TRANSFORMED NONPARAMETRIC FUNCTIONS ESTIMATION

By

THAGENDRA PRASAD TIMSINA

Bachelor of Science
University of Wisconsin
Milwaukee, WI, USA
2006

Masters of Science
University of Texas
Arlington, TX, USA
2008

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2015

TRANSFORMED NONPARAMETRIC FUNCTIONS ESTIMATION

Dissertation Approved:

Dr. Ibharim Ahmad

Dissertation Advisor

Dr. Joshua Habiger

Member

Dr. Lan Zhu

Member

Dr. Balabhaskar Balasundaram

Outside Member

ACKNOWLEDGMENTS

I would like to thank my Advisor Dr. Ibharim Ahmad for his guidence and support. His specialities and insights along with his inspiring behaviour helped me to successfully complete this work. I also express my gratitute to Dr. Lan Zhu, Dr. Balabhaskar (Baski) Balasundaram and Dr. Joshua Habiger for thier positive guidence.

A special thank to my colleague Dr. Ole Forsberg for his continous help in technical issues.

My loads of love and thanks to my wife for tirelessly taking care of calmless baby Ryan Timsina and for her patient and encouragement. Also I like to thank my friend Jhapendra Sapkota for entertaining me during boredorm time.[1]

[1]Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: Thagendra Timsina Date of Degree: May, 2015

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: TRANSFORMED NONPARAMETRIC FUNCTIONS ESTIMATION

Major Field: Statistics

The estimation of an unknown probability density functions of a random variable or its distribution function or a function related to it using standard kernel density estimate is the most popular technique among many density estimation methods. This is due to its favorable features such as it does not assume any functional form, data guide the underlying density and it accurately detects any multimodality present in the target density. Often, the standard kernel chosen has its support on whole Euclidean space. However, in many situations such as in survival, reliability, social and ecological analyses, the random variables have support only on positive half of the real line or on a compact interval and using standard kernel to estimate the density of these random variables assigns positive probabilities outside the support of the target density. Ignoring the probability mass outside the support of random variables will result in erroneous bias. To circumvent this problem, transformed kernel density and distribution functions estimates are proposed. A similar approach is used to estimate the density and distribution functions of data from weighted distribution. These estimates are used to estimate failure rate and regression functions. The asymptotic properties of these estimators are studied including the most crucial bandwidth selection. These new estimators have the same support as the data and preserve the fundamental properties of the random variables. Simulation studies and some real data examples are presented.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

This dissertation focuses on nonparametric density estimations of unbiased and biased sampling. The regression and failure rate functions are also estimated as examples of applications of both unbiased and biased sampling. For the unbiased case, we apply our technique to nonnegative random variables and for the biased case, samples from weighted distribution are considered.

1.1 Nonparametric Kernel Function Estimation

Nonparametric functions estimations are a major field of study in nonparametric statistics. It is an important data analytic tool in preliminary data analysis since it provides a very effective way of showing structure of data. This is specially important when the data structure has multimodality, skewed shape and long or heavy tails since parametric models are inadequate in these situations. The nonparametric estimators do not assume any fixed form of the target functions and depend upon random variables on hand to reach an estimate. The obvious example of function estimation is the density estimation of random variables. The most basic nonparametric example of the density estimate is the histogram. The histogram has several drawbacks such as it does not provide smooth estimate, it depends upon the starting point and number of bins grows exponentially with the number of dimensions in multivariate setting. The kernel density estimator is a class of nonparametric density estimators and has received tremendous attention in the past six decades. It is widely used in theoretical and applied fields, particularly in exploratory data analysis when parametric models are inadequate. It is the most popular technique of density estimations due to its simple assumptions and its accuracy in complex situations.

1.2 Kernel Density Estimation(KDE)

Let $X_1, X_2, ..., X_n$ be nonnegative random variables with density function $f(x)$. The standard kernel estimate(KDE) of $f(x)$ is

$$\hat{f}_n(x) = \frac{1}{na}\sum_{i=1}^{n} K\left(\frac{x - X_i}{a}\right), \quad -\infty < x < \infty \tag{1.1}$$

where a is called bandwidth and K is a known probability density that is symmetric with zero mean and finite variance.

Similarly,

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{K}\left(\frac{x - x_i}{a}\right). \tag{1.2}$$

Where $\mathbb{K}$ is the distribution function for corresponding K.

