\beta) - F_n(t) - F(t+\beta) + F(t).$$

Since $F_n(t+\beta) - F_n(t)$ can be expanded as

$$\begin{aligned} F_n(t+\beta) - F_n(t) &= \frac{1}{n}\sum_{i=1}^n \frac{\mu}{X_i} I\{t < X_i \le t+\beta\} \\ &\quad -[F(t+\beta) - F(t)]\Big(\frac{1}{n}\sum_{i=1}^n \frac{\mu}{X_i} - 1\Big) \\ &\quad +o\Big([F(t+\beta) - F(t)]\Big(\frac{1}{n}\sum_{i=1}^n \frac{\mu}{X_i} - 1\Big)\Big) \ \ a.s., \end{aligned} \tag{2.18}$$

we have

$$|H_n(t,\beta)| \le J_{n1}(t,\beta) + J_{n2}(t,\beta) + o(J_{n2}(t,\beta)) \ \ a.s. \tag{2.19}$$

where

$$J_{n1}(t,\beta) = \frac{1}{n}\Big|\sum_{i=1}^n U_i(t,\beta)\Big| \tag{2.20}$$

and

$$J_{n2}(t,\beta) = \Big|[F(t+\beta) - F(t)]\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} - 1\Big)\Big|. \tag{2.21}$$

For (2.20), first we consider that t is fixed. Using Lemma 2.3, following the proof of Lemma 1 of Bahadur (1966), we can claim that

$$\sup_{|\beta|\le b_n}\{|J_{n1}(t,\beta)|\} = O(b_n^{\frac{1}{2}}n^{-\frac{1}{2}}(\log n)^{1+\theta})\ \ a.s.$$

Furthermore, since $O(b_n^{\frac{1}{2}}n^{-\frac{1}{2}}(\log n)^{1+\theta})$ does not depend on t and $f'(t)$ is bounded, using the same technique as in Sen and Ghosh (1971), we can extend the result for t to the whole real line, that is

$$\sup_{t\in R^+}\sup_{|\beta|\le b_n}\{|J_{n1}(t,\beta)|\} = O(b_n^{\frac{1}{2}}n^{-\frac{1}{2}}(\log n)^{1+\theta})\ \ a.s. \tag{2.22}$$

At the same time, in Lemma 2.2, let $t = 0$ and $\beta \to +\infty$, then we have

$$\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} - 1\Big) = o(n^{-1/2}(\log n)^{(1+\theta)/2})\ \ a.s. \tag{2.23}$$

Since $f(t)$ is bounded(because $E(X_1^{-2}) < \infty$) as well, we have

$$\sup_{t\in R^+}\sup_{|\beta|\le b_n}\Big|F(t+\beta) - F(t)\Big| = O(b_n). \tag{2.24}$$

For (2.21), by (2.23) and (2.24), we have

$$\sup_{t\in R^+}\sup_{|\beta|\le b_n}\{|J_{n2}(t,\beta)|\} = o(b_n n^{-\frac{1}{2}}(\log n)^{(1+\theta/2)})\ \ a.s. \tag{2.25}$$

By (2.19), (2.22) and (2.25), we can establish the Lemma 2.1.

After all of these preparations, we can prove the Theorem 2.4.

Proof of Theorem 2.4: By the proof of Theorem 4.1 of Chaubey and Sen (1996), we just need to show that, when t belongs to some finite interval [0,C], we have (2.4), since we can deliberately choose C such that when t belongs to interval (C,$+\infty$), $\tilde{f}_n(t)$ and $f(t)$ can both be made sufficiently small.

We can write

$$\begin{aligned}
\tilde{f}_n(x) &= \lambda_n\{\sum_{k\geq 0} p_k(x\lambda_n)[F(\frac{k+1}{\lambda_n}) - F(\frac{k}{\lambda_n})] \\
&\quad + \sum_{k\geq 0} p_k(x\lambda_n)[F_n(\frac{k+1}{\lambda_n}) - F_n(\frac{k}{\lambda_n}) - F(\frac{k+1}{\lambda_n}) + F(\frac{k}{\lambda_n})]\} \\
&= T_{n1}(x) + T_{n2}(x). \qquad (2.26)
\end{aligned}$$

Using Lemma (2.1) by taking $b_n = 1/\lambda_n$, we have

$$\begin{aligned}
\sup_{k\geq 0}\{|F_n(\frac{k+1}{\lambda_n}) - F_n(\frac{k}{\lambda_n}) &- F(\frac{k+1}{\lambda_n}) + F(\frac{k}{\lambda_n})|\} \\
&= O(\lambda_n^{-1/2} n^{-1/2} (\log n)^{1+\theta}) \ \ a.s. \qquad (2.27)
\end{aligned}$$

By (2.27) and the fact that $\sum_{k\geq 0} p_k(x\lambda_n) = 1$, we have

$$\sup_{x\in R^+}\{|T_{n2}(x)|\} = O(\lambda_n^{1/2} n^{-1/2} (\log n)^{1+\theta}) \ \ a.s. \qquad (2.28)$$

which tends to 0 almost surely as $n \to \infty$ provided that $\lambda_n = O(n^\alpha)(0 < \alpha < 1)$.

At the same time, according to the proof of Theorem 4.1 of Chaubey and Sen (1996), under the assumption of boundedness of $f'(x)$, we have

$$\sup_{t\in[0,C]}\{|T_{n1}(x) - f(x)|\} \to 0 \ \ a.s. \qquad (2.29)$$

By (2.28) and (2.29), we obtain the theorem. The proof is complete.

Proof of Theorem 2.5: By (2.5), we have

$$\tilde{f}_n(x) = f(x) + \frac{1}{2\lambda_n} f'(x) + T_{n2}(x) + O(\lambda_n^{-1-\alpha}). \qquad (2.30)$$

Using Taylor's expansion which is similar to (2.18), we can write

$$
\begin{aligned}
T_{n2}(x) &= \lambda_n \sum_{k\geq 0} p_k(x\lambda_n)\big\{\big(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} I\{\frac{k}{\lambda_n} < X_i \leq \frac{k+1}{\lambda_n}\} - [F(\frac{k+1}{\lambda_n}) - F(\frac{k}{\lambda_n})]\big)\big\} \\
&\quad -\lambda_n \sum_{k\geq 0} p_k(x\lambda_n)\big\{[F(\frac{k+1}{\lambda_n}) - F(\frac{k}{\lambda_n})]\big(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} - 1\big)\big\} + o(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} - 1) \\
&= T_{n3}(x) - T_{n4}(x) + o(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} - 1) \;\; a.s. \qquad (2.31)
\end{aligned}
$$

For the leading term $T_{n3}(x)$, following the proof of Theorem 4.2 of Chaubey and Sen (1996), we can show that

$$V(T_{n3}(x)) \approx \frac{\mu}{2}(\pi x^3)^{-1/2} f(x)(\lambda_n^{1/2}/n) \qquad (2.32)$$

and, for $s \neq t$, as $n \to \infty$,

$$Cov[T_{n3}(s), T_{n3}(t)] = O(\frac{1}{n}). \qquad (2.33)$$

Moreover, since $T_{n4}(x) = O(\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{X_i} - 1) = o(n^{-1/2}(\log n)^{(1+\theta)/2})$, the order of $T_{n2}(x)$ is determined by the order of $T_{n3}(x)$.

From (2.30), we can see that the asymptotic normality of $T_{n2}(x)$ leads to the asymptotic normality of $\tilde{f}_n(x)$. by proper choice of λ_n. By (2.30), (2.31), (2.32) and (2.33), following the proof of Theorems 4.1 and 4.2 of Chaubey and Sen (1996), we can complete the proof of the theorem.

2.3 Estimators of Distribution and Density Functions with Asymmetric Kernels

2.3.1 Smooth Estimator of Distribution Function with Asymmetric Kernels

As in Chaubey, Sen and Sen (2007), let $Q_{v_n}(x)$ be a family of distributions on $[0, \infty)$ with mean 1 and variance v_n^2 where $v_n \to 0$ as $n \to \infty$. Substituting $F_n(t)$ and $Q_{v_n}(t/x)$ for $u(t)$ and $G_{x,v}(t)$ in Lemma 1.2 respectively, we have the following smooth estimator of $F(x)$:

$$\widetilde{F}_n^+(x) = \int_0^\infty F_n(t) dQ_{v_n}(t/x). \tag{2.34}$$

An alternative formula of (2.34) is given by

$$\widetilde{F}_n^+(x) = 1 - \frac{\sum_{i=1}^n \frac{1}{X_i} Q_{v_n}(\frac{X_i}{x})}{\sum_{i=1}^n \frac{1}{X_i}}, \tag{2.35}$$

where Q_{v_n} is a family of distributions as described earlier.

2.3.1.1 Asymptotic Properties

By the uniform strong convergence of (1.3)

$$\sup_{x \geq 0} |F_n(x) - F(x)| \xrightarrow{a.s.} 0$$

and the form of $\widetilde{F}_n^+(x)$ (2.34), it is easy to obtain the uniform strong convergence of $\widetilde{F}_n^+(x)$ as follows.

Thoerem 2.6 *If* $0 < E(X_1^{-1}) < \infty$ *and* $F(x)$ *is continuous (a.e.), then, as* $v_n \to 0$,

$$\|\widetilde{F}_n^+(x) - F(x)\| = \sup_{x \in R^+} \{|\widetilde{F}_n^+(x) - F(x)|\} \xrightarrow{a.s.} 0$$

The asymptotic normality of $\widetilde{F}_n^+(x)$ is given by the following theorem.

Thoerem 2.7 *If $E(X_1^{-2}) < \infty$, $\sqrt{n}v_n^2 \to 0$, $f(x)$ is absolutely continuous with bounded derivative $f'(x)$ on $\mathbf{R}^+$, then, as $n \to \infty$,*

$$\sqrt{n}(\widetilde{F}_n^+(x) - F(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x))$$

where

$$\delta^2(x) = \mu[\int_0^x \frac{1}{t} f(t)dt - 2F(x)\int_0^x \frac{1}{t} f(t)dt + \bar{\mu}F^2(x)]$$

where $\bar{\mu} = E_f(\frac{1}{X_1})$.

Proof: First, by (2.35), we write

$$\widetilde{F}_n^+(x) = \frac{\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}[1 - Q_{v_n}(\frac{X_i}{x})]}{\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}}, \tag{2.36}$$

then we can expand $\widetilde{F}_n^+(x)$ as

$$\begin{aligned}
\widetilde{F}_n^+(x) &\approx F(x) + \big(\frac{\mu}{n}\sum_{i=1}^n \frac{1}{X_i}[1 - Q_{v_n}(\frac{X_i}{x})] - F(x)\big) \\
&\quad -\mu F(x)\big(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} - \frac{1}{\mu}\big) \\
&= F(x) + \frac{1}{n}\sum_{i=1}^n \frac{\mu}{X_i}[1 - Q_{v_n}(\frac{X_i}{x}) - F(x)]. \qquad (2.37)
\end{aligned}$$

Let

$$\xi_i = \frac{\mu}{X_i}[1 - Q_{v_n}(\frac{X_i}{x}) - F(x)]. \tag{2.38}$$

In order to obtain the theorem, it is sufficient to show that, as $v_n \to 0$, $E(\sqrt{n}\xi_1) \to 0$ and $V(\xi_1) \to \delta^2(x)$.

$$
\begin{aligned}
E(\xi_1) &= \int_0^\infty \frac{\mu}{t}[1 - Q_{v_n}(\frac{t}{x}) - F(x)]g(t)dt \\
&= \int_0^\infty [1 - Q_{v_n}(\frac{t}{x})]f(t)dt - F(x) \\
&= \int_0^\infty F(t)q_{v_n}(\frac{t}{x})\frac{dt}{x} - F(x) \\
&= \int_0^\infty F(xy)q_{v_n}(y)dy - F(x). \qquad (2.39)
\end{aligned}
$$

Using the Taylor's expansion of $F(xy)$ at the point $y = 1$

$$
F(xy) = F(x) + xf(x)(y-1) + \frac{x^2 f'(\eta x)}{2}(y-1)^2 \qquad (2.40)
$$

where η is between 1 and y, and the fact that $f'(x)$ is bounded, we can show that

$$
E(\xi_1) = O(v_n^2). \qquad (2.41)
$$

This means that $E(\sqrt{n}\xi_1) \to 0$.

On the other hand, we have

$$
\begin{aligned}
E(\xi_1^2) &= E\big(\frac{\mu}{X_1}[1 - Q_{v_n}(\frac{X_1}{x}) - F(x)]\big)^2 \\
&= E\big(\frac{\mu^2}{X_1^2}[1 - Q_{v_n}(\frac{X_i}{x}]^2\big) \\
&\quad -2F(x)E\big(\frac{\mu^2}{X_1^2}[1 - Q_{v_n}(\frac{X_i}{x}]\big) + F^2(x)E\big(\frac{\mu^2}{X_1^2}\big) \qquad (2.42)
\end{aligned}
$$

Furthermore, we have

$$
\begin{aligned}
E\big(\frac{\mu^2}{X_1^2}[1 - Q_{v_n}(\frac{X_i}{x})]^2\big) &= \mu \int_0^\infty \frac{f(t)}{t}[1 - Q_{v_n}(t/x)]^2 dt \\
&= 2\mu \int_0^\infty H(t)[1 - Q_{v_n}(t/x)]q_{v_n}(t/x)\frac{dt}{x} \\
&= 2\mu \int_0^\infty H(xy)[1 - Q_{v_n}(y)]q_{v_n}(y)dy \qquad (2.43)
\end{aligned}
$$

where $H(x) = \int_0^x \frac{f(t)}{t} dt$. Using the Taylor's expansion of $H(xy)$ with respect to y at the point $y_0 = 1$

$$\begin{aligned} H(xy) &= H(x) + x\frac{f(\eta x)}{\eta x}(y-1) \\ &= H(x) + x\frac{f(\eta x) - f(0)}{\eta x}(y-1) \\ &= H(x) + xf'(\tau)(y-1) \end{aligned} \tag{2.44}$$

where η is between 1 and y and $\tau \in (0, x\eta)$. In the step above, we use a fact $f(0) = 0$, because $E(\frac{1}{X_1^2}) < \infty$. Since $f'(x)$ is bounded, we have

$$\begin{aligned} E\big(\frac{\mu^2}{X_1^2}[1 - Q_{v_n}(\frac{X_i}{x})]^2\big) &= \mu H(x) \\ &\quad + O(\int_0^\infty [1 - Q_{v_n}(y)] q_{v_n}(y)(y-1)dy). \end{aligned} \tag{2.45}$$

Note that

$$\begin{aligned} O(|\int_0^\infty [1 - Q_{v_n}(y)] q_{v_n}(y)(y-1)dy|) &\leq O(2\int_0^\infty q_{v_n}(y)|y-1|dy) \\ &\leq O(2[\int_0^\infty q_{v_n}(y)(y-1)^2 dy)]^{1/2}) \\ &= O(v_n), \end{aligned} \tag{2.46}$$

so, we have, as $v_n \to 0$,

$$E\big(\frac{\mu^2}{X_1^2}[1 - Q_{v_n}(\frac{X_i}{x})]^2\big) \to \mu H(x). \tag{2.47}$$

Similarly, we have

$$\begin{aligned} 2F(x)E\big(\frac{\mu^2}{X_1^2}[1 - Q_{v_n}(\frac{X_i}{x}]\big) &= 2\mu F(x)\int_0^\infty H(xy) q_{v_n}(y)dy \\ &= 2\mu F(x)H(x) + O(\int_0^\infty q_{v_n}(y)|y-1|dy) \\ &\to 2\mu F(x)H(x). \end{aligned} \tag{2.48}$$

By (2.42), (2.47) and (2.48), we have

$$E(\xi_1^2) \to \mu[H(x) - 2F(x)H(x) + \bar{\mu}F^2(x)]. \tag{2.49}$$

The proof is complete.

2.3.1.2 MSE

According to the proof of Theorem 2.3, we have

$$Bias(\widetilde{F}_n^+(x)) = \frac{x^2}{2} f'(x) v_n^2 + o(v_n^2) \tag{2.50}$$

and

$$V(\widetilde{F}_n^+(x)) = \frac{\mu}{n}[\int_0^x \frac{1}{t} f(t)dt - 2F(x) \int_0^x \frac{1}{t} f(t)dt + \bar{\mu} F^2(x)] + o(\frac{1}{n}). \tag{2.51}$$

So

$$\begin{aligned} MSE(\widetilde{F}_n^+(x)) &= \frac{\mu}{n}[\int_0^x \frac{1}{t} f(t)dt - 2F(x) \int_0^x \frac{1}{t} f(t)dt + \bar{\mu} F^2(x)] \\ &\quad + \frac{x^4}{2} f'^2(x) v_n^4 + o(\frac{1}{n} + v_n^4) \end{aligned} \tag{2.52}$$

2.3.2 Density Estimator using Asymmetric Kernels

We we can use the derivative of (2.35)

$$\tilde{f}_n(x) = \frac{\frac{1}{x^2} \sum_{i=1}^n q_{v_n}(\frac{X_i}{x})}{\sum_{i=1}^n \frac{1}{X_i}} \tag{2.53}$$

as a smooth estimator of $f(x)$ where $q_{v_n}(t) = \frac{d}{dt} Q_{v_n}(t)$.

However, (2.53) may not be defined at $x = 0$, except in cases where $\lim_{x \to 0} \tilde{f}_n(x)$ exists. Moreover, this limit is zero, which is acceptable only we are estimating $f(x)$ with $f(0) = 0$. This situation also occurs in estimating density with direct data [see Chaubey, Sen and Sen (2007)]. In their paper, they considered a perturbed version of the density estimator, replacing $Q_{v_n}(./x)$ by $Q_{v_n}(./(x + \epsilon))$, $\epsilon_n \downarrow 0$ as $n \to \infty$. This is equivalent to choosing $G_{x,n}$ such that the corresponding mean is $x + \epsilon_n \to x$ and the variance is $(x+\epsilon_n)^2 v_n^2 \to 0$. Motivated by their idea, the perturbed version of (2.53) is given by

$$\tilde{f}_n^+(x) = \frac{\frac{1}{(x+\varepsilon_n)^2}\sum_{i=1}^n q_{v_n}(\frac{X_i}{x+\varepsilon_n})}{\sum_{i=1}^n \frac{1}{X_i}}. \tag{2.54}$$

2.3.2.1 Asymptotic property of $\tilde{f}_n^+(x)$

Thoerem 2.8 *If*

A. $f(\cdot)$ *is Lipschitz continuous on* $[0,\infty)$ *and* $E(X_1^{-1}) < \infty$*;*

B. $\sup_{x\geq 0}\int_0^\infty |\frac{d}{dx}[\frac{1}{x+\varepsilon_n}q_{v_n}(\frac{t}{x+\varepsilon_n})]|dt = o((\frac{\log\log n}{n^{1/2}})^{-1})$*;*

C. $\sup_{u>0,v>0} uq_v(u) < \infty$*;*

D. $v_n \to 0$*,* $\varepsilon_n \to 0$ *as* $n \to \infty$*;*

then we have

$$\sup_{x\geq 0} |\tilde{f}_n^+(x) - f(x)| \xrightarrow{a.s.} 0$$

as $n \to \infty$*.*

Proof: We can write

$$\begin{aligned} \tilde{f}_n^+(x) &= \frac{d}{dx}\int F_n(t)[\frac{1}{x+\varepsilon_n}q_{v_n}(\frac{t}{x+\varepsilon_n})]dt \\ &= \int F_n(t)\frac{d}{dx}[\frac{1}{x+\varepsilon_n}q_{v_n}(\frac{t}{x+\varepsilon_n})]dt \end{aligned} \tag{2.55}$$

Using the uniform strong convergence of $F_n(x)$

$$\sup_{x\geq 0}|F_n(x) - F(x)| \xrightarrow{a.s.} 0. \tag{2.56}$$

and following the proof of Theorem 3 of Chaubey, Sen and Sen (2007), we can obtain the theorem.

Thoerem 2.9 *If*

E. $f(x)$ *is Lipschtiz continuous on* $[0,\infty)$ *and* $E(X_1^{-2})<\infty$*;*

F. $I_2(q) \triangleq \lim_{v_n\to 0} v_n\int_0^\infty (q_{v_n}(t))^2dt$ *exists;*

G1. for $1\le m\le 3$, $\int_0^\infty (q_{v_n}(t))^m dt = O(v^{1-m})$ *as* $v\to 0$*;*

G2. with $q^*_{m,v_n}(t)=\frac{(q_{v_n}(t))^m}{\int_0^\infty (q_{v_n}(w))^m dw}$, $1\le m\le 3$, *and as* $v_n\to 0$,

$$
\begin{aligned}
(i) \qquad & \mu_{m,v_n} = \int_0^\infty t q^*_{m,v_n}(t)dt = 1+O(v_n),\\
(ii) \qquad & \sigma^2_{m,v_n} = \int_0^\infty (t-\mu_{m,v_n})^2 q^*_{m,v_n}(t)dt = O(v_n^2)\\
(iii) \qquad & \sup_{0<v_n<\varepsilon}\int_0^\infty t^{4+\delta} q^*_{m,v_n}(t)dt<\infty, \textit{for some } \delta>0,\ \varepsilon>0;
\end{aligned}
$$

Then

(a) If $nv_n\to\infty$, $nv_n^3\to 0$, $nv_n\varepsilon_n^2\to 0$ *as* $n\to\infty$, *we have*

$$\sqrt{nv_n}(f_n^+(x)-f(x)) \to N\big(0, I_2(q)\frac{\mu f(x)}{x^2}\big), \textit{for } x>0.$$

(b) If $nv_n\epsilon_n^2\to\infty$ *and* $nv_n\varepsilon_n^4\to 0$ *as* $n\to\infty$, *we have*

$$\sqrt{nv_n\epsilon_n^2}(f_n^+(0)-f(0)) \to N\big(0, I_2(q)f(0)\big).$$

Proof: (a) Using Taylor's expansion, we can write

$$\tilde{f}_n^+(x) = \frac{1}{n}\sum_{i=1}^n Y_{in} - f(x)\big(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}-\frac{1}{\mu}\big) + o\big(\frac{1}{n}\sum_{i=1}^n\frac{1}{X_i}-\frac{1}{\mu}\big) \tag{2.57}$$

where

$$Y_{in} = \frac{\mu}{(x+\varepsilon_n)^2} q_{v_n}\big(\frac{X_i}{x+\varepsilon_n}\big) \tag{2.58}$$

For $x>0$, since, by the Law of the Iterated Logarithm,

$$\limsup_{n\to\infty}\big[\sqrt{nv_n}\big(\frac{1}{n}\sum_{i=1}^n\frac{1}{X_i}-\frac{1}{\mu}\big)\big] \overset{a.s.}{=} O(\sqrt{v_n}\log\log n) \overset{a.s.}{\to} 0,$$

it is sufficient to consider the first term in (2.57). Under the conditions of the theorem, we can show that, as $v_n \to 0$ and $\varepsilon_n \to 0$,

$$E(Y_{1n}^3) = O(v_n^{-2}) \tag{2.59}$$

and

$$v_n E(Y_{1n}^2) \to I_2(q)\frac{\mu f(x)}{x^2}. \tag{2.60}$$

Since Y_{in} is nonnegative, we have

$$\begin{aligned}
\frac{\sum_{i=1}^n E|Z_{in}|^3}{[\sum_{i=1}^n E(Z_{in}^2)]^{3/2}} &\leq \frac{E[Y_{1n} + E(Y_{1n})]^3}{\sqrt{n}[var(Y_{1n})]^{3/2}} \\
&= \frac{E(Y_{1n}^3) + 3E(Y_{1n}^2)E(Y_{1n}) + 4(E(Y_{1n}))^3}{\sqrt{n}[E(Y_{1n}^2) - (E(Y_{1n}))^2]^{3/2}}.
\end{aligned} \tag{2.61}$$

By (2.59), (2.60) and (2.61), we can claim that

$$\frac{\sum_{i=1}^n E|Z_{in}|^3}{[\sum_{i=1}^n E(Z_{in}^2)]^{3/2}} = O(\frac{1}{\sqrt{nv_n}}) \to 0. \tag{2.62}$$

Then by the Theorem 7.1.2 of Chung (1974), we have

$$\frac{\sum_{i=1}^n Z_{in}}{(\sum_{i=1}^n Z_{in}^2)} \to N(0,1).$$

This means

$$\sqrt{nv_n}\big[\frac{1}{n}\sum_{i=1}^n Y_{in} - E(Y_{1n})\big] \to N\big(0, I_2(q)\frac{\mu f(x)}{x^2}\big). \tag{2.63}$$

Further,

$$\begin{aligned}
\sqrt{nv_n}|E(Y_{1n}) - f(x)| &= \sqrt{nv_n}\Big|\int_0^\infty [f(t(x+\varepsilon_n)) - f(x)]tq_{v_n}(t)dt\Big| \\
&\leq \sqrt{nv_n}L\int_0^\infty |(t-1)x + t\varepsilon_n|tq_{v_n}(t)dt \\
&\leq \sqrt{nv_n}xL\int_0^\infty |t-1|tq_{v_n}(t)dt + \sqrt{nv_n\varepsilon_n^2}L\int_0^\infty t^2 q_{v_n}(t)dt \\
&\leq \sqrt{nv_n}xL[\int_0^\infty (t-1)^2 q_{v_n}(t)dt]^{1/2}[\int_0^\infty t^2 q_{v_n}(t)dt]^{1/2} \\
&\quad +O(\sqrt{nv_n\varepsilon_n^2}) \\
&= O(\sqrt{nv_n^3}) + O(\sqrt{nv_n\varepsilon_n^2})
\end{aligned} \tag{2.64}$$

Using (2.57), (2.63) and (2.64), we can establish the part (a) of the theorem.

(b) Similar to (2.57), we have

$$\tilde{f}_n^+(0) = \frac{1}{n}\sum_{i=1}^{n} Y'_{in} - f(0)\big(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}\big) + o\big(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}\big) \tag{2.65}$$

where

$$Y'_{in} = \frac{\mu}{(\varepsilon_n)^2} q_{v_n}\big(\frac{X_i}{\varepsilon_n}\big). \tag{2.66}$$

We can show that

$$E(Y'^3_{1n}) = O(\epsilon_n^4 v_n^{-2}) \tag{2.67}$$

and

$$\epsilon_n^2 v_n E(Y'^2_{1n}) \to I_2(q) f(0). \tag{2.68}$$

Let

$$Z'_{in} = \sqrt{\frac{v_n \epsilon_n^2}{n}} \left[Y'_{in} - E(Y'_{in})\right]. \tag{2.69}$$

Then using (2.67), (2.68) and a similar inequality to(2.61) , we can claim that

$$\frac{\sum_{i=1}^n E|Z'_{in}|^3}{[\sum_{i=1}^n E(Z'^2_{in})]^{3/2}} \le O\big(\frac{1}{\sqrt{n v_n \epsilon_n^2}}\big) \to 0. \tag{2.70}$$

Then by the Theorem 7.1.2 of Chung (1974), we have

$$\frac{\sum_{i=1}^n Z'_{in}}{(\sum_{i=1}^n Z'^2_{in})} \to N(0,1).$$

This means

$$\sqrt{n v_n \epsilon_n^2}\big[\frac{1}{n}\sum_{i=1}^{n} Y'_{in} - E(Y'_{1n})\big] \to N\big(0, I_2(q) f(0)\big). \tag{2.71}$$

Furthermore,

$$\begin{aligned}
\sqrt{n v_n \epsilon_n^2}|E(Y'_{1n}) - f(0)| &= \sqrt{n v_n \epsilon_n^2}\Big|\int_0^\infty [f(t\varepsilon_n) - f(0)] t q_{v_n}(t) dt\Big| \\
&\le \sqrt{n v_n \epsilon_n^2} L \int_0^\infty |\varepsilon_n| t^2 q_{v_n}(t) dt \\
&= \sqrt{n v_n \epsilon_n^2} L[\epsilon_n (v_n^2 + 1)] \\
&= O(\sqrt{n v_n \varepsilon_n^4}) \qquad (2.72)
\end{aligned}$$

Using (2.65), (2.71) and (2.72), we can establish the part (b) of the theorem.

Remark 2.4: Just as in Chaubey, Sen and Sen (2007), in this thesis we consider $q_{v_n}(x)$ to be a family of gamma density such that

$$q_{v_n}(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$$

where $\alpha = \frac{1}{v_n^2}$ and $\alpha\beta = 1$.

2.3.2.2 AMISE

In order to obtain MSE of $\tilde{f}_n^+$, we first compute the bias of $\tilde{f}_n^+$. According to (2.57), we have

$$\begin{aligned} Bias[\tilde{f}_n^+(x)] &\approx E(Y_{1n}) - f(x) \\ &= \int_0^\infty \frac{\mu}{(x+\epsilon_n)^2} q_{v_n}\left(\frac{t}{x+\epsilon_n}\right) g(t)dt - f(x) \\ &= \int_0^\infty \frac{t}{(x+\epsilon_n)^2} q_{v_n}\left(\frac{t}{x+\epsilon_n}\right) f(t)dt - f(x) \end{aligned}$$

Let $t/(x+\epsilon_n) = y$, then

$$Bias[\tilde{f}_n^+(x)] = \int_0^\infty y q_{v_n}(y) f[y(x+\epsilon_n)]dy - f(x).$$

Note that we have

$$\begin{aligned} f[y(x+\epsilon_n)] &= f(x+\epsilon_n) + (x+\epsilon_n)f'(x+\epsilon_n)(y-1) \\ &\quad + \frac{(x+\epsilon_n)^2}{2} f''(x+\epsilon_n)(y-1)^2 + o(y-1)^2 \\ &= f(x) + \epsilon_n f'(x) + x f'(x)(y-1) \\ &\quad + \frac{x^2}{2} f''(x)(y-1)^2 + o\left((y-1)^2\right) + o(\epsilon_n), \end{aligned}$$

$$
\begin{aligned}
Bias[\tilde{f}_n^+(x)] &= \epsilon_n f'(x) + x f'(x) \int_0^\infty y(y-1) q_{v_n}(y) dy \\
&\quad + \frac{x^2}{2} f''(x) \int_0^\infty y(y-1)^2 q_{v_n}(y) dy + o(v_n^2 + \epsilon_n) \\
&= \epsilon_n f'(x) + x f'(x) T_1 + \frac{x^2}{2} f''(x) T_2 + o(v_n^2 + \epsilon_n).
\end{aligned}
$$

Furthermore, we have

$$
\begin{aligned}
T_1 &= \int_0^\infty (y-1)^2 q_{v_n}(y) dy + \int_0^\infty (y-1) q_{v_n}(y) dy \\
&= v_n^2
\end{aligned}
$$

and

$$
\begin{aligned}
T_2 &= \int_0^\infty (y-1)^2 q_{v_n}(y) dy + \int_0^\infty (y-1)^3 q_{v_n}(y) dy \\
&= v_n^2 + o(v_n^2),
\end{aligned}
$$

So

$$
Bias[\tilde{f}_n^+(x)] = (x v_n^2 + \varepsilon_n) f'(x) + \frac{x^2}{2} f''(x) v_n^2 + o(v_n^2 + \varepsilon_n). \tag{2.73}
$$

By the proof of Theorem 2.9, it is easy to show that

$$
Var[\tilde{f}_n^+(x)] = \frac{I_2(q)\mu f(x)}{n v_n (x+\varepsilon_n)^2} + o((n v_n)^{-1}). \tag{2.74}
$$

By (2.73) and (2.74), we have

$$
MSE[\tilde{f}_n^+(x)] = [(x v_n^2 + \varepsilon_n) f'(x) + \frac{x^2}{2} f''(x) v_n^2]^2 + \frac{I_2(q)\mu}{n v_n} \frac{f(x)}{(x+\varepsilon_n)^2}.
$$

So

$$
AMISE[\tilde{f}_n^+] = \int_0^\infty [(x v_n^2 + \varepsilon_n) f'(x) + \frac{x^2}{2} f''(x) v_n^2]^2 dx + \frac{I_2(q)\mu}{n v_n} \int_0^\infty \frac{f(x)}{(x+\varepsilon_n)^2} dx. \tag{2.75}
$$

For a given f, the above expression may be technically used to find the optimum value of the smoothing parameter. However, the expressions are too complicated and in practice we use them for cross validation to obtain data dependent value for the smoothing parameter(s).

2.3.3 Corrected Density Estimator

Note that if we integrate (2.54) from 0 to ∞, we will obtain

$$\int_0^\infty \tilde{f}_n^+(x)dx = \frac{\sum_{i=1}^n \frac{Q_{v_n}(X_i/\epsilon_n)}{X_i}}{\sum_{i=1}^n \frac{1}{X_i}}. \tag{2.76}$$

If $\epsilon_n \neq 0$, (2.76) is not equal to 1. In this case, f_n^+ is not a real density estimator which is integrated to unity. In order to overcome this defect, we divide $f_n^+(x)$ by $\int_0^\infty f_n^+(x)dx$, which leads to a corrected estimator

$$\tilde{f}_n^*(x) = \frac{\frac{1}{(x+\varepsilon_n)^2}\sum_{i=1}^n q_{v_n}\left(\frac{X_i}{x+\varepsilon_n}\right)}{\sum_{i=1}^n \frac{Q_{v_n}(X_i/\epsilon_n)}{X_i}}. \tag{2.77}$$

Since $\sum_{i=1}^n \frac{Q_{v_n}(X_i/\epsilon_n)}{X_i} \to \sum_{i=1}^n \frac{1}{X_i}$ for a given sample, as $\epsilon_n \to 0$, most of the asymptotic properties of $\tilde{f}_n^+$ still hold for $\tilde{f}_n^*$. We can establish the same theorems as Theorem 2.8 and 2.9 for $\tilde{f}_n^*$. But the biases of the two estimator are slightly different. Note that

$$\begin{aligned} \tilde{f}_n^*(x) &= \frac{\tilde{f}_n^+}{1-\widetilde{F}_n^+(\epsilon_n)} \\ &\approx \frac{\tilde{f}_n^+}{1-F(\epsilon_n)}, \end{aligned}$$

then it is easy to show that

$$Bias(\tilde{f}_n^*(x)) = Bias(\tilde{f}_n^+(x)) + \epsilon_n f(0)f(x) + o(\epsilon_n). \tag{2.78}$$

Later, we will see that this boundary correction is very useful in reducing bias at the border and improving global performance of density estimator in some cases.

2.4 Other Density Estimators with Asymmetric Kernels

In this section, we will apply Chen's and Scaillet's idea to obtain some density estimators for LB data. Here we will mainly give two kinds of such density estimators. One kind is, motivated by the idea of Chen (2000), with gamma kernels and the other is, inspired by Scaillet (2004), with inverse and reciprocal inverse Gaussian kernels.

2.4.1 Chen Density Estimators for Length Biased Data

Note that if let $F_n'(x) = n^{-1} I\{X_i' \le x\}$, an alternative form of (1.12) is

$$\hat{f}(x) = \int_0^\infty K_{\rho_b(x),b}(t) dF_n'(t). \tag{2.79}$$

Recall that the empirical distribution for LB data is $F_n(x) = \frac{\sum_{i=1}^n \frac{1}{X_i} I\{X_i \le x\}}{\sum_{i=1}^n \frac{1}{X_i}}$. Substituting $F_n'(x)$ with $F_n(x)$ in (2.79) will give us Chen density estimators for LB data as follows.

$$\hat{f}_C(x) = \int_0^\infty K_{\rho_b(x),b}(t) dF_n(t), \tag{2.80}$$

which can also be written as

$$\hat{f}_C(x) = \frac{\sum_{i=1}^n \frac{1}{X_i} K_{\rho_b(x),b}(X_i)}{\sum_{i=1}^n \frac{1}{X_i}}. \tag{2.81}$$

Furthermore, by (1.5), we can have

$$\begin{aligned} E\left(\hat{f}_C(x)\right) &\approx \int_0^\infty \frac{\mu}{y} K_{\rho_b(x),b}(y) g(y) dy \\ &= \int_0^\infty K_{\rho_b(x),b}(y) f(y) dy \\ &= E\left(f(\xi_x)\right) \end{aligned} \tag{2.82}$$

where ξ_x is a $\Gamma(\rho_b(x), b)$ random variable. Similar to Chen estimator for direct data, it is easy to show that $E\left(\hat{f}_C(x)\right) \to f(x)$ as $b \to 0$.

We use $\hat{f}_{C1}(x)$ and $\hat{f}_{C2}(x)$ to denote the density estimator under the $\rho_b(x)$'s choices (1.13) and (1.14) respectively.

2.4.2 Scaillet Density Estimators for Length Biased Data

Replacing gamma kernels $K_{\rho(x),b}$ proposed by Chen with inverse or reciprocal inverse Gaussian kernels proposed by Scaillet in (2.80) , we can derive Scaillet density estimators for LB data as follows.

$$\tilde{f}_{IG}(x) = \frac{\sum_{i=1}^{n} \frac{1}{X_i} K_{IG(x,1/b)}(X_i)}{\sum_{i=1}^{n} \frac{1}{X_i}} \tag{2.83}$$

and

$$\tilde{f}_{RIG}(x) = \frac{\sum_{i=1}^{n} \frac{1}{X_i} K_{RIG(1/(x-b),1/b)}(X_i)}{\sum_{i=1}^{n} \frac{1}{X_i}}. \tag{2.84}$$

Remark 2.5: Note that Chen and Scaillet density estimators for LB data can also be obtain by generalized Hille's lemma. We use $f(t)$ and $k_{x,b}(t)dt$ [$k_{x,b}(.)$ represents the kernels proposed by Chen or Scaillet] to replace $u(t)$ and $dQ_{x,n}(t)$ in Hille's lemma respectively. Actually, the their proposed kernel $k_{x,b}(t)$ is a density of random variable ξ_x such that $E(\xi_x) \to x$ and $V(\xi_x) \to 0$ as $b \to 0$. This means the distribution function of ξ_x satisfies the conditions of distribution function $G_{x,n}$ in generalized Hille's Lemma 1.2. Then we have

$$\int_0^\infty f(t) k_{x,b}(t) dt \to f(x).$$

An alternative form of the right above is

$$\int_0^\infty k_{x,b}(t)dF(t).$$

So Chen and Scaillet density estimator are easy to be established by replacing distribution function $F(t)$ with Cox's estimator. That is

$$\int_0^\infty k_{x,b}(t)dF_n(t)$$

which is the same as (2.80).

Remark 2.6: The asymptotic distributions of estimators may be generally used for inference purpose. However, the expressions for asymptotic variance derived here are quite complicated, hence in practice Bootstrap procedures may be useful in this context. However, we have not considered such procedures in the thesis.

Remark 2.7: The asymptotic properties are quite different for $x > 0$ and $x = 0$. To study the properties of the estimators more carefully, we may consider x as a boundary point where $x/b \to k$ for some $k > 0$ and an interior point where $x/b \to \infty$. This will be investigated in future research.

Chapter 3

Smooth Estimators of Density and Distribution Functions Based on Empirical Distribution Function

3.1 Introduction

Note that an alternative form of length biased model (1.2) is given by

$$f(x) = \frac{g(x)/x}{\mu}. \tag{3.1}$$

This formula gives us an alternative strategy to estimate $f(x)$. We can first obtain an estimator of $g(x)$, say $\hat{g}(x)$, then, by (3.1), a natural estimator of $f(x)$ is given by

$$\hat{f}(x) = \frac{\hat{g}(x)/x}{\mu}. \tag{3.2}$$

By now μ is unknown. Note that we want to obtain an estimator $\hat{f}(x)$, which should satisfy the most basic property being integrated to unity. So integrating on both sides of (3.2) gives us an estimator of μ based on $\hat{g}(x)$, which can be defined as

$$\hat{\mu} = \int_0^\infty \frac{\hat{g}(x)}{x} dx. \tag{3.3}$$

Therefore, a valid estimator of $f(x)$ is

$$\hat{f}(x) = \frac{\hat{g}(x)/x}{\int (\hat{g}(x)/x) dx} \tag{3.4}$$

where $\hat{g}(x)$ must satisfy the following conditions:

(i) $\hat{g}(x) = 0$ for $x \leq 0$;

(ii) $\hat{g}(x)/x$ is integrable on $[0, \infty)$.

Bhattacharyya *et al.* (1988) use (3.2) with $\hat{\mu}$ being the harmonic mean estimator to establish a density estimator . However, since their estimator does not satisfy condition (ii) and even condition (i) under certain circumstances, their estimator does not perform very well [see Chaubey *et al.* (2010), Jones (1991) and Wu and Mao (1996)]. Therefore, it seems that formula (3.4) might give us some valid density estimators.

In this chapter, we will follow formula (3.4) to obtain some density estimators. Similar to previous chapter, we will use Hille's lemma in Poisson weights and generalized version to build two kinds of estimators, one using Poisson weights and the other using asymmetric kernels. However the smooth technique motivated by Hille's lemma in Poisson weights is not suitable to be applied directly in this case. Some necessary modifications to the smooth technique should be made. The route we follow in this chapter is the opposite of that in previous chapter. Here we first obtain smooth density estimator. Then, by integrating the density estimator on interval $[0, x)$, we can have smooth estimator of distribution function. In Section 3.2, the modified smooth technique in Poisson weights is applied to find a new smooth density estimator. The integration of this estimator gives us a distribution estimator. Their asymptotic properties are studied. Without any modification, the smooth technique using asymmetric kernels can be directly applied to the formula (3.4). And the perturbation and boundary correction are still necessary to be used in new density estimator. Therefore, in Section 3.3, new density and distribution estimator using asymmetric kernels are found and their the asymptotic properties are studied as well.

3.2 Estimators of Density and Distribution Functions with Poisson Weights

3.2.1 Smooth Density Estimator

Define

$$G_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \le x\}. \tag{3.5}$$

Using the poisson weights

$$p_k(x\lambda_n) = e^{-x\lambda_n}\frac{(x\lambda_n)^k}{k!}, \tag{3.6}$$

we can obtain a smooth estimator of $G(x)$. Since the smooth estimator is differentiable, we take its derivative as a smooth estimator of density function $g(x)$. In order to let the smooth estimator of density satisfy conditions (i) and (ii), we attach the Poisson weight $p_k(x\lambda_n)$ to $G_n((k-1)/\lambda_n)$. This results the following smooth estimator of $G(x)$

$$\widehat{G}_n(x) = \sum_{k\ge 0} p_k(x\lambda_n) G_n\big(\frac{k-1}{\lambda_n}\big). \tag{3.7}$$

As in (2.4) differentiating the above expression gives us the following smooth estimator

$$\hat{g}_n(x) = \lambda_n \sum_{k\ge 1} p_k(x\lambda_n)\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)\big] \tag{3.8}$$

such that $\hat{g}(x) = 0$ when $x \le 0$ and $\hat{g}(x)/x$ is integrable, since

$$\begin{aligned}\int_0^\infty \frac{\hat{g}_n(x)}{x}dx &= \lambda_n \sum_{k\ge 1}\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)\big]\int_0^\infty \frac{1}{x}p_k(x\lambda_n)dx \\ &= \lambda_n \sum_{k\ge 1}\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)\big]\frac{1}{k}.\end{aligned}$$

The new estimator of $f(x)$ is given by

$$\hat{f}_n(x) = \lambda_n \frac{\sum_{k\ge 1}\frac{p_{k-1}(x\lambda_n)}{k}\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)\big]}{\sum_{k\ge 1}\frac{1}{k}\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)\big]}. \tag{3.9}$$

3.2.1.1 Asymptotic Property of $\hat{f}_n(x)$

Lemma 3.1 *If $0 < E(X_1^{-1}) < \infty$, $\lambda_n = O(n^\alpha)(0 < \alpha < 1)$ and $g(x)$ is absolutely continuous with a bounded derivative $g'(x)$ on $\mathbf{R}^+$, then*

$$\int_0^\infty \frac{\hat{g}_n(x)}{x}dx \xrightarrow{a.s.} \frac{1}{\mu}. \tag{3.10}$$

Proof: First we can write

$$\begin{aligned}\int_0^\infty \frac{\hat{g}_n(x)}{x}dx &= \lambda_n \sum_{k\ge 1}\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big) - G\big(\frac{k}{\lambda_n}\big) + G\big(\frac{k-1}{\lambda_n}\big)\big]\frac{1}{k} \\ &\quad + \lambda_n \sum_{k\ge 1}\big[G\big(\frac{k}{\lambda_n}\big) - G\big(\frac{k-1}{\lambda_n}\big)\big]\frac{1}{k} \\ &= T_{n1}(\lambda_n) + T_{n2}(\lambda_n).\end{aligned} \tag{3.11}$$

First, we want to show that

$$T_{n1}(\lambda_n) \xrightarrow{a.s.} 0. \tag{3.12}$$

Note that for any $0 < \alpha < 1$, we can find a $\beta(0 < \beta < 1/2)$ such that

$$\alpha < \frac{1/2}{1/2 + \beta}. \tag{3.13}$$

For any fixed k, we apply Lemma 1 of Bahadur (1966) in the interval $[(k-1)/\lambda_n, k/\lambda_n)$. However, We can not use the lemma directly. Here we make some slight modifications. Let $a_n = \lambda_n^{-1}(\lambda_n/k)^{2\beta}$, $b_n = n^{1/2}/(\log n)^{1+\theta}$,

$\gamma_n = \lambda_n^{-1/2}(\lambda_n/k)^{\beta} n^{-1/2}(\log n)^{1+\theta}$ and $\eta_{r,n} = \frac{k-1}{\lambda_n} + \gamma_n r$.

First of all, we need verify a fact that is, for any $s, t \in [(k-1)/\lambda_n, k/\lambda_n)$, there exits a c_2 such that

$$|G(s) - G(t)| \le c_2 a_n. \tag{3.14}$$

This is because

$$|G(s) - G(t)| \le g(\eta)\lambda_n^{-1} = (k/\lambda_n)^{2\beta} g(\eta)\lambda_n^{-1}(\lambda_n/k)^{2\beta} \tag{3.15}$$

where $\eta \in ((k-1)/\lambda_n, k/\lambda_n)$. Note that $(k/\lambda_n)^{2\beta} g(\eta) \approx (k/\lambda_n)^{2\beta} g(k/\lambda_n)$. Since $g(x)$ is a density function, it is easy to know that $x^{2\beta} g(x)$ is bounded on $\mathbf{R}^+$. Then we can find a c_2 which is finite and greater than $(k/\lambda_n)^{2\beta} g(\eta)$. So (3.14) holds, which means under our modifications, we can still have the inequality $z_{r,n} \le c_2 a_n (0 \le r \le b_n)$ in the proof of Bahadur's lemma.

Following the proof Lemma 1 of Bahadur (1966), we can claim that

$$\Big|G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big) - G\big(\frac{k}{\lambda_n}\big) + G\big(\frac{k-1}{\lambda_n}\big)\Big| \le \gamma_n \ \ a.s. \tag{3.16}$$

Then we have

$$|T_{n1}(\lambda_n)| \le \lambda_n^{1/2+\beta} n^{-1/2}(\log n)^{1+\theta} \sum_{k \ge 1} \frac{1}{k^{1+\beta}} \ \ a.s. \tag{3.17}$$

If $\lambda_n = O(n^{\alpha})(0 < \alpha < 1)$, by (3.13) and (3.17), we can see that (3.12) holds.

On the other hand, we have, as $\lambda_n \uparrow \infty$,

$$T_{n2}(\lambda_n) = \sum_{k \ge 1} (\lambda_n/k) g(k/\lambda_n) \frac{1}{\lambda_n} + \sum_{k \ge 1} (\lambda_n/k) g'(\xi_k) \frac{1}{2\lambda_n^2} \to \frac{1}{\mu} \tag{3.18}$$

where $\xi_k \in ((k-1)/\lambda_n, k/\lambda_n)$. By (3.12) and (3.18), the lemma follows. The proof is complete.

Using the same method in the proof of Theorem 4.1 of Chaubey and Sen (1996), we can show that when $\lambda_n = O(n^\alpha)(0 < \alpha < 1)$ and $g(x)$ is absolutely continuous with a bounded derivative $g'(x)$ a.e. on $\mathbf{R}^+$,

$$||\hat{g}_n(x) - g(x)|| = \sup_{x \in R^+} \{|\hat{g}_n(x) - g(x)|\} \xrightarrow{a.s.} 0 \tag{3.19}$$

as $n \uparrow \infty$. By Lemma 3.1, (3.19) and (3.9), we can obtain the following theorem.

Thoerem 3.1 *If* $0 < E(X_1^{-1}) < \infty$ *and* $g(x)$ *is absolutely continuous with a bounded derivative* $g'(x)$ *on* $\mathbf{R}^+$ *and* $\lambda_n = O(n^\alpha)(0 < \alpha < 1)$, *then*

$$||\hat{f}_n(x) - f(x)|| = \sup_{x \in R^+} \{|\hat{f}_n(x) - f(x)|\} \xrightarrow{a.s.} 0. \tag{3.20}$$

Now we suppose that $g'(x)$ satisfies Lipschitz order α condition, for some $\alpha > 0$, there exits a positive $K(< \infty)$, such that

$$|g'(t) - g'(s)| \leq K|t - s|^\alpha. \tag{3.21}$$

We can write

$$\hat{g}_n(x) = \lambda_n \sum_{k \geq 1} p_k(x\lambda_n)\big[G\big(\frac{k}{\lambda_n}\big) - G\big(\frac{k-1}{\lambda_n}\big)\big] + T_n'(x) \tag{3.22}$$

where $T_n'(x) = \lambda_n \sum_{k \geq 1} p_k(x\lambda_n)\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big) - G\big(\frac{k}{\lambda_n}\big) + G\big(\frac{k-1}{\lambda_n}\big)\big]$. Using (3.21) and Taylor's expansions of $G(k/\lambda_n)$ and $G((k-1)/\lambda_n)$ at point x, then we can rewrite the first term of (3.22) and establish

$$\hat{g}_n(x) = g(x) - \frac{1}{2\lambda_n} g'(x) + O(\lambda_n^{-1-\alpha}) + T_n'(x). \tag{3.23}$$

So

$$\int_0^\infty \frac{\hat{g}_n(x)}{x} dx = \frac{1}{\mu} - \frac{1}{2\lambda_n} \int_0^\infty \frac{g'(x)}{x} dx + O(\lambda_n^{-1-\alpha}) + T_{n1}(\lambda_n) \tag{3.24}$$

It is easy to see that

$$
\begin{aligned}
V(T_{n1}(\lambda_n)) &= \frac{1}{n}\{\sum_{k\geq 1}(\lambda_n/k)^2[G(\frac{k}{\lambda_n})-G(\frac{k-1}{\lambda_n})] \\
&\quad -\big(\sum_{k\geq 1}(\lambda_n/k)[G(\frac{k}{\lambda_n})-G(\frac{k-1}{\lambda_n})]\big)^2\} \\
&\approx \frac{1}{n}V(\frac{1}{X}). \qquad (3.25)
\end{aligned}
$$

By (3.23), we have

$$
\begin{aligned}
\frac{\hat{g}_n(x)}{x} &= \frac{g(x)}{x}-\frac{1}{2\lambda_n}\frac{g'(x)}{x}+O(\lambda_n^{-1-\alpha})+T_n(x) \\
&= \frac{f(x)}{\mu}-\frac{1}{2\lambda_n}\frac{g'(x)}{x}+O(\lambda_n^{-1-\alpha})+T_n(x) \qquad (3.26)
\end{aligned}
$$

where

$$T_n(x)=\frac{\lambda_n}{x}\sum_{k\geq 1}p_k(x\lambda_n)\big[G_n\big(\frac{k}{\lambda_n}\big)-G_n\big(\frac{k-1}{\lambda_n}\big)-G\big(\frac{k}{\lambda_n}\big)+G\big(\frac{k-1}{\lambda_n}\big)\big]. \qquad (3.27)$$

Using (3.23) and (3.24), we can write

$$
\begin{aligned}
\hat{f}_n(x) &= \frac{\mu g(x)}{x}+\mu(\frac{\hat{g}_n(x)}{x}-\frac{g(x)}{x})-\mu^2\frac{g(x)}{x}(\int_0^\infty\frac{\hat{g}_n(x)}{x}dx-\frac{1}{\mu}) \\
&\quad +o(\int_0^\infty\frac{\hat{g}_n(x)}{x}dx-\frac{1}{\mu}) \\
&= f(x)+\frac{1}{2\lambda_n}[\bar{\mu}f(x)-f(0)f(x)-\frac{f(x)}{x}-f'(x)]+\mu T_n(x)+\mu T_{n1}(\lambda_n) \\
&\quad +O(\lambda_n^{-1-\alpha})\ a.s. \qquad (3.28)
\end{aligned}
$$

where $\bar{\mu}=\int(f(x)/x)dx$. According to Chaubey and Sen (1996), we have $V(T_n'(x))\approx \frac{1}{2}(2\pi x)^{-1/2}g(x)(\lambda_n^{1/2}/n)$ and, if $x\neq y$, $Cov[T_n'(x),T_n'(y)]=O(n^{-1})$. Note that $T_n(x)=T_n'(x)/x$, then we have

$$V(T_n(x))\approx\frac{1}{2}(\pi x)^{-1/2}\frac{g(x)}{x^2}(\lambda_n^{1/2}/n) \qquad (3.29)$$

and, if $x\neq y$,

$$Cov[T_n(x),T_n(y)]=O(n^{-1}). \qquad (3.30)$$

By (3.25) and (3.29), we can see that the variance of $\hat{f}_n(x)$ should have an order $O(\lambda_n^{1/2}/n)$ far greater than $O(1/n)$ the order of variance of $T_{n1}(\lambda_n)$. So we can dispense with $T_{n1}(\lambda_n)$ and then $\hat{f}_n(x)$ behaves like $\mu T_n(x)$. Thus from equations (3.27), (3.29) and (3.30) we can establish the following theorem.

Thoerem 3.2 *Under the same assumptions on $g(x)$ and $g'(x)$ in Theorem 3.1, if (3.21) holds and $E(X_1^{-2}) < \infty$, when $\lambda_n = O(n^{2/5})$(nonstochastic), we have, for a compact set $\mathcal{C} \subset \mathbf{R}^+$,*

$$\Big\{\Big(n^{2/5}[\hat{f}_n(x) - f(x)] - \frac{1}{2\delta^2}[\bar{\mu} f(x) - f(0)f(x) - \frac{f(x)}{x} - f'(x)]\Big), x \in \mathcal{C}\Big\} \xrightarrow{\mathcal{D}} \text{ Gaussian process}$$

with covariance function $\gamma_x^2 \delta_{xy}$ where $\gamma_x^2 = \frac{\mu}{2}(\pi x^3)^{-1/2} f(x)\delta$, $\delta_{xy} = 0$ for $x \neq y$ and 1 for $x = y$ and $\delta = \lim_{n\to\infty}(n^{-1/5}\lambda_n^{1/2})$.

3.2.1.2 MSE

For $\hat{f}_n(x)$, we have

$$Bias(\hat{f}_n(x)) = \frac{1}{2\lambda_n}[\bar{\mu} f(x) - f(0)f(x) - \frac{f(x)}{x} - f'(x)] + o(\lambda_n^{-1}) \tag{3.31}$$

and

$$V(\hat{f}_n(x)) = \sqrt{\lambda_n}\frac{\mu}{2\sqrt{\pi}n}\frac{f(x)}{x^{3/2}} + o(\frac{\sqrt{\lambda_n}}{n}). \tag{3.32}$$

So

$$\begin{aligned} MSE(\hat{f}_n(x)) &= \frac{1}{4\lambda_n^2}[\bar{\mu} f(x) - f(0)f(x) - \frac{f(x)}{x} - f'(x)]^2 \\ &\quad + \sqrt{\lambda_n}\frac{\mu}{2\sqrt{\pi}n}\frac{f(x)}{x^{3/2}} + o(\frac{\sqrt{\lambda_n}}{n} + \frac{1}{\lambda_n}). \end{aligned} \tag{3.33}$$

3.2.2 Distribution Function Estimator

By (3.9), we can see that the corresponding smooth estimator of distribution function is

$$\widehat{F}_n(x) = \frac{\sum_{k\geq 1}(1/k)W_k(x\lambda_n)[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)]}{\sum_{k\geq 1}\frac{1}{k}[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big)]} \tag{3.34}$$

where

$$W_k(\lambda_n x) = \frac{1}{\Gamma(k)}\int_0^{\lambda_n x} e^{-y}y^{k-1}dy = \sum_{j\geq k} p_j(\lambda_n x).$$

Next, we will discuss asymptotic property of $\widehat{F}_n(x)$. By (3.26), if $g'(x)$ exits, we have

$$\frac{\hat{g}_n(x)}{x} = \frac{f(x)}{\mu} + O(\lambda_n^{-1}) + T_n(x). \tag{3.35}$$

Integrating from 0 to x, we have

$$\int_0^x \frac{\hat{g}_n(t)}{t}dt = \frac{F(x)}{\mu} + \int_0^x T_n(t)dt + O(\lambda_n^{-1}). \tag{3.36}$$

Note that

$$\begin{aligned}
|\int_0^x T_n(t)dt| &\leq \int_0^x \frac{\lambda_n}{t}\sum_{k\geq 1}p_k(t\lambda_n)|\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big) - G\big(\frac{k}{\lambda_n}\big) + G\big(\frac{k-1}{\lambda_n}\big)\big]|dt \\
&\leq \int_0^\infty \frac{\lambda_n}{t}\sum_{k\geq 1}p_k(t\lambda_n)|\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big) - G\big(\frac{k}{\lambda_n}\big) + G\big(\frac{k-1}{\lambda_n}\big)\big]|dt \\
&= \lambda_n\sum_{k\geq 1}|\big[G_n\big(\frac{k}{\lambda_n}\big) - G_n\big(\frac{k-1}{\lambda_n}\big) - G\big(\frac{k}{\lambda_n}\big) + G\big(\frac{k-1}{\lambda_n}\big)\big]|\frac{1}{k}.
\end{aligned}$$

Then using (3.16), for any x, we have

$$|\int_0^x T_n(t)dt| \leq \lambda_n^{1/2+\beta}n^{-1/2}(\log n)^{1+\theta}\sum_{k\geq 1}\frac{1}{k^{1+\beta}} \;\; a.s. \tag{3.37}$$

So, as $\lambda_n = O(n^\alpha)(0 < \alpha < 1) \uparrow \infty$,

$$\sup_{x\in R^+}\{|\int_0^x T_n(t)dt|\} \xrightarrow{a.s.} 0. \tag{3.38}$$

By (3.36) and (3.38), we have

$$\sup_{x\in R^+}\{|\int_0^x \frac{\hat{g}_n(t)}{t} - \frac{F(x)}{\mu}|\} \xrightarrow{a.s.} 0. \tag{3.39}$$

Combining with Lemma 3.1, we have the following theorem.

Thoerem 3.3 *If* $\lambda_n = O(n^\alpha)(0 < \alpha < 1)$, $0 < E(X_1^{-1}) < \infty$ *and* $g(x)$ *is absolutely continuous with a bounded derivative* $g'(x)$ *a.e. on* $\mathbf{R}^+$*, then, as* $n \uparrow \infty$,

$$||\widehat{F}_n(x) - F(x)|| = \sup_{x\in R^+}\{|\widehat{F}_n(x) - F(x)|\} \xrightarrow{a.s.} 0. \tag{3.40}$$

Next, we will discuss the weak asymptotic properties of $\widehat{F}(x)$.

Thoerem 3.4 *If* $\lambda_n = O(n^\alpha)(1/2 < \alpha < 1)$, $E(X_1^{-2}) < \infty$ *and under the same assumptions of Theorem 3.3 on* $g(x)$ *and* $g'(x)$*, when* $n \uparrow \infty$*, we have*

$$\sqrt{n}(\widehat{F}_n(x) - F(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x))$$

where $\delta^2(x) = \mu\left[\int_0^x \frac{1}{t}f(t)dt - 2F(x)\int_0^x \frac{1}{t}f(t)dt + \bar{\mu}F^2(x)\right]$.

Proof: According to (3.9), we have

$$\begin{aligned}\widehat{F}_n(x) &= \frac{\sum_{k\geq 1}(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n}{k}\left[G_n\left(\frac{k}{\lambda_n}\right) - G_n\left(\frac{k-1}{\lambda_n}\right)\right]}{\sum_{k\geq 1}\frac{1}{k}[G_n\left(\frac{k}{\lambda_n}\right) - G_n\left(\frac{k-1}{\lambda_n}\right)]} \\ &= \frac{\sum_{k\geq 1}(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k}\left[G_n\left(\frac{k}{\lambda_n}\right) - G_n\left(\frac{k-1}{\lambda_n}\right)\right]}{\sum_{k\geq 1}\frac{\lambda_n}{k}[G_n\left(\frac{k}{\lambda_n}\right) - G_n\left(\frac{k-1}{\lambda_n}\right)]} \\ &= \frac{T_{1n}}{T_{2n}}. \end{aligned} \tag{3.41}$$

Note that, by (3.36), (3.37) and Lemma 3.1, we have $T_{1n} \xrightarrow{a.s.} \int_0^x \frac{1}{t}g(t)dt$ and $T_{2n} \xrightarrow{a.s.} \frac{1}{\mu}$. Using the Taylor expansion of $\widehat{F}_n(x)$ at the point (T_{01}, T_{02}) where $T_{01} = \int_0^x \frac{1}{t}g(t)dt$ and $T_{02} = \frac{1}{\mu}$, $\widehat{F}_n(x)$ can be approximated by

$$\widehat{F}_n(x) \approx \frac{T_{01}}{T_{02}} + \frac{1}{T_{02}}(T_{1n} - T_{01}) - \frac{T_{01}}{T_{02}^2}(T_{2n} - T_{02}). \tag{3.42}$$

So

$$\begin{aligned} E(\widehat{F}_n(x)) &\approx \frac{T_{01}}{T_{02}} + \frac{1}{T_{02}} E(T_{1n} - T_{01}) - \frac{T_{01}}{T_{02}^2} E(T_{2n} - T_{02}) \\ &\rightarrow F(x) \text{ as } n \rightarrow \infty. \end{aligned} \tag{3.43}$$

Actually, we can show that

$$E(\widehat{F}_n(x)) = F(x) + O(\lambda_n^{-1}). \tag{3.44}$$

From (3.44), we can see that only when $1/2 < \alpha < 1$, $\sqrt{n}[\widehat{F}_n(x) - F(x)] \rightarrow 0$.

Now we discuss the variance of $\widehat{F}_n(x)$. By (3.41) and (3.43), we have

$$\begin{aligned} \widehat{F}_n(x) - F(x) \approx \sum_{k\geq 1} & \{ \quad [\mu(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k} - \mu^2(\int_0^x \frac{1}{t} g(t)dt)\frac{\lambda_n}{k}] \\ & [G_n(\frac{k}{\lambda_n}) - G_n(\frac{k-1}{\lambda_n}]\} \\ = \sum_{k\geq 1} & \{ \quad [\mu(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k} - \mu F(x)\frac{\lambda_n}{k}] \\ & [G_n(\frac{k}{\lambda_n}) - G_n(\frac{k-1}{\lambda_n}]\}. \end{aligned} \tag{3.45}$$

Then

$$\begin{aligned} V(\widehat{F}_n(x)) \approx \frac{1}{n}\sum_{k\geq 1} & \{ \quad [\mu(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k} - \mu F(x)\frac{\lambda_n}{k}]^2[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n}]\} \\ -\frac{1}{n}(\sum_{k\geq 1} & \{ \quad [\mu(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k} - \mu F(x)\frac{\lambda_n}{k}][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n}]\})^2 \\ = \frac{1}{n}\sum_{k\geq 1} & \{ \quad [\mu(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k}]^2[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})]\} \\ -\frac{2}{n}\sum_{k\geq 1} & \{ \quad [\mu^2(\int_0^x p_{k-1}(t\lambda_n)dt)F(x)\frac{\lambda_n^3}{k^2}][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})]\} \\ +\frac{1}{n}\sum_{k\geq 1} & \{ \quad [\mu F(x)\frac{\lambda_n}{k}]^2[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})]\} \\ -\frac{1}{n}(\sum_{k\geq 1} & \{ \quad [\mu(\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k} - \mu F(x)\frac{\lambda_n}{k}][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n}]\})^2 \\ = \quad & \frac{1}{n} \quad \big(T_{1n} - 2T_{2n} + T_{3n} - T_{4n}\big). \end{aligned} \tag{3.46}$$

We write

$$\begin{aligned} T_{1n} &= \sum_{k\geq 1}\mu^2(\lambda_n\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k^2}[G(\frac{k}{\lambda_n})-G(\frac{k-1}{\lambda_n})] \\ &\quad -\sum_{k\geq 1}\mu^2(1-\lambda_n\int_0^x p_{k-1}(t\lambda_n)dt)(\lambda_n\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k^2}[G(\frac{k}{\lambda_n})-G(\frac{k-1}{\lambda_n})] \\ &= S_1 - S_2. \end{aligned} \tag{3.47}$$

Using Hille's Lemma, we can easily show that, as $\lambda_n \uparrow \infty$,

$$\sum_{k\geq 1} p_{k-1}(t\lambda_n)\frac{\lambda_n^2}{k^2}g(\frac{k}{\lambda_n}) \to \frac{1}{t^2}g(t)$$

uniformly in the finite interval $[0, x]$, then we have

$$S_1 \approx \mu^2\int_0^x \big(\sum_{k\geq 1} p_{k-1}(t\lambda_n)\frac{\lambda_n^2}{k^2}g(\frac{k}{\lambda_n})\big)dt \to \mu^2\int_0^x \frac{1}{t^2}g(t)dt. \tag{3.48}$$

Next, we will show that $S_2 \to 0$ as $\lambda_n \uparrow \infty$.

Let $\mathbf{N} = \{1, 2, \ldots, n, \ldots\}$ and $b_n = \lambda_n^{-1/2}(\log n)^{\frac{1+\delta}{2}}$ where $\delta > 0$. Denote $\mathbf{N}_x^1 = \{k \big| k/\lambda_n - x < -b_n, k \in \mathbf{N}\}$, $\mathbf{N}_x^2 = \{k \big| |k/\lambda_n - x| \leq b_n, k \in \mathbf{N}\}$ and $\mathbf{N}_x^3 = \{k \big| k/\lambda_n - x > b_n, k \in \mathbf{N}\}$.

Let

$$a_k = \mu^2(1-\lambda_n\int_0^x p_{k-1}(t\lambda_n)dt)(\lambda_n\int_0^x p_{k-1}(t\lambda_n)dt)\frac{\lambda_n^2}{k^2}[G(\frac{k}{\lambda_n})-G(\frac{k-1}{\lambda_n})], \tag{3.49}$$

then we can write

$$S_2 = \sum_{k\in\mathbf{N}_x^1} a_k + \sum_{k\in\mathbf{N}_x^2} a_k + \sum_{k\in\mathbf{N}_x^3} a_k. \tag{3.50}$$

For any $k \in \mathbf{N}_x^1$, by the proof of Lemma 3.1 of Chaubey and Sen (1996), we can claim that $[1-\lambda_n\int_0^x p_{k-1}(t\lambda_n)] = \sum_0^{k-1} p_i(x\lambda_n) < \frac{1}{n}$. Then

$$0 < \sum_{k\in\mathbf{N}_x^1} a_k < \frac{1}{n}S_1. \tag{3.51}$$

For any $k \in \mathbf{N}_x^3$, by the same lemma above, we can claim that $\lambda_n \int_0^x p_{k-1}(t\lambda_n) = \sum_{i\geq k}^{\infty} p_i(x\lambda_n) < \frac{1}{n}$. At the same time, we have $[1-\lambda_n \int_0^x p_{k-1}(t\lambda_n)] < 1$ and $(k/\lambda_n)^2 < (x+b_n)^{-2}$. Then

$$0 < \sum_{k\in\mathbf{N}_x^3} a_k < \frac{\mu^2}{n}(x+b_n)^{-2}. \tag{3.52}$$

For any $k \in \mathbf{N}_x^2$, by the facts $[1-\lambda_n \int_0^x p_{k-1}(t\lambda_n)] < 1$, $\lambda_n \int_0^x p_{k-1}(t\lambda_n) < 1$ and $(k/\lambda_n)^2 < (x-b_n)^{-2}$, we have

$$0 < \sum_{k\in\mathbf{N}_x^2} a_k < \mu^2(x-b_n)^{-2}[G(x+b_n) - G(x-b_n)]. \tag{3.53}$$

From expressions of (3.51), (3.52) and (3.53), we can see that they all tend to zero as $\lambda_n = O(n^\alpha) \uparrow \infty$, which means

$$S_2 \to 0. \tag{3.54}$$

By (3.48) and (3.54), we have

$$T_{1n} \to \mu^2 \int_0^x \frac{1}{t^2} g(t)dt = \mu \int_0^x \frac{1}{t} f(t)dt. \tag{3.55}$$

At the same time, by Hille's lemma, we have

$$\begin{aligned} T_{2n} &\approx \sum_{k\geq 1} \{[\mu^2(\int_0^x p_{k-1}(t\lambda_n)dt)F(x)\frac{\lambda_n^3}{k^2}]\frac{1}{\lambda_n} g(\frac{k}{\lambda_n}) \\ &= \sum_{k\geq 1} \{[\mu^2(\int_0^x p_{k-1}(t\lambda_n)dt)F(x)\frac{\lambda_n^2}{k^2}] g(\frac{k}{\lambda_n}) \\ &\to \mu F(x) \int_0^x \frac{1}{t} f(t)dt, \end{aligned} \tag{3.56}$$

$$
\begin{aligned}
T_{3n} &\approx \sum_{k\geq 1}\{[\mu F(x)\frac{\lambda_n}{k}]^2\frac{1}{\lambda_n}g(\frac{k}{\lambda_n}) \\
&= \sum_{k\geq 1}\{[\mu F(x)]^2\frac{1}{\lambda_n}(\frac{k}{\lambda_n})^{-2}g(\frac{k}{\lambda_n}) \\
&\to \mu F^2(x)\int_0^\infty \frac{1}{t}f(t)dt \\
&= \mu\bar{\mu}F^2(x) \qquad (3.57)
\end{aligned}
$$

where $\bar{\mu} = \int_0^\infty \frac{1}{t}f(t)dt$, and, by (3.43),

$$T_{4n} = [E(\widehat{F}_n(x) - F(x))]^2 \to 0. \qquad (3.58)$$

By (3.46), (3.55), (3.56), (3.57) and (3.58), we have

$$V(\sqrt{n}\widehat{F}_n(x)) \to \mu\left[\int_0^x \frac{1}{t}f(t)dt - 2F(x)\int_0^x \frac{1}{t}f(t)dt + \bar{\mu}F^2(x)\right]. \qquad (3.59)$$

The proof is complete.

3.3 Estimators of Density and Distribution Functions with Asymmetric Kernel

3.3.1 Smooth Density Estimator

Using generalized Hille's lemma, we can obtain a smooth estimator of $g(x)$ [see Chaubey, Sen and Sen (2007)] which is given by

$$g_n(x) = \frac{1}{nx^2}\sum_{i=1}^n X_i q_{v_n}(\frac{X_i}{x}). \qquad (3.60)$$

Note that $\int_0^\infty \frac{g_n(x)}{x}dx = \frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}$, which is an estimator of $1/\mu$. By (3.4), a smooth estimator of $f(x)$ can be formed as

$$
\begin{aligned}
f_n(x) &= \frac{g_n(x)/x}{\int_0^\infty (g_n(x)/x)dx} \\
&= \frac{\frac{1}{x^3}\sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{x}\right)}{\sum_{i=1}^n \frac{1}{X_i}}. \qquad (3.61)
\end{aligned}
$$

However, (3.61) may be reasonable for a density $f(x)$ with $f(0) = 0$. For general density functions, by the idea of perturbation, the acceptable estimator is given by

$$
\hat{f}_n^+(x) = \frac{\frac{1}{(x+\varepsilon_n)^3}\sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{x+\varepsilon_n}\right)}{\sum_{i=1}^n \frac{1}{X_i}} \qquad (3.62)
$$

3.3.1.1 Asymptotic Properties

Thoerem 3.5 *If*

A. $v_n \to 0$, $\varepsilon_n \to 0$ *as* $n \to \infty$ *and* $E(X_1^{-1}) < \infty$;

B. $\sup_{x\geq 0}\int_0^\infty |\frac{d}{dx}[\frac{1}{x+\varepsilon_n}q_{v_n}(\frac{t}{x+\varepsilon_n})]|dt = o((\frac{\log\log n}{n^{1/2}})^{-1})$;

C. $\sup_{u>0,v>0} uq_v(u) < \infty$;

D. $g(\cdot)$ *is Lipschitz continuous on* $[0,\infty)$;

then we have, as $n \to \infty$,

$$
\sup_{x>0}|\hat{f}_n^+(x) - f(x)| \xrightarrow{a.s.} 0 \qquad (3.63)
$$

Proof: By Theorem 3 of Chaubey, Sen and Sen (2007), under the conditions of Theorem 3.5, we have

$$
\sup_{x>0}|g_n^+(x) - g(x)| \xrightarrow{a.s.} 0 \qquad (3.64)
$$

where $g_n^+(x) = \frac{1}{n(x+\epsilon_n)^2}\sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{x+\varepsilon_n}\right)$.

On the other hand, by the strong law of large number, we have

$$
\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} \xrightarrow{a.s.} \frac{1}{\mu} \qquad (3.65)
$$

By (3.62), (3.64) and (3.65), we can obtain the theorem.

Thoerem 3.6 *If*

E. $g(\cdot)$ *is Lipschitz continuous on* $[0, \infty)$ *and* $E(X_1^{-2}) < \infty$*;*

F. $I_2(q) \triangleq \lim_{v \to 0} v \int_0^\infty (q_{v_n}(t))^2 dt$ *exists;*

G1. for $1 \le m \le 3$, $\int_0^\infty (q_{v_n}(t))^m dt = O(v^{1-m})$ *as* $v \to 0$*;*

G2. with $q^*_{m,v_n}(t) = \frac{(q_{v_n}(t))^m}{\int_0^\infty (q_{v_n}(w))^m dw}$, $1 \le m \le 3$, *and as* $v_n \to 0$,

$$\begin{aligned}
&(i) \qquad \mu_{m,v_n} = \int_0^\infty t q^*_{m,v_n}(t) dt = 1 + O(v_n),\\
&(ii) \qquad \sigma^2_{m,v_n} = \int_0^\infty (t - \mu_{m,v_n})^2 q^*_{m,v_n}(t) dt = O(v_n^2)\\
&(iii) \qquad \sup_{0 < v_n < \varepsilon} \int_0^\infty t^{4+\delta} q^*_{m,v_n}(t) dt < \infty, \textit{for some } \delta > 0,\ \varepsilon > 0;
\end{aligned}$$

Then

(a) If $n v_n \to \infty$, $n v_n \varepsilon_n \to \infty$, $n v_n^3 \to 0$, $n v_n \varepsilon_n^2 \to 0$ *as* $n \to \infty$, *we have*

$$\sqrt{n v_n}(\hat{f}_n^+(x) - f(x)) \to N\big(0, I_2(q)\frac{\mu f(x)}{x^2}\big), \textit{for } x > 0.$$

(b) If $n v_n \epsilon_n^2 \to \infty$ *and* $n v_n \varepsilon_n^4 \to 0$ *as* $n \to \infty$, *we have*

$$\sqrt{n v_n \epsilon_n^2}(f_n^+(0) - f(0)) \to N\big(0, I_2(q) f(0)\big).$$

Proof: (a) Since $g_n^+(x)$ is a density obtained by the method in Chaubey, Sen and Sen (2007), according to the proof of Theorem 4 in Chaubey, Sen and Sen (2007), we have

$$\sqrt{n v_n}(g_n^+(x) - g(x)) \to N\left(0, I_2(q)\frac{g(x)}{x}\right), \text{for } x > 0. \tag{3.66}$$

Note that $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} - \frac{1}{\mu}\right) \to \textit{normal}$. So the asymptotic normality of

$$\hat{f}_n^+(x) = \frac{g_n^+(x)}{(x + \epsilon_n)\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}} \tag{3.67}$$

is equivalent to the the asymptotic normality of

$$\frac{\mu g_n^+(x)}{(x + \epsilon_n)}. \tag{3.68}$$

Using (3.66), it is easy to show that

$$\sqrt{nv_n}\left(\frac{\mu g_n^+(x)}{(x+\epsilon_n)} - \frac{\mu g(x)}{x}\right) \to N\left(0, I_2(q)\frac{\mu f(x)}{x^2}\right), \text{for } x > 0. \tag{3.69}$$

Then the part (a) of theorem follows.

(b)we have

$$\hat{f}_n^+(0) = \frac{1}{n}\sum_{i=1}^{n} Y'_{in} - f(0)\big(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}\big) + o\big(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} - \frac{1}{\mu}\big) \tag{3.70}$$

where

$$Y'_{in} = \frac{\mu X_i}{(\varepsilon_n)^3} q_{v_n}\big(\frac{X_i}{\varepsilon_n}\big). \tag{3.71}$$

We can show that

$$E(Y_{1n}^{'3}) = O(\epsilon_n^4 v_n^{-2}) \tag{3.72}$$

and

$$\epsilon_n^2 v_n E(Y_{1n}^{'2}) \to I_2(q) f(0). \tag{3.73}$$

Using (3.72) and (3.73) and following the lines of proof of Theorem 2.8 of part (b), we can obtain part (b) of this theorem.

3.3.1.2 MSE and AMISE

We can show that

$$\begin{aligned} Bias[\hat{f}_n^+(x)] &= v_n^2 f(x) + (2v_n^2 x + \epsilon_n) f'(x) \\ &\quad + v_n^2 \frac{x^2}{2} f''(x) + o(v_n^2 + \epsilon_n) \end{aligned} \tag{3.74}$$

and

$$Var[\hat{f}_n^+(x)] = \frac{I_2(q)\mu f(x)}{nv_n(x+\epsilon_n)^2} + o\big((nv_n)^{-1}\big). \tag{3.75}$$

So we have

$$MSE[\hat{f}_n^+(x)] \approx \frac{I_2(q)\mu f(x)}{nv_n x^2} + [v_n^2 f(x) + (2v_n^2 x + \epsilon_n)f'(x) + v_n^2 \frac{x^2}{2} f''(x)]^2. \tag{3.76}$$

Furthermore, we have

$$\begin{aligned} AMISE[\hat{f}_n^+(x)] &= \frac{I_2(q)\mu}{nv_n} \int_0^\infty \frac{f(x)}{(x+\epsilon_n)^2} dx \\ &+ \int_0^\infty [v_n^2 f(x) + (2v_n^2 x + \epsilon_n)f'(x) + v_n^2 \frac{x^2}{2} f''(x)]^2 dx \end{aligned} \tag{3.77}$$

This expression is useful in cross validation method for obtaining data dependent values of smoothing parameter(s).

3.3.1.3 Corrected Density Estimator

Note that if $\epsilon_n > 0$ the integral of (3.62) is less than 1. In this case, the density estimator seems a little left-shifted and slightly " lose " some weights. In order to get the " lost " weights back, we divide (3.62) by its integral $\int_0^\infty \hat{f}_n^+(x)dx$ and obtain the following corrected density estimator

$$\hat{f}_n^*(x) = \frac{\frac{1}{(x+\varepsilon_n)^3} \sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{x+\varepsilon_n}\right)}{\int_0^\infty \frac{1}{(x+\varepsilon_n)^3} \sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{x+\varepsilon_n}\right) dx}. \tag{3.78}$$

Since, as $\epsilon_n \to 0$, $\int_0^\infty \frac{1}{(x+\varepsilon_n)^3} \sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{x+\varepsilon_n}\right) dx \to \sum_{i=1}^n \frac{1}{X_i}$, we can have the same theorem as Theorem 3.5 and 3.6 for $\hat{f}_n^*(x)$. Furthermore, it is easy to show that

$$Bias(\hat{f}_n^*) = Bias(\hat{f}_n^+) + \epsilon_n f(0) f(x) + o(\epsilon_n). \tag{3.79}$$

3.3.2 Smooth Estimator of Distribution Function

According to (3.61), an acceptable distribution function estimator is

$$\widehat{F}_n^+(x) = \frac{\int_0^x \frac{1}{t^3} \sum_{i=1}^n X_i q_{v_n}\left(\frac{X_i}{t}\right) dt}{\sum_{i=1}^n \frac{1}{X_i}}. \tag{3.80}$$

If q_{v_n} is a gamma density, an alternative form of (3.80) is given by

$$\widehat{F}_n^+(x) = \frac{\sum_{i=1}^n \frac{1}{X_i}\left(1 - F_{1/v_n^2+1, v_n^2}\left(\frac{X_i}{x}\right)\right)}{\sum_{i=1}^n \frac{1}{X_i}}$$

where $F_{1/v_n^2+1, v_n^2}(x)$ is a gamma distribution function with density function $\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$ with $\alpha = 1/v_n^2 + 1$ and $\beta = v_n^2$.

Thoerem 3.7 *If $E(X_1^{-1}) < \infty$, $v_n \to 0$ as $n \to \infty$ and $f(x)$ is absolutely continuous with bounded derivative $f'(x)$ on $\mathbf{R}^+$, then*

$$\sup_{x \geq 0} |\widehat{F}_n^+(x) - F(x)| \xrightarrow{a.s.} 0.$$

Proof: We can write (3.80) as

$$\widehat{F}_n^+(x) = \frac{\frac{1}{n}\sum_{i=1}^n \int_0^x \frac{1}{t^3} X_i q_{v_n}\left(\frac{X_i}{t}\right) dt}{\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}}. \tag{3.81}$$

Let $\xi_i = \int_0^x \frac{1}{t^3} X_i q_{v_n}\left(\frac{X_i}{t}\right) dt$, then

$$E(\xi_i) = \int_0^\infty \left[\int_0^x \frac{1}{t^3} y q_{v_n}\left(\frac{y}{t}\right) dt\right] g(y) dy. \tag{3.82}$$

Let $y/t = z$, then

$$\begin{aligned} E(\xi_i) &= \int_0^\infty \left[\int_{y/x}^\infty z q_{v_n}(z) dz\right] \frac{g(y)}{y} dy \\ &= \frac{1}{\mu}\int_0^\infty \left[\int_{y/x}^\infty z q_{v_n}(z) dz\right] f(y) dy \\ &= \frac{1}{\mu}\left[F(y)\int_{y/x}^\infty z q_{v_n}(z) dz\right]_0^\infty + \frac{1}{\mu}\int_0^\infty F(y)\frac{y}{x} q_{v_n}\left(\frac{y}{x}\right)\frac{dy}{x}. \end{aligned} \tag{3.83}$$

Let $\frac{y}{x} = u$, then

$$E(\xi_i) = \frac{1}{\mu}\int_0^\infty F(ux)uq_{v_n}(u)\,du. \tag{3.84}$$

Note that

$$F(ux) = F(x) + xf(x)(u-1) + O(u-1)^2,$$

then

$$E(\xi_i) = F(x) + O(v_n^2). \tag{3.85}$$

By the strong law of large number, we have that the numerator of (3.81) converges uniformly to $\frac{1}{\mu}F(x)$ and the denominator of (3.81) to $\frac{1}{\mu}$. Since $\widehat{F}_n^+(x)$ is nondecreasing, the uniform strong convergency of $\widehat{F}_n^+(x)$ follows.

Thoerem 3.8 *If $E(X_1^{-2}) < \infty$, $\sqrt{n}v_n^2 \to 0$, $f(x)$ is absolutely continuous with bounded derivative $f'(x)$ on $\mathbf{R}^+$, then, as $n \to \infty$,*

$$\sqrt{n}(\widehat{F}_n^+(x) - F(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x))$$

where

$$\delta^2(x) = \mu[\int_0^x \frac{1}{t}f(t)dt - 2F(x)\int_0^x \frac{1}{t}f(t)dt + \bar{\mu}F^2(x)]$$

where $\bar{\mu} = E_f(\frac{1}{X})$.

Proof: We can expand (3.80) as

$$\begin{aligned}
\widehat{F}_n^+(x) &= F(x) + \left[\frac{\mu}{n}\sum_{i=1}^n \int_0^x \frac{1}{t^3}X_i q_{v_n}(\frac{X_i}{t})dt - F(x)\right] \\
&\quad -\mu F(x)\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} - \frac{1}{\mu}\right) + o\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} - \frac{1}{\mu}\right) \\
&= F(x) + \frac{1}{n}\sum_{i=1}^n \left[\int_0^x \frac{\mu}{t^3}X_i q_{v_n}(\frac{X_i}{t})dt - \frac{\mu F(x)}{X_i}\right] \\
&\quad + o\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} - \frac{1}{\mu}\right)
\end{aligned} \tag{3.86}$$

Let $\eta_i = \int_0^x \frac{\mu}{t^3} X_i q_{v_n}\left(\frac{X_i}{t}\right) dt - \frac{\mu F(x)}{X_i}$. Since $E\left(\frac{1}{X_i}\right) = \frac{1}{\mu}$, we can show that as in the proof of Theorem 3.7,

$$E(\eta_i) = O(v_n^2). \tag{3.87}$$

Furthermore, we have

$$\begin{aligned} E(\eta_i^2) &= \int_0^\infty \left[\int_0^x \frac{\mu}{t^3} y q_{v_n}\left(\frac{y}{t}\right) dt\right]^2 g(y)dy - 2\mu F(x) \int_0^\infty \left[\int_0^x \frac{\mu}{t^3} q_{v_n}\left(\frac{y}{t}\right) dt\right] g(y)dy \\ &\quad + \mu F^2(x) \int_0^\infty \frac{\mu}{t^2} g(t)dt \\ &= T_1 - 2\mu F(x) T_2 + \mu F^2(x)\bar{\mu}. \end{aligned} \tag{3.88}$$

Let $y/t = z$ and $H(x) = \int_0^x \frac{f(t)}{t} dt$, for T_1, we have

$$\begin{aligned} T_1 &= \int_0^\infty \left[\int_{y/x}^\infty \mu z q_{v_n}(z)dz\right]^2 \frac{g(y)}{y^2} dy \\ &= \mu \int_0^\infty \left[\int_{y/x}^\infty z q_{v_n}(z)dz\right]^2 \frac{f(y)}{y} dy \\ &= \mu \left[H(y) \left[\int_{y/x}^\infty z q_{v_n}(z)dz\right]^2\right]_0^\infty \\ &\quad - \mu \int_0^\infty H(y) d\left[\int_{y/x}^\infty z q_{v_n}(z)dz\right]^2 \\ &= -\mu \int_0^\infty H(y) d\left[\int_{y/x}^\infty z q_{v_n}(z)dz\right]^2. \end{aligned} \tag{3.89}$$

Let $y/x = u$, then we have

$$T_1 = -\mu \int_0^\infty H(xu) d\left[\int_u^\infty z q_{v_n}(z)dz\right]^2. \tag{3.90}$$

Note that we have

$$H(xu) = H(x) + f(x)(u-1) + o(u-1),$$

then

$$\begin{aligned} T_1 &\approx \mu H(x) - \mu f(x) \int_0^\infty (u-1) d\left[\int_u^\infty z q_{v_n}(z)dz\right]^2 \\ &= \mu H(x) - \mu f(x) T_3. \end{aligned} \tag{3.91}$$

For T_3, we have

$$\begin{aligned}
|T_3| &= 2 \mid \int_0^\infty (u-1) \left[\int_u^\infty z q_{v_n}(z) dz \right] u q_{v_n}(u) du \mid \\
&\le 2 \int_0^\infty |u(u-1)| q_{v_n}(u) du \\
&\le o(v_n). \qquad (3.92)
\end{aligned}$$

So, as $v_n \to 0$,

$$T_1 \to \mu H(x). \qquad (3.93)$$

Similarly we have

$$\begin{aligned}
T_2 &= \int_0^\infty H(xu) u q_{v_n}(u) du \\
&= H(x) + o(v_n) \qquad (3.94)
\end{aligned}$$

By (3.86), (3.87), (3.88), (3.93) and (3.94), we can obtain the theorem.

Chapter 4

A Numerical Study of the New Estimators

4.1 Introduction

In this chapter, we propose to compare various density estimators described in the previous chapters through extensive simulation. The basic criteria are mean squared error (MSE) and mean integrated squared error ($MISE$) of the estimator f_n given by

$$MSE(f_n(x)) = E[(f_n(x) - f(x))^2] \tag{4.1}$$

and

$$MISE(f_n, f) = E\left[\int_0^\infty (f_n(x) - f(x))^2 dx\right] \tag{4.2}$$

Note that MSE may be considered to measure the local performance of the estimator f_n and $MISE$ may be considered to measure the global performance. In practical applications, since f is to be estimated, data-dependent choices corresponding to the above criteria are considered. These are commonly known as "*cross validation*" methods that attempt to estimate above quantities based on the observed sample which are in turn minimized (numerically) as a function of smoothing parameters. Such cross-validation method may consider another measure of departure of the estimator f_n from f instead of the integrated squared error and that would give a different choice of the parameters. So the question may be which measure of departure may be better suited to amplify the differences between the estimators and the true density?

Hence we first study this question in the next section where the candidate estimator is the Poisson based smoothing of the Cox estimator. The conclusion from the simulation studies points towards the conjecture that the the data dependent integrated error (ISE) cross-validation methods provide optimal choice of the smoothing parameter(s) for large samples in the sense of minimizing the MISE and that the choice of departure measure is not of much relevance.

The next section, Section 4.3 therefore considers ISE cross-validation methods for all the density estimators and presents a comparison of MISE and MSE for some known standard densities and Section 4.4 presents the conclusions.

In these expositions, we will mainly proposed two kinds of data-driven methods, one being unbiased cross-validation method, the other being biased cross-validation method as commonly used in the literature dealing with kernel density estimation [see Scott and Terrell (1987)].

It seen that the performance of an estimator based on F_n may be better than that based on G_n over some region but not on another region. Hence, in Section 4.5, we propose a linear combination of two competing density estimators and investigate its properties numerically through simulation.

4.2 A Comparison of Different Criteria for Selecting Smoothing Parameters: The Case of $\tilde{f}_n$

Selection of smoothing parameters is an old and challenging topic in nonparametric functional estimator. Since smoothing parameter determines the performances of estimator under finite samples, it is an important issue in practice and many methods have been proposed proposed for this purpose. In this area, the early work was done by Kronmal and Tarter (1968). Rudemo (1982) proposed a least squares cross-validation method. Bowman (1984), using Kullback-Liebler divergence, proposed an alternative cross-validation method. Using asymptotic $MISE$, Scott and Terrell (1987) proposed biased cross-validation method. The further modification of this method was made by Park and Marron (1990). Here we will propose and study several selection methods for our proposed density estimator of LB data.

A convenient stochastic choice of λ_n was proposed by Chaubey and Sen (1996) as

$$\lambda_{n(1)} = \frac{n}{\max\{X_1, \ldots, X_n\}} \tag{4.3}$$

provided that $E(X) < \infty$ and X has an infinite support. However, if X has a finite support, Chaubey and Sen (1998) noticed that the choice (4.3) will not satisfy that $n^{-1}\lambda_n \to 0$ as $n \to \infty$. To cover these cases they proposed the choice

$$\lambda_{n(2)} = \frac{n}{X_{n-r_n+1:n} \log\log n} \tag{4.4}$$

where $r_n = o(\log\log n)$. Based on the asymptotic property of MSE of $\tilde{f}_n(x)$, a non-stochastic choice is

$$\lambda_{n(3)} = cn^{2/5} \tag{4.5}$$

These choices are based on the asymptotic theory, however, in finite sample case they may not be satisfactory. In the procedure of using direct data to estimate density, Chaubey and Sen (2009) find that the choices $\lambda_{n(1)}$ and $\lambda_{n(2)}$ may be very large so that they create problems in computation. Our study shows that they may also cause the same problems in the procedure of using LB data to estimate density. The purpose of this subsection is to give the choices of λ_n for finite samples. We will investigate two kinds of cross-validation methods, one is unbiased cross-validation method, the other is biased cross-validation method.

4.2.1 Unbiased Cross-Validation Method

Here we investigate two unbiased cross-validation methods, one being based on Kullback-Liebler divergence, the other being based on integrated squared error. We also use the Hellinger distance defined between two densities to compare the closeness of a density estimator to its true population density.

4.2.1.1 Kullback-Liebler Divergence Cross Validation

The Kullback-Liebler divergence between the two density functions $f(x)$ and $g(x)$ is defined as

$$KL(f, g) = \int f(x) \log \frac{f(x)}{g(x)} dx \tag{4.6}$$

So the Kullback-Liebler divergence between the estimator $\tilde{f}_n(x)$ and the true density $f(x)$ is given by

$$KL(f, \tilde{f}_n) = \int f(x) \log f(x) dx - \int f(x) \log \tilde{f}_n(x) dx \tag{4.7}$$

By the fact that $f(x) = \mu g(x)/x$, we can also write (4.7) as

$$KL(f, \tilde{f}_n) = \int f(x) \log f(x) dx - \int (\mu g(x)/x) \log \tilde{f}_n(x) dx \tag{4.8}$$

The leave-one-out estimator of the second term is

$$-\frac{\mu}{n} \sum_{i=1}^{n} \log \frac{\tilde{f}_{n-1}(X_i, \lambda_n; \mathcal{D}_i)}{X_i} \tag{4.9}$$

where $\mathcal{D}_i$ denotes data with X_i removed from $\mathcal{D}$. Since the first term in (4.8) does not depend on λ_n, dispensing with the constant μ and n in (4.9), minimizing (4.8) is equivalent to minimize

$$CV_{KL}(\lambda_n) = -\sum_{i=1}^{n} \log(\tilde{f}_{n-1}(X_i, \lambda_n; \mathcal{D}_i)/X_i) \tag{4.10}$$

The solution of the above minimization problem will be denoted by λ_{nKL}.

4.2.1.2 Integrated Squared Error Cross Validation

The integrated squared error between $\tilde{f}_n(x)$ and $f(x)$ is given by

$$\begin{aligned} ISE(f, \tilde{f}_n) &= \int (f(x) - \tilde{f}_n(x))^2 dx \\ &= \int f^2(x) dx - 2 \int \tilde{f}_n(x) f(x) dx + \int \tilde{f}_n^2(x) dx \end{aligned} \tag{4.11}$$

In the studies of bandwidth choice for kernel density estimates with selection biased data, Wu (1997) used the leave-one-out estimator

$$2 \sum_{i=1}^{n} \tilde{f}_{n-1}(X_i, \lambda_n; \mathcal{D}_i)/Z_i \tag{4.12}$$

to estimate the second term in (4.11) where $Z_i = \sum_{j \neq i} \frac{X_i}{X_j}$. Substituting the leave-one-out estimator (4.12) for the second term in (4.11) and subtracting the first term which does not depend on λ_n gives us the following cross-validation function for ISE criteria

$$CV_{ISE}(\lambda_n) = \int_0^{\infty} \tilde{f}_n^2(x)dx - 2\sum_{i=1}^{n} \tilde{f}_{n-1}(X_i, \lambda_n; \mathcal{D}_i)/Z_i \tag{4.13}$$

The solution of the above minimization problem will be denoted by λ_{nISE}.

4.2.1.3 Hellinger Distance

The Hellinger distance between two density functions $f(x)$ and $g(x)$ is given by

$$H(f, g) = \int \left(\sqrt{f(x)} - \sqrt{g(x)}\right)^2 dx \tag{4.14}$$

This measure has a good property, as shown in Chaubey and Sen (2009), that is

$$0 \leq H(f, g) \leq 2 \tag{4.15}$$

We will use this measure to establish the closeness of the estimated density to the true density in finite samples.

4.2.1.4 Simulation Studies for Optimal Smoothing Parameter: The Case of $\tilde{f}_n$

Lognormal Density

To understand the possible numerical intricacies in obtaining the value of the smoothing parameter λ_n, we simulate samples from a standard Lognormal density for sample size n=10, 20, 30, 40, 50, 100. For each sample we obtain the optimum choice of λ_n by KL and ISE cross validation methods. To judge the closeness between the estimated density and the true density we list the Hellinger distance $H(\tilde{f}_n, f)$ for each choice of λ_n. Here we use the routine optimise of R language to obtain the optimum solution of λ_n for KL and ISE cross validation methods. However, we must be careful because the function $CV(\lambda_n)$ is a rough function [For details see Chaubey and Sen (2009)].

For 100 samples, we use routine optimise with an interval $(1, 20)$ to obtain optimum solutions 6.482327 and 5.770364 for KL and ISE criterion, respectively. In this case, the optimum solution of $H(\tilde{f}_n, f)$ is 5.750105. To make sure of these solutions, we plot the $CV(\lambda_n)$ functions [see Figure 4.1]. Checking these plots, the solutions see reasonable. At the same time, Chaubey-Sen choice is 3.445616. The Hellinger distance of the estimated density using Chaubey-Sen, KL, ISE with the true lognormal density are given by 0.05391589, 0.05036821 and 0.04514759 respectively which are close to the true distance 0.04298512 if we know the density.

Here we also plot the estimated densities and compare with the histogram of Lognormal distribution. The histogram estimator of Lognormal distribution is given by

$$\hat{f}_{his}(x) = \frac{F_n(x_i) - F_n(x_j)}{x_i - x_j} \;\; if \;\; x \in (x_i, x_j)$$

where $F_n(x)$ is defined as in (1.3). Looking at the Figure 4.2, we can find that there is almost no difference in them qualitatively. It may conclude that as long as the value of λ_n is in the close neighborhood of minima, the estimated density does not differ very much from the optimum choice.

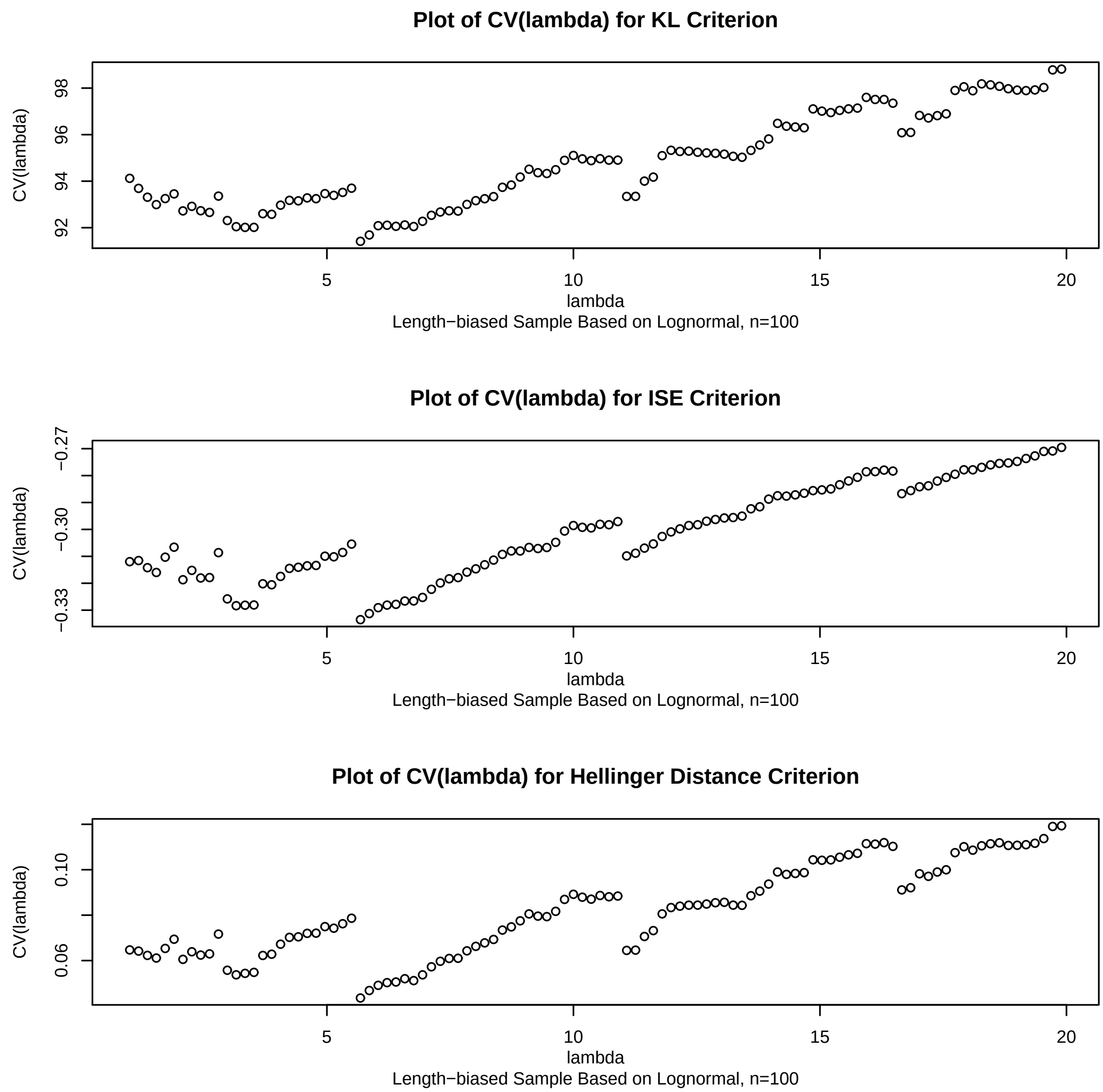


Figure 4.1: $CV(\lambda)$ Plots, Sample Size=100

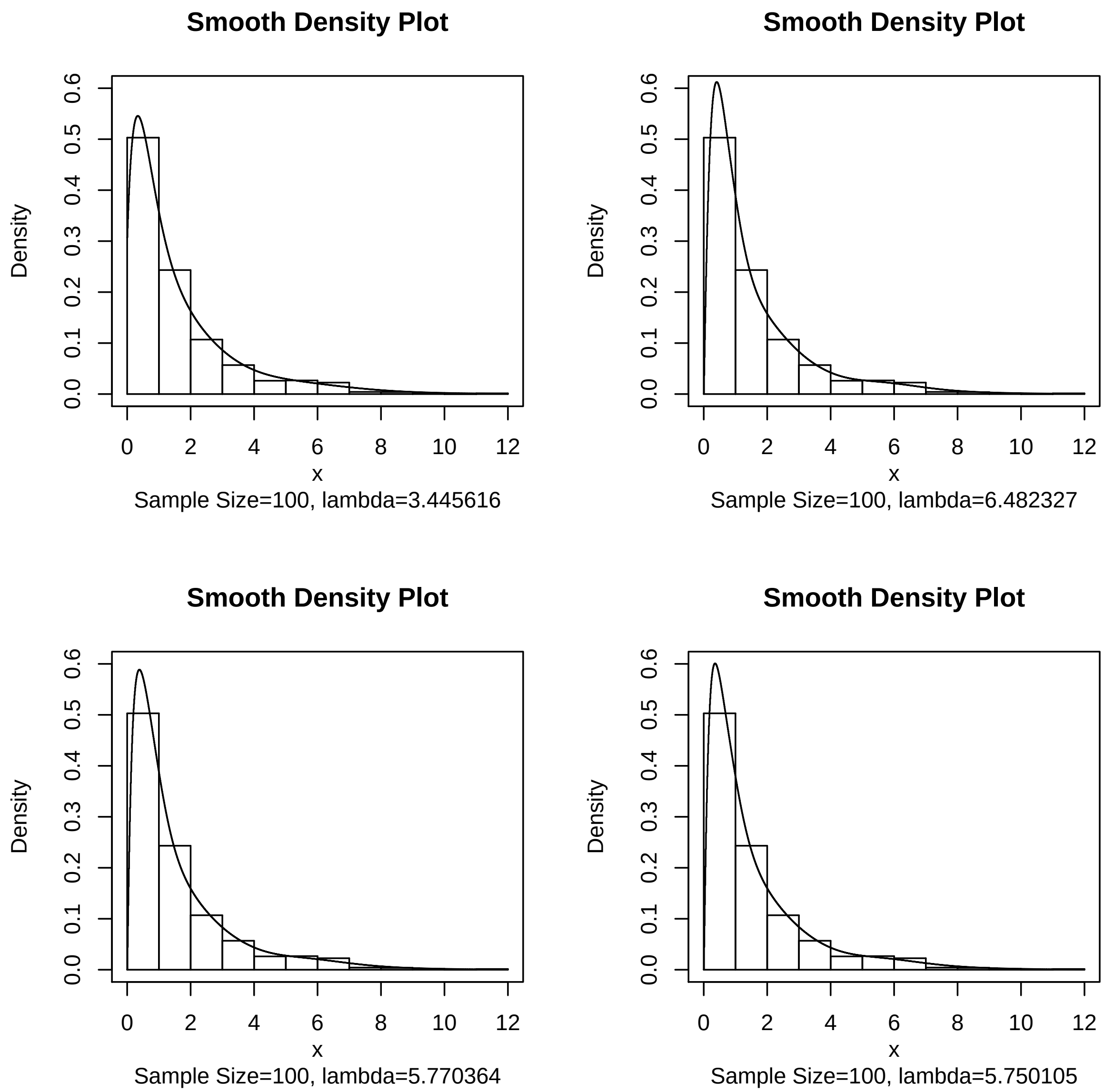


Figure 4.2: Smooth Density Plots, Sample Size=100

Some Other Standard Distributions

Next we consider the following densities in place of the Lognormal density and repeat the steps described earlier in selecting the smoothing parameter:

(i). Exponential Distribution

$$f(x) = \exp(-x)I\{x > 0\}$$

(ii). Lognormal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi}x} \exp\{-(\log x - \mu)^2/2\}I\{x > 0\}$$

(iii). Gamma Distribution

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

(iv). Weibull Distribution

$$f(x) = \alpha x^{\alpha-1} \exp(-x^{\alpha})I\{x > 0\}$$

(v). Mixtures of Two Exponential Distribution

$$f(x) = [\pi\frac{1}{\theta_1} \exp(-x/\theta_1) + (1-\pi)\frac{1}{\theta_2} \exp(-x/\theta_2]I\{x > 0\}$$

The methods of generating corresponding LB data are given by, respectively,

(i'). $X \sim \Gamma(2, 1)$;

(ii'). $X = e^Y$ where $Y \sim N(\mu + 1, 1)$;

(iii'). $X \sim \Gamma(\alpha + 1, 1)$;

(iv'). $X = Y^{1/\alpha}$ where $Y \sim \Gamma(1+\frac{1}{\alpha}, 1)$;

(v'). $X = \pi Y_1 + (1-\pi)Y_2$, where $Y_1 \sim \Gamma(2, \theta_1)$ and $Y_2 \sim \Gamma(2, \theta_2)$.

Remark 4.1: The methods of generating LB data (i'), (iii'), (v') are straightforward. Here we give brief proofs of (ii') and (iv'). For (ii'), if let $f_w(x)$ denote the density of LB data, then we need to show that $f_w(x) \propto \exp\{-(\log x - \mu)^2/2\} I\{x > 0\}$. Let $Y \sim N(\mu+1, 1)$, then

$$
\begin{aligned}
F_w(x) = P(X \le x) &= P(e^Y \le x) \\
&= P(Y \le \log x) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\log x} e^{-(z-\mu-1)^2/2} dz.
\end{aligned}
$$

So $f_w(x) \propto e^{-(\log x - \mu - 1)^2/2}/x \propto e^{-(\log x - \mu)^2/2}$.

For (iv'), let $Y \sim \Gamma(1+\frac{1}{\alpha}, 1)$, $F^w(x)$ be the distribution function of X, then

$$
\begin{aligned}
F_w(x) = P(X \le x) &= P(Y^{\frac{1}{\alpha}} \le x) \\
&= P(Y \le x^\alpha) \\
&= \frac{1}{\Gamma(1+\frac{1}{\alpha})} \int_0^{x^\alpha} z^{1/\alpha} e^{-z} dx.
\end{aligned}
$$

So we have the density of LB $f_w(x) \propto x^\alpha e^{-x^\alpha}$.

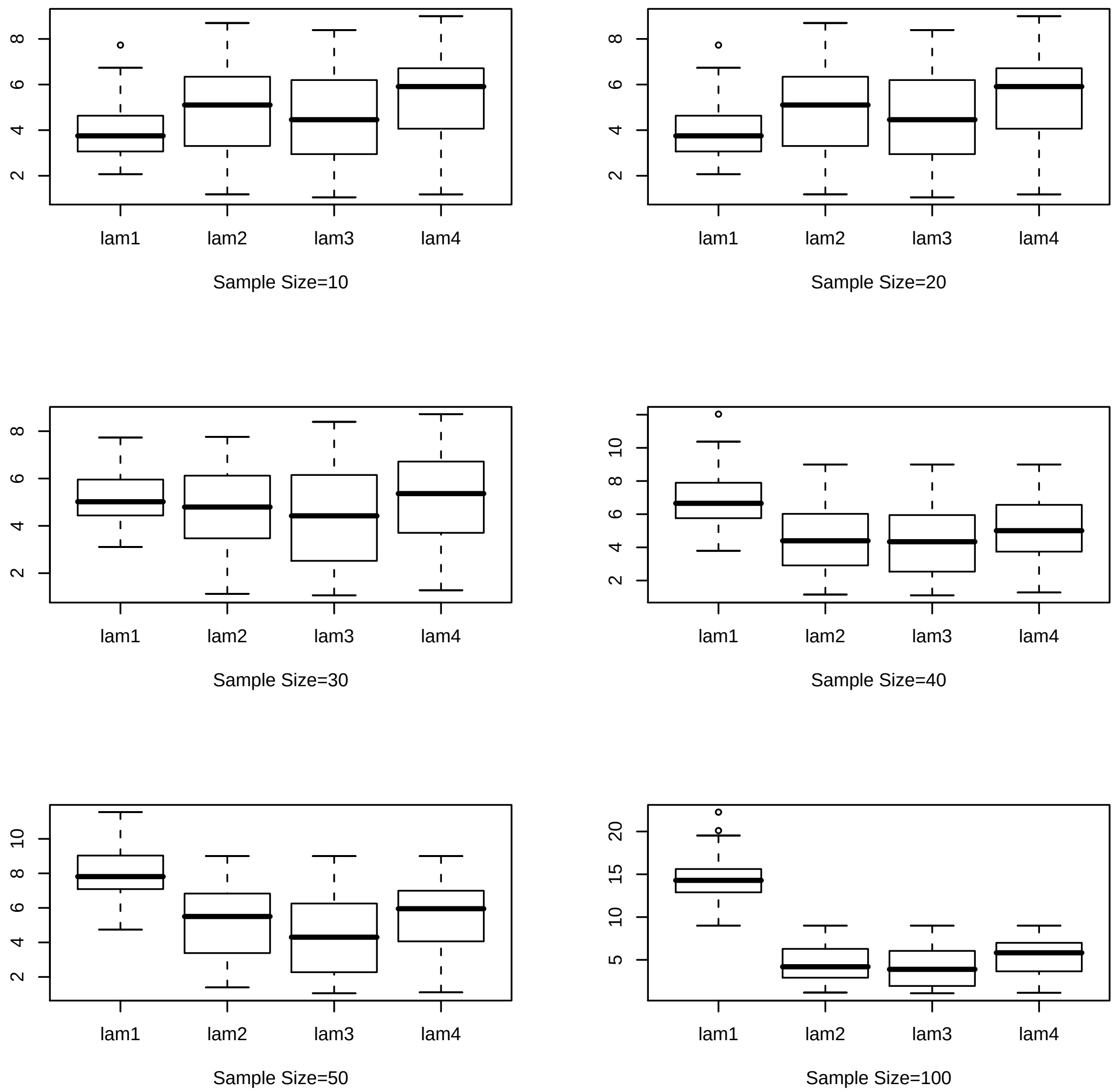


Figure 4.3: Box Plot for λ_n for 100 Samples, Underlying Density: Exponential, lam1: For Chaubey-Sen Choice, lam2: KL Cross Validation, lam3: ISE Cross Validation, lam4: Optimum Hellinger Distance

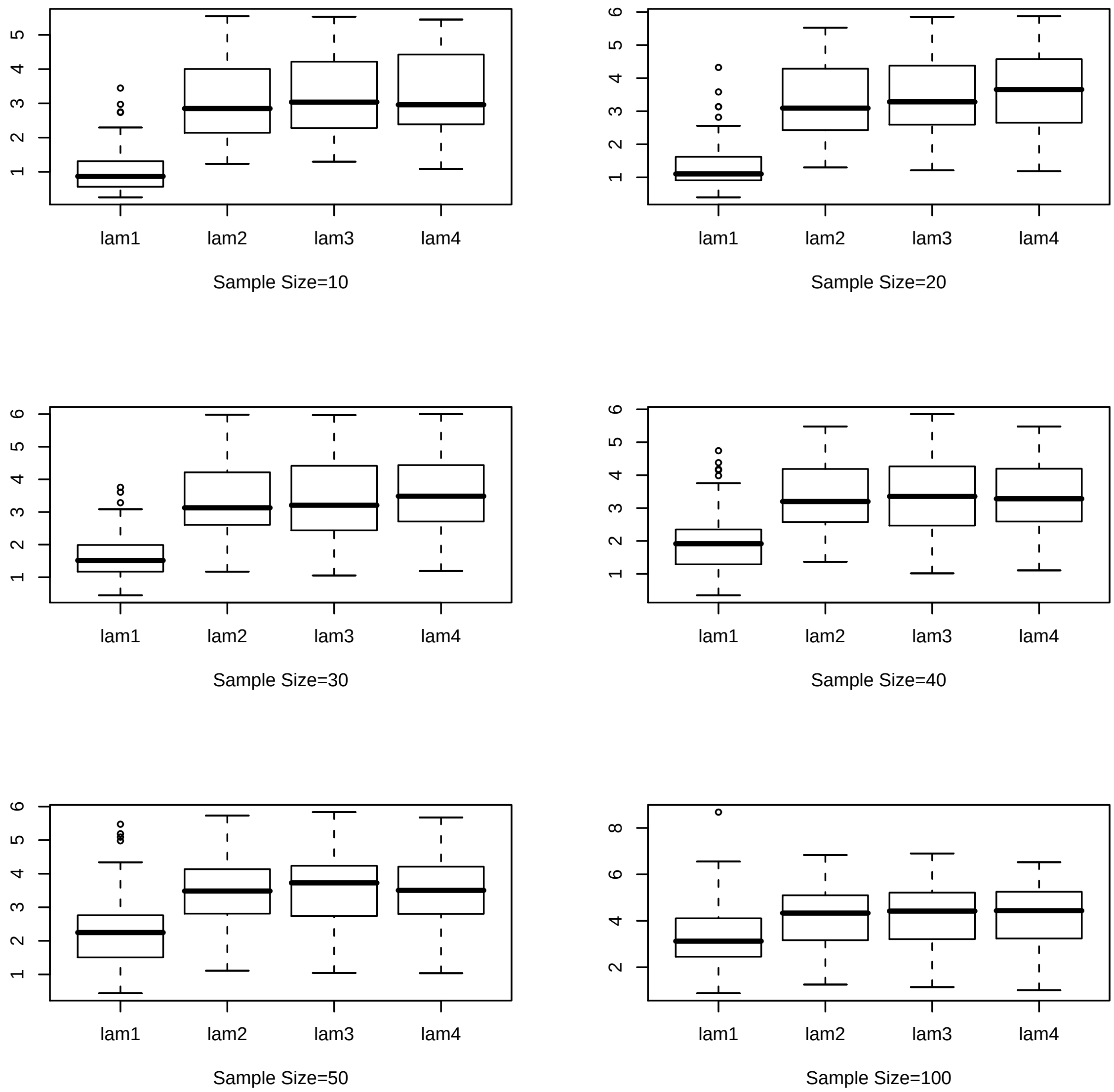


Figure 4.4: Box Plot for for λ_n for 100 Samples, Underlying Density: Lognormal, lam1: For Chaubey-Sen Choice, lam2: KL Cross Validation, lam3: ISE Cross Validation, lam4: Optimum Hellinger Distance

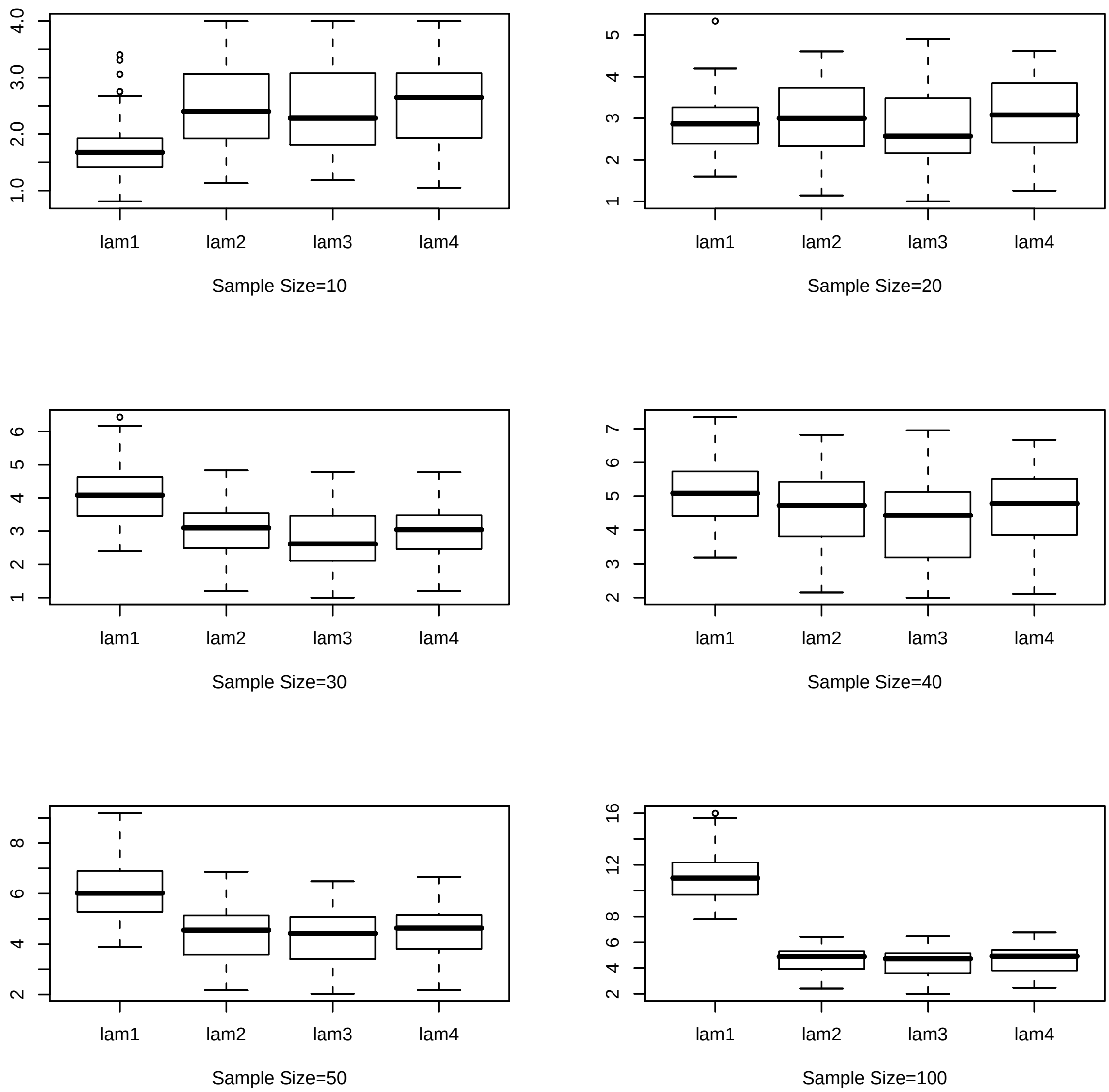


Figure 4.5: Box Plot for for λ_n for 100 Samples, Underlying Density: Gamma(2,1), lam1: For Chaubey-Sen Choice, lam2: KL Cross Validation, lam3: ISE Cross Validation, lam4: Optimum Hellinger Distance

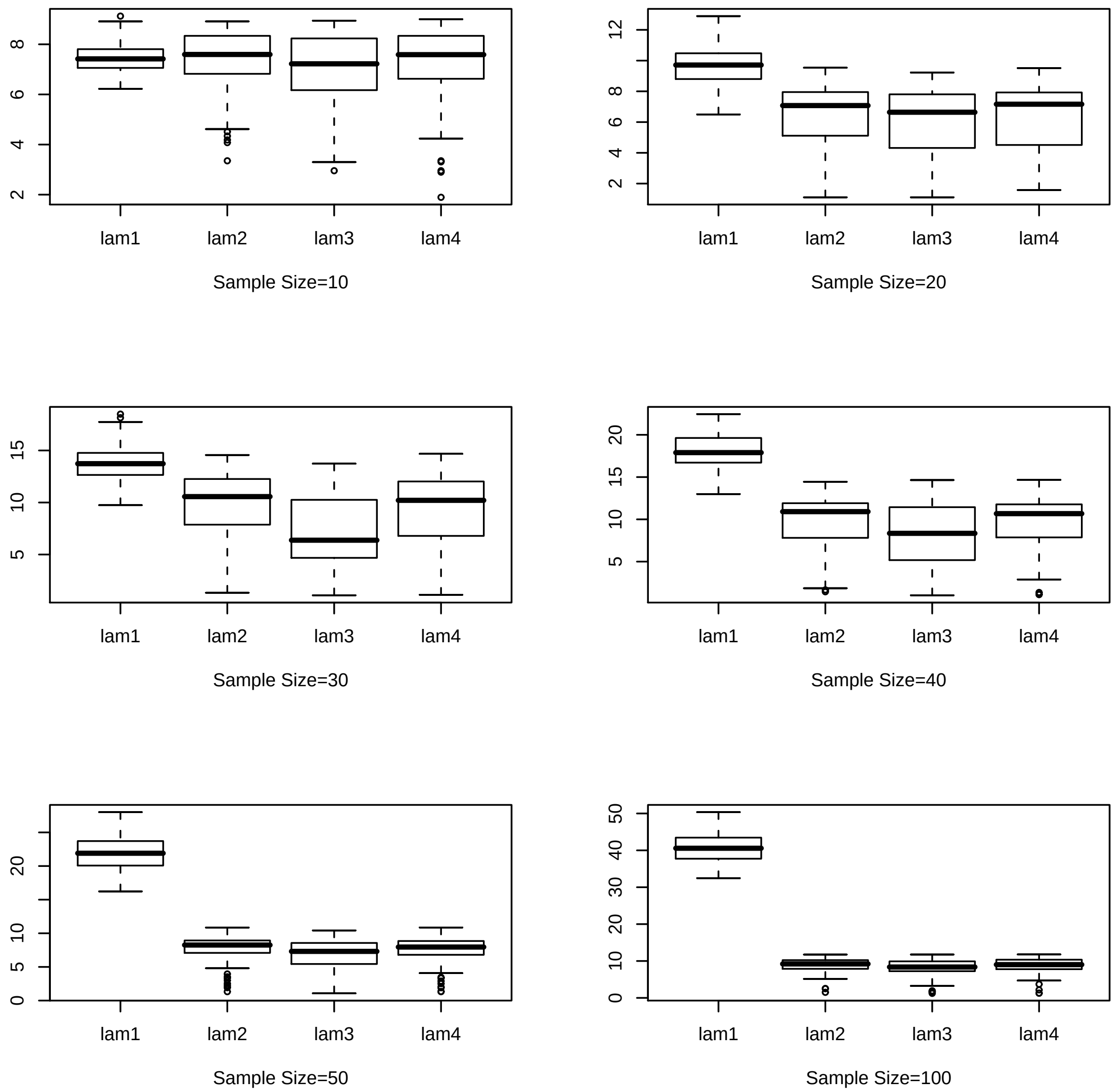


Figure 4.6: Box Plot for for λ_n for 100 Samples, Underlying Density: Weibull, $\alpha =$ 2, , lam1: For Chaubey-Sen Choice, lam2: KL Cross Validation, lam3: ISE Cross Validation, lam4: Optimum Hellinger Distance

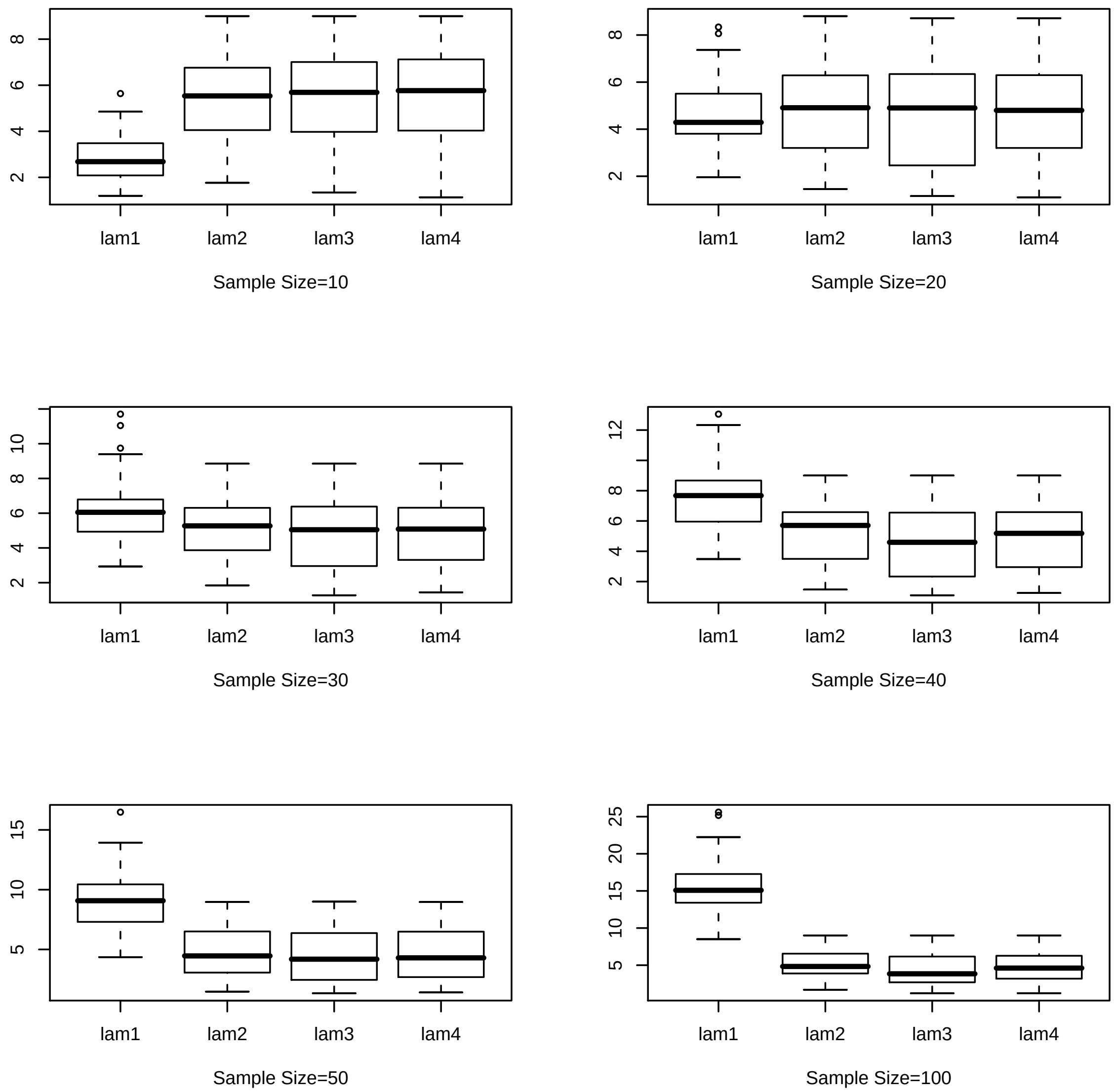


Figure 4.7: Box Plot for for λ_n for 100 Samples, Underlying Density: Exponential Mixture, $\theta_1 = 2$, $\theta_2 = 1$, $\pi = 0.4$, lam1: For Chaubey-Sen Choice, lam2: KL Cross Validation, lam3: ISE Cross Validation, lam4: Optimum Hellinger Distance

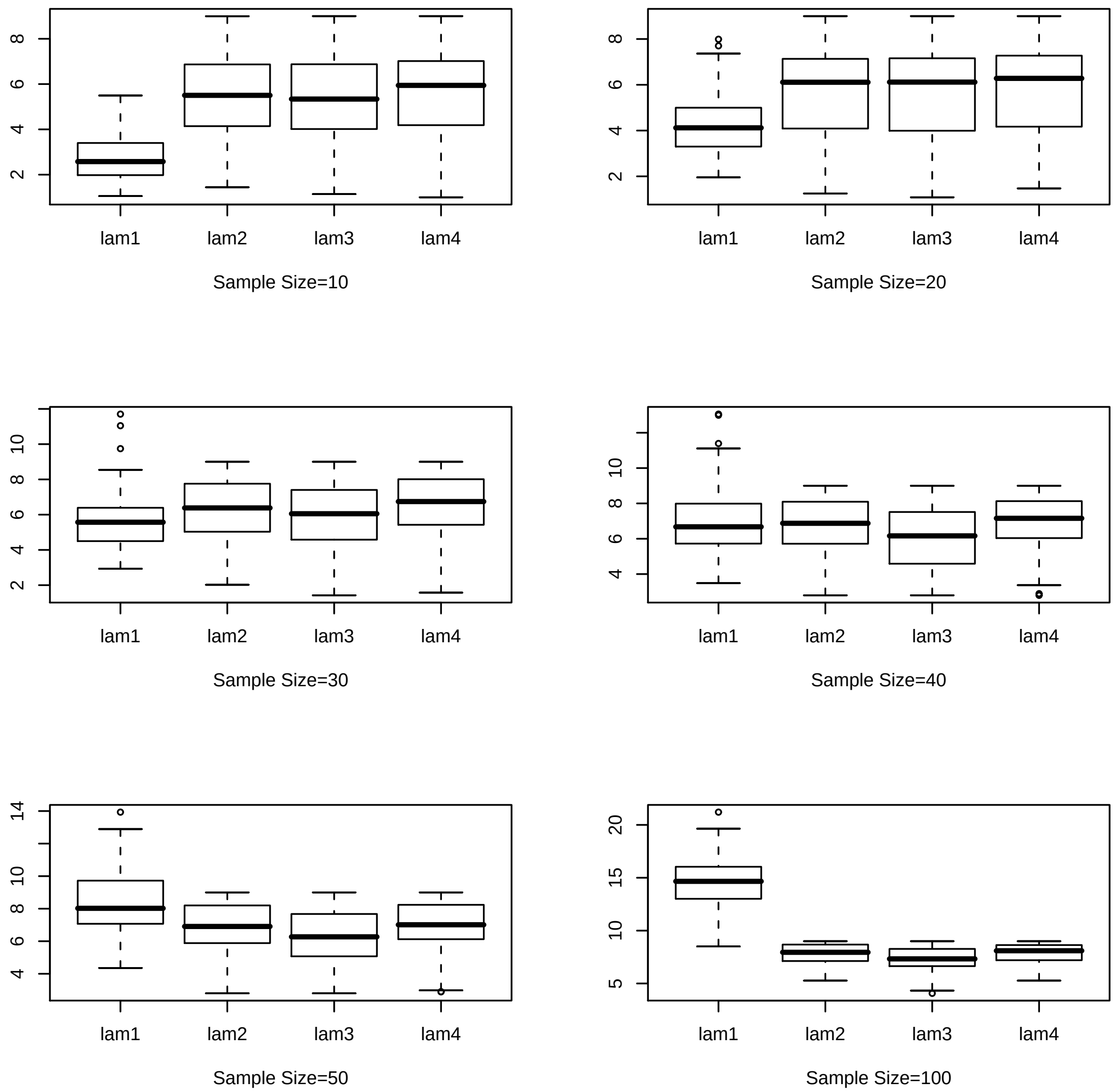


Figure 4.8: Box Plot for for λ_n for 100 Samples, Underlying Density: Exponential Mixture, $\theta_1 = 10$, $\theta_2 = 1$, $\pi = 0.2$, lam1: For Chaubey-Sen Choice, lam2: KL Cross Validation, lam3: ISE Cross Validation, lam4: Optimum Hellinger Distance

Conclusions

Denote by λ_{1O} the value which minimizing the expected Kullback-Liebler divergence

$$KL(\lambda_n) = \mathbb{E} \int \log \frac{f(x)}{\tilde{f}_n(x)} dF(x)$$

λ_{2O} the minimizer of

$$MISE(\lambda_n) = \mathbb{E} \int (\tilde{f}_n(x) - f(x))^2 dx$$

and λ_{3O} the minimizer of the expected Hellinger distance

$$H(\lambda_n) = \mathbb{E} \int \left(\sqrt{\tilde{f}_n(x)} - \sqrt{f(x)}\right)^2 dx$$

We will have the following conclusions which are the same as conclusions in using direct data to estimate density.

1. Chaubey-Sen choice usually produces large values of the smoothing parameters, especially, for large samples.

2. Chaubey-Sen choice is much more variable when samples are large even in the cases on an average it is close to the true optimum.

3. The two cross-validation criteria generally produce similar results, especially for larger samples and they converge to the true optimum under the known density.

4. We conjecture that suppose λ_{iO} denotes the true value of λ_n which minimizes criterion i, $i = 1, 2, 3$ and λ_{in} is the minima based on the data, then

$$(i) \qquad \lim_{n\to\infty} \frac{\lambda_{in}}{\lambda_{iO}} = 1 \text{ a.s.}$$

$$(ii) \qquad \lim_{n\to\infty} \frac{\lambda_{1O}}{\lambda_{3O}} = \lim_{n\to\infty} \frac{\lambda_{2O}}{\lambda_{3O}} = 1 \text{ a.s.}$$

4.2.2 Biased Cross-Validation Method

Regarding $\tilde{f}_n(x)$, we have the following asymptotic mean integrated square error(AMISE)

$$AMISE(\tilde{f}_n) = \lambda_n^{-2} \int_0^\infty (f'(x)/2)^2 dx + \sqrt{\lambda_n} \frac{\mu}{2\sqrt{\pi} n} \int_0^\infty \frac{f(x)}{x^{3/2}} dx. \tag{4.16}$$

We can write the integral in the last term as

$$\begin{aligned} \int_0^\infty \frac{f(x)}{x^{3/2}} dx &= \mu \int_0^\infty x^{-5/2} \frac{x f(x)}{\mu} dx \\ &= \mu \int_0^\infty x^{-5/2} g(x) dx = \mu E_g(X^{-5/2}). \end{aligned} \tag{4.17}$$

So we can use Monte Carlo method to estimate the integral, that is

$$\int_0^\infty \frac{f(x)}{x^{3/2}} dx \approx \frac{\mu}{n} \sum_{i=1}^n X_i^{-5/2} = \mu MCE_n. \tag{4.18}$$

Note that $\tilde{f}_n^+(x)$ is differentiable with respect to x. In (4.16), replacing $f'(x)$ and $\int_0^\infty \frac{f(x)}{x^{3/2}} dx$ with their estimators $\tilde{f}_n'(x)$ and MCE_n respectively, we can obtain the following biased cross-validation function

$$BCV(\lambda_n) = \lambda_n^{-2} \int_0^\infty (\tilde{f}_n'(x)/2)^2 dx + \sqrt{\lambda_n} \frac{\mu^2}{2\sqrt{\pi} n} MCE_n. \tag{4.19}$$

Since μ is unknown, we can substitute μ with its estimator $\frac{n}{\sum_{i=1}^n \frac{1}{X_i}}$ in the procedure of computation.

For now, we have two crossed validation methods related to ISE (4.11). One is based on asymptotic mean ISE and has the form as (4.19) referred as to BCV method. The other is based on ISE and has the form as (4.13) referred as to UCV method. In this thesis, we mainly use $MISE = \mathbb{E} \int (f_n(x) - f(x))^2 dx$ to judge the global performance of estimators. We will pay more attention to parameter selection methods UCV and BCV related to ISE. We certainly concern which method is better. We will use extensive simulation to answer this question.

4.2.2.1 Simulation Studies

In this subsection, we will do extensive simulation with diverse sample size to compare UCV and BCV methods. We have simulated from the following underlying densities with sample size 30, 50, 100, 200, 300, and 500. For each sample size, we obtain 1000 samples of smooth parameter. Under each chosen parameter, we computer the ISE as well and take the average of ISEs as the approximation of $MISE$ to evaluate UCV and BCV methods.

(i). Chi-Square Distribution

$$f(x) = \frac{1}{2^{\frac{\alpha}{2}}\Gamma(\frac{\alpha}{2})} x^{\frac{\alpha}{2}-1} \exp(-x/2) I\{x > 0\}$$

(ii). Lognormal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi x}} \exp\{-(\log x - \mu)^2/2\} I\{x > 0\}$$

(iii). Weibull Distribution

$$f(x) = \alpha x^{\alpha-1} \exp(-x^{\alpha}) I\{x > 0\}$$

(iv). Mixtures of Two Exponential Distribution

$$f(x) = [\pi \frac{1}{\theta_1} \exp(-x/\theta_1) + (1-\pi)\frac{1}{\theta_2} \exp(-x/\theta_2] I\{x > 0\}$$

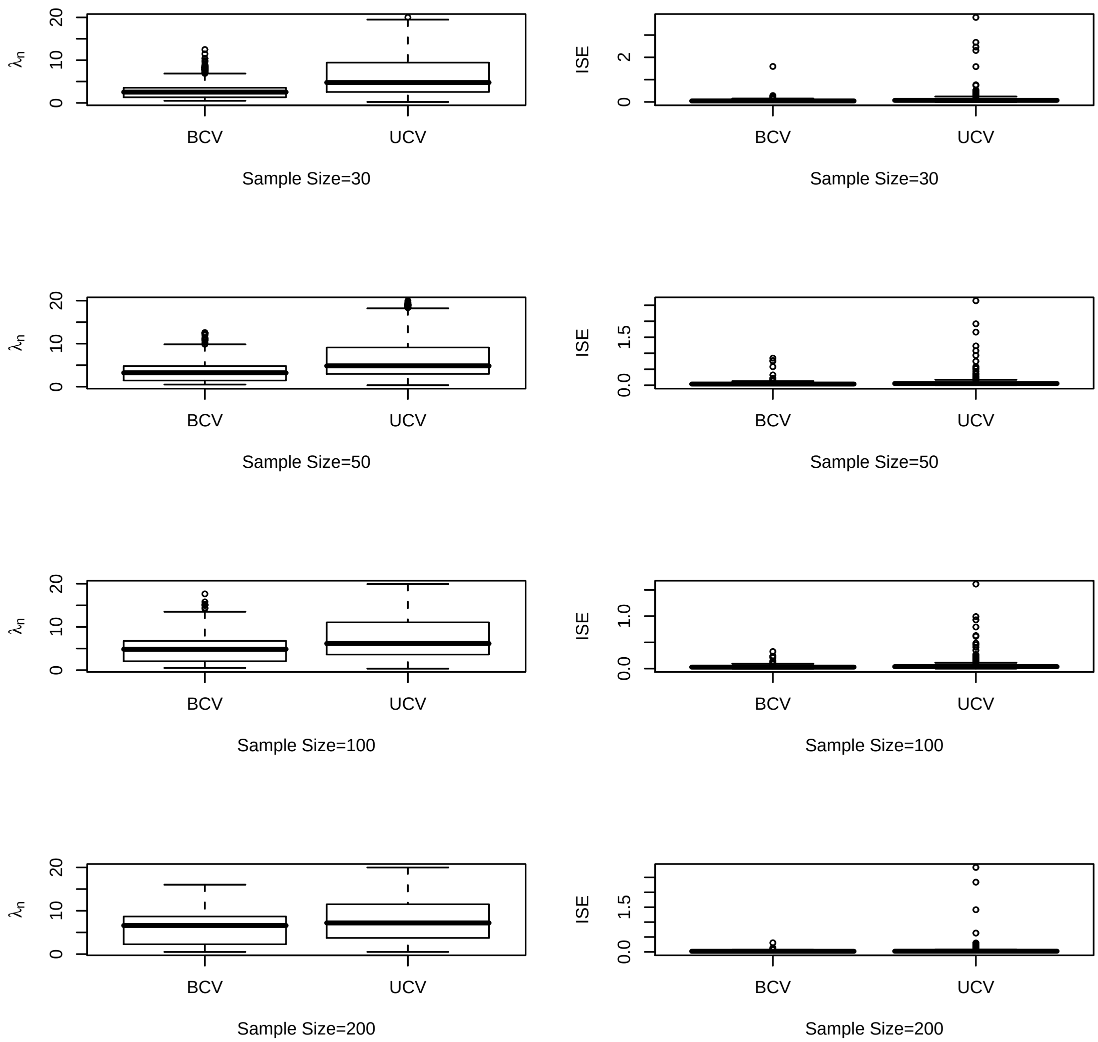
The methods of generating corresponding LB data are given by, respectively,

(i'). $X \sim \chi^2_{\alpha+2}$;

(ii'). $X = e^Y$ where $Y \sim N(\mu+1, 1)$;

$(iii').\ X = Y^{1/\alpha}$ where $Y \sim \Gamma(1 + \frac{1}{\alpha}, 1)$;

$(iv').\ X = \pi Y_1 + (1 - \pi)Y_2$, where $Y_1 \sim \Gamma(2, \theta_1)$ and $Y_2 \sim \Gamma(2, \theta_2)$.



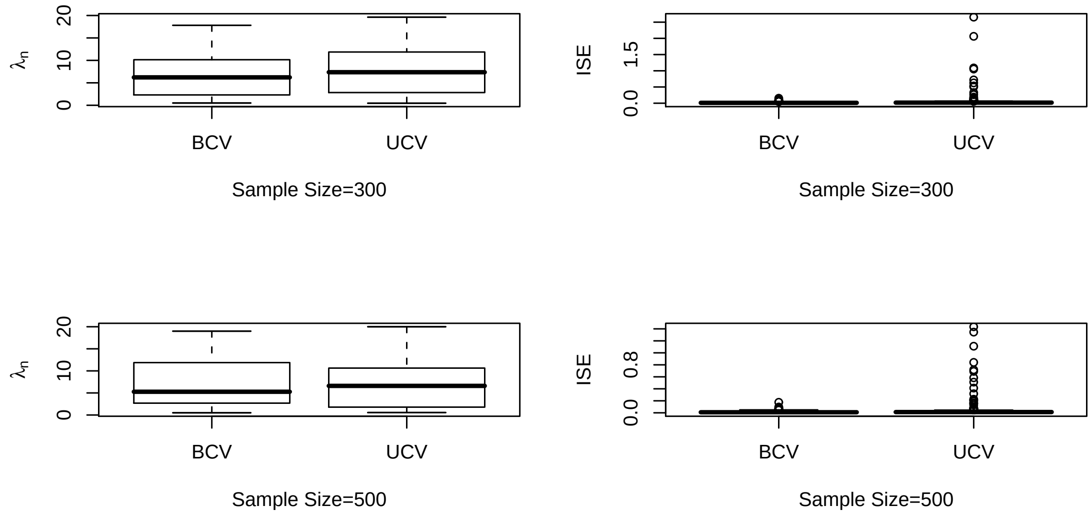


Figure 4.9: Boxplots of parameter and ISE for χ_2^2.

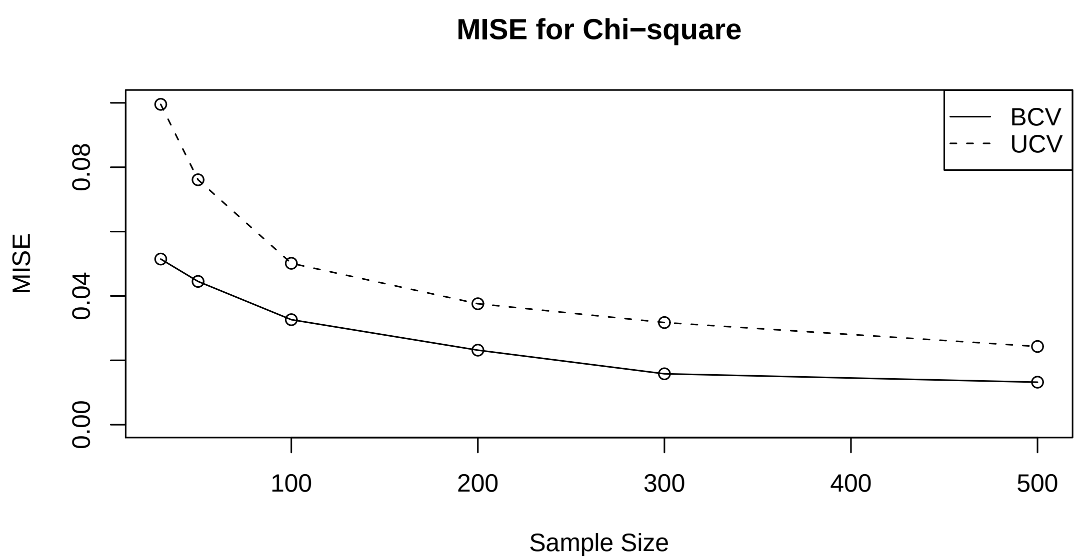
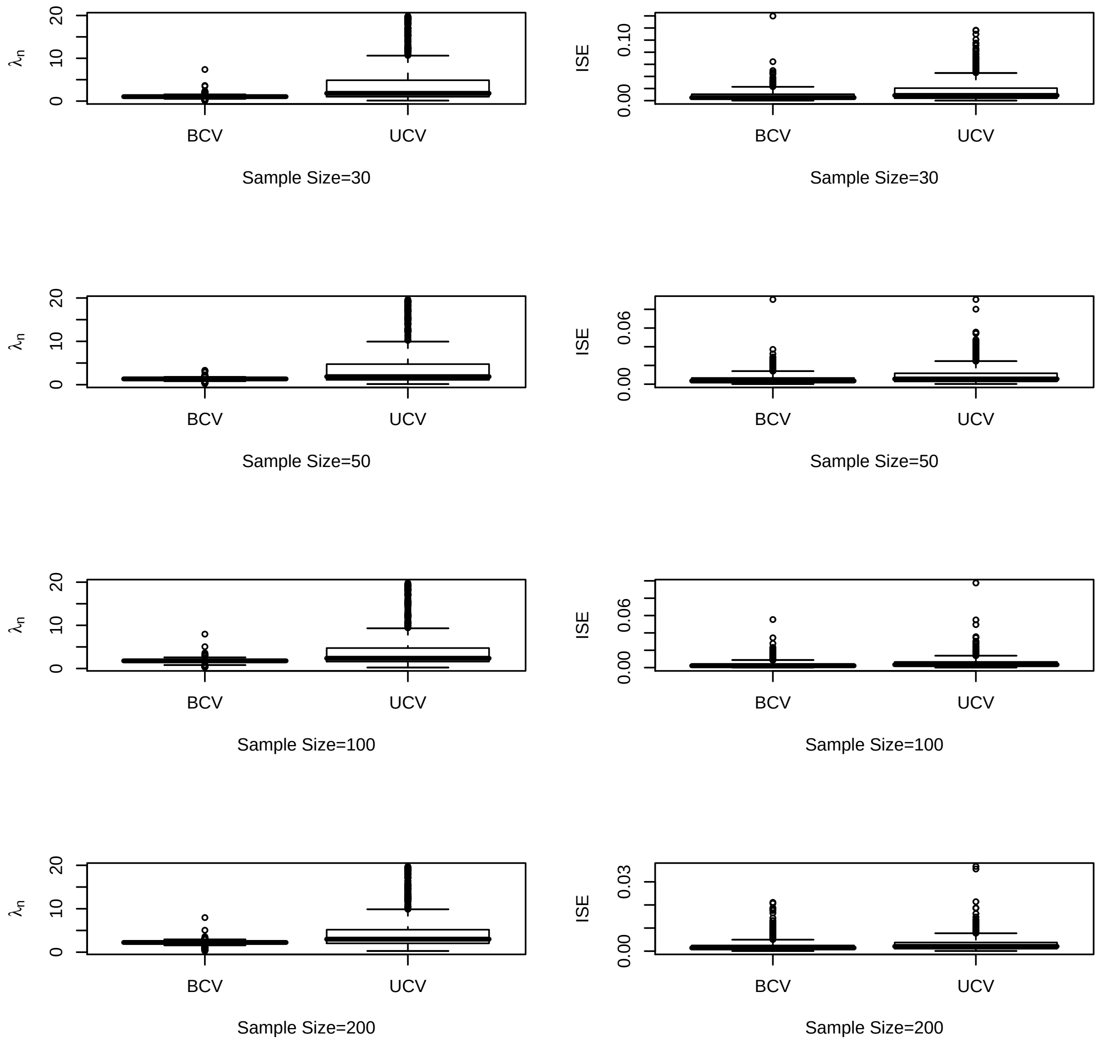


Figure 4.10: Plots of BCV and UCV MISE for χ_2^2.


λn
ISE
0
10
20
0.00
0.03
0.06
0.10
BCV
UCV
Sample Size=30
Sample Size=50
Sample Size=100
Sample Size=200

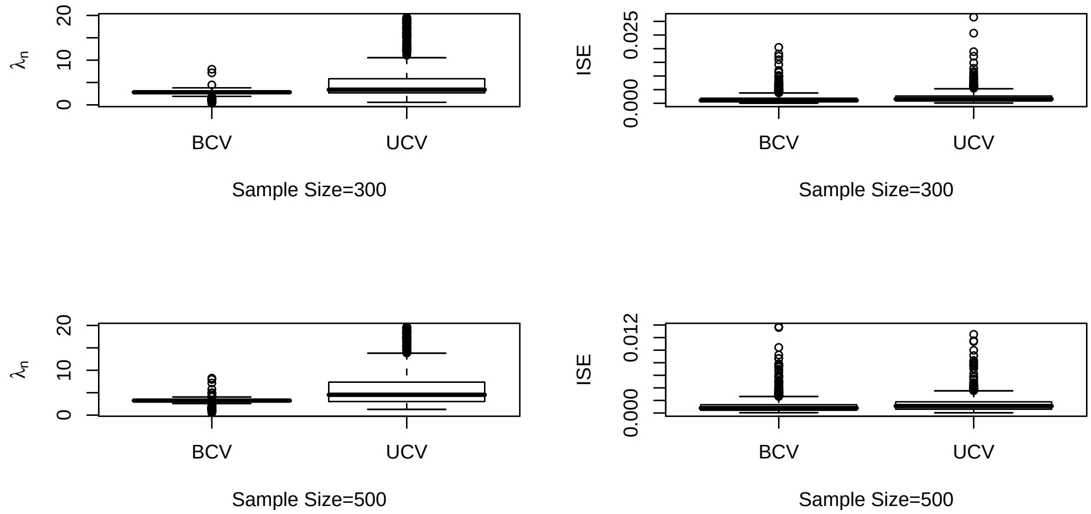


Figure 4.11: Boxplots of parameter and ISE for χ^2_6.

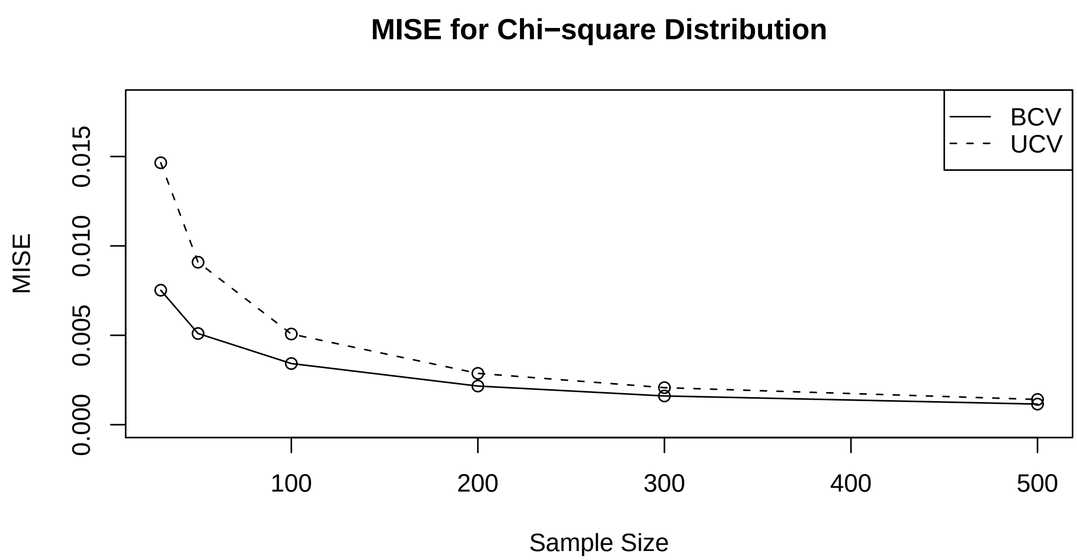
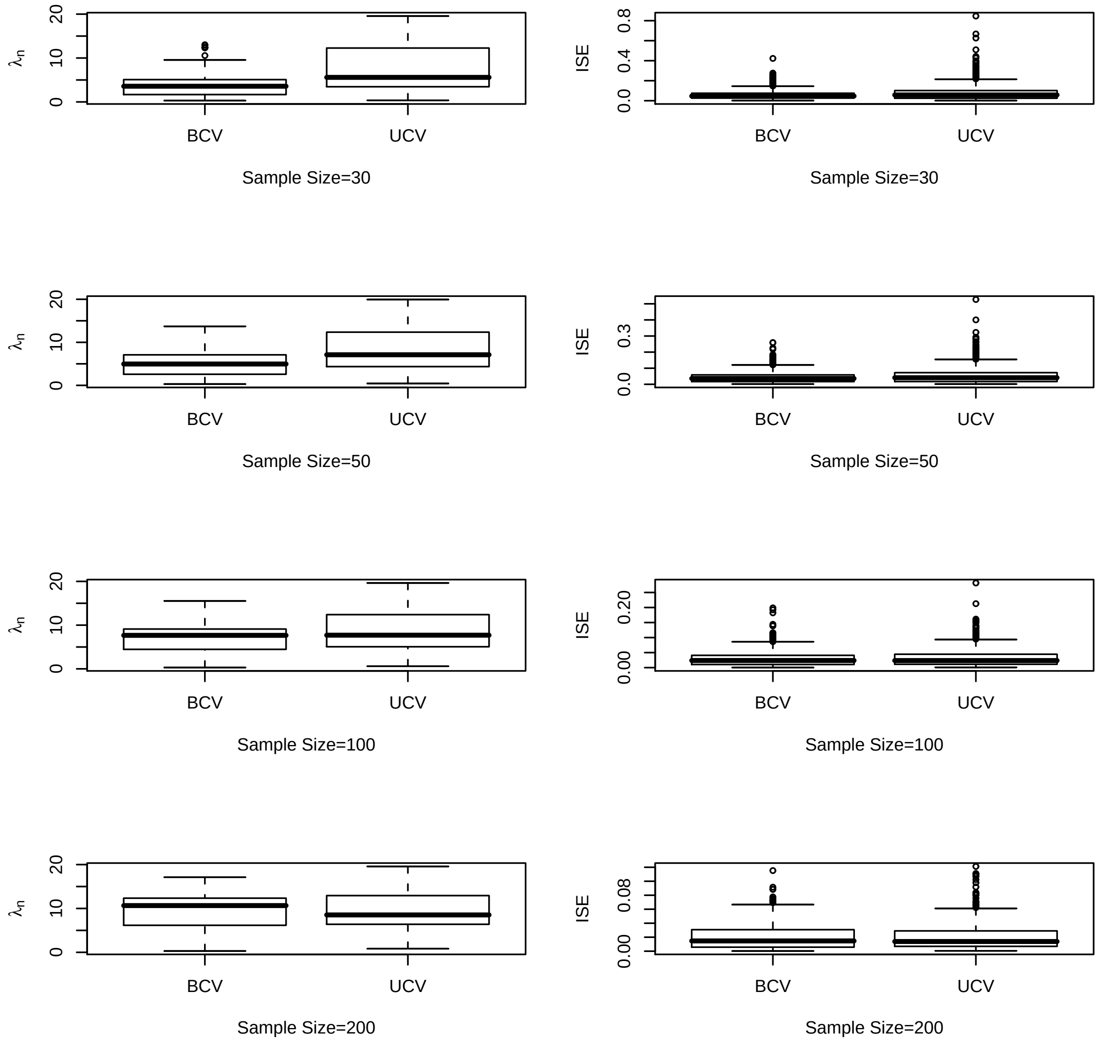


Figure 4.12: Plots of BCV and UCV MISE for χ^2_6.



λ_n
0
10
20
BCV
UCV
Sample Size=30
ISE
0.0
0.4
0.8
BCV
UCV
Sample Size=30
λ_n
0
10
20
BCV
UCV
Sample Size=50
ISE
0.0
0.3
BCV
UCV
Sample Size=50
λ_n
0
10
20
BCV
UCV
Sample Size=100
ISE
0.00
0.20
BCV
UCV
Sample Size=100
λ_n
0
10
20
BCV
UCV
Sample Size=200
ISE
0.00
0.08
BCV
UCV
Sample Size=200

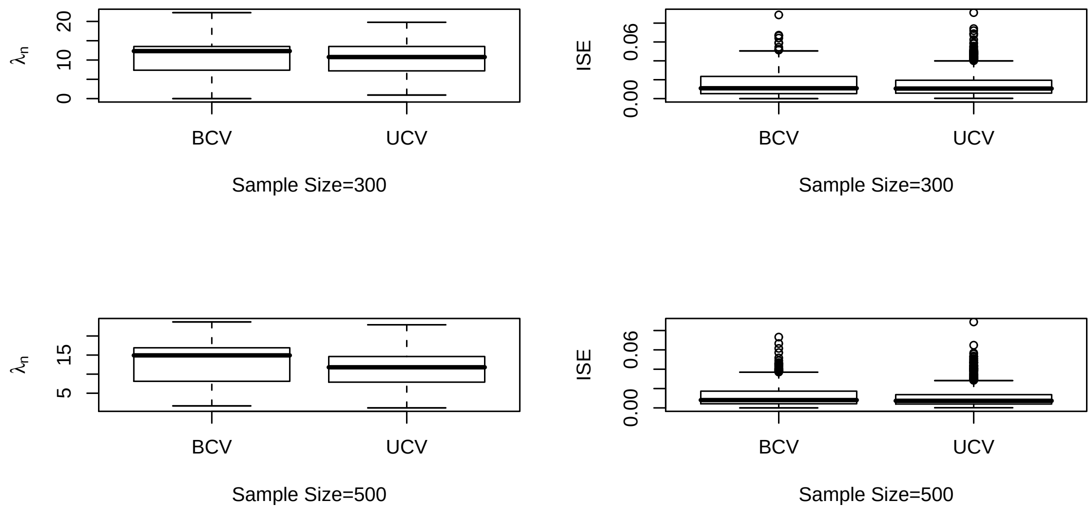


Figure 4.13: Boxplots of parameter and ISE for Lognormal with parameter 1.

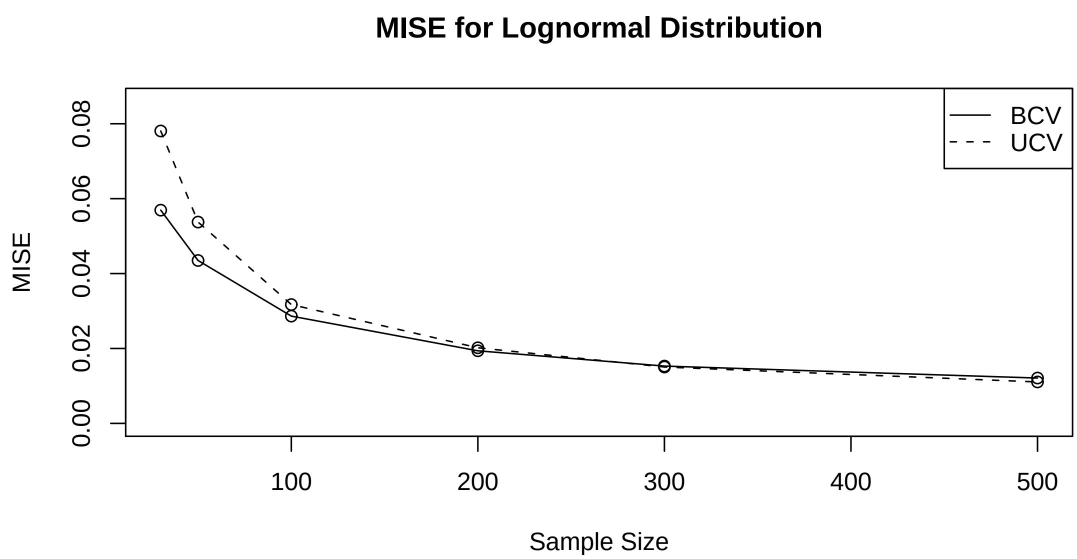
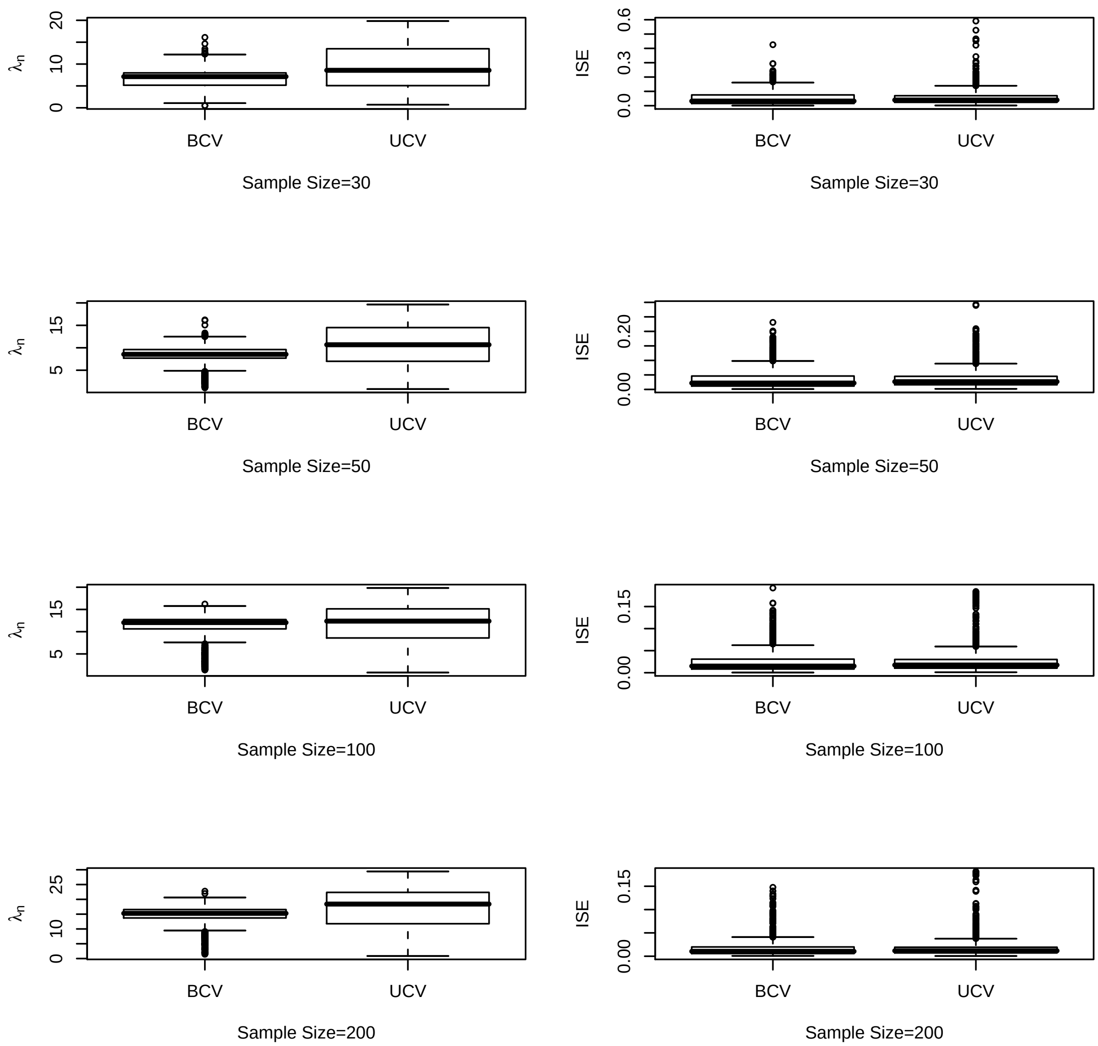


Figure 4.14: Plots of BCV and UCV MISE for Lognormal with parameter 1.



λn
0
10
20
BCV
UCV
Sample Size=30
ISE
0.0
0.3
0.6
BCV
UCV
Sample Size=30
λn
5
15
BCV
UCV
Sample Size=50
ISE
0.00
0.20
BCV
UCV
Sample Size=50
λn
5
15
BCV
UCV
Sample Size=100
ISE
0.00
0.15
BCV
UCV
Sample Size=100
λn
0
10
25
BCV
UCV
Sample Size=200
ISE
0.00
0.15
BCV
UCV
Sample Size=200

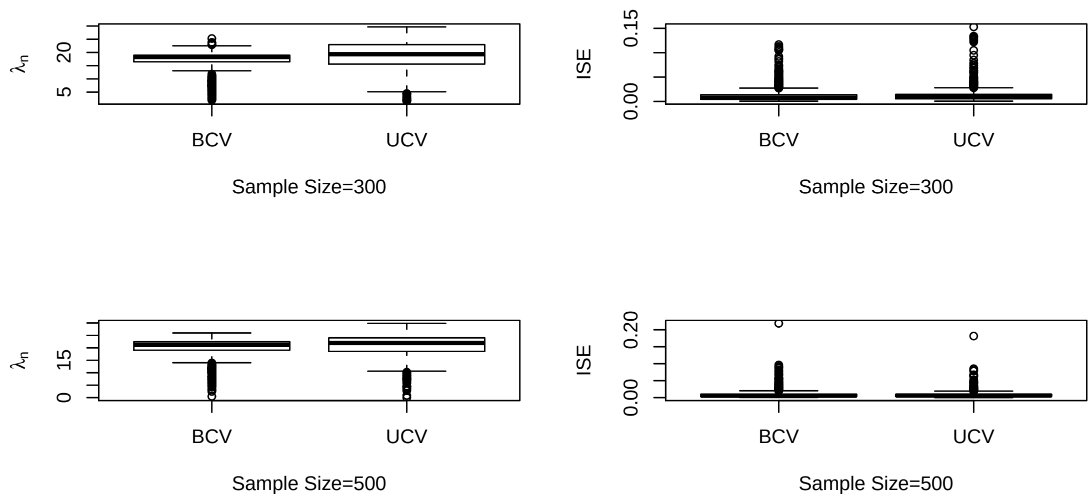


Figure 4.15: Boxplots of parameter and ISE for Weibull with parameter 2.

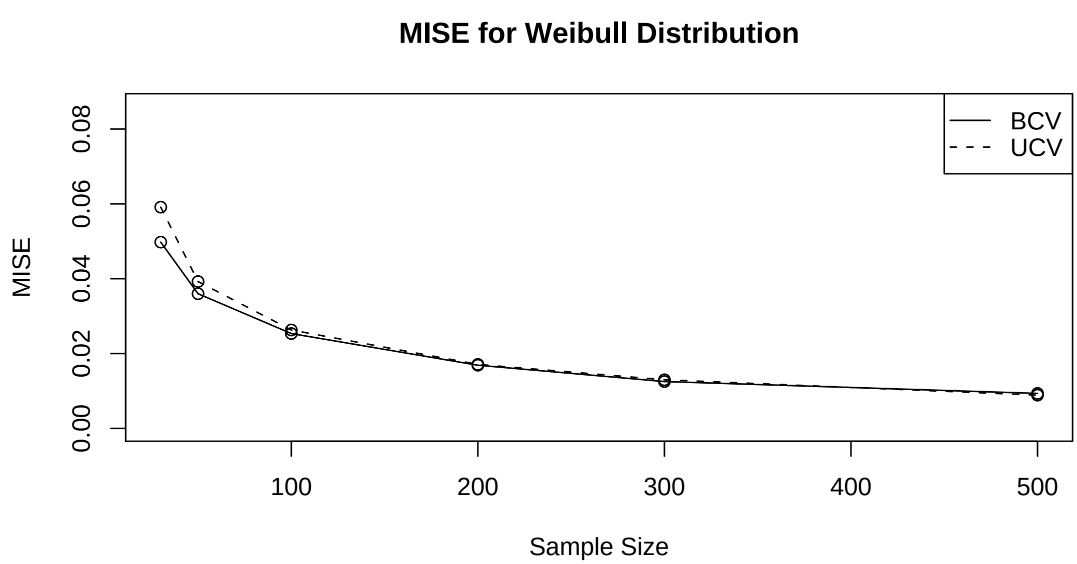
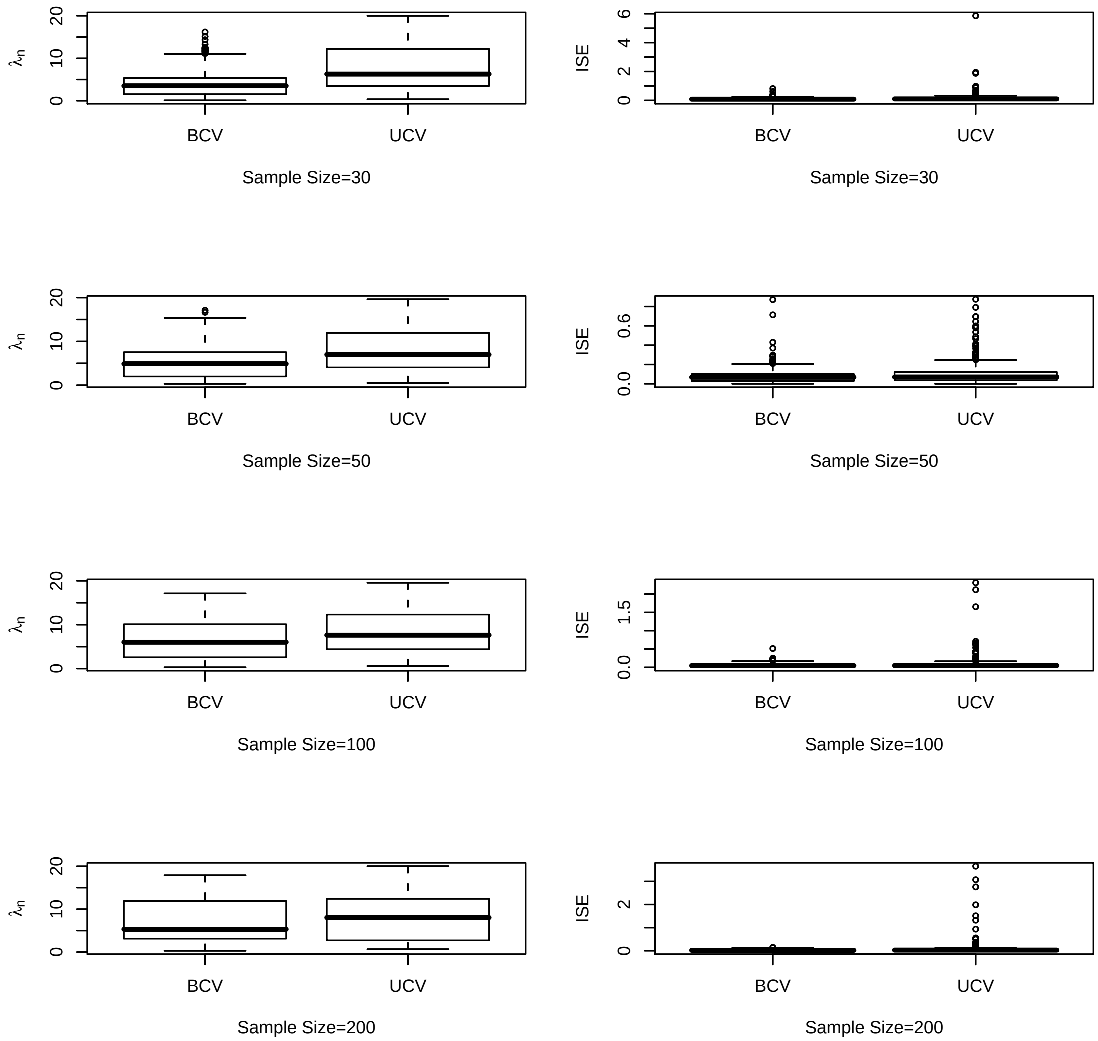


Figure 4.16: Plots of BCV and UCV MISE for Weibull with parameter 2.


λ_n
ISE
BCV
UCV
Sample Size=30
Sample Size=50
Sample Size=100
Sample Size=200

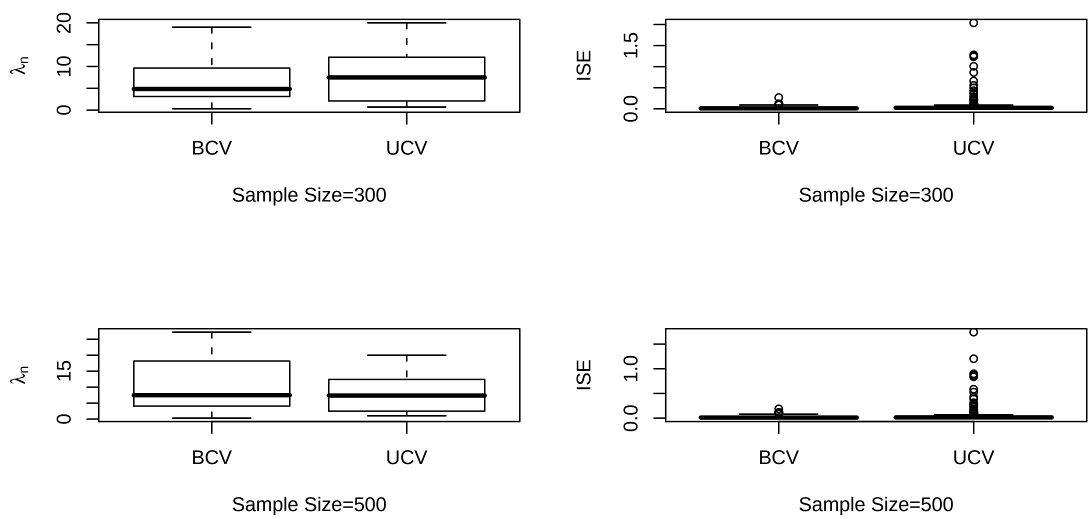


Figure 4.17: Boxplots of parameter and ISE for mixtures of two exponential distributions, $\pi = 0.4$, $\theta_1 = 2$, $\theta_2 = 1$.

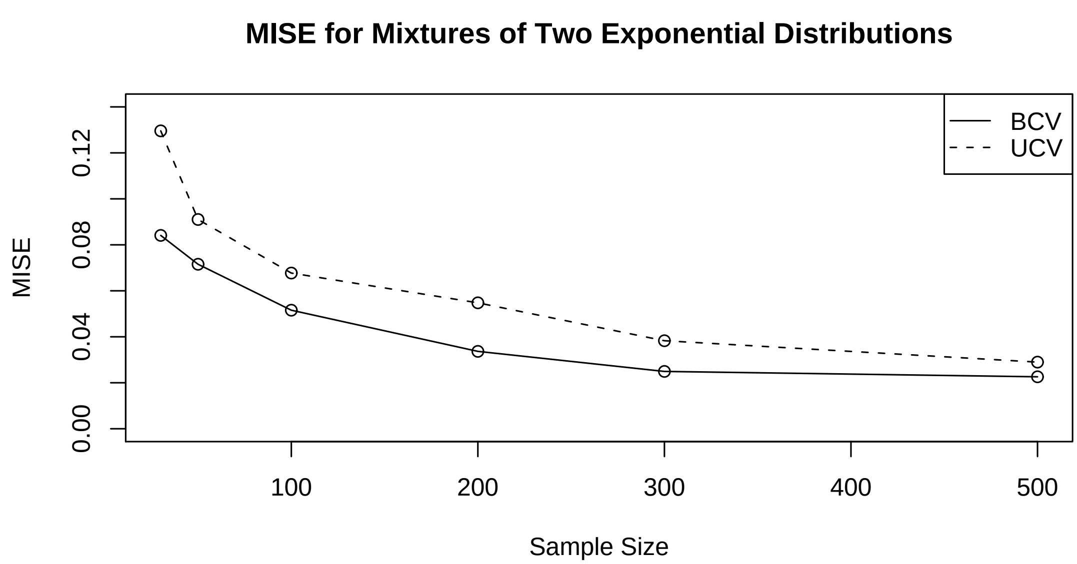


Figure 4.18: Plots of BCV and UCV MISE for mixtures of two exponential distributions, $\pi = 0.4$, $\theta_1 = 2$, $\theta_2 = 1$.

4.2.3 Choice Between UCV and BCV Methods

From the extensive simulation studies, we find that biased cross-validation method usually produces smaller MISE than unbiased cross-validation method, particularly when the sample size is small. So measured by MISE, BCV method performs better than UCV method when sample size is small. This is partly due to the facts that UCV function is rougher, which cause more difficulty in searching optimal solution. Furthermore, we find that with small sample size UCV function sometimes might gives us a great parameter value [see Figure 4.19], which causes the estimator rough [see Figure 4.21] and produces larger ISE [see Table 4.1]. This is why we can see more outliers on the ISE boxplots of UCV. However, under this circumstance, BCV function is smoother [see Figure 4.20] and gives us an acceptable optimal choice which generate much smaller ISE and smoother estimator.

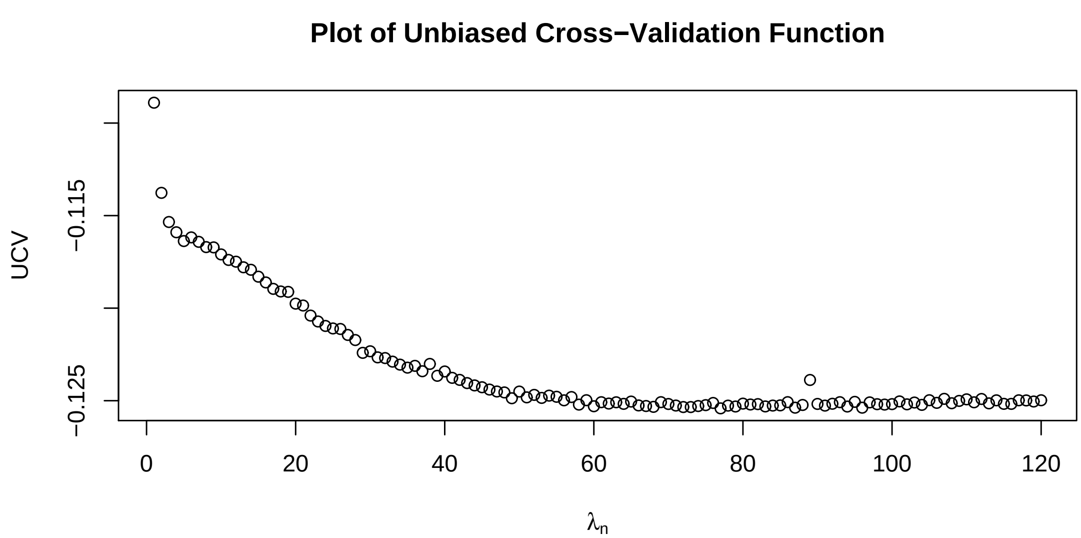


Figure 4.19: UCV function for χ^2_6, sample size=100.

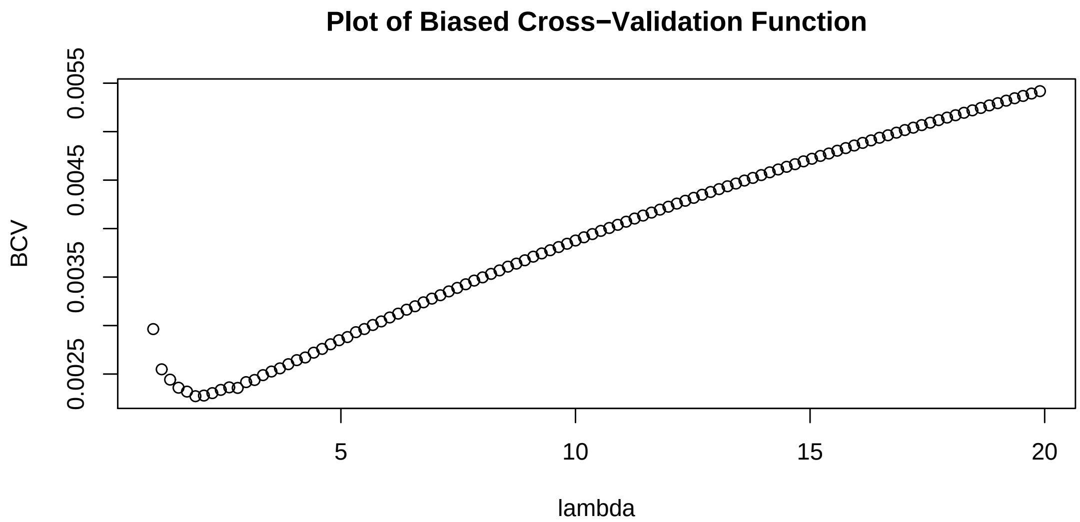


Figure 4.20: BCV function for χ^2_6, sample size=100.

Table 4.1: Parameter and ISE

Method	λ_n	Value of CV	ISE
UCV	74.04	-0.1254	0.02694
BCV	1.884	0.002267	0.0006223

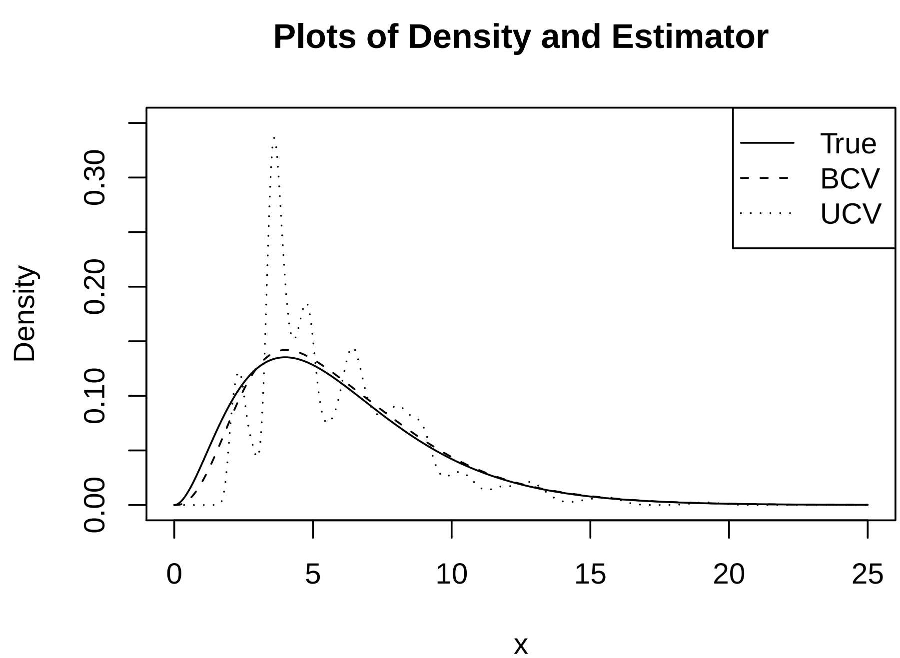


Figure 4.21: Density and estimators for χ^3_6.

Through our studies, we may obtain the following conclusions.

1. UCV function is rougher than BCV, which cause more difficulty in searching optimal solution by *optimise* and the choices are more variable;

2. The leave-one-out estimator in UCV is complicated, which causes to take more time to search optimal solution of UCV function, particularly when sample size is great;

3. BCV function is smoother, which is easier to search optimal solution and saves time in the procedure of computation;

4. BCV method produces smaller MISEs than UCV method when sample size is small. When sample size is great enough, the MISEs generated by two methods are very close;

5. If we denote by λ_{n_o} the optimal solution

$$MISE(\lambda_n) = \mathbb{E}\int(\tilde{f}_n(x) - f(x))^2 dx,$$

$\lambda_{n_{BCV}}$ the optimal solution of BCV function and $\lambda_{n_{UCV}}$ the optimal solution of UCV function, we might conjecture that

$$\lim_{n\to\infty} \frac{\lambda_{n_{BCV}}}{\lambda_{n_o}} = \lim_{n\to\infty} \frac{\lambda_{n_{UCV}}}{\lambda_{n_o}} = 1 \text{ a.s.}$$

4.3 Parameter Selection for Other Density Estimators

4.3.1 Parameter Selection for $\tilde{f}_n^+(x)$ and $\tilde{f}_n^*(x)$

In $\tilde{f}_n^+(x)$ and $\tilde{f}_n^*(x)$, there are two parameters. One is v_n controlling the smoothness of the estimator and the other is ε_n controlling the bias of the estimator at boundary. In order to find a proper choice of (v_n, ε_n), for $\tilde{f}_n^+(x)$ we investigate two cross-validation methods. One is Biased Cross-Validation based on AMISE of $\tilde{f}_n^+$ and the other is Unbiased Cross-Validation based on ISE of $\tilde{f}_n^+$.

4.3.1.1 Biased Cross-Validation

For $\tilde{f}_n^+(x)$, we have

$$\begin{aligned} AMISE[\tilde{f}_n^+] &= \int_0^\infty [(xv_n^2 + \varepsilon_n)f'(x) + \frac{x^2v_n^2}{2}f''(x)]^2dx \\ &\quad + \frac{I_2(q)\mu}{nv_n}\int_0^\infty \frac{f(x)}{(x+\varepsilon_n)^2}dx \end{aligned} \tag{4.20}$$

In the AMISE of $\tilde{f}_n$ (4.20), replacing $f(x)$, $f'(x)$ and $f''(x)$ with $\tilde{f}_n^+(x)$, $\tilde{f}_n^{+'}(x)$ and $\tilde{f}_n^{+''}(x)$ respectively, we obtain the following Biased Cross-Validation function

$$BCV(v_n, \varepsilon_n) = \int_0^\infty [(xv_n^2 + \varepsilon_n)\tilde{f}_n^{+'}(x) + \frac{x^2v_n^2}{2}\tilde{f}_n^{+''}(x)]^2dx + \frac{I_2(q)\hat{\mu}}{nv_n}\int_0^\infty \frac{\tilde{f}_n^+(x)}{(x+\varepsilon_n)^2}dx. \tag{4.21}$$

where $\hat{\mu} = \left(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}\right)^{-1}$, being an estimator of μ. We can minimize (4.21) with respect to (v_n, ε_n) to find a choice of (v_n, ε_n).

4.3.1.2 Unbiased Cross-Validation

Let us consider the Integrated Squared Error

$$\begin{aligned} ISE(v_n, \varepsilon_n) &= \int_0^\infty [\tilde{f}_n(x) - f(x)]^2dx \\ &= \int_0^\infty \tilde{f}_n^2(x)dx - 2\int_0^\infty \tilde{f}_n(x)f(x)dx + \int_0^\infty f^2(x)dx. \end{aligned}$$

Disregarding the last constant term, substituting the second term with its Leave-One-Out estimator, we obtain the following Unbiased Cross-Validation function

$$UCV(v_n, \varepsilon_n) = \int_0^\infty \tilde{f}_n^{+2}(x)dx - 2\sum_{i=1}^n \tilde{f}_{n-1}^+(X_i; \mathcal{D}_i)/Z_i \tag{4.22}$$

where $\mathcal{D}_i$ denotes data set with X_i removed from the original complete data set $\mathcal{D}$, $\tilde{f}^+_{n-1}(x; \mathcal{D}_i)$ denotes the density estimator built on $\mathcal{D}_i$ and $Z_i = \sum_{j \neq i} \frac{X_i}{X_j}$. Minimizing (4.22) will give us a choice of (v_n, ε_n).

4.3.1.3 Numerical Comparison

In order to compare the two methods, we simulate for χ^2_2 and χ^2_{12} with sample size 100, 300, 500 and 1000. For each sample, we minimize (2.76) and (4.22) to obtain the choices of (v_n^2, ε_n). At the same, for comparison, we also compute the ISE under each choice. We repeat the procedure 1000 times and obtain 1000 samples of (v_n^2, ε_n) and ISE.

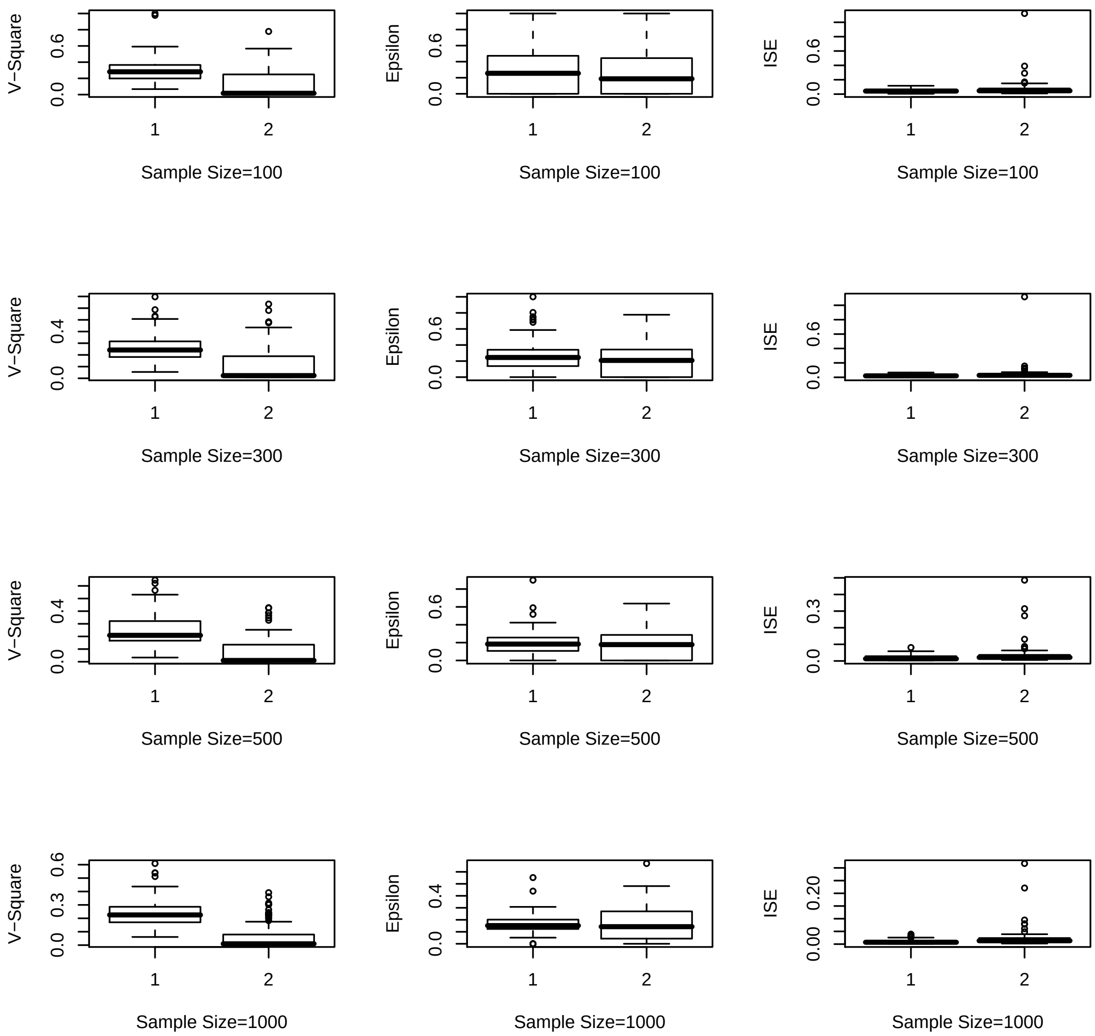


Figure 4.22: Plots of samples of (v_n^2, ε_n) and ISE for χ_2^2.

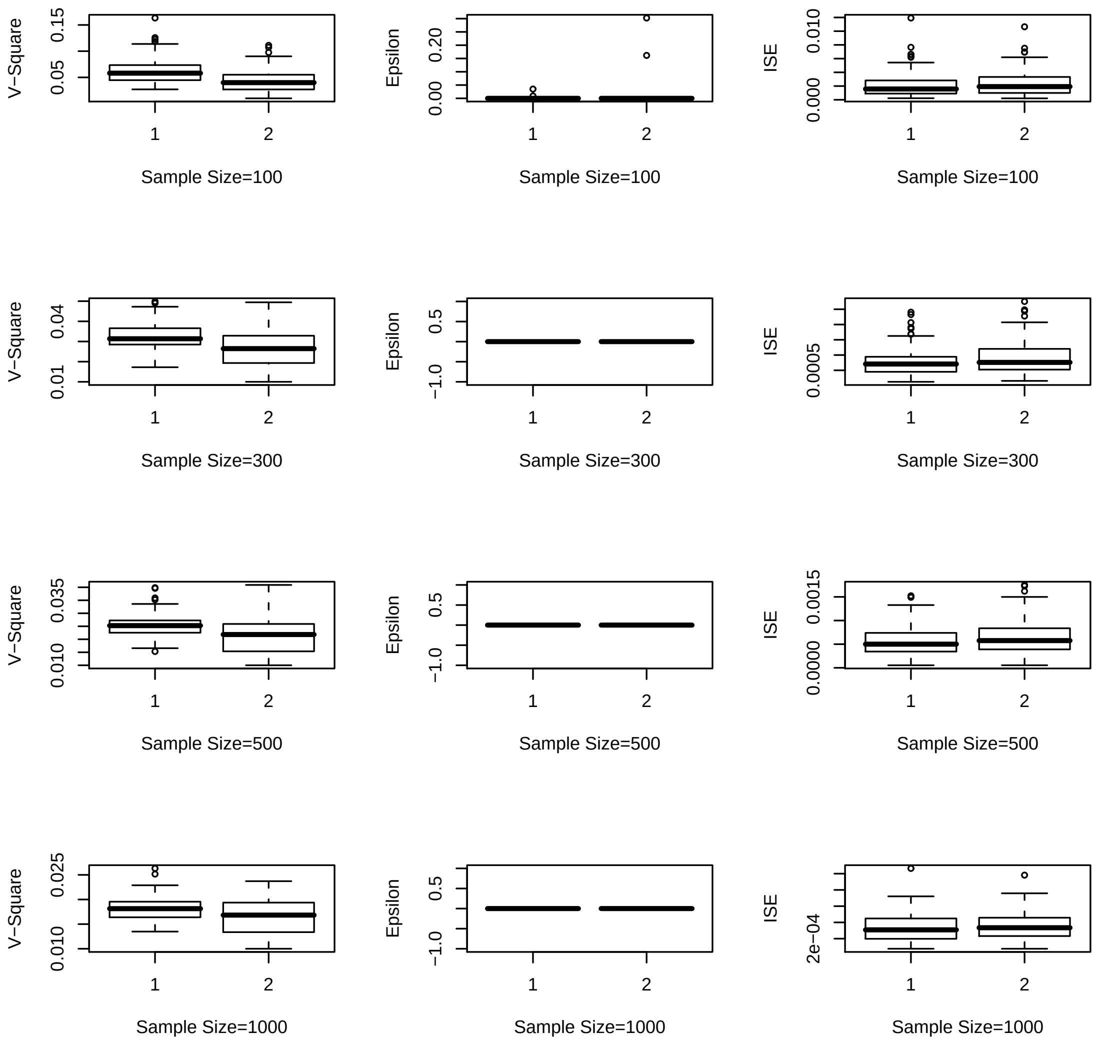


Figure 4.23: Plots of samples of (v_n^2, ε_n) and ISE for χ_6^2..

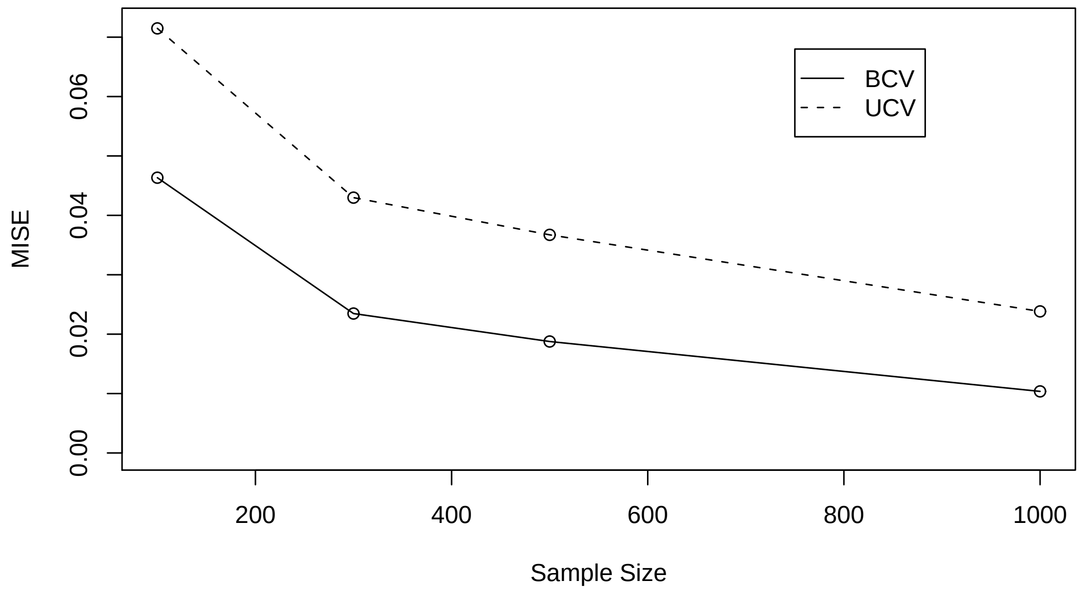


Figure 4.24: Plot of MISE for χ_2^2.

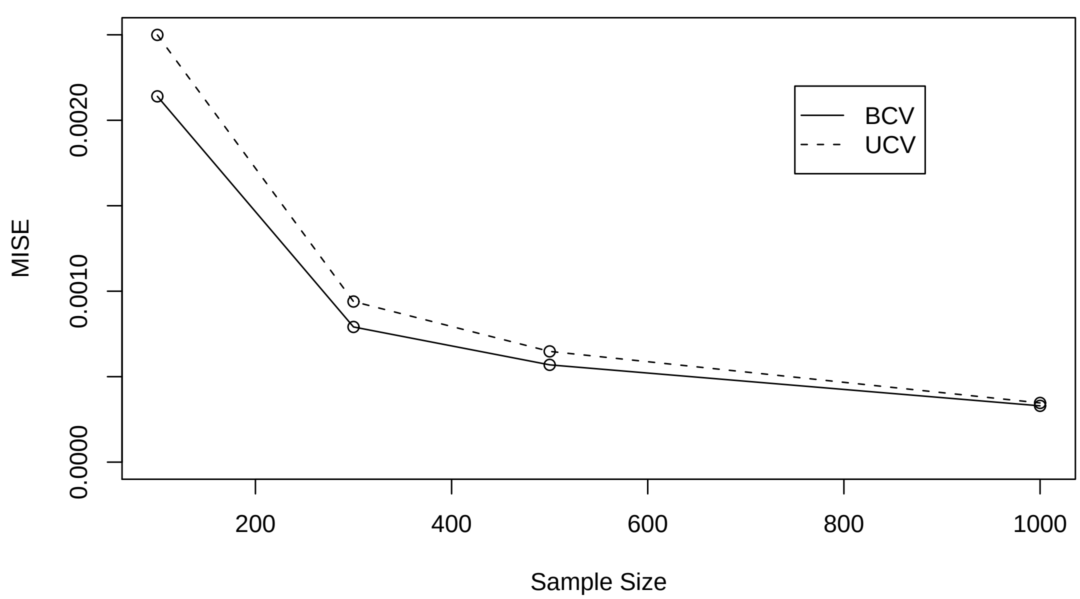


Figure 4.25: Plot of MISE for χ_6^2.

Through the simulation, we can see that, in the first example in which the true density satisfies $f(0) \neq 0$, unbiased and biased cross-validation method both give similar optimal solutions for ϵ_n. However, unbiased cross-validation method usually produces smaller optimal solutions for v_n than biased cross-validation method does, which generates greater MISE [see Figure 4.24]. In the second example where $f(0) = 0$, two methods have very similar results. Technically, when sample size is small, BCV is slightly better than UCV according to MISE. In our opinion, we prefer BCV method to choose parameter.

4.3.2 Parameter Selection for $\tilde{f}_n^*(x)$

Note that $\tilde{f}_n^*$ have the same asymptotic normality as $\tilde{f}_n^+(x)$ and slightly different bias. Therefore, according to (2.78), the BCV function for $\tilde{f}_n^*$ seems to be

$$\begin{aligned} BCV^*(v_n, \varepsilon_n) &= \frac{I_2(q)\hat{\mu}}{nv_n}\int_0^\infty \frac{\tilde{f}_n^+(x)}{(x+\varepsilon_n)^2}dx \\ &\quad + \int_0^\infty [(xv_n^2+\epsilon_n)\tilde{f}_n^{+'}(x) + \frac{x^2v_n^2}{2}\tilde{f}_n^{+''}(x) + \epsilon_n\tilde{f}_n^+(0)\tilde{f}_n^+(x)]^2 dx. \end{aligned} \tag{4.23}$$

4.3.3 Parameter Selection for $\hat{f}_n(x)$

For estimator with Poisson weights based on G_n, we have the following BCV function

$$\begin{aligned} BCV(\lambda_n) &= \sqrt{\lambda_n}\frac{\hat{\mu}^2}{2\sqrt{\pi}n}MCE_n \\ &\quad + \frac{1}{2\lambda_n}\int_0^\infty [\hat{\hat{\mu}}\hat{f}_n(x) - \hat{f}_n(0)\hat{f}_n(x) - \frac{\hat{f}_n(x)}{x} - \hat{f}_n'(x)]^2 dx \end{aligned} \tag{4.24}$$

where $\hat{\mu} = \frac{1}{\frac{1}{n}\sum_{i=1}^n X_i^{-1}}$, $MCE_n = \frac{1}{n}\sum_{i=1}^n X_i^{-5/2}$ being an estimator of $\int_0^\infty \frac{f(x)}{\mu x^{3/2}}dx$ and $\hat{\hat{\mu}} = \frac{\frac{1}{n}\sum_{i=1}^n X_i^{-2}}{\hat{\mu}}$ being an estimator of $\int_0^\infty \frac{f(x)}{x}dx$

4.3.4 Parameter Selection for $\hat{f}_n^+(x)$ and $\hat{f}_n^*(x)$

According to the $AMISE$ of $\hat{f}_n^+(x)$, we can obtain the following BCV for $\hat{f}_n^+(x)$

$$\begin{aligned} BCV(v_n, \epsilon_n) &= \frac{I_2(q)\hat{\mu}}{nv_n}\int_0^\infty \frac{\hat{f}_n^+(x)}{(x+\epsilon_n)^2}dx \\ &+ \int_0^\infty [v_n^2\hat{f}_n^+(x) + (2v_n^2x+\epsilon_n)\hat{f}_n^{+\prime}(x) + v_n^2\frac{x^2}{2}\hat{f}_n^{+\prime\prime}(x)]^2dx. \end{aligned} \tag{4.25}$$

Furthermore, using the relation of bias between $\hat{f}_n^+(x)$ and $\hat{f}_n^*(x)$, we can establish the following BCV function for $\hat{f}_n^*(x)$

$$\begin{aligned} BCV^*(v_n, \epsilon_n) &= \frac{I_2(q)\hat{\mu}}{nv_n}\int_0^\infty \frac{\hat{f}_n^+(x)}{(x+\epsilon_n)^2}dx + \int_0^\infty \Big[v_n^2\hat{f}_n^+(x) \\ &+ (2v_n^2x+\epsilon_n)\hat{f}_n^{+\prime}(x) + v_n^2\frac{x^2}{2}\hat{f}_n^{+\prime\prime}(x) + \epsilon_n\hat{f}_n^+(0)\hat{f}_n^+(x)\Big]^2dx. \end{aligned} \tag{4.26}$$

4.3.5 Parameter Selection for Chen and Scaillet Estimators

In both Chen and Scaillet estimators there is a parameter b which controls the smoothness of density estimator. The way to choose the parameter is UCV method. Plugging in the corresponding Chen or Scaillet estimators and minimizing

$$UCV(b) = \int_0^\infty f_n^2(x)dx - 2\sum_{i=1}^n f_{n-1}(X_i, b; \mathcal{D}_i)/Z_i$$

,where $Z_i = \sum\limits_{j\neq i}\frac{X_i}{X_j}$, will give us the optimal solution of b. BCV method which involves the derivative of density estimator is not applicable for them.

4.4 A Comparison Between Different Estimators: Simulation Studies

In this section, we will compare these different density estimators through extensive simulations. First we generate LB data. Based on generated data, we choose the values of parameters. For our proposed estimators, we use BCV method and minimize BCV functions by `optimise` or `optim` in R to obtain the optimal solutions of parameters. For density estimators motivated by Chen and Scaillet's idea, we use UCV criterion to select parameters. Under the chosen parameters, we compute

$$ISE(f_n, f) = \int_0^\infty [f_n(x) - f(x)]^2 dx$$

and

$$SE\left(f_n(x), f(x)\right) = [f_n(x) - f(x)]^2$$

at some chosen points. We obtain 1000 samples of ISE and SE and use the averages of them as approximations of $MISE$ and MSE. Here, $MISE$ give us the global performance of density estimator. MSE let us to see how the density estimator performs locally at the points in which we might be interested. It is no doubt that we particularly want to know the behavior of density estimators near the lower boundary.

4.4.1 Simulation for χ_2^2 and χ_6^2

First we simulate for χ^2 distribution

$$f(x) = \frac{1}{2^{\alpha/2}\Gamma(2)} x^{\alpha/2-1} \exp\{-x/2\} I\{x > 0\}$$

with $\alpha = 2$ and $\alpha = 6$. When $\alpha = 2$, $f(x)$ is also an exponential distribution. The LB data has distributions of χ_4^2 and χ_8^2 respectively. Note that estimator with inverse Gaussian kernel does not perform very well for direct data [see Kulasekera and Padgett (2006)]. Our computation show that similar things happens to LB data. Here we do not include the simulation for IG estimator.

Table 4.2: Simulated MISE for χ_2^2

Distribution	Estimator	Sample Size					
		30	50	100	200	300	500
χ_2^2	Chen-1	0.13358	0.08336	0.07671	0.03900	0.03056	0.02554
	Chen-2	0.11195	0.08592	0.05642	0.03990	0.03301	0.02298
	RIG	0.14392	0.11268	0.07762	0.06588	0.05466	0.04734
	Poisson(F)	0.04562	0.03623	0.02673	0.01888	0.01350	0.01220
	Poisson(G)	0.08898	0.06653	0.04594	0.03127	0.02487	0.01885
	Gamma(F)	0.06791	0.05863	0.03989	0.03135	0.02323	0.01589
	Gamma*(F)	0.02821	0.01964	0.01224	0.00796	0.00609	0.00440
	Gamma(G)	0.09861	0.07663	0.05168	0.03000	0.02007	0.01317
	Gamma*(G)	0.02370	0.01244	0.00782	0.00537	0.00465	0.00356

Table 4.3: Simulated MSE for χ^2_2

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=30	I	0.1307	0.2040	0.0556	0.0181	0.0044	0.0003	1.6×10^{-5}
	II	0.1187	0.2499	0.0590	0.0173	0.0045	0.0012	8.7×10^{-5}
	III	0.2222	0.1823	0.0975	0.0250	0.0074	0.0022	3.2×10^{-5}
	IV	0.1487	0.1001	0.0239	0.0049	0.0015	0.0005	3.5×10^{-5}
	V	0.3003	0.2438	0.0695	0.0286	0.0148	0.0013	2.4×10^{-5}
	VI	0.1936	0.1447	0.0324	0.0117	0.0042	0.0002	7.2×10^{-5}
	VI*	0.0329	0.0286	0.0181	0.0090	0.0030	9.8×10^{-5}	1.5×10^{-4}
	VII	0.1893	0.1720	0.0637	0.0209	0.0066	0.0003	6.1×10^{-6}
	VII*	0.0528	0.0410	0.0150	0.0032	0.0020	8.4×10^{-5}	1.9×10^{-5}
n=50	I	0.1370	0.1493	0.0420	0.0121	0.0030	0.0002	9.0×10^{-6}
	II	0.1279	0.1894	0.0370	0.0112	0.0032	0.0008	4.6×10^{-5}
	III	0.2193	0.1774	0.0640	0.0161	0.0046	0.0046	1.6×10^{-5}
	IV	0.1393	0.0885	0.0159	0.0034	0.0012	0.0003	2.2×10^{-5}
	V	0.2939	0.1924	0.0438	0.0218	0.0094	0.0007	1.3×10^{-5}
	VI	0.1808	0.1365	0.0249	0.0101	0.0036	0.0001	5.7×10^{-5}
	VI*	0.0196	0.0172	0.0118	0.0070	0.0024	6.8×10^{-5}	1.4×10^{-4}
	VII	0.1584	0.1440	0.0420	0.0168	0.0060	0.0002	3.6×10^{-6}
	VII*	0.0322	0.0236	0.0063	0.0014	0.0012	4.8×10^{-5}	1.5×10^{-5}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI-Gamma(F), VI*-Corrected Gamma(F), VII-Gamma(G), VII*-Corrected Gamma(G)

Table 4.4: Simulated MSE for χ_2^2

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=100	I	0.1442	0.8201	0.0232	0.0070	0.0017	0.0001	4.5×10^{-6}
	II	0.1142	0.1391	0.0253	0.0054	0.0019	0.0005	2.3×10^{-5}
	III	0.2151	0.1631	0.0435	0.0091	0.0030	0.0020	7.8×10^{-6}
	IV	0.1335	0.0724	0.0094	0.0023	0.0008	0.0002	1.2×10^{-5}
	V	0.2498	0.1267	0.0245	0.0116	0.0050	0.0003	6.5×10^{-6}
	VI	0.1332	0.0823	0.0177	0.0090	0.0032	0.0001	3.6×10^{-5}
	VI*	0.0105	0.0094	0.0070	0.0047	0.0015	5.9×10^{-5}	1.0×10^{-4}
	VII	0.1078	0.0980	0.0255	0.0121	0.0051	0.0002	1.9×10^{-6}
	VII*	0.0280	0.0184	0.0025	0.0006	0.0007	3.8×10^{-5}	9.4×10^{-6}
n=200	I	0.3327	0.0901	0.0112	0.0046	0.0012	6.6×10^{-5}	2.5×10^{-6}
	II	0.2111	0.0943	0.0131	0.0027	0.0012	0.0003	1.3×10^{-5}
	III	0.2127	0.1896	0.0229	0.0067	0.0019	0.0080	5.9×10^{-6}
	IV	0.1139	0.0545	0.0054	0.0015	0.0005	0.0001	8.8×10^{-6}
	V	0.1908	0.0703	0.0116	0.0056	0.0026	0.0001	3.0×10^{-6}
	VI	0.0995	0.0782	0.0120	0.0065	0.0024	7.4×10^{-5}	2.9×10^{-5}
	VI*	0.0137	0.0125	0.0046	0.0031	0.0010	5.8×10^{-5}	8.2×10^{-5}
	VII	0.0636	0.0560	0.0123	0.0072	0.0038	0.0002	1.3×10^{-6}
	VII*	0.0217	0.0134	0.0013	0.0002	0.0005	2.9×10^{-5}	6.7×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI-Gamma(F), VI*-Corrected Gamma(F), VII-Gamma(G), VII*-Corrected Gamma(G)

Table 4.5: Simulated MSE for χ^2_2

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=300	I	0.2607	0.0690	0.0092	0.0033	0.0009	4.9×10^{-5}	1.9×10^{-6}
	II	0.1565	0.0897	0.0090	0.0018	0.0010	0.0002	1.0×10^{-5}
	III	0.2101	0.1643	0.0172	0.0051	0.0026	0.0020	5.1×10^{-6}
	IV	0.0771	0.0406	0.0036	0.0011	0.0003	0.0001	7.2×10^{-6}
	V	0.1724	0.0510	0.0070	0.0042	0.0018	0.0001	2.0×10^{-6}
	VI	0.0737	0.0577	0.0086	0.0050	0.0019	4.8×10^{-5}	2.6×10^{-5}
	VI*	0.0194	0.0118	0.0011	0.0001	0.0004	2.5×10^{-5}	5.5×10^{-6}
	VII	0.0449	0.0363	0.0067	0.0049	0.0030	0.0001	1.2×10^{-6}
	VII*	0.0420	0.0167	0.0012	0.0006	0.0003	1.5×10^{-5}	6.1×10^{-6}
n=500	I	0.2320	0.0732	0.0059	0.0024	0.0006	3.1×10^{-5}	1.2×10^{-6}
	II	0.1124	0.0723	0.0048	0.0012	0.0007	0.0001	6.7×10^{-6}
	III	0.2087	0.1681	0.0144	0.0039	0.0040	6.4×10^{-5}	1.8×10^{-6}
	IV	0.1014	0.0345	0.0029	0.0008	0.0002	6.9×10^{-5}	5.4×10^{-6}
	V	0.1321	0.0309	0.0039	0.0025	0.0013	9.1×10^{-5}	1.4×10^{-6}
	VI	0.0507	0.0365	0.0058	0.0038	0.0015	2.9×10^{-5}	2.4×10^{-5}
	VI*	0.0038	0.0034	0.0024	0.0016	0.0005	4.7×10^{-5}	5.1×10^{-5}
	VII	0.0310	0.0205	0.0040	0.0033	0.0023	0.0001	1.1×10^{-6}
	VII*	0.0146	0.0088	0.0009	0.0001	0.0003	2.0×10^{-5}	4.4×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI-Gamma(F), VI*-Corrected Gamma(F), VII-Gamma(G), VII*-Corrected Gamma(G)

Table 4.6: Simulated MISE for χ^2_6

Distribution	Estimator	Sample Size					
		30	50	100	200	300	500
χ^2_6	Chen-1	0.01592	0.01038	0.00578	0.00338	0.00246	0.00165
	Chen-2	0.01419	0.00973	0.00528	0.00303	0.00224	0.00153
	RIG	0.01438	0.00871	0.00482	0.00281	0.00208	0.00148
	Poisson(F)	0.00827	0.00582	0.00382	0.00241	0.00178	0.00119
	Poisson(G)	0.00834	0.00562	0.00356	0.00216	0.00166	0.00117
	Gamma(F)	0.01109	0.00805	0.00542	0.00327	0.00249	0.00181
	Gamma*(F)	0.01141	0.00844	0.00578	0.00345	0.00264	0.00193
	Gamma(G)	0.01536	0.01063	0.00688	0.00398	0.00303	0.00213
	Gamma*(G)	0.01536	0.01063	0.00688	0.00398	0.00303	0.00213

Table 4.7: Simulated MSE for χ^2_6

Sample Size	Estimator	x = 0	0.1	1	4	6	10	20
n=30	I	0.0017	0.0018	0.0018	0.0019	0.0011	0.0001	5.6×10^{-6}
	II	0.0018	0.0017	0.0011	0.0017	0.0008	0.0002	2.6×10^{-6}
	III	5.6×10^{-5}	6.7×10^{-5}	0.0006	0.0017	0.0008	0.0002	1.1×10^{-5}
	IV	0.0016	0.0016	0.0012	0.0012	0.0006	0.0001	1.9×10^{-5}
	V	0.0000	2.6×10^{-5}	0.0017	0.0008	0.0004	0.0001	2.9×10^{-6}
	VI	0.0011	0.0010	0.0019	0.0012	0.0012	8.5×10^{-5}	4.9×10^{-5}
	VI*	0.0015	0.0021	0.0020	0.0012	0.0012	7.9×10^{-5}	4.2×10^{-5}
	VII	0.0000	3.6×10^{-7}	0.0058	0.0008	0.0006	0.0001	4.1×10^{-6}
	VII*	0.0000	3.6×10^{-7}	0.0058	0.0008	0.0006	0.0001	4.1×10^{-6}
n=50	I	0.0012	0.0013	0.0015	0.0012	0.0007	0.0001	3.2×10^{-6}
	II	0.0013	0.0012	0.0008	0.0011	0.0005	0.0001	1.3×10^{-5}
	III	4.6×10^{-5}	5.7×10^{-5}	0.0006	0.0011	0.0005	0.0001	6.2×10^{-6}
	IV	0.0011	0.0011	0.0010	0.0008	0.0004	8.4×10^{-5}	1.0×10^{-5}
	V	0.0000	6.7×10^{-5}	0.0012	0.0005	0.0003	0.0001	1.7×10^{-6}
	VI	0.0005	0.0005	0.0016	0.0008	0.0009	5.5×10^{-5}	3.1×10^{-5}
	VI*	0.0006	0.0015	0.0016	0.0008	0.0009	5.3×10^{-5}	3.2×10^{-5}
	VII	0.0000	4.3×10^{-6}	0.0037	0.0004	0.0004	7.9×10^{-5}	3.0×10^{-5}
	VII*	0.0000	4.3×10^{-6}	0.0037	0.0004	0.0004	7.9×10^{-5}	3.0×10^{-5}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI-Gamma(F), VI*-Corrected Gamma(F), VII-Gamma(G), VII*-Corrected Gamma(G)

Table 4.8: Simulated MSE for χ^2_6

Sample Size	Estimator	x						
		0	0.1	1	4	6	10	20
n=100	I	0.0008	0.0009	0.0012	0.0006	0.0003	4.8×10^{-5}	1.4×10^{-6}
	II	0.0008	0.0008	0.0007	0.0006	0.0002	9.3×10^{-5}	5.5×10^{-6}
	III	3.4×10^{-5}	4.5×10^{-5}	0.0006	0.0006	0.0002	9.9×10^{-5}	2.8×10^{-6}
	IV	0.0008	0.0008	0.0007	0.0005	0.0002	4.9×10^{-5}	4.7×10^{-6}
	V	0.0000	0.0002	0.0008	0.0003	0.0002	5.9×10^{-5}	8.7×10^{-7}
	VI	0.0001	0.0001	0.0012	0.0006	0.0006	2.8×10^{-5}	2.2×10^{-5}
	VI*	0.0001	0.0014	0.0012	0.0006	0.0006	2.8×10^{-5}	2.2×10^{-5}
	VII	0.0000	0.0022	0.0023	0.0002	0.0002	4.9×10^{-5}	2.0×10^{-6}
	VII*	0.0000	0.0022	0.0023	0.0002	0.0002	4.9×10^{-5}	2.0×10^{-6}
n=200	I	0.0003	0.0004	0.0007	0.0003	0.0002	2.7×10^{-5}	6.5×10^{-7}
	II	0.0003	0.0003	0.0004	0.0003	0.0001	5.3×10^{-5}	1.9×10^{-6}
	III	2.0×10^{-5}	2.9×10^{-5}	0.0003	0.0003	0.0001	5.6×10^{-5}	1.2×10^{-6}
	IV	0.0005	0.0005	0.0004	0.0003	0.0001	3.0×10^{-5}	2.2×10^{-6}
	V	0.0000	0.0001	0.0004	0.0002	0.0001	3.4×10^{-5}	4.6×10^{-7}
	VI	3.0×10^{-5}	7.0×10^{-5}	0.0007	0.0003	0.0003	1.4×10^{-5}	1.2×10^{-5}
	VI*	2.0×10^{-5}	0.0005	0.0007	0.0003	0.0003	1.5×10^{-5}	1.2×10^{-5}
	VII	0.0000	0.0007	0.0012	0.0001	0.0001	3.0×10^{-5}	1.4×10^{-6}
	VII*	0.0000	0.0007	0.0012	0.0001	0.0001	3.0×10^{-5}	1.4×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI-Gamma(F), VI*-Corrected Gamma(F), VII-Gamma(G), VII*-Corrected Gamma(G)

Table 4.9: Simulated MSE for χ^2_6

Sample Size	Estimator	x = 0	0.1	1	4	6	10	20
n=300	I	0.0002	0.0003	0.0006	0.0002	0.0001	1.9×10^{-5}	4.0×10^{-7}
	II	0.0002	0.0002	0.0004	0.0002	0.0001	3.8×10^{-5}	1.0×10^{-6}
	III	1.4×10^{-5}	2.2×10^{-5}	0.0003	0.0003	0.0001	4.0×10^{-5}	8.1×10^{-7}
	IV	0.0003	0.0003	0.0003	0.0002	0.0001	2.3×10^{-5}	1.3×10^{-6}
	V	0.0000	8.3×10^{-5}	0.0003	0.0001	9.7×10^{-5}	2.4×10^{-5}	3.7×10^{-7}
	VI	1.4×10^{-5}	0.0001	0.0005	0.0002	0.0002	1.0×10^{-5}	8.4×10^{-6}
	VI*	1.3×10^{-5}	0.0004	0.0005	0.0002	0.0002	1.1×10^{-5}	8.8×10^{-6}
	VII	0.0000	0.0011	0.0008	0.0001	0.0001	2.3×10^{-5}	1.0×10^{-6}
	VII*	0.0000	0.0011	0.0008	0.0001	0.0001	2.3×10^{-5}	1.0×10^{-6}
n=500	I	0.0001	0.0002	0.0004	0.0001	9.5×10^{-5}	1.3×10^{-5}	2.3×10^{-6}
	II	0.0001	0.0001	0.0003	0.0001	6.6×10^{-5}	2.5×10^{-5}	4.7×10^{-7}
	III	9.4×10^{-6}	1.8×10^{-5}	0.0003	0.0001	6.5×10^{-5}	2.6×10^{-5}	4.5×10^{-7}
	IV	0.0002	0.0002	0.0002	0.0001	7.6×10^{-5}	1.5×10^{-5}	6.4×10^{-7}
	V	0.0000	6.1×10^{-5}	0.0002	0.0001	6.5×10^{-5}	1.6×10^{-5}	2.0×10^{-7}
	VI	9.2×10^{-6}	0.0001	0.0004	0.0001	0.0001	6.8×10^{-6}	5.2×10^{-6}
	VI*	3.9×10^{-5}	0.0002	0.0004	0.0001	0.0001	7.5×10^{-6}	5.9×10^{-6}
	VII	0.0000	0.0006	0.0006	8.1×10^{-5}	9.9×10^{-5}	1.6×10^{-5}	7.7×10^{-7}
	VII*	0.0000	0.0006	0.0006	8.1×10^{-5}	9.9×10^{-5}	1.6×10^{-5}	7.7×10^{-7}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI-Gamma(F), VI*-Corrected Gamma(F), VII-Gamma(G), VII*-Corrected Gamma(G)

From Tables 4.2-4.9, we can see that, for χ_2^2 density, two Chen estimators are slightly different. $\hat{f}_{C2}$ has smaller MSEs at the boundary and $MISE$s than $\hat{f}_{C1}$. This means $\hat{f}_{C2}$ performs better locally and globally than $\hat{f}_{C1}$. This adapts to Chen (2000) in direct data case which shows that $\hat{f}_2$ should have smaller $MISE$ than $\hat{f}_1$. Overall, the density estimators motivated by Chen or Scaillet's idea do not perform very well either globally or locally near lower boundary. Generally, estimators using Chen's idea have similar MSEs at the origin to Poisson estimator based on F_n and their performances at the lower boundary are comparable. However, in some cases for example $n = 200$ and $n = 300$, PWE is much better. Poisson weights estimator based on F_n behave much well at the rest points. So it has much smaller $MISE$s than Chen and Scaillet estimators. Scaillet estimator has huge MSEs at the boundary and the largest $MISE$s. Although Poisson weight estimator based on G_n has relatively smaller $MISE$s, it has great MSEs at the boundary as well, just like Scaillet estimator. Therefore, they might not be suitable for estimating the density whose value does not equal to zero at the boundary. Two original gamma estimator perform similarly to PWE based F_n. Even though they have two parameters, their behaviors quite differ from what are expected. This is due to the fact that, in this case the parameter ϵ_ns are usually not zero, which causes the estimators to " lose " some weights and not to be a valid density estimators [Their integrals from 0 to ∞ is less than 1]. In this example, two " stars " are two corrected gamma estimators, which perform best locally and globally. The boundary corrections are very necessary and effective. They reduce dramatically estimators' MSEs near the boundary and relatively slightly at the rest points. Therefore, two estimators have the smallest $MISE$s. The corrected gamma estimator based on F_n behave better near the boundary than the corrected gamma estimator based on G_n. However, at the points away from the origin, it is just the opposite. The estimator based G_n is better than the estimator based on F_n. Overall, estimator based on G_n has slightly smaller $MISE$s.

For χ_6^2, all estimators have comparable global results. At the lower boundary, RIG estimator, gamma estimators, Poisson weights estimator based on G_n have similar results. Chen estimators behave like PWE based on F_n. They are slightly worse than previous estimators. Tow Poisson weights estimators, which have the smallest $MISE$s, perform similarly and very well globally. In this case, original gamma estimators are almost as same as the corrected gamma estimators. This is because, in this example, ϵ_ns are almost zero. When ϵ_n is 0, the corrected gamma estimators are the same as original gamma estimators which has a value of zero at the lower boundary.

In the first example with density such that $f(0) > 0$, the corrected gamma estimators perform much better than the original estimators. In the second example with density such that $f(0) = 0$, the corrected estimators have similar local and global behaviors to the original ones. So we can use the corrected estimators to replace the original estimators without hesitation.

4.4.2 Simulation for Some Other Standard Distributions

We have simulated for the following standard distribution as well.

(i). Lognormal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi}x} \exp\{-(\log x - \mu)^2/2\} I\{x > 0\};$$

(ii). Weibull Distribution

$$f(x) = \alpha x^{\alpha-1} \exp(-x^{\alpha}) I\{x > 0\};$$

(iii). Mixtures of Two Exponential Distribution

$$f(x) = [\pi \frac{1}{\theta_1} \exp(-x/\theta_1) + (1-\pi)\frac{1}{\theta_2} \exp(-x/\theta_2] I\{x > 0\}.$$

Table 4.10: Simulated MISE for Lognormal with $\mu = 0$

Distribution	Estimator	Sample Size					
		30	50	100	200	300	500
	Chen-1	0.12513	0.08416	0.05109	0.03450	0.02514	0.01727
	Chen-2	0.12327	0.08886	0.05200	0.03545	0.02488	0.01717
	RIG	0.14371	0.09733	0.05551	0.03308	0.02330	0.01497
Lognormal	Poisson(F)	0.05559	0.04379	0.02767	0.01831	0.01346	0.01001
	Poisson(G)	0.06952	0.04820	0.03158	0.01470	0.01474	0.01061
	Gamma*(F)	0.06846	0.05614	0.03963	0.02640	0.01998	0.01470
	Gamma*(G)	0.16365	0.12277	0.07568	0.04083	0.029913	0.02035

Table 4.11: Simulated MSE for Lognormal with $\mu = 0$

Sample Size	Estimator	x						
		0	0.1	e^{-1}	0.5	1	5	8
n=30	I	0.1108	0.0618	0.1682	0.1765	0.0211	2.0×10^{-4}	2.8×10^{-5}
	II	0.1045	0.0441	0.1616	0.1303	0.0196	6.0×10^{-4}	6.4×10^{-5}
	III	0.0026	0.0494	0.2555	0.1957	0.0207	3.0×10^{-4}	3.2×10^{-5}
	IV	0.1307	0.0485	0.0767	0.0545	0.0071	3.5×10^{-4}	3.8×10^{-5}
	V	0.0009	0.0810	0.0908	0.0532	0.0126	2.2×10^{-4}	2.6×10^{-5}
	VI*	0.0090	0.1546	0.0607	0.0451	0.0133	2.0×10^{-4}	9.1×10^{-5}
	VII*	0.0007	0.7321	0.1230	0.0608	0.0121	8.2×10^{-5}	9.4×10^{-6}
n=50	I	0.1123	0.0535	0.1199	0.1105	0.0158	1.4×10^{-4}	1.1×10^{-5}
	II	0.1056	0.0442	0.1391	0.1158	0.0110	3.7×10^{-4}	2.7×10^{-5}
	III	0.0027	0.0436	0.1719	0.1198	0.0133	2.0×10^{-4}	1.5×10^{-5}
	IV	0.1398	0.0482	0.0564	0.0389	0.0050	1.8×10^{-4}	1.6×10^{-5}
	V	0.0020	0.0641	0.0602	0.0333	0.0090	1.3×10^{-4}	1.3×10^{-5}
	VI*	0.0035	0.1349	0.0446	0.0320	0.0111	1.7×10^{-4}	6.6×10^{-5}
	VII*	0.0000	0.5482	0.0810	0.0386	0.0080	4.7×10^{-5}	4.9×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.12: Simulated MSE for Lognormal with $\mu = 0$

Sample Size	Estimator	x						
		0	0.1	e^{-1}	0.5	1	5	8
n=100	I	0.1044	0.0486	0.1122	0.0428	0.0086	5.1×10^{-5}	8.2×10^{-6}
	II	0.1038	0.0412	0.1248	0.0416	0.0059	1.5×10^{-4}	1.3×10^{-5}
	III	0.0028	0.0383	0.0917	0.0407	0.0064	7.2×10^{-5}	7.2×10^{-6}
	IV	0.1053	0.0424	0.0311	0.0187	0.0029	4.9×10^{-5}	4.7×10^{-6}
	V	0.0018	0.0541	0.0352	0.0202	0.0060	5.7×10^{-5}	6.8×10^{-6}
	VI*	0.0033	0.1011	0.0261	0.0218	0.0086	1.1×10^{-4}	4.7×10^{-5}
	VII*	0.0000	0.3237	0.0417	0.0202	0.0044	1.9×10^{-5}	2.1×10^{-6}
n=200	I	0.0854	0.0422	0.0310	0.0235	0.0054	2.7×10^{-5}	2.9×10^{-6}
	II	0.0871	0.0387	0.0367	0.0199	0.0036	6.5×10^{-5}	4.6×10^{-6}
	III	0.0024	0.0318	0.0385	0.0180	0.0042	3.5×10^{-5}	3.3×10^{-6}
	IV	0.0663	0.0320	0.0196	0.0113	0.0019	2.3×10^{-5}	2.4×10^{-6}
	V	0.0019	0.0309	0.0151	0.0085	0.0027	2.0×10^{-5}	2.5×10^{-6}
	VI*	0.0058	0.0717	0.0167	0.0153	0.0058	9.2×10^{-5}	3.3×10^{-5}
	VII*	0.0015	0.1780	0.0210	0.0102	0.0026	1.0×10^{-5}	1.1×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.13: Simulated MSE for Lognormal with $\mu = 0$

Sample Size	Estimator	x						
		0	0.1	e^{-1}	0.5	1	5	8
n=300	I	0.0670	0.0394	0.0230	0.0175	0.0034	1.7×10^{-5}	2.0×10^{-6}
	II	0.0739	0.0328	0.0256	0.0140	0.0023	4.0×10^{-5}	3.0×10^{-6}
	III	0.0021	0.0295	0.0243	0.0134	0.0027	2.2×10^{-5}	2.2×10^{-6}
	IV	0.0436	0.0270	0.0137	0.0078	0.0014	1.5×10^{-5}	1.6×10^{-6}
	V	0.0028	0.0319	0.0147	0.0089	0.0026	1.9×10^{-5}	2.3×10^{-6}
	VI*	0.0033	0.1011	0.0261	0.0218	0.0086	1.1×10^{-4}	4.7×10^{-5}
	VII*	0.0013	0.1330	0.0152	0.0074	0.0019	6.4×10^{-5}	8.2×10^{-6}
n=500	I	0.0519	0.0359	0.0158	0.0122	0.0021	1.1×10^{-5}	1.3×10^{-6}
	II	0.0561	0.0277	0.0169	0.0099	0.0017	2.1×10^{-5}	1.6×10^{-6}
	III	0.0018	0.0262	0.0167	0.0100	0.0019	1.4×10^{-5}	1.3×10^{-6}
	IV	0.0361	0.0199	0.0100	0.0057	0.0010	1.0×10^{-5}	1.1×10^{-6}
	V	0.0033	0.0244	0.0108	0.0065	0.0018	1.3×10^{-5}	1.5×10^{-6}
	VI*	0.0054	0.0577	0.0125	0.0119	0.0042	7.5×10^{-5}	2.4×10^{-5}
	VII*	0.0018	0.0902	0.0098	0.0048	0.0012	4.5×10^{-6}	6.2×10^{-7}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.14: Simulated MISE for Weibull with $\alpha = 2$

Distribution	Estimator	Sample Size					
		30	50	100	200	300	500
Weibull	Chen-1	0.10495	0.06636	0.03884	0.02312	0.01700	0.01167
	Chen-2	0.08651	0.05719	0.03595	0.02225	0.01611	0.01111
	RIG	0.08530	0.05532	0.03227	0.01984	0.01470	0.01045
	Poisson(F)	0.04993	0.03658	0.02432	0.01459	0.01179	0.00856
	Poisson(G)	0.05288	0.03548	0.02268	0.01392	0.01106	0.00810
	Gamma*(F)	0.08358	0.06671	0.04935	0.03169	0.02652	0.01694
	Gamma*(G)	0.12482	0.08526	0.05545	0.03402	0.02731	0.02188

Table 4.15: Simulated MSE for Weibull with $\alpha = 2$

Sample Size	Estimator	x						
		0	0.1	0.5	$1/\sqrt{2}$	1	2	3
n=30	I	0.0856	0.1343	0.0796	0.0720	0.0588	0.0030	1.9×10^{-4}
	II	0.0949	0.0555	0.0802	0.0571	0.0398	0.0116	1.4×10^{-4}
	III	0.0025	0.0802	0.0843	0.0549	0.0394	0.0095	4.6×10^{-4}
	IV	0.0844	0.0548	0.0239	0.0389	0.0280	0.0086	6.9×10^{-4}
	V	0.0068	0.0636	0.0370	0.0275	0.0186	0.0031	3.1×10^{-5}
	VI*	0.0019	0.1049	0.0287	0.0600	0.0682	0.0053	0.0022
	VII*	0.0000	0.2852	0.0514	0.0293	0.0336	0.0011	1.8×10^{-4}
n=50	I	0.0644	0.0576	0.0469	0.0514	0.0349	0.0020	1.0×10^{-4}
	II	0.0679	0.0431	0.0464	0.0382	0.0223	0.0077	7.1×10^{-4}
	III	0.0021	0.0208	0.0483	0.0391	0.0218	0.0063	2.2×10^{-4}
	IV	0.0682	0.0427	0.0158	0.0286	0.0217	0.0059	3.7×10^{-4}
	V	0.0025	0.0453	0.0246	0.0186	0.0138	0.0018	1.6×10^{-5}
	VI*	1.1×10^{-6}	0.0763	0.0195	0.0475	0.0560	0.0048	0.0018
	VII*	0.0000	0.1865	0.0353	0.0194	0.0251	0.0008	1.4×10^{-4}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.16: Simulated MSE for Weibull with $\alpha = 2$

Sample Size	Estimator	x						
		0	0.1	0.5	$1/\sqrt{2}$	1	2	3
n=100	I	0.0468	0.0500	0.0229	0.0261	0.0204	0.0012	4.3×10^{-5}
	II	0.0477	0.0332	0.0239	0.0233	0.0129	0.0045	3.2×10^{-4}
	III	0.0017	0.0154	0.0254	0.0238	0.0121	0.0039	1.0×10^{-4}
	IV	0.0516	0.0330	0.0095	0.0180	0.0141	0.0035	1.1×10^{-4}
	V	0.0071	0.0342	0.0152	0.0119	0.0084	0.0009	7.0×10^{-6}
	VI*	0.0003	0.0577	0.0127	0.0338	0.0425	0.0041	0.0013
	VII*	0.0033	0.1282	0.0215	0.0107	0.0180	0.0005	1.0×10^{-4}
n=200	I	0.0281	0.0349	0.0128	0.0144	0.0109	0.0007	1.6×10^{-5}
	II	0.0287	0.0273	0.0148	0.0140	0.0072	0.0024	1.1×10^{-4}
	III	0.0011	0.0161	0.0160	0.0142	0.0067	0.0022	3.8×10^{-5}
	IV	0.0317	0.0213	0.0063	0.0106	0.0078	0.0020	3.5×10^{-5}
	V	0.0117	0.0225	0.0099	0.0075	0.0049	0.0005	3.7×10^{-6}
	VI*	0.0004	0.0460	0.0082	0.0181	0.0246	0.0032	7.5×10^{-4}
	VII*	0.0106	0.0845	0.0135	0.0070	0.0121	0.0004	7.5×10^{-5}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.17: Simulated MSE for Weibull with $\alpha = 2$

Sample Size	Estimator	x						
		0	0.1	0.5	$1/\sqrt{2}$	1	2	3
n=300	I	0.0228	0.0303	0.0089	0.0099	0.0077	0.0004	8.8×10^{-6}
	II	0.0233	0.0237	0.0107	0.0098	0.0050	0.0017	5.6×10^{-5}
	III	0.0009	0.0164	0.0115	0.0101	0.0048	0.0016	2.6×10^{-5}
	IV	0.0293	0.0175	0.0048	0.0086	0.0064	0.0015	2.5×10^{-5}
	V	0.0156	0.0194	0.0073	0.0057	0.0036	0.0004	2.4×10^{-6}
	VI*	3.3×10^{-7}	0.0400	0.0062	0.0144	0.0197	0.0026	5.7×10^{-4}
	VII*	0.0158	0.0668	0.0103	0.0061	0.0102	0.0003	6.6×10^{-5}
n=500	I	0.0160	0.0218	0.0068	0.0065	0.0050	0.0003	4.5×10^{-6}
	II	0.0156	0.0189	0.0078	0.0067	0.0033	0.0010	2.5×10^{-5}
	III	0.0006	0.0148	0.0083	0.0071	0.0033	0.0011	1.6×10^{-5}
	IV	0.0293	0.0133	0.0036	0.0059	0.0044	0.0010	1.2×10^{-5}
	V	0.0076	0.0150	0.0056	0.0040	0.0025	0.0002	1.4×10^{-6}
	VI*	1.6×10^{-9}	0.0311	0.0042	0.0068	0.0101	0.0019	2.3×10^{-4}
	VII*	0.0283	0.0552	0.0078	0.0059	0.0089	0.0002	6.5×10^{-5}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.18: Simulated MSE for Mixture of Two Exponential Distributions with $\pi = 0.4$, $\theta_1 = 2$ and $\theta_2 = 1$

Distribution	Estimator	Sample Size					
		30	50	100	200	300	500
	Chen-1	0.22876	0.17045	0.08578	0.06718	0.05523	0.03811
	Chen-2	0.17564	0.15083	0.07331	0.08029	0.04931	0.03808
	RIG	0.25284	0.20900	0.13843	0.10879	0.09344	0.07776
Mixture	Poisson(F)	0.06838	0.05746	0.04116	0.02612	0.01896	0.01179
	Poisson(G)	0.11831	0.09274	0.06863	0.05019	0.03881	0.03044
	Gamma*(F)	0.04147	0.02645	0.01375	0.00758	0.00532	0.00361
	Gamma*(G)	0.02534	0.01437	0.01091	0.01223	0.01132	0.00994

Table 4.19: Simulated MSE for Mixtures of Two Exponential Distributions with $\pi = 0.4$, $\theta_1 = 2$ and $\theta_2 = 1$

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=30	I	0.3499	0.3075	0.1033	0.0249	0.0037	1.1×10^{-4}	2.6×10^{-6}
	II	0.3190	0.3181	0.0825	0.0245	0.0071	4.5×10^{-4}	1.3×10^{-5}
	III	0.5610	0.4423	0.1245	0.0564	0.0056	1.9×10^{-4}	2.9×10^{-6}
	IV	0.3778	0.1907	0.0181	0.0057	0.0027	1.3×10^{-4}	1.7×10^{-6}
	V	0.6409	0.3237	0.0388	0.0156	0.0043	1.1×10^{-4}	2.1×10^{-6}
	VI*	0.0652	0.0549	0.0271	0.0098	0.0006	3.4×10^{-4}	1.1×10^{-4}
	VII*	0.0696	0.0539	0.0070	0.0065	0.0009	5.4×10^{-5}	1.4×10^{-5}
n=50	I	0.3158	0.7921	0.0511	0.0128	0.0023	6.4×10^{-5}	1.1×10^{-6}
	II	0.2848	0.7600	0.0551	0.0143	0.0051	2.3×10^{-4}	2.3×10^{-6}
	III	0.5582	0.8473	0.0797	0.0364	0.0041	9.1×10^{-5}	1.3×10^{-6}
	IV	0.3840	0.1633	0.0136	0.0051	0.0020	7.7×10^{-5}	1.0×10^{-6}
	V	0.6228	0.2673	0.0263	0.0121	0.0028	7.1×10^{-5}	1.3×10^{-6}
	VI*	0.0489	0.0414	0.0187	0.0066	0.0004	2.9×10^{-4}	7.7×10^{-5}
	VII*	0.0500	0.0336	0.0032	0.0030	0.0007	3.4×10^{-5}	9.2×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.20: Simulated MSE for Mixtures of Two Exponential Distributions with $\pi = 0.4$, $\theta_1 = 2$ and $\theta_2 = 1$

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=100	I	0.3253	0.2465	0.0257	0.0072	0.0012	4.0×10^{-5}	5.0×10^{-7}
	II	0.2581	0.1926	0.0235	0.0082	0.0031	1.1×10^{-4}	8.3×10^{-7}
	III	0.5433	0.4810	0.0505	0.0633	0.0021	4.8×10^{-5}	5.8×10^{-7}
	IV	0.3378	0.1163	0.0094	0.0034	0.0011	4.2×10^{-5}	4.6×10^{-7}
	V	0.6336	0.1893	0.0181	0.0073	0.0015	3.9×10^{-5}	6.5×10^{-7}
	VI*	0.0238	0.0200	0.0101	0.0036	0.0001	2.0×10^{-4}	4.4×10^{-5}
	VII*	0.0647	0.0317	0.0016	0.0014	0.0006	1.4×10^{-5}	4.4×10^{-6}
n=200	I	0.8107	0.1730	0.0149	0.0040	0.0008	2.6×10^{-5}	3.2×10^{-7}
	II	0.6067	0.1642	0.0158	0.0052	0.0019	4.8×10^{-5}	3.5×10^{-7}
	III	0.5376	0.3734	0.0269	0.0589	0.0011	2.7×10^{-5}	3.4×10^{-7}
	IV	0.2520	0.0722	0.0060	0.0019	0.0005	2.3×10^{-5}	2.6×10^{-7}
	V	0.6395	0.1232	0.0124	0.0044	0.0008	2.7×10^{-5}	3.8×10^{-7}
	VI*	0.0125	0.0105	0.0055	0.0020	0.0001	1.3×10^{-4}	2.4×10^{-5}
	VII*	0.0870	0.0380	0.0011	0.0010	0.0007	7.1×10^{-6}	1.7×10^{-6}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G),
VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

Table 4.21: Simulated MSE for Mixtures of Two Exponential Distributions with $\pi = 0.4$, $\theta_1 = 2$ and $\theta_2 = 1$

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=300	I	0.5566	0.0958	0.0110	0.0033	0.0006	2.1×10^{-5}	2.8×10^{-7}
	II	0.3110	0.1213	0.0107	0.0044	0.0015	3.3×10^{-5}	2.6×10^{-7}
	III	0.3110	0.1213	0.0107	0.0044	0.0015	3.3×10^{-5}	2.6×10^{-7}
	IV	0.1921	0.0531	0.0038	0.0014	0.0004	1.6×10^{-5}	2.7×10^{-7}
	V	0.5974	0.0901	0.0090	0.0034	0.0005	2.0×10^{-5}	3.4×10^{-7}
	VI*	0.0092	0.0075	0.0038	0.0014	7.6×10^{-5}	0.0001	1.5×10^{-5}
	VII*	0.0806	0.0353	0.0011	0.0007	0.0006	6.2×10^{-6}	1.0×10^{-6}
n=500	I	0.4798	0.0723	0.0075	0.0024	0.0004	2.0×10^{-5}	2.3×10^{-7}
	II	0.2553	0.0950	0.0077	0.0034	0.0009	1.8×10^{-5}	1.7×10^{-7}
	III	0.5273	0.2300	0.0176	0.0045	0.0012	3.0×10^{-5}	3.9×10^{-7}
	IV	0.1147	0.0338	0.0020	0.0009	0.0002	1.1×10^{-5}	1.6×10^{-7}
	V	0.5613	0.0588	0.0071	0.0023	0.0003	1.8×10^{-5}	2.8×10^{-7}
	VI*	0.0074	0.0047	0.0025	0.0010	5.0×10^{-5}	7.1×10^{-5}	1.0×10^{-5}
	VII*	0.0698	0.0311	0.0010	0.0006	0.0006	6.1×10^{-6}	7.0×10^{-7}

I-Chen-1, II-Chen-2, III-RIG, IV-Poisson(F), V-Poisson(G), VI*-Corrected Gamma(F), VII*-Corrected Gamma(G)

4.4.3 Discussions and Conclusions

From the simulation results given in tables, we can see that two Chen estimators have similar performances at the edge. Usually $\hat{f}_{C2}$ have smaller $MISE$ than $\hat{f}_{C1}$. For direct data, Chen (2000) show that density estimator under parameter choice (1.14) has a better global performance than that under choice (1.13). This property might be adapted to LB data. The simulated MSEs show that the $\hat{f}_{C2}$ preform better than $\hat{f}_{C1}$ in the neighborhood of origin, so $MISE$ of $\hat{f}_{C2}$ is lower. However, the two Chen density estimators do not perform very well globally and locally near the origin comparing with other density estimators. They have relatively great $MISE$s and MSEs near the lower edge. Judged by the simulated MSEs at the boundary, the two estimators can not completely remove the bias at the edge, even in the case that underlying density such that $f(0) = 0$ [see simulation for χ^2_6, Lognormal].

Replacing gamma kernels with RIG kernels, Scaillet estimator have a great advance in reducing MSEs at the origin for underlying density such that $f(0) = 0$ [see simulation for χ^2_6, Lognormal and Weibull distributions]. In some cases, the estimator even has zero error at the origin. So, under this circumstance, Scaillet estimator have smaller $MISE$s than Chen estimators. It seems that RIG density are more suitable as kernels than gamma density in these cases. However, the advantage becomes disadvantage in estimating underlying density such $f(0) \neq 0$. In this kind of cases, Scaillet estimator has huge MSEs at the origin [see simulation for χ^2_2 and mixture of two exponential distributions]. According to the examples we have here, it seems to be concluded that Scaillet estimator might be just suitable to estimate density with zero value at the border.

For χ^2_6, Lognormal and Weibull distributions, Poisson weights estimator (PWE) based on F_n have the smallest $MISE$s. If we just consider $MISE$, this estimator is perfect in these examples. However, if looking into MSEs, we find that it is not that perfect. It still has relatively great MSEs at the origin. But the MSEs at these points away from origin are much smaller. So it has smallest $MISE$s. From this example, we can see that MSEs give us a valid method to observe the local performance of estimator, especially the performance at these points in which we are interested. For χ^2_2 and mixtures distributions, although this estimator performs better than Chen estimators, it still has relatively great MSEs at the border. Note that the Scaillet estimators' value at the lower boundary is always zero or close to zero. If we look into the plots of Poisson weights [Figure 4.26] near zero and the gamma kernels used in Chen estimators [see Chen (2000)], we find that PWE and Chen estimators use a similar strategy to

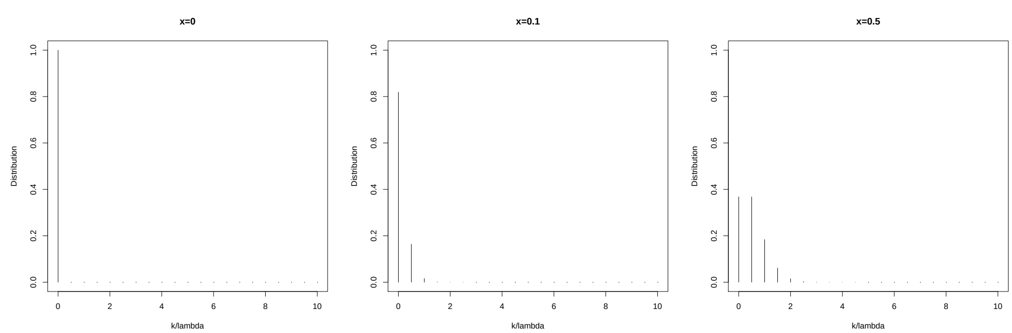


Figure 4.26: Plots of distribution of Poisson weights with $\lambda = 2$.

avoid the defect in Scaillet estimators, that is they both change the shape of kernels or weights near zero into exponential-like density shape. Although changing the shape of kernels or weights is a valid way to avoid such a defect in Scaillet estimators, it is not a perfect strategy to remove the bias at the boundary. This strategy might be more suitable for the true underlying density which has an exponential-like shape. For some other kinds of true underlying density, they may have less efficiency in exploring the true underlying density's character of $f(0) = 0$ when sample size is small [see simulated MSEs at the boundary for Lognormal and Weibull distributions].

For PWE based on G_n, in the simulation for χ_6^2, Lognormal and Weibull distributions, it has similar $MISE$s to PWE based on F_n and at the boundary has smaller MSEs. However, it does not perform very well in the simulation for χ_2^2 and mixtures of two exponential distributions, especially at the boundary. Therefore, it may be suitable to estimate true underlying density with $f(0) = 0$.

The corrected gamma estimator based on F_n performs very well locally and globally in the simulation for χ_2^2 and mixtures distributions. The parameter ϵ_n and boundary correction effectively reduce the bias at the boundary and result in the dramatic decrease of $MISE$s. For the rest distributions, this estimator has comparable $MISE$s to other estimators and satisfactory MSEs at the boundary. The BCV method is valid to decide whether the optimal solution of ϵ_n is zero or not. So that this estimator can accurately explore the characters of underly density at the boundary behind the data. Inspired by Scaillet estimator, in order to further reduce $MISE$, we can substitute gamma kernels with RIG or IG kernels. Actually, RIG or IG kernels are completely adapted to our estimator.

The corrected gamma estimator based on G_n has the smallest $MISE$s in the simulation for χ_2^2. However, the MSEs at the boundary are little worse than corrected gamma estimator based on F_n. Further simulation shows that, although this estimator has zero error at the boundary in some cases, it has relatively great MSEs in the neighborhood of origin. This estimator may not be very stable in this area. Recalling PWE based on G_n, we may conclude that, using the same smooth technique, the density estimator obtained by smoothing F_n performs better than that obtained by smoothing G_n. This seems to be true for kernel method as well, since Jones estimator is better than Bhattacharyya *et al.* estimator. If we first obtain $g_n(x)$ the estimator of weighted density, then, we divide the estimator by x to obtain the unweighted density. Because the existence of bias of $g_n(x)$, it is not easy to control the ratio $g_n(x)/x$ to be close to $f(x)/\mu$ near the lower border. Besides, because of the term $1/x$, the bias will be enlarged and even blows up [see Bhattacharyya *et al.* estimator]. Therefore, the estimators based on G_n may have more difficulties in exploring characters of underlying density near boundary in some cases. For the bias data, if the weight function is more complicated than x, or has a term with a higher order than x, say x^2, the situation will become more worse. For LB data or biased data, a better way is smoothing Cox estimator to estimate unweighted density. So, through this point, we can see that the Cox's estimator plays an important role in estimating density function for LB data.

4.5 A Linear Combination of Two Density Estimators

Through the simulation studies given in the previous section, we can see that two corrected gamma estimator perform well. Looking into SEs, we find that gamma estimator based on F_n has smaller bias near the lower boundary. However, gamma estimator based on G_n has smaller error at the tail. In order to take a full advantage of the two estimators' merits, we consider the linear combination of the two estimators as follows.

$$f_n^C(x) = a\tilde{f}_n^*(x) + (1-a)\hat{f}_n^*(x). \tag{4.27}$$

where $0 \le a \le 1$. Note that five parameters are involved in f_n^C: two pairs of (υ_n, ϵ_n) in $\tilde{f}_n^*$ and $\hat{f}_n^*$ respectively; one parameter a connecting two density estimators. We consider to use two steps to choose these parameters. In first step, using BCV methods described above to select the parameters in $\tilde{f}_n^*$ and $\hat{f}_n^*$, say $(\upsilon_{n1}, \epsilon_{n1})$ and $(\upsilon_{n2}, \epsilon_{n2})$; then select parameter a. We hope that the chosen parameter a would make the variance of f_n^C as small as possible. Note that

$$V(f_n^C(x)) = a^2 V(\tilde{f}_n^*(x)) + 2a(1-a)Cov(\tilde{f}_n^*(x), \hat{f}_n^*(x)) + (1-a)^2 V(\hat{f}_n^*(x)). \tag{4.28}$$

According to (4.23) and (4.26), we have

$$V(\tilde{f}_n^*(x)) \approx \frac{I_2(q)\hat{\mu}}{n\upsilon_{n1}} \frac{\tilde{f}_n^+(x)}{(x+\epsilon_{n1})^2} \tag{4.29}$$

and

$$V(\hat{f}_n^*(x)) \approx \frac{I_2(q)\hat{\mu}}{n\upsilon_{n2}} \frac{\hat{f}_n^+(x)}{(x+\epsilon_{n2})^2}. \tag{4.30}$$

Furthermore, we have $\tilde{f}_n^*(x) \approx \frac{\mu}{n(x+\epsilon_{n1})^2} \sum_{i=1}^n q_{\upsilon_{n1}}\left(\frac{X_i}{x+\epsilon_{n1}}\right)$, and
$\hat{f}_n^*(x) \approx \frac{\mu}{n(x+\epsilon_{n1})^3} \sum_{i=1}^n X_i q_{\upsilon_{n2}}\left(\frac{X_i}{x+\epsilon_{n1}}\right)$, then we can compute

$$Cov(\tilde{f}_n^*(x), \hat{f}_n^*(x)) \approx \frac{1}{n}\left[\frac{f(x)}{(x+\epsilon_{n1})^2}\int_0^\infty t q_{\upsilon_{n1}}(t) q_{\upsilon_{n2}}(t) dt - (f(x))^2\right].$$

So we can estimate $Cov(\tilde{f}_n^*(x), \hat{f}_n^*(x))$ by

$$\widehat{Cov}(\tilde{f}_n^*(x), \hat{f}_n^*(x)) = \frac{1}{n}\left[\frac{\tilde{f}_n^+(x)}{(x+\epsilon_{n1})^2}\int_0^\infty t q_{\upsilon_{n1}}(t) q_{\upsilon_{n2}}(t) dt - (\tilde{f}_n^+(x))^2\right]. \tag{4.31}$$

So integrated variance of f_n^C can be approximated by

$$\begin{aligned} AIV(a) &= a^2 \frac{I_2(q)\hat{\mu}}{n\upsilon_{n1}} \int_0^\infty \frac{\tilde{f}_n^+(x)}{(x+\epsilon_{n1})^2} dx \\ &\quad + \frac{2a(1-a)}{n}\left[\hat{\mu}\int_0^\infty \frac{\tilde{f}_n^+(x)}{(x+\epsilon_n 1)^2} dx \int_0^\infty t q_{\upsilon_{n1}}(t) q_{\upsilon_{n2}}(t) dt - \int_0^\infty (\tilde{f}_n^+(x))^2 dx\right] \\ &\quad + (1-a)^2 \frac{I_2(q)\hat{\mu}}{n\upsilon_{n2}} \int_0^\infty \frac{\hat{f}_n^+(x)}{(x+\epsilon_{n2})^2} dx. \end{aligned} \tag{4.32}$$

Minimizing (4.32) between 0 and 1 will give us an optimal solution of parameter a.

We also present some results of the new estimator's $MISE$ and SE based on a simulation study.

Table 4.22: Simulated MISE for Standard Distributions

Distribution	Estimator	Sample Size					
		30	50	100	200	300	500
χ_2^2	Gamma*(F)	0.02821	0.01964	0.01224	0.00796	0.00609	0.00440
	Gamma*(G)	0.02370	0.01244	0.00782	0.00537	0.00465	0.00356
	Combination	0.01638	0.00927	0.00597	0.00440	0.00363	0.00286
χ_6^2	Gamma*(F)	0.01141	0.00844	0.00578	0.00345	0.00264	0.00193
	Gamma*(G)	0.01536	0.01063	0.00688	0.00398	0.00303	0.00213
	Combination	0.01133	0.00838	0.00574	0.00343	0.00263	0.00192
Lognormal	Gamma*(F)	0.06846	0.05614	0.03963	0.02640	0.01998	0.01470
	Gamma*(G)	0.16365	0.12277	0.07568	0.04083	0.029913	0.02035
	Combination	0.06845	0.05616	0.03962	0.02555	0.01950	0.01444
Weibull	Gamma*(F)	0.08358	0.06671	0.04935	0.03169	0.02652	0.01694
	Gamma*(G)	0.12482	0.08526	0.05545	0.03402	0.02731	0.02188
	Combination	0.07868	0.06351	0.04727	0.03038	0.02429	0.01861
Mixture	Gamma*(F)	0.04147	0.02645	0.01375	0.00758	0.00532	0.00361
	Gamma*(G)	0.02534	0.01437	0.01091	0.01223	0.01132	0.00994
	Combination	0.01861	0.01165	0.00689	0.00458	0.00346	0.00242

Table 4.23: Simulated MSE for χ^2_2 Distribution

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=30	VI*	0.0329	0.0286	0.0181	0.0090	0.0030	9.8×10^{-5}	1.5×10^{-4}
	VII*	0.0528	0.0410	0.0150	0.0032	0.0020	8.4×10^{-5}	1.9×10^{-5}
	VIII	0.0325	0.0242	0.0095	0.0029	0.0017	6.9×10^{-5}	6.5×10^{-5}
n=50	VI*	0.0196	0.0172	0.0118	0.0070	0.0024	6.8×10^{-5}	1.4×10^{-4}
	VII*	0.0322	0.0236	0.0063	0.0014	0.0012	4.8×10^{-5}	1.5×10^{-5}
	VIII	0.0172	0.0121	0.0043	0.0020	0.0013	3.2×10^{-5}	5.4×10^{-5}
n=100	VI*	0.0105	0.0094	0.0070	0.0047	0.0015	5.9×10^{-5}	1.0×10^{-4}
	VII*	0.0280	0.0184	0.0025	0.0006	0.0007	3.8×10^{-5}	9.4×10^{-6}
	VIII	0.0079	0.0054	0.0023	0.0019	0.0010	1.8×10^{-5}	5.5×10^{-5}
n=200	VI*	0.0137	0.0125	0.0046	0.0031	0.0010	5.8×10^{-5}	8.2×10^{-5}
	VII*	0.0217	0.0134	0.0013	0.0002	0.0005	2.9×10^{-5}	6.7×10^{-6}
	VIII	0.0044	0.0031	0.0018	0.0016	0.0008	1.8×10^{-5}	4.9×10^{-5}
n=300	VI*	0.0194	0.0118	0.0011	0.0001	0.0004	2.5×10^{-5}	5.5×10^{-6}
	VII*	0.0420	0.0167	0.0012	0.0006	0.0003	1.5×10^{-5}	6.1×10^{-6}
	VIII	0.0034	0.0024	0.0015	0.0013	0.0006	1.9×10^{-5}	4.3×10^{-5}
n=500	VI*	0.0038	0.0034	0.0024	0.0016	0.0005	4.7×10^{-5}	5.1×10^{-5}
	VII*	0.0146	0.0088	0.0009	0.0001	0.0003	2.0×10^{-5}	4.4×10^{-6}
	VIII	0.0025	0.0019	0.0012	0.0010	0.0004	2.1×10^{-5}	3.5×10^{-5}

VI*-Corrected Gamma(F), VII*-Corrected Gamma(G), VIII-Combination

Table 4.24: Simulated MSE for χ^2_6 Distribution

Sample Size	Estimator	x						
		0	0.1	1	4	6	10	20
n=30	VI*	0.0015	0.0021	0.0020	0.0012	0.0012	7.9×10^{-5}	4.2×10^{-5}
	VII*	0.0000	3.6×10^{-7}	0.0058	0.0008	0.0006	0.0001	4.1×10^{-6}
	VIII	0.0015	0.0021	0.0021	0.0011	0.0012	7.9×10^{-5}	4.2×10^{-5}
n=50	VI*	0.0006	0.0015	0.0016	0.0008	0.0009	5.3×10^{-5}	3.2×10^{-5}
	VII*	0.0000	4.3×10^{-6}	0.0037	0.0004	0.0004	7.9×10^{-5}	3.0×10^{-5}
	VIII	0.0006	0.0015	0.0017	0.0008	0.0009	5.3×10^{-5}	2.8×10^{-5}
n=100	VI*	0.0001	0.0014	0.0012	0.0006	0.0006	2.8×10^{-5}	2.2×10^{-5}
	VII*	0.0000	0.0022	0.0023	0.0002	0.0002	4.9×10^{-5}	2.0×10^{-6}
	VIII	0.0001	0.0014	0.0012	0.0006	0.0006	2.8×10^{-5}	2.0×10^{-5}
n=200	VI*	2.0×10^{-5}	0.0005	0.0007	0.0003	0.0003	1.5×10^{-5}	1.2×10^{-5}
	VII*	0.0000	0.0007	0.0012	0.0001	0.0001	3.0×10^{-5}	1.4×10^{-6}
	VIII	1.9×10^{-5}	0.0005	0.0008	0.0003	0.0003	1.4×10^{-5}	1.1×10^{-5}
n=300	VI*	1.3×10^{-5}	0.0004	0.0005	0.0002	0.0002	1.1×10^{-5}	8.8×10^{-6}
	VII*	0.0000	0.0011	0.0008	0.0001	0.0001	2.3×10^{-5}	1.0×10^{-6}
	VIII	1.3×10^{-5}	0.0004	0.0006	0.0002	0.0002	1.0×10^{-5}	8.0×10^{-6}
n=500	VI*	3.9×10^{-5}	0.0002	0.0004	0.0001	0.0001	7.5×10^{-6}	5.9×10^{-6}
	VII*	0.0000	0.0006	0.0006	8.1×10^{-5}	9.9×10^{-5}	1.6×10^{-5}	7.7×10^{-7}
	VIII	4.2×10^{-5}	0.0002	0.0004	0.0001	0.0001	7.2×10^{-6}	5.4×10^{-6}

VI*-Corrected Gamma(F), VII*-Corrected Gamma(G), VIII-Combination

Table 4.25: Simulated MSE for Lognormal Distribution

Sample Size	Estimator	x						
		0	0.1	e^{-1}	0.5	1	5	8
n=30	VI*	0.0090	0.1546	0.0607	0.0451	0.0133	2.0×10^{-4}	9.1×10^{-5}
	VII*	0.0007	0.7321	0.1230	0.0608	0.0121	8.2×10^{-5}	9.4×10^{-6}
	VIII	0.0097	0.1548	0.0607	0.0447	0.0134	2.0×10^{-4}	8.9×10^{-5}
n=50	VI*	0.0035	0.1349	0.0446	0.0320	0.0111	1.7×10^{-4}	6.6×10^{-5}
	VII*	0.0000	0.5482	0.0810	0.0386	0.0080	4.7×10^{-5}	4.9×10^{-6}
	VIII	0.0035	0.1349	0.0445	0.0318	0.0112	1.6×10^{-4}	6.5×10^{-5}
n=100	VI*	0.0033	0.1011	0.0261	0.0218	0.0086	1.1×10^{-4}	4.7×10^{-5}
	VII*	0.0000	0.3237	0.0417	0.0202	0.0044	1.9×10^{-5}	2.1×10^{-6}
	VIII	0.0033	0.1011	0.0259	0.0218	0.0086	1.1×10^{-4}	4.7×10^{-5}
n=200	VI*	0.0058	0.0717	0.0167	0.0153	0.0058	9.2×10^{-5}	3.3×10^{-5}
	VII*	0.0015	0.1780	0.0210	0.0102	0.0026	1.0×10^{-5}	1.1×10^{-6}
	VIII	0.0073	0.0717	0.0165	0.0152	0.0057	9.2×10^{-5}	3.3×10^{-5}
n=300	VI*	0.0033	0.1011	0.0261	0.0218	0.0086	1.1×10^{-4}	4.7×10^{-5}
	VII*	0.0013	0.1330	0.0152	0.0074	0.0019	6.4×10^{-5}	8.2×10^{-6}
	VIII	0.0067	0.0579	0.0124	0.0119	0.0042	7.5×10^{-5}	2.4×10^{-5}
n=500	VI*	0.0054	0.0577	0.0125	0.0119	0.0042	7.5×10^{-5}	2.4×10^{-5}
	VII*	0.0018	0.0902	0.0098	0.0048	0.0012	4.5×10^{-6}	6.2×10^{-7}
	VIII	0.0074	0.0450	0.0094	0.00092	0.0030	5.8×10^{-5}	1.7×10^{-5}

VI*-Corrected Gamma(F), VII*-Corrected Gamma(G), VIII-Combination

Table 4.26: Simulated MSE for Weibull Distribution

Sample Size	Estimator	x						
		0	0.1	0.5	$1/\sqrt{2}$	1	2	3
n=30	VI*	0.0019	0.1049	0.0287	0.0600	0.0682	0.0053	0.0022
	VII*	0.0000	0.2852	0.0514	0.0293	0.0336	0.0011	1.8×10^{-4}
	VIII	0.0014	0.1052	0.0315	0.0592	0.0667	0.0043	0.0019
n=50	VI*	1.1×10^{-6}	0.0763	0.0195	0.0475	0.0560	0.0048	0.0018
	VII*	0.0000	0.1865	0.0353	0.0194	0.0251	0.0008	1.4×10^{-4}
	VIII	1.1×10^{-6}	0.0763	0.0219	0.0465	0.0544	0.0040	0.0016
n=100	VI*	0.0003	0.0577	0.0127	0.0338	0.0425	0.0041	0.0013
	VII*	0.0033	0.1282	0.0215	0.0107	0.0180	0.0005	1.0×10^{-4}
	VIII	0.0034	0.0591	0.0136	0.0322	0.0409	0.0036	0.0012
n=200	VI*	0.0004	0.0460	0.0082	0.0181	0.02460	0.0032	7.5×10^{-4}
	VII*	0.0106	0.0845	0.0135	0.0070	0.0121	0.0004	7.5×10^{-5}
	VIII	0.0186	0.0116	0.0068	0.0045	0.0034	0.0024	7.0×10^{-4}
n=300	VI*	3.3×10^{-7}	0.0400	0.0062	0.0144	0.0197	0.0026	5.7×10^{-4}
	VII*	0.0158	0.0668	0.0103	0.0061	0.0102	0.0003	6.6×10^{-5}
	VIII	0.0158	0.0471	0.0066	0.0145	0.0199	0.0023	5.0×10^{-4}
n=500	VI*	1.6×10^{-9}	0.0311	0.0042	0.0068	0.0101	0.0019	2.3×10^{-4}
	VII*	0.0283	0.0552	0.0078	0.0059	0.0089	0.0002	6.5×10^{-5}
	VIII	0.0283	0.0435	0.0055	0.0134	0.0180	0.0019	4.7×10^{-4}

VI*-Corrected Gamma(F), VII*-Corrected Gamma(G), VIII-Combination

Table 4.27: Simulated MSE for Mixtures of Two Exponential Distributions with $\pi = 0.4$, $\theta_1 = 2$ and $\theta_2 = 1$

Sample Size	Estimator	x						
		0	0.1	0.5	1	2	5	10
n=30	VI*	0.0652	0.0549	0.0271	0.0098	0.0006	3.4×10^{-4}	1.1×10^{-4}
	VII*	0.0696	0.0539	0.0070	0.0065	0.0009	5.4×10^{-5}	1.4×10^{-5}
	VIII	0.0426	0.0335	0.0098	0.0044	0.0007	0.0001	3.8×10^{-5}
n=50	VI*	0.0489	0.0414	0.0187	0.0066	0.0004	2.9×10^{-4}	7.7×10^{-5}
	VII*	0.0500	0.0336	0.0032	0.0030	0.0007	3.4×10^{-5}	9.2×10^{-6}
	VIII	0.0234	0.0181	0.0069	0.0033	0.0004	1.0×10^{-4}	3.3×10^{-5}
n=100	VI*	0.0238	0.0200	0.0101	0.0036	0.0001	2.0×10^{-4}	4.4×10^{-5}
	VII*	0.0647	0.0317	0.0016	0.0014	0.0006	1.4×10^{-5}	4.4×10^{-6}
	VIII	0.0128	0.0089	0.0045	0.0023	0.0002	1.0×10^{-4}	2.5×10^{-5}
n=200	VI*	0.0125	0.0105	0.0055	0.0020	0.0001	1.3×10^{-4}	2.4×10^{-5}
	VII*	0.0870	0.0380	0.0011	0.0010	0.0007	7.1×10^{-6}	1.7×10^{-6}
	VIII	0.0076	0.0051	0.0031	0.0017	0.0001	8.1×10^{-5}	1.6×10^{-5}
n=300	VI*	0.0092	0.0075	0.0038	0.0014	7.6×10^{-5}	0.0001	1.5×10^{-5}
	VII*	0.0806	0.0353	0.0011	0.0007	0.0006	6.2×10^{-6}	1.0×10^{-6}
	VIII	0.0060	0.0041	0.0023	0.0012	0.0001	6.5×10^{-5}	1.1×10^{-5}
n=500	VI*	0.0074	0.0047	0.0025	0.0010	5.0×10^{-5}	7.1×10^{-5}	1.0×10^{-5}
	VII*	0.0698	0.0311	0.0010	0.0006	0.0006	6.1×10^{-6}	7.0×10^{-7}
	VIII	0.0038	0.0027	0.0016	0.0009	8.6×10^{-5}	4.8×10^{-5}	8.1×10^{-6}

VI*-Corrected Gamma(F), VII*-Corrected Gamma(G), VIII-Combination

From the simulation results given in above tables, we can see that the combination estimator has the merits of two corrected Gamma estimators at the same time. The combination estimator performs very well for χ_2^2 and mixtures of two exponential distributions. Table 4.22 shows that the $MISE$s decrease obviously. At the same time, if we look into the table of MSEs, we will find that MSEs at each point are improved in a certain extent as well. For χ_6^2 the two estimators perform very well separately and $MISE$ may be very close to the lowest bound. So, the $MISE$s of combination estimator are not improved very much. For Lognormal and Weibull distribution, the corrected gamma estimator based on G_n is little worse than gamma estimator on F_n. The combination estimator has slightly better $MISE$s than gamma estimator on F_n. This seems that the combination estimator will choose the best automatically for us. So, the recently introduced parameter a, which combines the two gamma estimators, seems to improve the performance of combination estimator and make the combination estimator have the goodness of two gamma estimators.

Chapter 5

Smooth Estimators of Some Functionals of the Distribution Function

5.1 Introduction

Survival analysis is a branch of statistics. In engineering, economics or sociology, it is called reliability theory. In survival analysis, cumulative hazard function

$$H(x) = -\log(S(x)) \tag{5.1}$$

where survival function $S(x) = 1 - F(x)$ and hazard function

$$h(x) = \frac{f(x)}{S(x)} \tag{5.2}$$

occupy an important position. They have many applications in engineering, industrial reliability, biomedical science, economic, life insurance and so on. In survival analysis, mean residual life (MRL) function

$$m(x) = E(X - x|X > x) \tag{5.3}$$

also has some important applications [see Abdous and Berred (2005)]. In some situations, it is more useful than hazard function [see Calabria and Pulcini (1987)].

In this chapter, we will propose some smooth estimators of cumulative hazard, hazard, and MRL functions using Hille's lemma in Poisson weights and generalized version. In Section 5.2, we will first study the estimators of hazard function theoretically, presenting some properties of the proposed estimators, such as strong consistency and asymptotic normality. These properties shows the behaviors of estimators with infinite samples. In order to show the performances of these estimators under finite samples, numerical results of a simulation study are presented as well. The comparison of different estimators is carried out based on MSE. In Section 5.3, we will propose three smooth estimators of MRL function and investigate their asymptotic properties. At the same time, results of the simulation study are given as well.

5.2 Smooth Estimators of Hazard Function

5.2.1 Estimators with Poisson Weights

Define the estimator of survival function as

$$\widetilde{S}_n(x) = 1 - \widetilde{F}_n(x).$$

It is easy to see that the smooth estimator of survival function $\widetilde{S}_n(x)$ has the same asymptotic properties as the smooth estimator of distribution function $\widetilde{F}_n(x)$. Taking advantage of the relationship between cumulative hazard, hazard and survival, density function, a natural thought is that, using

$$\widetilde{H}_n(x) = -\log \widetilde{S}_n(x) \tag{5.4}$$

and

$$\tilde{h}_n(x) = \tilde{f}_n(x)/\widetilde{S}_n(x) \tag{5.5}$$

to estimate $H(x)$ and $h(x)$ respectively.

5.2.1.1 Asymptotic Property of $\widetilde{H}_n(x)$ and $\tilde{h}_n(x)$

Note that, because of the strong convergence of $\widetilde{S}_n(x)$ and $\tilde{f}_n(x)$, if $S(x) \neq 0$, we have

$$\widetilde{H}_n(x) = H(x) + \frac{1}{S(x)}(\widetilde{S}_n(x) - S(x)) + o(\widetilde{S}_n(x) - S(x)) \;\; a.s. \tag{5.6}$$

and

$$\begin{aligned}\tilde{h}_n(x) &= h(x) + \frac{1}{S(x)}(\tilde{f}_n(x) - f(x)) - \frac{f(x)}{S^2(x)}(\widetilde{S}_n(x) - S(x)) \\ &\quad + o\big(\frac{1}{S(x)}(\tilde{f}_n(x) - f(x)) - \frac{f(x)}{S^2(x)}(\widetilde{S}_n(x) - S(x))\big) \;\; a.s. \end{aligned} \tag{5.7}$$

By (5.6) and (5.7), we can see that the strong convergence of $\widetilde{S}_n(x)$ and $\tilde{f}_n(x)$ leads to the the strong convergence of $\widetilde{H}_n(x)$ and $\tilde{h}_n(x)$. So we have following theorem.

Thoerem 5.1 *Under the same assumptions on $f(x)$ and $f'(x)$ in Theorem 2.4, if $\lambda_n = O(n^\alpha)$ and $0 < \alpha < 1$, $E(X_1^{-2}) < \infty$ and $\mathcal{C} \subset \mathbf{R}^+$ is a compact set such that when $x \in \mathcal{C}$, $S(x) \neq 0$, then, as $n \to \infty$, we have*

$$\|\widetilde{H}_n(x) - H(x)\|_{\mathcal{C}} = \sup_{t\in\mathcal{C}}|\widetilde{H}_n(x) - H(x)| \xrightarrow{a.s.} 0 \tag{5.8}$$

and

$$\|\tilde{h}_n(x) - h(x)\|_{\mathcal{C}} = \sup_{x\in\mathcal{C}}|\tilde{h}_n(x) - h(x)| \xrightarrow{a.s.} 0 \tag{5.9}$$

From (5.6), we also note the weak convergence of $\widetilde{H}_n(x)$ led by $\widetilde{S}_n(x)$.

Thoerem 5.2 *Under the same assumptions on $f(x)$ and $f'(x)$ in Theorem 2.3, if $E(X_1^{-2}) < \infty$, $n^{-1}\lambda_n \to 0$ and $\mathcal{C} \subset \mathbf{R}^+$ is a compact set such that when $t \in \mathcal{C}$, $S(x) \neq 0$, then, as $n \to \infty$, we have*

$$\sqrt{n}(\widetilde{H}_n(x) - H(x)) \xrightarrow{\mathcal{D}} N(0, \delta^{'}(x)) \tag{5.10}$$

where $\delta'(x) = \dfrac{\delta^2(x)}{S^2(x)}$ and $\delta^2(x)$ is defined same as in Theorem 2.3.

Now we suppose $f'(x)$ satisfies the Lipschitz order α condition (2.5). Under this assumption, we can write

$$S(k/\lambda_n) - S(x) = -f(x)(k/\lambda_n - x) - \frac{f'(x)}{2}(k/\lambda_n - x)^2 + O\big([(k/\lambda_n - x)^2]^{1+\alpha}\big) \tag{5.11}$$

Furthermore, using (5.11), we can also write

$$\widetilde{S}_n(x) - S_n(x) = T_n'(x) - \frac{f'(x)}{2\lambda_n} + O(\lambda_n^{-1-\alpha}) \tag{5.12}$$

where

$$T_n'(x) = \sum_{k\geq 0} p_k(x\lambda_n)\left[S_n(\lambda_n/k) - S_n(x) - S(\lambda_n/k) + S(x)\right].$$

Following along the lines of the proof of Theorem 3.2 in Chaubey and Sen (1996) using Lemma 2.1 with $b_n = \lambda_n^{-\frac{1}{2}}(\log n)^{\frac{1+\theta}{2}}$, we can show that

$$\sup_{x\geq 0}|T_n'(x)| = O(\lambda_n^{-1/4}n^{-1/2}(\log n)^{1+\theta}).$$

Then the variance $V(T_n'(x)) \leq O(\lambda_n^{-1/2}n^{-1}(\log n)^{1+\theta})$). By (2.30), (5.7) and (5.12), we have

$$\begin{aligned} \tilde{h}_n(x) - h(x) \quad &\sim \quad \frac{f'(x)}{2\lambda_n S^2(x)}(S(x) + f(x)) + \frac{T_{n2}(x)}{S(x)} && (5.13)\\ &+\frac{f(x)}{S^2(x)}[T_n'(x) + S_n(x) - S(x)] && (5.14)\end{aligned}$$

where $T_{n2}(x)$ is defined as in (2.26). Since the variance of (5.14) does not exceed $O(n^{-1})$ and covariance with $T_{n2}(x)$ not exceed $O(\lambda_n^{1/4}n^{-1})$, the order of variance of $\tilde{f}_n(x)$ is determined by the order of variance of $T_{n2}(x)/S(x)$. So we have

$$V(\tilde{h}_n(x)) \approx \frac{V(T_{n2}(x))}{S^2(x)} \approx \frac{\mu}{2}(\pi x^3)^{-1/2}\frac{f(x)}{S^2(x)}(\lambda_n^{1/2}/n) \tag{5.15}$$

and

$$Cov[\tilde{h}_n(s), \tilde{h}_n(x)] \approx \frac{Cov[T_{n2}(s), T_{n2}(x)]}{S(s)S(x)} = O(\frac{1}{n}) \tag{5.16}$$

From the previous analysis about $T_{n2}(x)$, we can obtain the following theorem.

Thoerem 5.3 *If $\lambda_n = O(n^{2/5})$(nonstochatic) and (2.5) holds, and the set $\mathcal{C} \subset \mathbf{R}^+$ is a compact set such that when $x \in \mathcal{C}$, $S(x) \neq 0$, then, as $n \to \infty$,*

$$\left\{\left(n^{2/5}[\tilde{h}_n(x) - h(x)] - \frac{(S(x) + xf(x))}{2\delta^2 S^2(x)}f'(x)\right), x \in \mathcal{C}\right\} \xrightarrow{\mathcal{D}} \textit{ Gaussian process}$$

with covariance function $\gamma_x^2\delta_{sx}$ where $\gamma_x^2 = \frac{\mu}{2}(\pi x^3)^{-1/2}\frac{f(x)}{S^2(x)}\delta$, $\delta_{sx} = 0$ for $s \neq x$ and 1 for $s = x$ and $\delta = \lim_{n\to\infty}(n^{-1/5}\lambda_n^{1/2})$

5.2.1.2 MSE

Similar to density function, we have

$$MSE(\tilde{h}_n(x)) \approx \lambda_n^{-2}\Big[\frac{f'(x)(S(x) + f(x))}{2S^2(x)}\Big]^2 + \frac{\mu\lambda_n^{1/2}}{2n}(\pi x^3)^{-1/2}\frac{f(x)}{S^2(x)} \tag{5.17}$$

5.2.2 Estimator with Asymmetric Kernels

Using the definition of hazard function $h(x) = f(x)/S(x)$, a natural smooth estimator of hazard function with asymmetric weights is given by

$$\tilde{h}_n^*(x) = \frac{\tilde{f}_n^*(x)}{\int_x^\infty \tilde{f}_n^*(t)dt}. \tag{5.18}$$

5.2.2.1 Asymptotic Properties of $\tilde{h}_n^*(x)$

Using Theorem 2.6 and 2.8, it is easy to obtain the following theorem regarding strong convergence of $\tilde{h}_n^*(x)$.

Thoerem 5.4 *Under the assumption of Theorem 2.6 and 2.8, for a compact set $\mathcal{C} \subset$ $\mathbf{R}^+$ such that when $x \in \mathcal{C}$, $S(x) \neq 0$, we have*

$$\|\tilde{h}_n^*(x) - h(x)\|_{\mathcal{C}} = \sup_{x\in\mathcal{C}} |\tilde{h}_n(x) - h(x)| \xrightarrow{a.s.} 0$$

Using the Taylor expansion of (5.18)

$$\tilde{h}_n^*(x) \approx h(x) + \frac{1}{S(x)}(\tilde{f}_n^*(x) - f(x)) - \frac{f(x)}{S^2(x)}(\widetilde{S}_n^+(x) - S(x)) \tag{5.19}$$

where $\widetilde{S}_n^+(x) = 1 - \widetilde{F}_n^+(x)$, we can show the following theorem regard weak convergence of $\tilde{h}_n^*(x)$.

Thoerem 5.5 *Under the assumptions of Theorem 2.9, we have*

$$\sqrt{nv_n}(\tilde{h}_n^*(x) - h(x)) \rightarrow N\big(0, I_2(q)\frac{\mu f(x)}{x^2 S^2(x)}\big), \textit{for } x > 0.$$

5.2.2.2 MSE

According to the proofs of Theorem 2.7 and 2.9, we have

$$\begin{aligned} Bias(\tilde{h}_n^*(x)) &= [(xv_n^2 + \varepsilon_n)f'(x) + \frac{x^2}{2}f''(x)v_n^2 + \epsilon_n f(0)f(x)]/S(x) \\ &\quad - \frac{x^2 f'(x) f(x)}{2S^2(x)} v_n^2 + o(v_n^2 + \varepsilon_n). \end{aligned} \tag{5.20}$$

So

$$\begin{aligned} MSE(\tilde{h}_n^*(x)) &= \Big[\frac{(xv_n^2 + \varepsilon_n)f'(x) + \frac{x^2}{2}f''(x)v_n^2 + \epsilon_n f(0)f(x)}{S(x)} \\ &\quad - \frac{x^2 f'(x) f(x)}{2S^2(x)} v_n^2\Big]^2 + \frac{I_2(q)\mu f(x)}{nv_n(x+\epsilon_n)^2 S^2(x)} + o(v_n^2 + \varepsilon_n) \end{aligned} \tag{5.21}$$

5.2.3 Numerical Comparison

In this section, we compare the two proposed smooth hazard function estimators through the simulation for the following standard distributions.

(i). Chi-Square Distribution

$$f(x) = \frac{1}{2^{\frac{\alpha}{2}}\Gamma(\frac{\alpha}{2})} x^{\frac{\alpha}{2}-1} \exp(-x/2) I\{x > 0\}$$

(ii). Lognormal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi x}} \exp\{-(\log x - \mu)^2/2\} I\{x > 0\}$$

(iii). Gamma Distribution

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

(iv). Weibull Distribution

$$f(x) = \alpha x^{\alpha-1} \exp(-x^{\alpha}) I\{x > 0\}$$

We use the same selection methods of parameters in $\tilde{f}_n^*(x)$ and $\tilde{f}_n(x)$ to choose the parameters in $\tilde{h}_n^*(x)$ and $\tilde{h}_n(x)$ respectively. Under the chosen parameters, we compute

$$SE\left(h_n(x)\right) = \left[h_n(x) - h(x)\right]^2$$

at some fixed points where $h_n(x)$ could be $\tilde{h}_n^*(x)$ or $\tilde{h}_n(x)$. The fixed points are Q_qs ($q = 0, 0.10, 0.25, 0.50, 0.75, 0.90$). [We refer to Q_0 as 0, $Q_{0.50}$ as the median, $Q_{0.25}$, $Q_{0.75}$ as the the first and third quartiles and $Q_{0.10}$, $Q_{0.90}$ as the first and ninth deciles]. We present the simulation results in the following tables.

Table 5.1: Simulated MSE for χ_2^2

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{h}_n^*(x)$	0.03294	0.03429	0.03879	0.05216	0.08115	0.10952
	$\tilde{h}_n(x)$	0.17339	0.10311	0.04620	0.01707	0.01013	0.01008
50	$\tilde{h}_n^*(x)$	0.01967	0.02139	0.02743	0.04363	0.07183	0.10042
	$\tilde{h}_n(x)$	0.18529	0.09515	0.03242	0.01076	0.00711	0.00794
100	$\tilde{h}_n^*(x)$	0.01052	0.01226	0.01733	0.03145	0.05575	0.08239
	$\tilde{h}_n(x)$	0.17330	0.06228	0.01924	0.00674	0.00474	0.00519
200	$\tilde{h}_n^*(x)$	0.00625	0.00767	0.01143	0.02199	0.04123	0.06420
	$\tilde{h}_n(x)$	0.14181	0.03928	0.01050	0.00442	0.00338	0.00367
300	$\tilde{h}_n^*(x)$	0.00489	0.00587	0.00874	0.01732	0.03368	0.05416
	$\tilde{h}_n(x)$	0.12227	0.02746	0.00768	0.00337	0.00276	0.00301
500	$\tilde{h}_n^*(x)$	0.00359	0.00434	0.00642	0.01285	0.02594	0.04338
	$\tilde{h}_n(x)$	0.10463	0.01788	0.00519	0.00247	0.00217	0.00247

Table 5.2: Simulated MSE for χ_6^2

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{h}_n^*(x)$	0.00151	0.00190	0.00191	0.00750	0.02048	0.03697
	$\tilde{h}_n(x)$	0.00046	0.00151	0.00194	0.00347	0.00519	0.00659
50	$\tilde{h}_n^*(x)$	0.00070	0.00138	0.00123	0.00564	0.01653	0.03103
	$\tilde{h}_n(x)$	0.00094	0.00091	0.00137	0.00279	0.00443	0.00591
100	$\tilde{h}_n^*(x)$	0.00025	0.00078	0.00073	0.00397	0.01235	0.02414
	$\tilde{h}_n(x)$	0.00061	0.00048	0.00085	0.00195	0.00321	0.00426
200	$\tilde{h}_n^*(x)$	1.2×10^{-5}	0.00049	0.00049	0.00237	0.00777	0.01612
	$\tilde{h}_n(x)$	0.00033	0.00028	0.00056	0.00131	0.00212	0.00286
300	$\tilde{h}_n^*(x)$	1.4×10^{-5}	0.00040	0.00038	0.00169	0.00578	0.01238
	$\tilde{h}_n(x)$	0.00025	0.00022	0.00045	0.00100	0.00162	0.00221
500	$\tilde{h}_n^*(x)$	4.5×10^{-5}	0.00028	0.00027	0.00120	0.00420	0.00922
	$\tilde{h}_n(x)$	0.00018	0.00015	0.00032	0.00072	0.00115	0.00159

Table 5.3: Simulated MSE for Lognormal(0,1)

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{h}_n^*(x)$	0.05778	0.10281	0.10571	0.09779	0.07517	0.04962
	$\tilde{h}_n(x)$	0.13073	0.10778	0.12258	0.06855	0.02499	0.01121
50	$\tilde{h}_n^*(x)$	0.05294	0.08226	0.07455	0.07455	0.06015	0.03986
	$\tilde{h}_n(x)$	0.13272	0.08194	0.08175	0.03930	0.01381	0.00754
100	$\tilde{h}_n^*(x)$	0.03970	0.05107	0.04717	0.05571	0.04651	0.03092
	$\tilde{h}_n(x)$	0.10536	0.05460	0.04460	0.01850	0.00704	0.00515
200	$\tilde{h}_n^*(x)$	0.02459	0.03009	0.03168	0.04069	0.03391	0.02248
	$\tilde{h}_n(x)$	0.06632	0.03623	0.02693	0.01041	0.00426	0.00331
300	$\tilde{h}_n^*(x)$	0.01697	0.02099	0.02459	0.03212	0.02679	0.01774
	$\tilde{h}_n(x)$	0.04367	0.02659	0.01867	0.00697	0.00337	0.00246
500	$\tilde{h}_n^*(x)$	0.01092	0.01407	0.01879	0.02443	0.02018	0.01341
	$\tilde{h}_n(x)$	0.03619	0.01877	0.01313	0.00495	0.00228	0.00168

Table 5.4: Simulated MSE for $\Gamma(2, 1)$

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{h}_n^*(x)$	0.01239	0.02497	0.01822	0.04918	0.12183	0.20781
	$\tilde{h}_n(x)$	0.03585	0.02587	0.02037	0.02219	0.02522	0.02898
50	$\tilde{h}_n^*(x)$	0.01022	0.01820	0.01230	0.04152	0.10741	0.18688
	$\tilde{h}_n(x)$	0.03017	0.01707	0.01359	0.01588	0.01883	0.02271
100	$\tilde{h}_n^*(x)$	0.00705	0.01045	0.00843	0.03328	0.08966	0.16072
	$\tilde{h}_n(x)$	0.02213	0.00909	0.00837	0.01028	0.01270	0.01577
200	$\tilde{h}_n^*(x)$	0.00443	0.00607	0.00584	0.02573	0.07163	0.13236
	$\tilde{h}_n(x)$	0.01712	0.00520	0.00515	0.00710	0.00911	0.01069
300	$\tilde{h}_n^*(x)$	0.00295	0.00447	0.00458	0.02044	0.05945	0.11310
	$\tilde{h}_n(x)$	0.01684	0.00360	0.00446	0.00713	0.00960	0.01168
500	$\tilde{h}_n^*(x)$	0.00261	0.00346	0.00311	0.01417	0.04338	0.08653
	$\tilde{h}_n(x)$	0.01269	0.00271	0.00295	0.00434	0.00536	0.00668

Table 5.5: Simulated MSE for Weibull(4)

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{h}_n^*(x)$	2.0×10^{-7}	0.10168	0.07904	1.09960	7.13035	22.6114
	$\tilde{h}_n(x)$	0.00353	0.04884	0.10895	1.92473	9.71302	27.2550
50	$\tilde{h}_n^*(x)$	0.00000	0.07516	0.05898	0.64748	4.92443	17.1557
	$\tilde{h}_n(x)$	0.00164	0.04943	0.05817	1.40755	7.87478	23.2188
100	$\tilde{h}_n^*(x)$	0.00000	0.04147	0.04663	0.24293	2.37609	9.86292
	$\tilde{h}_n(x)$	0.00017	0.04085	0.02726	0.93694	5.94016	18.6429
200	$\tilde{h}_n^*(x)$	0.00000	0.02565	0.03284	0.10989	1.20060	5.72528
	$\tilde{h}_n(x)$	0.00000	0.02798	0.01287	0.51189	3.82420	13.1161
300	$\tilde{h}_n^*(x)$	0.00000	0.01980	0.02481	0.07362	0.82973	4.19847
	$\tilde{h}_n(x)$	0.00000	0.02221	0.00908	0.37122	2.99963	10.7567
500	$\tilde{h}_n^*(x)$	0.00000	0.01313	0.01737	0.04767	0.55880	3.01171
	$\tilde{h}_n(x)$	0.00000	0.01538	0.00643	0.23763	2.12960	8.10190

The results of simulation show that $\tilde{h}_n^*(x)$ perform better than $\tilde{h}_n(x)$ between Q_0 and $Q_{0.5}$. This is because the density estimator $\tilde{f}_n^*(x)$ perform much better than $\tilde{f}_n(x)$ near the lower boundary. At the tail, it is the opposite, which leads hazard function estimator $\tilde{h}_n(x)$ is better than $\tilde{h}_n^*(x)$. But the difference is not significant.

5.3 Smooth Estimator of Mean Residual Life

In this section, we propose three smooth estimators of MRL function, two using Poisson weights ,one using gamma kernels. Numerical comparison is given at the end of this section.

5.3.1 Smooth Estimator of MRL with Poisson Weights Based on F_n

If we define

$$S^G(x) = \int_x^\infty g(t)dt \tag{5.22}$$

and

$$S^F(x) = \int_x^\infty f(t)dt \tag{5.23}$$

then the mean residual life function is given by

$$\begin{aligned} m(x) &= \frac{\int_x^\infty tf(t)dt}{\int_x^\infty f(t)dt} - x \\ &= \frac{\mu S^G(x)}{S^F(x)} - x \\ &= M(x) - x \end{aligned} \tag{5.24}$$

The empirical estimators of $S^G(x)$ and $S^F(x)/\mu$ are given by, respectively, $S_n^G(x) = \frac{1}{n}\sum_{i=1}^n I\{X_i > x\}$ and $D_n^F(x) = \frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} I\{X_i > x\}$. Using the discrete version of Hille's lemma, we can obtain the following two smooth estimators

$$\widetilde{S}_n^G(x) = \sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n) \tag{5.25}$$

and

$$\widetilde{D}_n^F(x) = \sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n) \tag{5.26}$$

If we substitute the two smooth estimators for the corresponding functions in (5.24), then we have the following smooth estimator of $m(x)$

$$\tilde{m}_n(x) = \frac{\sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n)}{\sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n)} - x \tag{5.27}$$

Thoerem 5.6 *If $\lambda_n \to \infty$, $E(X_1^{-1}) < \infty$, then for any compact set $\mathcal{C}$ such that $S(x) \neq 0$ when $x \in \mathcal{C}$, as $n \uparrow \infty$,*

$$||\tilde{m}_n(x) - m(x)||_{\mathcal{C}} = \sup_{x\in\mathcal{C}}\{|\tilde{m}_n(x) - m(x)|\} \xrightarrow{a.s.} 0$$

Proof: The proof is straightforward. As $\lambda_n \uparrow \infty$, we have

$$\sup_{x\in\mathcal{C}}\{|\sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n) - \int_x^\infty g(t)dt|\} \xrightarrow{a.s.} 0 \tag{5.28}$$

and

$$\sup_{x\in\mathcal{C}}\{|\sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n) - \frac{1}{\mu}\int_x^\infty f(t)dt|\} \xrightarrow{a.s.} 0 \tag{5.29}$$

By (5.28) and (5.29), we have

$$\begin{aligned} \sup_{x\in\mathcal{C}}\{|\tilde{m}(x) - m(x)|\} &= \sup_{x\in\mathcal{C}}\Big\{\Big|\frac{\sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n)}{\sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n)} - \frac{\int_x^\infty g(t)dt}{\frac{1}{\mu}\int_x^\infty f(t)dt}\Big|\Big\} \\ &\xrightarrow{a.s.} 0 \end{aligned} \tag{5.30}$$

Thoerem 5.7 *If $\sqrt{n}\lambda_n^{-1} \to 0$, $E(X_1^{-2}) < \infty$ and $f(x)$ is absolutely continuous with a bounded derivative $f'(x)$ a.e. on $\mathbf{R}^+$, then*

$$\sqrt{n}(\tilde{m}_n(x) - m(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x))$$

where $\delta^2(x) = \dfrac{\mu}{(S^F(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]$

Proof: Using Taylor expansion, we can approximate $\tilde{m}_n(x)$ by

$$\begin{aligned}
\tilde{m}_n(x) &\approx m(x) + \frac{\mu}{S^F(x)}[\sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n) - \int_x^\infty g(t)dt \\
&\quad - \frac{\mu M(x)}{S^F(x)}[\sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n) - \frac{1}{\mu} S(x)] \\
&= m(x) + \frac{\mu}{S^F(x)}\big[\sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n) \\
&\quad - M(x) \sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n)\big] \qquad (5.31)
\end{aligned}$$

Actually, we can write

$$\sum_{k\geq 0} p_k(x\lambda_n) S_n^G(k/\lambda_n) = \sum_{k\geq 0} (\int_x^\infty \lambda_n p_k(t\lambda_n)dt)[S_n^G(k/\lambda_n) - S_n^G((k+1)/\lambda_n)] \qquad (5.32)$$

and

$$\sum_{k\geq 0} p_k(x\lambda_n) D_n^F(k/\lambda_n) = \sum_{k\geq 0} (\int_x^\infty \lambda_n p_k(t\lambda_n)dt)[D_n^F(k/\lambda_n) - D_n^F((k+1)/\lambda_n)] \qquad (5.33)$$

Let

$$\xi_i = \sum_{k\geq 0} (\int_x^\infty \lambda_n p_k(t\lambda_n)dt)[1 - \frac{M(x)}{X_i}] I\{\frac{k}{\lambda_n} \leq X_i < \frac{k+1}{\lambda_n}\} \qquad (5.34)$$

then

$$\begin{aligned}
E(\xi_i) &= \sum_{k\geq 0} (\int_x^\infty \lambda_n p_k(t\lambda_n)dt)[S^G((k)/\lambda_n) - S^G((k+1)/\lambda_n)] \\
&\quad - \frac{M(x)}{\mu} \sum_{k\geq 0} (\int_x^\infty \lambda_n p_k(t\lambda_n)dt)[S^F((k)/\lambda_n) - S^F((k+1)/\lambda_n)] \\
&= \sum_{k\geq 0} p_k(x\lambda_n) S^G(k/\lambda_n) - \frac{M(x)}{\mu} \sum_{k\geq 0} p_k(x\lambda_n) S^F(k/\lambda_n) \\
&\to S^G(x) - \frac{M(x)}{\mu} S^F(x) = 0 \qquad (5.35)
\end{aligned}$$

So, by (5.31), we have

$$E(\tilde{m}_n(x)) = m(x) + E\xi_i \to m(x) \qquad (5.36)$$

On the other hand, we have

$$
\begin{aligned}
\xi_i^2 &= \sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2[1-\frac{M(x)}{X_i}]^2 I\{\frac{k}{\lambda_n}\leq X_i<\frac{k+1}{\lambda_n}\} \\
&= \sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2 I\{\frac{k}{\lambda_n}\leq X_i<\frac{k+1}{\lambda_n}\} \\
&\quad -2M(x)\sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2\frac{1}{X_i}I\{\frac{k}{\lambda_n}\leq X_i<\frac{k+1}{\lambda_n}\} \\
&\quad +M^2(x)\sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2\frac{1}{X_i^2}I\{\frac{k}{\lambda_n}\leq X_i<\frac{k+1}{\lambda_n}\} \qquad (5.37)
\end{aligned}
$$

then

$$
\begin{aligned}
E(\xi_i^2) &= \sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2[S^G(k/\lambda_n)-S^G((k+1)/\lambda_n)] \\
&\quad -2\frac{M(x)}{\mu}\sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2[S^F(k/\lambda_n)-S^F((k+1)/\lambda_n)] \\
&\quad +\frac{M^2(x)}{\mu}\sum_{k\geq 0}(\int_x^\infty \lambda_n p_k(t\lambda_n)dt)^2\int_{k/\lambda_n}^{(k+1)/\lambda_n}\frac{1}{t}f(t)dt \\
&= T_1-2\frac{M(x)}{\mu}T_2+\frac{M^2(x)}{\mu}T_3 \qquad (5.38)
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
T_1 &= \sum_{k\geq 0}\int_x^\infty \lambda_n p_k(t\lambda_n)dt[S^G(k/\lambda_n)-S^G((k+1)/\lambda_n)] \\
&\quad -\sum_{k\geq 0}(1-\int_x^\infty \lambda_n p_k(t\lambda_n)dt)\int_x^\infty \lambda_n p_k(t\lambda_n)dt[S^G(k/\lambda_n)-S^G((k+1)/\lambda_n)] \\
&= S_1-S_2 \qquad (5.39)
\end{aligned}
$$

Calculating the integration in S_1 and rearranging the sum give us

$$S_1=\sum_{k\geq 0}p_k(x\lambda_n)S^G(k/\lambda_n) \qquad (5.40)$$

By Hiller's theorem, we can claim that, as $\lambda_n\uparrow\infty$,

$$S_1=\sum_{k\geq 0}p_k(x\lambda_n)S^G(k/\lambda_n)\to S^G(x). \qquad (5.41)$$

Next, we will show that $S_2 \to 0$.

Let $\mathbf{N} = \{0, 1, \ldots, n, \ldots\}$ and $b_n = \lambda_n^{-1/2}(\log n)^{\frac{1+\delta}{2}}$ where $\delta > 0$. Denote $\mathbf{N}_x^1 = \{k \big| k/\lambda_n - x < -b_n, k \in \mathbf{N}\}$, $\mathbf{N}_x^2 = \{k \big| |k/\lambda_n - x| \le b_n, k \in \mathbf{N}\}$ and $\mathbf{N}_x^3 = \{k \big| k/\lambda_n - x > b_n, k \in \mathbf{N}\}$.

Let

$$a_k = (1 - \lambda_n \int_x^\infty p_k(t\lambda_n)dt)(\lambda_n \int_x^\infty p_k(t\lambda_n)dt)[S^G(k/\lambda_n) - S^G((k+1)/\lambda_n)], \quad (5.42)$$

then we can write

$$S_2 = \sum_{k \in \mathbf{N}_x^1} a_k + \sum_{k \in \mathbf{N}_x^2} a_k + \sum_{k \in \mathbf{N}_x^3} a_k. \quad (5.43)$$

For any $k \in \mathbf{N}_x^1$, by the proof of Lemma 3.1 of Chaubey and Sen(1996), we can claim that$(\lambda_n \int_x^\infty p_k(t\lambda_n)dt) = [1 - \lambda_n \int_0^x p_k(t\lambda_n)dt] = \sum_0^k p_i(x\lambda_n) < \frac{1}{n}$. Then

$$0 < \sum_{k \in \mathbf{N}_x^1} a_k < \frac{1}{n}(1 + S_1). \quad (5.44)$$

For any $k \in \mathbf{N}_x^3$, by the same lemma above, we can claim that $[1 - \lambda_n \int_x^\infty p_k(t\lambda_n)dt] = \sum_{i \ge k+1}^\infty p_i(x\lambda_n) < \frac{1}{n}$. At the same time, we have $[1 - \lambda_n \int_0^x p_k(t\lambda_n)dt] < 1$. Then

$$0 < \sum_{k \in \mathbf{N}_x^3} a_k < \frac{1}{n} S_1. \quad (5.45)$$

For any $k \in \mathbf{N}_x^2$, by the facts $[1 - \lambda_n \int_x^\infty p_k(t\lambda_n)dt] < 1$, $\lambda_n \int_x^\infty p_k(t\lambda_n)dt < 1$, we have

$$0 < \sum_{k \in \mathbf{N}_x^2} a_k < [S^G(x + b_n) - S^G(x - b_n)]. \quad (5.46)$$

By (5.44), (5.45) and (5.46), we can see that as $\lambda_n \uparrow \infty$, $\sum_{k \in \mathbf{N}_x^i} a_k (i = 1, 2, 3)$ all tend to 0. This means

$$S_2 \to 0 \quad (5.47)$$

By (5.41) and (5.47), we have

$$T_1 \to S^G(x) \quad (5.48)$$

Similarly, we have

$$T_2 \to S^F(x) \tag{5.49}$$

and

$$T_3 \to \int_x^\infty \frac{f(t)}{t} dt \tag{5.50}$$

By (5.35), (5.38), (5.48),(5.49) and (5.50) and a little work of algebra, we have

$$V(\xi_i) \to \delta^2(x). \tag{5.51}$$

Finally, by the fact that $V(\sqrt{n}\tilde{m}_n(x)) = \frac{\mu^2}{S^2(x)} V(\xi_i)$, we can obtain the theorem.

Remark 5.1: Note that an empirical estimator of $m(x) = \frac{\int_x^\infty tf(t)dt}{\int_x^\infty f(t)dt} - x$ is

$$m_n(x) = \frac{\sum_{i=1}^n I\{X_i > x\}}{\sum_{i=1}^n \frac{1}{X_i}\{X_i > x\}} - x. \tag{5.52}$$

Smoothing this empirical estimator will also give us an alternative smooth estimator with Poisson weights as follows.

$$\tilde{m}_n'(x) = \sum_{k \geq 0} p_k(x\lambda_n) m_n(k/\lambda_n) - x. \tag{5.53}$$

Using Taylor's expansion, we can expand (5.53) as (5.31). This means that $\tilde{m}_n'$ has the same asymptotic properties as $\tilde{m}_n$. So for $\tilde{m}_n'$, we can still establish theorems as Theorem 5.6 and 5.7.

Remark 5.2: According to the proof, we have

$$Bias(\tilde{m}_n(x)) = \frac{x^2}{2\mu\lambda_n}[M(x)f'(x) - xf'(x) - f(x)] + o(\lambda_n^{-1})$$

and

$$V(\tilde{m}_n(x)) = \frac{\mu}{n(S^F(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right] + o(\frac{1}{n}).$$

So

$$
\begin{aligned}
MSE(\tilde{m}_n(x)) &= \left\{\frac{x^2}{2\mu\lambda_n}[M(x)f'(x) - xf'(x) - f(x)]\right\}^2 \\
&\quad + \frac{\mu}{n(S^F(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right] + o(\frac{1}{n} + \lambda_n^{-2})
\end{aligned}
\tag{5.54}
$$

5.3.2 Smooth Estimator of MRL with Poisson Weights Based on G_n

Note that, for LB data, the mean residual life function can be defined as

$$
m(x) = \frac{\int_x^\infty yf(y)dy}{S(x)} - x \tag{5.55}
$$

where $S(x) = \int_x^\infty f(t)dt$. Using the smooth estimators based on G_n with Poisson weights to replace the corresponding functions in (5.55) gives us the following smooth estimator of $m(x)$

$$
\hat{m}_n(x) = \frac{1 - \sum_{k\geq 1} p_k(x\lambda_n) G_n\left(\frac{k-1}{\lambda_n}\right)}{\lambda_n \sum_{k\geq 1} G_n\left(\frac{k}{\lambda_n}\right)\left[\frac{P_{k-1}(x\lambda_n)}{k} - \frac{P_k(x\lambda_n)}{k+1}\right]} - x. \tag{5.56}
$$

where

$$
P_{k-1}(\lambda_n x) = \frac{1}{\Gamma(k)}\int_{\lambda_n x}^\infty e^{-y}y^{k-1}dy = \sum_{0\leq j<k} p_j(\lambda_n x).
$$

Note that a computational version of (5.56) is given by

$$
\hat{m}_n(x) = \frac{p_0(\lambda_n x) + \sum_{k=1}^N p_k(\lambda_n x) S_n^G\left(\frac{k-1}{\lambda_n}\right)}{\lambda_n \sum_{k=1}^N \frac{P_{k-1}(\lambda_n x)}{k}\left[S_n^G\left(\frac{k-1}{\lambda_n}\right) - S_n^G\left(\frac{k}{\lambda_n}\right)\right]} - x
$$

where $S_n^G(x) = \frac{1}{n}\sum_{i=1}^n I\{X_i > x\}$ and $N = [\lambda_n X_{n:n}] + 1$.

Regarding the strong consistence of $\hat{m}_n(x)$, we have the following theorem.

Thoerem 5.8 *If* $\lambda_n = O(n^\alpha)(0 < \alpha < 1)$, $E(X_1^{-1}) < \infty$ *and* $f(x)$ *is absolutely continuous with a bounded derivative* $f'(x)$ *a.e. on* $\mathbf{R}^+$*, then for any compact set* $\mathcal{C}$ *such that* $S(x) \neq 0$ *when* $x \in \mathcal{C}$*, as* $n \uparrow \infty$*,*

$$||\hat{m}_n(x) - m(x)||_{\mathcal{C}} = \sup_{x\in\mathcal{C}}\{|\hat{m}_n(x) - m(x)|\} \xrightarrow{a.s.} 0.$$

Proof: The proof is straight forward. Uniformly in any compact set $\mathcal{C}$ such that $S(x) \neq 0$ when $x \in \mathcal{C}$, we have

$$1 - \sum_{k\geq 1} p_k(x\lambda_n)G_n\big(\frac{k-1}{\lambda_n}\big) \xrightarrow{a.s.} \int_x^\infty g(y)dy \tag{5.57}$$

and

$$\lambda_n \sum_{k\geq 1} G_n\big(\frac{k}{\lambda_n}\big)\big[\frac{P_{k-1}(x\lambda_n)}{k} - \frac{P_k(x\lambda_n)}{k+1}\big] \xrightarrow{a.s.} \frac{1}{\mu}S(x). \tag{5.58}$$

So

$$\hat{m}_n(x) \xrightarrow{a.s.} \frac{\int_x^\infty g(y)dy}{\frac{1}{\mu}S(x)} - x = \frac{\int_x^\infty yf(y)dy}{S(x)} - x \tag{5.59}$$

uniformly in any compact set $\mathcal{C}$. The proof is complete.

Regarding the weak convergence of $\hat{m}_n(x)$, we have the following theorem.

Thoerem 5.9 *If* $\sqrt{n}\lambda_n^{-1} \to 0$, $E(X_1^{-2}) < \infty$ *and* $f(x)$ *is absolutely continuous with a bounded derivative* $f'(x)$ *a.e. on* $\mathbf{R}^+$*, then, as* $\lambda_n \uparrow \infty$*,*

$$\sqrt{n}(\hat{m}_n(x) - m(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x))$$

where $\delta^2(x) = \dfrac{\mu}{(S(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]$.

Proof: First, we have

$$
\begin{aligned}
\hat{m}_n(x) &\approx m(x) + \frac{\mu}{S(x)}[(1 - \sum_{k\geq 1} p_k(x\lambda_n) G_n(\frac{k-1}{\lambda_n})) - (1 - G(x))] \\
&\quad -\frac{\mu M(x)}{S(x)}[(\lambda_n \sum_{k\geq 1} G_n(\frac{k}{\lambda_n})[\frac{P_{k-1}(x\lambda_n)}{k} - \frac{P_k(x\lambda_n)}{k+1}]) - \frac{1}{\mu} S(x)] \\
&= m(x) + \frac{\mu}{S(x)}\{\sum_{k\geq 1}(\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt)[G_n(\frac{k}{\lambda_n}) - G_n(\frac{k-1}{\lambda_n})] - (1 - G(x))\} \\
&\quad -\frac{\mu M(x)}{S(x)}\{\sum_{k\geq 1}(\int_x^\infty \frac{\lambda_n^2}{k} p_{k-1}(t\lambda_n)dt)[G_n(\frac{k}{\lambda_n}) - G_n(\frac{k-1}{\lambda_n})] - \frac{1}{\mu} S(x)\} \\
&= m(x) + \frac{\mu}{S(x)}\{\sum_{k\geq 1}(\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt)[G_n(\frac{k}{\lambda_n}) - G_n(\frac{k-1}{\lambda_n})]\} \\
&\quad -\frac{\mu M(x)}{S(x)}\{\sum_{k\geq 1}(\int_x^\infty \frac{\lambda_n^2}{k} p_{k-1}(t\lambda_n)dt)[G_n(\frac{k}{\lambda_n}) - G_n(\frac{k-1}{\lambda_n})]\}. \qquad (5.60)
\end{aligned}
$$

Let

$$\xi_i = \sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt - M(x)\int_x^\infty \frac{\lambda_n^2}{k} p_{k-1}(t\lambda_n)dt] I\{\frac{k-1}{\lambda_n} < X_i \leq \frac{k}{\lambda_n}\}. \quad (5.61)$$

Note that, as $\lambda_n \uparrow \infty$,

$$E(\xi_i) \to [(1 - G(x)) - M(x)\frac{S(x)}{\mu}] = 0 \qquad (5.62)$$

and using (3.23), we can show

$$E(\xi_i) = O(\lambda_n^{-1}). \qquad (5.63)$$

So

$$E(\hat{m}_n(x)) = m(x) + O(\lambda_n^{-1}). \qquad (5.64)$$

Moreover,

$$\begin{aligned}
E(\xi_i^2) &= \sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt - M(x)\int_x^\infty \frac{\lambda_n^2}{k}p_{k-1}(t\lambda_n)dt]^2[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&= \sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt]^2[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&\quad -2M(x)\sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt\int_x^\infty \frac{\lambda_n^2}{k}p_{k-1}(t\lambda_n)dt][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&\quad +M^2(x)\sum_{k\geq 1}[\int_x^\infty \frac{\lambda_n^2}{k}p_{k-1}(t\lambda_n)dt]^2[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&= T_{1n} - 2M(x)T_{2n} + M^2(x)T_{3n}. \qquad (5.65)
\end{aligned}$$

$$\begin{aligned}
T_{1n} &= \sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&\quad -\sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt][1-\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&= S_1 - S_2. \qquad (5.66)
\end{aligned}$$

It is obvious that, as $\lambda_n \uparrow \infty$,

$$S_1 \longrightarrow [1 - G(x)]. \qquad (5.67)$$

Using the same method in the proof of (3.54), we can show that $S_2 \longrightarrow 0$. So

$$T_{1n} \longrightarrow [1 - G(x)]. \qquad (5.68)$$

At the same time, we have

$$\begin{aligned}
T_{2n} &= \sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt]\frac{\lambda_n}{k}[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&\quad -\sum_{k\geq 1}[\int_x^\infty \lambda_n p_{k-1}(t\lambda_n)dt\int_0^x \frac{\lambda_n^2}{k}p_{k-1}(t\lambda_n)dt][G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})].
\end{aligned} \qquad (5.69)$$

we can similarly claim that

$$T_{2n} \longrightarrow \int_x^\infty \frac{1}{t}g(t)dt. \qquad (5.70)$$

Since

$$
\begin{aligned}
T_{3n} &= \sum_{k\geq 1}[\lambda_n \int_x^\infty p_{k-1}(t\lambda_n)dt]\frac{\lambda_n^2}{k^2}[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})] \\
&\quad - \sum_{k\geq 1}[\lambda_n \int_x^\infty p_{k-1}(t\lambda_n)dt][\lambda_n \int_0^x p_{k-1}(t\lambda_n)dt]\frac{\lambda_n^2}{k^2}[G(\frac{k}{\lambda_n}) - G(\frac{k-1}{\lambda_n})],
\end{aligned}
\tag{5.71}
$$

we can also claim

$$T_{3n} \to \int_x^\infty \frac{1}{t^2} g(t)dt. \tag{5.72}$$

So

$$E(\xi_i^2) \to \frac{1}{\mu}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]. \tag{5.73}$$

By (5.62) and (5.73), we have

$$V(\xi_i) \to \frac{1}{\mu}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]. \tag{5.74}$$

By (5.60), we can see that $\hat{m}_n(x) = m(x) + \frac{\mu}{S(x)}(\frac{1}{n}\sum_{i=1}^n \xi_i)$, so

$$V(\sqrt{n}\hat{m}_n(x)) = \frac{\mu^2}{S^2(x)}V(\xi_i). \tag{5.75}$$

Then, as $\lambda_n \uparrow \infty$, $V(\sqrt{n}\hat{m}_n(x)) \to \delta^2(x)$. Combining with (5.64), we can establish the theorem. The proof is complete.

Remark 5.3: According to the proof, we have

$$Bias(\hat{m}_n(x)) = \frac{1}{2\mu\lambda_n}\left\{xf(x) + M(x)[f(x) - \int_x^\infty \frac{f(t)}{t}dt]\right\} + o(\lambda_n^{-1})$$

and

$$V(\hat{m}_n(x)) = \frac{\mu}{n(S(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right] + o(n^{-1}).$$

So

$$MSE(\hat{m}_n(x)) = \frac{1}{4\mu^2\lambda_n^2}\left\{xf(x) + M(x)[f(x) - \int_x^\infty \frac{f(t)}{t}dt]\right\}^2 + \frac{\mu}{n(S(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right] + o(n^{-1} + \lambda_n^{-2}) \tag{5.76}$$

5.3.3 Smooth Estimator of MRL with Asymmetric Kernels

If we apply generalized Hille's lemma to smooth $S_n^G(x) = \frac{1}{n}\sum_{i=1}^n I\{X_i > x\}$ and $D_n^F(x) = \frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} I\{X_i > x\}$ and combine them, we can obtain the following smooth estimator of MRL function.

$$\tilde{m}_n^*(x) = \frac{\sum_{i=1}^n Q_{v_n}(\frac{X_i}{x})}{\sum_{i=1}^n \frac{1}{X_i} Q_{v_n}(\frac{X_i}{x})} - x. \tag{5.77}$$

Thoerem 5.10 *If $\lambda_n \to \infty$ and $0 < E(X_1^{-1}) < \infty$, then for any compact set $\mathcal{C}$ such that $S(x) \neq 0$ when $x \in \mathcal{C}$, as $n \uparrow \infty$,*

$$||\tilde{m}_n^*(x) - m(x)||_{\mathcal{C}} = \sup_{x\in\mathcal{C}}\{|\tilde{m}_n^*(x) - m(x)|\} \xrightarrow{a.s.} 0$$

Proof: Under the conditions of the theorem, using the facts $\frac{1}{n}\sum_{i=1}^n I\{X_i > x\} \xrightarrow{a.s.} S^G(x)$ and $\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} I\{X_i > x\} \xrightarrow{a.s.} S^F(x)/\mu$, it is easy to show that

$$\sup_{x\in\mathcal{C}}\{|\frac{1}{n}\sum_{i=1}^n Q_{v_n}(\frac{X_i}{x}) - S^G(x)|\} \xrightarrow{a.s.} 0 \tag{5.78}$$

and

$$\sup_{x\in\mathcal{C}}\{|\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i} Q_{v_n}(\frac{X_i}{x}) - S^F(x)/\mu|\} \xrightarrow{a.s.} 0. \tag{5.79}$$

By (5.78), (5.79) and $m(x) = \frac{S^G(x)}{S^F(x)/\mu} - x$, we can obtain the theorem.

The weak convergence of $\widetilde{F}_n(x)$ is given by the following theorem.

Thoerem 5.11 *If $E(X_1^{-2}) < \infty$, $\sqrt{n}v_n^2 \to 0$, $f(x)$ is absolutely continuous with bounded derivative $f'(x)$, then, as $n \to \infty$,*

$$\sqrt{n}(\tilde{m}_n^*(x) - m(x)) \xrightarrow{\mathcal{D}} N(0, \delta^2(x)) \tag{5.80}$$

where $\delta^2(x) = \dfrac{\mu}{(S^F(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]$.

Proof: Note that

$$\begin{aligned}
\tilde{m}_n^*(x) &\approx m(x) + \frac{\mu}{S^F(x)}\left[\frac{1}{n}\sum_{i=1}^n Q_{v_n}(\frac{X_i}{x}) - S^G(x)\right] \\
&\quad -\frac{\mu M(x)}{S^F(x)}\left[\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}Q_{v_n}(\frac{X_i}{x}) - \frac{1}{\mu}S^F(x)\right] \\
&= m(x) + \frac{\mu}{S^F(x)}\frac{1}{n}\left[\sum_{i=1}^n Q_{v_n}(\frac{X_i}{x}) - M(x)\sum_{i=1}^n \frac{1}{X_i}Q_{v_n}(\frac{X_i}{x})\right] \qquad (5.81)
\end{aligned}$$

Let $\xi_i = Q_{v_n}(\frac{X_i}{x}) - m(x)\frac{1}{X_i}Q_{v_n}(\frac{X_i}{x})$. In order to obtain the theorem, it is sufficient to show that $E(\xi_i) = O(v_n^2)$ and $E(\xi_i^2) \to \frac{1}{\mu}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]$.

Note that, for ξ_i we have

$$\begin{aligned}
E(\xi_i) &= \int_0^\infty Q(t/x)g(t)dt - \frac{M(x)}{\mu}\int_0^\infty Q(t/x)f(t)dt \\
&= T_1(x) - \frac{M(x)}{\mu}T_2(x). \qquad (5.82)
\end{aligned}$$

For $T_1(x)$, using integration by parts, we have

$$\begin{aligned}
T_1(x) &= \int_0^\infty S^G(xy)q_{v_n}(y)dy \\
&\approx \int_0^\infty \left[S^G(x) - xg(x)(y-1) - x^2g'(x)(y-1)^2/2\right] q_{v_n}(y)dy \\
&= S^G(x) + O(v_n^2). \qquad (5.83)
\end{aligned}$$

Similarly, for $T_2(x)$, we have

$$T_2(x) = S^F(x) + O(v_n^2). \tag{5.84}$$

By (5.82) and (5.84),

$$E(\xi_i) = O(v_n^2). \tag{5.85}$$

Furthermore,

$$\begin{aligned} E(\xi_i^2) &= \int_0^\infty Q_{v_n}^2(t/x)g(t)dt - 2\frac{M(x)}{\mu}\int_0^\infty Q_{v_n}^2(t/x)f(t)dt \\ &\quad + \frac{M^2(x)}{\mu}\int_0^\infty Q_{v_n}^2(t/x)\frac{f(t)}{t}dt \\ &= J_1(x) - 2\frac{M(x)}{\mu}J_2(x) + \frac{M^2(x)}{\mu}J_3(x). \end{aligned} \tag{5.86}$$

For $J_1(x)$, we have

$$\begin{aligned} J_1(x) &= 2\int_0^\infty S^G(xy)Q_{v_n}(y)q_{v_n}(y)dy \\ &= \int_0^\infty \left[S^G(x) - xg(x)(y-1) + o(y-1)\right] 2Q_{v_n}(y)q_{v_n}(y)dy \\ &= S^G(x) + xg(x)\int_0^\infty (y-1)2Q_{v_n}(y)q_{v_n}(y)dy \\ &\quad + o(\int_0^\infty (y-1)2Q_{v_n}(y)q_{v_n}(y)dy) \\ &= S^G(x) + J_4(x) + o(J_4(x)). \end{aligned} \tag{5.87}$$

By the fact that $O(|J_4(x)|) \le O(\sqrt{\int_0^\infty (y-1)^2 q_{v_n}(y)dy}) = O(v_n)$, we have, as $v_n \to 0$,

$$J_1(x) \to S^G(x). \tag{5.88}$$

Similarly, we have

$$J_2(x) \to S^F(x) \tag{5.89}$$

and

$$J_3(x) \to \int_x^\infty \frac{f(t)}{t}dt. \tag{5.90}$$

By (5.86), (5.88), (5.89) and (5.90), we have

$$E(\xi_i^2) \to \frac{1}{\mu}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right]. \tag{5.91}$$

By (5.85) ,(5.91) and $V(\tilde{m}_n^*(x)) = \frac{\mu^2}{(S^F(x))^2}V(\xi_i)$, the theorem follows.

Remark 5.4: According to the proof, we have

$$Bias(\tilde{m}_n^*(x)) = \frac{x^2}{2\mu}[M(x)f'(x) - xf'(x) - f(x)]v_n^2 + o(v_n^2)$$

and

$$V(\tilde{m}_n^*(x)) = \frac{\mu}{n(S^F(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right] + o(\frac{1}{n}).$$

So

$$\begin{aligned} MSE(\tilde{m}_n^*(x)) &= \left\{\frac{x^2}{2\mu}[M(x)f'(x) - xf'(x) - f(x)]v_n^2\right\}^2 \\ &+ \frac{\mu}{n(S^F(x))^2}\left[M^2(x)\int_x^\infty \frac{f(t)}{t}dt - \int_x^\infty tf(t)dt\right] + o(\frac{1}{n} + v_n^4) \end{aligned} \tag{5.92}$$

5.3.4 Numerical Comparison

To compare the proposed three MRL function estimators numerically, we simulate for the following distributions with sample sizes $n = 30, 50, 100, 200, 300, 500$.

(i). Chi-Square Distribution

$$f(x) = \frac{1}{2^{\frac{\alpha}{2}}\Gamma(\frac{\alpha}{2})}x^{\frac{\alpha}{2}-1}\exp(-x/2)I\{x > 0\}.$$

(ii). Lognormal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi}x}\exp\{-(\log x - \mu)^2/2\}I\{x > 0\}.$$

(iii). Gamma Distribution

$$f(x) = \frac{1}{\Gamma(\alpha)}x^{\alpha-1}\exp(-x)I\{x > 0\}.$$

(iv). Weibull Distribution

$$f(x) = \alpha x^{\alpha-1}\exp(-x^\alpha)I\{x > 0\}.$$

Note that it is not easy to develop a valid method to choose smoothing parameters like in the simulation for density estimators, since the $SE(m_n(x), m(x))$ defined in (5.93) is usually not integrable on $[0, \infty)$. Here we use the smoothing parameters selected by BCV methods for the density estimators $\tilde{f}_n(x)$, $\hat{f}_n(x)$ and $\tilde{f}_n^*(x)$ as the values of parameters in $\tilde{m}_n(x)$, $\hat{m}_n(x)$ and $\tilde{m}_n^*(x)$ respectively. Under the selected smoothing parameters, we computer

$$SE\left(m_n(x), m(x)\right) = \left[m_n(x) - m(x)\right]^2 \tag{5.93}$$

at points $Q_q(q = 0, 0.1, 0.25, 0.5, 0.75, 0.9)$ where $m_n(.)$ represents MRL function estimator and $m(.)$ the true MRL function . For each point, we obtain 1000 replications and take their average as simulated MSE. We present the results in the Tables from 5.6 to 5.10.

From the results of simulation, we can see that, overall, in the most cases two estimators using Poisson weights perform better than the estimator using gamma kernels. Similar things happen in simulation for density estimators, where although it does not perform very well at the boundary in some cases, density estimators with Poisson weights usually have smaller MSEs at most points. We may conclude that Hille's lemma in Poisson weights provide us a very valid smoothing technique. In most cases, it can give us some very satisfactory smooth estimators. If we look at the MSEs between Q_0 and $Q_{0.5}$, we find that three estimators have comparative MSEs. Specially at the point Q_0, the MSEs are very close. The main difference among these estimators is at the two rear points $Q_{0.75}$ and $Q_{0.9}$. $\hat{m}_n(x)$ perform much better than the two other estimators.

Table 5.6: Simulated MSE for χ^2_2

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{m}_n(x)$	0.28469	0.22589	0.19305	0.27717	0.67733	1.63454
	$\hat{m}_n(x)$	0.30650	0.19798	0.13439	0.11073	0.13713	0.22175
	$\tilde{m}^*_n(x)$	0.28469	0.17852	0.14188	0.36044	2.21863	8.78077
50	$\tilde{m}_n(x)$	0.18729	0.14154	0.12591	0.21236	0.54582	1.30629
	$\hat{m}_n(x)$	0.19665	0.11102	0.07220	0.06603	0.08365	0.14264
	$\tilde{m}^*_n(x)$	0.18729	0.10114	0.08132	0.27625	1.93104	7.98868
100	$\tilde{m}_n(x)$	0.10018	0.07458	0.07356	0.14225	0.38805	0.94850
	$\hat{m}_n(x)$	0.11175	0.05708	0.03816	0.03450	0.04234	0.06987
	$\tilde{m}^*_n(x)$	0.10018	0.04722	0.04147	0.18097	1.43907	6.38956
200	$\tilde{m}_n(x)$	0.06504	0.05121	0.05994	0.12484	0.33662	0.81487
	$\hat{m}_n(x)$	0.08125	0.03778	0.02582	0.02190	0.02317	0.03871
	$\tilde{m}^*_n(x)$	0.06504	0.02343	0.02087	0.11072	1.00699	4.82878
300	$\tilde{m}_n(x)$	0.04821	0.03891	0.04857	0.10582	0.29588	0.73288
	$\hat{m}_n(x)$	0.06917	0.03178	0.02072	0.01532	0.01587	0.02568
	$\tilde{m}^*_n(x)$	0.04821	0.01632	0.01478	0.08323	0.80137	4.01270
500	$\tilde{m}_n(x)$	0.03481	0.02920	0.03877	0.08811	0.24995	0.62073
	$\hat{m}_n(x)$	0.05406	0.02353	0.01486	0.01071	0.01104	0.01602
	$\tilde{m}^*_n(x)$	0.03481	0.00983	0.00897	0.05753	0.59842	3.14847

Table 5.7: Simulated MSE for χ^2_6

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{m}_n(x)$	0.56425	0.36577	0.33011	0.67652	2.25749	5.94781
	$\hat{m}_n(x)$	0.67831	0.45272	0.38014	0.32527	0.35531	0.55924
	$\tilde{m}^*_n(x)$	0.28469	0.17668	0.13708	0.28653	1.63022	6.71116
50	$\tilde{m}_n(x)$	0.33973	0.20905	0.19641	0.52031	1.90966	5.08080
	$\hat{m}_n(x)$	0.39580	0.24668	0.21103	0.19388	0.22311	0.33589
	$\tilde{m}^*_n(x)$	0.18729	0.10039	0.074943	0.20795	1.42166	6.15694
100	$\tilde{m}_n(x)$	0.16392	0.09440	0.09567	0.34788	1.39640	3.75535
	$\hat{m}_n(x)$	0.18459	0.11133	0.09888	0.09647	0.11318	0.17530
	$\tilde{m}^*_n(x)$	0.10018	0.04770	0.04087	0.16578	1.29348	5.76687
200	$\tilde{m}_n(x)$	0.08119	0.04947	0.04988	0.21372	0.90319	2.45378
	$\hat{m}_n(x)$	0.09166	0.05887	0.05206	0.04938	0.06035	0.09507
	$\tilde{m}^*_n(x)$	0.06504	0.02397	0.02293	0.13632	1.17829	5.38536
300	$\tilde{m}_n(x)$	0.05614	0.03425	0.03462	0.16518	0.71140	1.92673
	$\hat{m}_n(x)$	0.06177	0.04021	0.03390	0.03350	0.04183	0.06712
	$\tilde{m}^*_n(x)$	0.04821	0.01673	0.01694	0.11925	1.07117	4.98145
500	$\tilde{m}_n(x)$	0.03570	0.02129	0.02193	0.11576	0.50631	1.37372
	$\hat{m}_n(x)$	0.03883	0.02599	0.02162	0.02052	0.02623	0.03808
	$\tilde{m}^*_n(x)$	0.03481	0.01001	0.01078	0.09527	0.91658	4.41602

Table 5.8: Simulated MSE for $\Gamma(2)$

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{m}_n(x)$	0.11315	0.07632	0.06777	0.13851	0.47183	1.27887
	$\hat{m}_n(x)$	0.12788	0.07910	0.06046	0.05218	0.05906	0.09913
	$\tilde{m}^*_n(x)$	0.11315	0.06888	0.06526	0.18962	0.94600	3.23575
50	$\tilde{m}_n(x)$	0.07117	0.04564	0.04401	0.11070	0.39078	1.04265
	$\hat{m}_n(x)$	0.07562	0.04310	0.03318	0.03045	0.03640	0.06316
	$\tilde{m}^*_n(x)$	0.07117	0.03897	0.04060	0.15237	0.82954	2.91363
100	$\tilde{m}_n(x)$	0.03408	0.02047	0.02094	0.06516	0.24796	0.67109
	$\hat{m}_n(x)$	0.03562	0.02082	0.01622	0.01525	0.01851	0.03250
	$\tilde{m}^*_n(x)$	0.03408	0.01804	0.01995	0.10486	0.66127	2.44700
200	$\tilde{m}_n(x)$	0.01811	0.01045	0.01087	0.04129	0.16803	0.45966
	$\hat{m}_n(x)$	0.01970	0.01113	0.00854	0.00790	0.01028	0.01783
	$\tilde{m}^*_n(x)$	0.01811	0.00929	0.01077	0.07511	0.52142	2.01298
300	$\tilde{m}_n(x)$	0.01200	0.00696	0.00776	0.03189	0.13362	0.36695
	$\hat{m}_n(x)$	0.01248	0.00728	0.00560	0.00526	0.00652	0.01108
	$\tilde{m}^*_n(x)$	0.01200	0.00640	0.00760	0.05958	0.43353	1.72104
500	$\tilde{m}_n(x)$	0.00767	0.00407	0.00479	0.02258	0.09649	0.26412
	$\hat{m}_n(x)$	0.00803	0.00474	0.00374	0.00328	0.00402	0.00661
	$\tilde{m}^*_n(x)$	0.00767	0.00392	0.00445	0.03979	0.31560	1.31721

Table 5.9: Simulated MSE for Lognormal(0)

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{m}_n(x)$	0.16723	0.13876	0.12771	0.22797	0.70043	1.75035
	$\hat{m}_n(x)$	0.19097	0.14539	0.11928	0.13120	0.21252	0.57104
	$\tilde{m}_n^*(x)$	0.16723	0.12604	0.11635	0.25019	1.44315	7.29962
50	$\tilde{m}_n(x)$	0.09739	0.07739	0.07635	0.16257	0.50335	1.19923
	$\hat{m}_n(x)$	0.11168	0.08378	0.06897	0.07780	0.12630	0.34518
	$\tilde{m}_n^*(x)$	0.09739	0.07040	0.06657	0.17498	1.14222	6.03014
100	$\tilde{m}_n(x)$	0.04326	0.03233	0.03326	0.07551	0.21161	0.46069
	$\hat{m}_n(x)$	0.04782	0.03544	0.03016	0.03425	0.06165	0.17379
	$\tilde{m}_n^*(x)$	0.04326	0.03001	0.03093	0.11896	0.90914	4.96370
200	$\tilde{m}_n(x)$	0.02293	0.01550	0.01535	0.03664	0.09948	0.20858
	$\hat{m}_n(x)$	0.02645	0.01818	0.01566	0.01738	0.03235	0.08444
	$\tilde{m}_n^*(x)$	0.02293	0.01465	0.01582	0.08222	0.69572	3.89849
300	$\tilde{m}_n(x)$	0.01425	0.00969	0.00986	0.02257	0.05695	0.11898
	$\hat{m}_n(x)$	0.01765	0.01268	0.01093	0.01231	0.02224	0.05600
	$\tilde{m}_n^*(x)$	0.01425	0.00954	0.01057	0.06365	0.56762	3.25099
500	$\tilde{m}_n(x)$	0.00942	0.00631	0.00682	0.01609	0.04011	0.08261
	$\hat{m}_n(x)$	0.01159	0.00834	0.00699	0.00699	0.01289	0.03329
	$\tilde{m}_n^*(x)$	0.00942	0.00630	0.00743	0.05061	0.45736	2.63676

Table 5.10: Simulated MSE for Weibull(4)

n	Estimator	Quantile					
		$Q_{0.00}$	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$
30	$\tilde{m}_n(x)$	0.00284	0.00177	0.00226	0.00998	0.03402	0.07485
	$\hat{m}_n(x)$	0.00359	0.00401	0.00511	0.00653	0.00763	0.00820
	$\tilde{m}_n^*(x)$	0.00284	0.00154	0.00200	0.00574	0.02180	0.05265
50	$\tilde{m}_n(x)$	0.00160	0.00093	0.00132	0.00729	0.02680	0.06105
	$\hat{m}_n(x)$	0.00195	0.00231	0.00315	0.00426	0.00514	0.00562
	$\tilde{m}_n^*(x)$	0.00160	0.00084	0.00087	0.00354	0.01491	0.03831
100	$\tilde{m}_n(x)$	0.00079	0.00047	0.00063	0.00448	0.01850	0.04456
	$\hat{m}_n(x)$	0.00101	0.00135	0.00192	0.00270	0.00335	0.00371
	$\tilde{m}_n^*(x)$	0.00079	0.00044	0.00036	0.00118	0.00617	0.01858
200	$\tilde{m}_n(x)$	0.00042	0.00024	0.00029	0.00267	0.01221	0.03102
	$\hat{m}_n(x)$	0.00061	0.00103	0.00156	0.00230	0.00293	0.00332
	$\tilde{m}_n^*(x)$	0.00042	0.00022	0.00017	0.00051	0.00299	0.00999
300	$\tilde{m}_n(x)$	0.00029	0.00015	0.00018	0.00193	0.00935	0.02454
	$\hat{m}_n(x)$	0.00047	0.00093	0.00148	0.00222	0.00286	0.00325
	$\tilde{m}_n^*(x)$	0.00029	0.00014	0.00011	0.00032	0.00201	0.00703
500	$\tilde{m}_n(x)$	0.00018	9.9×10^{-5}	0.00011	0.00126	0.00650	0.01779
	$\hat{m}_n(x)$	0.00037	0.00086	0.00141	0.00216	0.00280	0.00318
	$\tilde{m}_n^*(x)$	0.00018	9.1×10^{-5}	7.0×10^{-5}	0.00020	0.00135	0.00493

Chapter 6

Future Research

The methods discussed earlier may be applied to other topics that we plan to investigate in future. Some of these topics are described in detail in the following sections.

6.1 Dependent Data

All the results we have obtained are based on the assumption that the samples are *i.i.d.* random variables. In some practice, we may have some dependent samples. Actually, our results are easy to extend to stationary φ-mixing process.

Definition 6.1 *A stationary stochastic process $\{X_i\}_{i=1}^{\infty}$ is called φ-mixing process, if, for all $B \in \mathfrak{M}_{k+n}^{\infty}$ with probability 1*

$$|P(B|\mathfrak{M}_1^k) - P(B)| \leq \varphi(n) \downarrow 0, \quad \textit{as } n \to \infty \tag{6.1}$$

where $\mathfrak{M}_b^a$ denotes the σ-algebra generated by $X_i (b \leq i \leq a)$.

Now we suppose that $\{X_i\}_{i=1}^{\infty}$ is a stationary φ-mixing process satisfying

$$\sum_n \left(\varphi(n)\right)^{1/2} < \infty \tag{6.2}$$

and

$$EX_i^{-2} < \infty. \tag{6.3}$$

Then we can slightly change the proof of of Lemma 2.3 and establish an almost same lemma for φ-mixing process $\{X_i\}_{i=1}^{\infty}$. Using this lemma, we can obtain similar results to *i.i.d.* case. The conditions of φ-mixing process might be too strong. We can consider some associated sequence with some slightly weak conditions as well. For example, Bagai and Prakasa Rao (1991) investigate strong and weak consistency of empirical function for stationary associated sequence. The dependence of samples is described by the covariances of samples instead of (6.1). This kind of conditions in their paper might be more universal and practical.

6.2 Censored Data

In analyzing times duration, LB data and censored data may emerge at the same time [see Asgharian *et al.* (2002), Uña-Álvarez (2002)]. The presence of censored data is very natural in many application of statistics. Here, we plan to consider random censorship. Suppose that $X_1, \ldots, X_n$ are *i.i.d.* random variables with distribution function $G(x)$. In practice, we may observe

$$Z_i = \min(X_i, Y_i) \text{ and } \delta_i = I\{X_i \leq Y_i\},\ 1 \leq i \leq n,$$

where $\{Y_i\}_{i=1}^{n}$ is another *i.i.d.* sequence with censoring pdf $H(x)$ being independent of the sequence $\{X_i\}_{i=1}^{n}$ as well and δ_i points out whether X_i has been observed or not. Then the well known product-limit estimator of $G(x)$, being nonparametric maximum likelihood estimator as well, proposed by Kaplan and Meier (1958) is given by

$$1 - G_n(x) = \prod_{i=1}^{n} \left[1 - \frac{\delta_{[i:n]}}{n-i+1}\right]^{I\{Z_{i:n} \le x\}} \tag{6.4}$$

where $\{Z_{i:n}\}_{i=1}^n$ are the order statistics of $\{Z_i\}_{i=1}^n$ and $\delta_{[i:n]}$ is the value of δ corresponding to $Z_{i:n}$. Stute and Wang (1993) studied the strong convergence of (6.4). Stute (1995) gave the central limit theorem of (6.4). Combining random censorship with length biased data may result in the following estimator for distribution function $F(x)$ which we are interested in.

$$F_n(x) = \frac{\int_0^x \frac{1}{t} dG_n(t)}{\int_0^\infty \frac{1}{t} dG_n(t)}. \tag{6.5}$$

An alternative form of (6.5) is given by

$$F_n(x) = \frac{\sum_{i=1}^n W_{in} Z_{i:n}^{-1} I\{Z_{i:n} \le x\}}{\sum_{i=1}^n W_{in} Z_{i:n}^{-1}} \tag{6.6}$$

where for $1 \le i \le n$,

$$W_{in} = \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n]}}.$$

Using Hille's lemma to smooth (6.6) will give us a smooth estimator of distribution function. We can obtain smooth estimator of density function by taking the advantage of the derivative of smooth pdf estimator. Furthermore, we can achieve other smooth estimator related to smooth density and distribution estimator.

Length biased data is a special case of biased data by taking the weight function $w(x) = x$ in biased data model

$$f_w(x) = \frac{w(x) f(x)}{\mu_w}.$$

If we are looking for method estimating density with general biased data, a valid method is first to smooth the generalized Cox estimator for distribution function in biased data case

$$F_n(x) = \frac{\sum_{i=1}^n \frac{1}{w(X_i)} I\{X_i \le x\}}{\sum_{i=1}^n \frac{1}{w(X_i)}} \tag{6.7}$$

where $\{X_i\}_{i=1}^n$ are *i.i.d.* random variables or associate sequence satisfying certain dependence conditions with the same weighted density $f_w(x)$. Then take the derivative of smooth estimator of distribution function as density estimator. For randomly censored biased data, the raw estimator of unweighted distribution function $F(x)$ might have the following form.

$$F_n(x) = \frac{\sum_{i=1}^n W_{in} \frac{1}{w(Z_{i:n})} I\{Z_{i:n} \le x\}}{\sum_{i=1}^n W_{in} \frac{1}{w(Z_{i:n})}}. \tag{6.8}$$

6.3 Unknown Weight Function

For now, we have all these discussions based on the assumption that the weight function is known. Lloyd and Jones (2000) gave a nonparametric density estimator for biased data with unknown weight function $w(x) \le 1$. In their article, they treated weight function $w(x)$ as a selection probability that the sample x_i is chosen with probability $w(x_i)$. They obtain two independent samples denoted as S_1 and S_2 from original population with nonrandom size. Each individual x_i belonging to S_1 or S_2 is with a selection probability $w(x_i)$. Then each individual x_i in $S_{11} = S_1 \bigcap S_2$ is with a selection probability $w^2(x_i)$. Using the samples in S_1 or S_2, it is easy to obtain density estimators of weighted density $f_w(x) = \mu_w^{-1} w(x) f(x)$. Since the selection probability in S_{11} is $w^2(x)$, a density estimator of weighted density $f_{w^2}(x) = \mu_{w^2}^{-1} w^2(x) f(x)$ can be built by using the samples in S_{11}. After having the estimators of f_w and f_{w^2}, the estimator of density function $f(x)$ and weight $w(x)$ can be found by the facts $(f_w)^2/f_{w^2} \propto f(x)$ and $f_{w^2}/f_w \propto w(x)$ respectively. However, their density estimators are obtained by the traditional kernel method. If using density estimators proposed in this thesis, we should obtain some better estimator for biased data with unknown weight function.

6.4 Estimation of Other Functionals and Their Integrals

In the area of nonparametric functional estimation, the estimation of derivatives of a density is an active field as well. Singh (1977a) mentioned that estimation of derivatives of a density has many applications, such as estimation of regression curves, estimation of *Fisher Information* and other quantities related to minimum expected loss estimation. Therefore, the estimation of derivatives of a density has drawn a lot of attention in statistical literature. Actually, the estimation of derivatives of a density has almost as long a history as nonparametric density estimation. Bhattacharya (1967) suggested using the pth derivative of traditional kernel density estimator as the estimator of the pth derivative of underlying density and studied their asymptotic properties. These properties were further investigated by Schuster (1969). Also Singh (1977b) studied asymptotic properties of the derivatives of kernel density estimator under some conditions weaker than that in Bhattacharya (1967) and Schuster (1969). Note that the smooth density estimators proposed in this thesis are differentiable. Hence, intuitively we can think of using these derivatives as estimators of the corresponding derivatives of underlying density. Besides the applications of estimators of derivatives mentioned in Singh (1977a), estimators of derivatives are also required for selecting smoothing parameter(s) in our proposed estimators. It might be proper to investigate asymptotic properties of all these procedures.

Another topic related to estimation of derivatives of a density is to estimate the integral involving density derivatives. Cheng (1997) considered estimation of integrated products of density derivatives in general and estimation of integrated squared density derivative in particular. Namely, he considered estimation of the following integral:

$$\theta_{\gamma,\nu} = \int f^{(\gamma)}(x) f^{(\nu)}(x) dx, \tag{6.9}$$

where $f^{(p)}(x)$ represents the pth derivative of density $f(x)$, γ and ν are two nonnegative integers such that $\gamma + \nu$ is an even number. The most direct application of the estimation of integral (6.9) is in bandwidth selection method for nonparametric functional estimators. The plug-in bandwidth selection method for density estimator has such an integral as (6.9) in the special case $\gamma = \nu = 2$ [see Scott and Terrell (1987), Park and Marron (1990)]. Given a better estimator of integral, a better optimal bandwidth can be obtained. So a lot of work regarding the estimation of integral (6.9) is going on. To estimate the integral (6.9) based on biased data, one way is that we plug-in the corresponding derivative estimators into (6.9) directly. An alternative way may be based on local polynomial fitting as proposed by Cheng (1997). The comparison of the two methods is an interesting future project.

References

Abdous, B. and Berred, A. (2005). Mean Residual Life Estimation. *Journal of Statistical Planning and Inference*, **132**, 3-19.

Asgharian, M., M'lan, C. E. and Wolfson, D. B. (2002). Length-Biased Sampling with Right Censoring: An Unconditional Approach. *Journal of the American Statistical Association*, **97**, 201-209.

Babu, G. J. and Singh, K. (1978). On Deviations Between Empirical and Quantile Process for Mixing Random Variables. *Journal of Multivariate Analysis*, **8**, 532-549.

Bagai, I. and Prakasa Rao, B. L. S. (1991). Estimation of the Survavial Function for Stationary Associated Process. *Statistics and Probability Letters*, **12**, 385-391.

Bahadur, R. R. (1966). A Note on Quantiles in Large Samples. *Annals of Mathematical Statistics*, **37**, 577-580.

Bhattacharya, P. K. (1967). Estimation of a Probability Density Function and Its Derivatives. *Sankhyā Ser. A*, **29**, 373-382.

Bhattacharyya, B. B., Franklin, L. A. and Richardson, G. D. (1988). A Comparison of Nonparametric Unweighted and Length-Biased Density Estimation of Fibres. *Communications in Statistics - Theory and Methods*, **17**, 3629-3644.

Bowman, A. W. (1984). An Alternative Method of Cross-Vlidation for the Smoothing of Denstiy Estimates. *Biometrika*, **71**, 353-360.

Calabria, R.and Pulcini, G. (1987). On the Asymptotic Behaviour of Mean Residual Life Function. *Reliability Engineering*, **19**, 165-170.

Chaubey, Y. P. , Sen, A. and Sen, P. K. (2007). A New Smooth Density Estimator for Non-Negtive Random Variables. *Technical Report No. 01/07, Deapartment of Mathematics and Statistics, Concordia University, Montreal.*

Chaubey, Y. P. and Sen, P. K. (1996). On Smooth Estimation of Survival and Density Function. *Statistics and Decision*, **14**, 1-22.

Chaubey, Y. P. and Sen, P. K. (1998). On Smooth Estimation of Hazard and Cumulative Hazard Functions. *In Frontiers of Probability and Statistics (S. P. Mukherjee et al., eds.)*, 91-99. Narosa Publishing House, New Delhi.

Chaubey, Y. P. and Sen, P. K. (2009). On the Selection of Smoothing Parameter in Poisson Smoothing of Histogram Estimator: Computational Aspects. *Pak. J. Statist.*, **25(4)**, 385-401.

Chaubey, Y. P. , Sen, P. K. and Li, J. (2010). Smooth Density Estimation for Length Biased Data. *Journal of the Indian Society of Agricultural Statistics*, **64(2)**, 145-155.

Chen, S. X. (2000). Probability Density Function Estimation Using Gamma Kernels. *Annals of the Institute of Statistical Mathematics*, **52**, 471-480.

Cheng, M. Y. (1997). Boundary Aware Estimators of Integrated Density Derivative Products. *J. R. Statist. Soc. B*, **59**, 191-203.

Chung, K. L. (1974). *A Course in Probability Theory (2nd ed.)*. Academic Press, New York and London.

Cook, R. G. and Martin, F. B. (1974). A Model for Quadrat Sampling with Visibility Bias. *J. Amer. Statist. Assoc.*, **69**, 345-349.

Cox, D. R. (1969). Some Sampling Problems in Technology in *New Developments in Survey Sampling*. 506-527 (N.L. Johnson and H. Smith, eds.) John Wiley, New York.

Eberhardt, L. L. (1978). Transect Methods for Populations Studies. *Journal of Wildlife Management*, **42**, 1-31.

Feller, W. (1966). *An Introduction to Probability Theory and Application II.* John Wiley, New York.

Fisher, R. A. (1934). The Effects of Methods of Ascertianment upon the Estimation of Frequencies. *Annals of Eugenics*, **6**, 13-25.

Jones, M. C. (1991). Kernel Density Estimation for Length Biased Data. *Biometrika*, **78**, 511-519.

Jones, M. C. and Kaunamuni, R. J. (1997). Fourier Series Estimation for Length Biased Data. *Australian Journal of Statistics*, **39**, 57-68.

Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *J. Amer. Statist. Assoc.*, **53**, 457-481.

Klein, R. and Sherman, R. (1997). Estimating New Product Demand from Biased Survey Data. *Journal of Econometrics*, **76**, 53-76.

Kronmal, R. and Tarter, M. E. (1968). The Estimation of Probability Densities and Cumulatives by Fourier Series Methods. *J. Amer. Statist. Assoc.*, **63**, 925-952.

Kulasekera, K. B. and Padgett, W. J. (2006). Bayes Bandwidth Selection in Kernel Density Estimation with Censored Data. *Journal of Nonparametric Statistics*, **18**, 129-143.

Lloyd, J. and Jones, M. C. (2000) Nonparametric Density Estimation from Biased Data with Unknow Biasing Function. *J. Amer. Statist. Assoc.*, **95**, 865-876.

Loève, M. (1977). *Probability Thoery I.* 4th Edition, Springer-Verlag.

Muttlak, H. A. and McDonald, L. L. (1990). Ranked Set Sampling with Size-biased Probability of Selection. *Biometrics,* **46,** 435-446.

Nair, V. N. and Wang, P. C. C. (1989). Maximum Likelihood Estimation under a Successive Sampling Discovery Model. *Technometrics*, **31**, 423-436.

Park, B. U. and Marron, J. S. (1990). Comparison of Data-Driven Bandwidth Selectors. *J. Amer. Statist. Assoc.*, **85**, 66-72.

Patil, G. P. (1984). Studies in Statistical Ecology Involving Weighted Distributions. In *Statistics: Applications and New Directions* . Indian Statistical Institute, 478-503.

Patil, G. P. and Rao, C. R. (1977). The Weighted Distributions: A Survey of Their Applications. In *Application of Statistics*, 383-405(R.R. Krishnaiah, ed.). North-Holland.

Patil, G. P. and Rao, C. R. (1978). Weight Distributions. In *Encyclopaedia of Statistical Sciences*, **9**, 565-571(S. Kotz and N.L. Johnson, eds.). John Wiley.

Rao, C. R. (1965). On Discrete Distributions Arising out of Methods of Ascertainment. In *Classical and Contagious Discrete Distribution*, 320-332(G.P.Patil,ed.). Pergamon Press and Statistical Publishing Society, Calcutta.

Rao, C. R. (1977). A Natural Example of a Weighted Binomial Distribution. *American Statistician*, **31**, 24-26.

Rudemo, M. (1982). Empirical Choice of Histogram and Kernel Density Estimators. *Scandinavian Journal of Statistics*, **9**, 65-78.

Scaillet, O. (2004). Density Estimation Using Inverse and Reciprocal Inverse Gaussian Kernels. *Journal of Nonparametric Statistics*, **16**, 217-226.

Schuster, E. F. (1969). Estimation of a Probability Density Function and Its Derivatives. *Ann. Math. Statis.*, **40**, 1187-1195.

Scott, D. W. and Terrell, G. R. (1987). Biased and Unbiased Cross-Validation in Density Estimation. *J. Amer. Statist. Assoc.*, **82**, 1131-1146.

Sen, P. K. (1984). On Asymptotic Representation of Reduced Quantiles in Sampling from Length-Biased Distribution. *Calcutta Statistical Association Bulletin*, **33**, 59-67.

Sen, P. K. and Ghosh, M. (1971). On Bounded Length Sequential Confidence Intervals Based on One Sample Rank Order Statistic. *Ann. Math. Statist.*, **42**, 189-203.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London.

Simon, R. (1980). Length Biased Sampling in Etiologic Studies. *American Journal of Epidemiology*, **111**, 444-452.

Singh, R. S. (1977a). Applications of Estimators of a Density and its Derivatives to Certain Statistical Problems. *J. R. Statist. Soc. B* , **39**, 357-363.

Singh, R. S. (1977b). Improvement on Some Known Nonparametric Uniformly Consistent Estimator of Derivatives of a Density. *Ann. Statist.*, **5**, 394-400.

Stute, W. (1995). The Central Limit Theorem under Random Censorship. *Ann. Statist.*, **23**, 422-439.

Stute, W. and Wang, J. L. (1993). The Strong Law under Random Censorship. *Ann. Statist.*, **21**, 1591-1607.

Uũa-Álvarez, J. D. (2002). Product-Limit Estimation for Length-biased Censored Data. *Test*, **11**, 109-125.

van der Waart, A. W. (1998). *Asymptotic Statistics.* Campridge University Press, New York.

Vardi, Y. (1982). Nonparametric Estimation in the Presence of Length Bias. *Ann. Statist.*, **10**, 616-620.

Vardi, Y. (1985). Empirical Distributions in Selection Bias Models. *J. Amer. Statist. Assoc.*, **83**, 183-197.

Winter, B. B. and Földes, A. (1988). A Product-Limit Estimator for Use with Length-Biased Data. *Canadian Journal of Statistics*, **16**, 337-355.

Wu, C. O. (1996). Kernel Smoothing of the Nonparametric Maximum Likelihood Estimators for Biased Sampleing Models. *Mathematical Methods in Statistics*, **5**, 275-298.

Wu, C. O. (1997). A Cross-Validation Bandwith Choice for Kernel Density Estimates with Slection Biased Data. *J. Multivariate Annalysis*, **61**, 38-60.

Wu, C. O. and Mao, A. Q. (1996). Minimax Kernel for Density Estimation with Biased Data. *Annals of Institute of Statistical Mathematics*, **48**, 451-467.