In many studies such as survival and reliability analysis, ecological and social sciences, data comes from Euclidean half space or from a compact interval. The classical kernel estimate(1.1) when used in estimating the density of nonnegative random variables or variables confined to finite support, however, suffers a major drawback as it assigns positive values outside the support of random variables. Ignoring the positive mass out side the support of the target density results in unnecessary bias and the estimate does not integrate out to unity. It is desirable to have the same support of random variables as the kernel density. To overcome this problem, we use in the current research transformation of nonnegative random variables to variables defined on the whole real line and propose a new kernel-type density estimator of transformed random variables on the entire real line. Then one can re-transform back to obtain the density estimates of the original data preserving fundamental property of random variables. The density function estimate is assessed both globally by Mean Integrated

Square Error (MISE) and locally by Mean Square Error(MSE). The Mean Squared Error is used in assessing distribution function to avoid integrability problems.

1.3 Regression Function Estimation

In many situations, the functional form of regression function is unknown and non-parametric estimation is used to estimate the regression function. Let Y and X be continuously distributed response and independent variables respectively with a joint density $f(y,x)$ and let $f(y|x) = f(x))$ be conditional density of Y given x. The regression function of Y on X is given by

$$m(x) = E(y|X = x) = \frac{\int y f(x,y)dy}{f(x)}.$$

For the nonnegative data, the standard kernel estimate of this function puts weights out side the support of target function. We used proposed transformed kernel density estimates in estimating the regression function.

1.4 Failure Rate Function Estimation

Let $F(x)$ be distribution function of non negative random variable X that represents time to failure of a subject , then the univariate failure rate function is defined as

$$h(x) = -\frac{d}{dx}\log(1 - F(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{\overline{F}(x)}, \quad x \geq 0. \tag{1.3}$$

where $\overline{F}(x) > 0$ is a survival function and is given by

$$\overline{F}(x) = P(X > x) = 1 - \int_0^x f(u)du.$$

Kernel failure rate is then estimated by plugging in the estimates of $f(x)$ and $\overline{F}(x)$ in equation (1.3).

Nonnegative random variables have major role in reliability theory and survival analysis. The density of nonnegative random variables has a support on the positive half of the real line so in estimating reliability functions such as Failure Rate Function, Mean Residual Life and Conditional Survival Function, it is desirable to use the density estimate that has nonnegative support. Thus, for the estimation of failure rate function, the density and distribution estimates of transformed random variables are used. To find the estimate of failure rate function of the original nonnegative random variables, one can re-transform the failure rate function estimate back. Asymptotic properties of the estimates have been studied. We examine univariate and bivariate cases and multivariate case can be generalized analogously from the bivariate estimation.

1.5 Nonparametric Weighted Kernel Function Estimation

The theory of weighted distributions provide a unifying approach in situations where the random variables of interest come from the non-experimental, non- replicated, and nonrandom categories such as in environmental and ecological study. The estimation of the density in this arrangement is important for data analysis. To consider the method of ascertainment, the weighted distributions adjusts the probabilities of actual occurrence of events to arrive at a specification of the probabilities of those events as observed and recorded. Failure to make such adjustments can result in erroneous conclusions.

1.6 Weighted Kernel Density Estimation(WKDE)

Weighted distributions are used to deal missing data, damaged data, sociological or ecological data. Let $X_i,\ i = 1, 2, ..., n$ are non-negative random variables with probability density function(pdf) f and distribution function F.The weighted density g is related to f as :

$$g(x) = \frac{w(x)f(x)}{\theta_w} \quad where \;\; \theta_w = E_f w(X).$$

Let G be the distribution function of g.

Suppose X_i is not observable but we observe another random variable $Y_i, \;\; i = 1, 2, ..., n$ from the weighted distribution G and we want to use these random variables to estimate f. Then the weighted Kernel Density Estimation is

$$\hat{f}_n(x) = \frac{1}{a}\left(\sum_i \frac{1}{w(Y_i)}\right)^{-1} \sum_i \frac{1}{w(Y_i)} K\left(\frac{x - Y_i}{a}\right) \quad , 0 < w(y))$$

Since X_i are non negative and if the standard kernel is used here, this estimate assigns positive probabilities to the left of the origin where no random variables exist.

1.7 Regression Function Estimation

The idea of transformed weighted kernel density estimator can be used to estimate the regression function of nonnegative random variables. Let Y and X be continuously distributed response and independent variables. The regression function of Y on X is given by

$$m(x) = E(y|X = x) = \frac{\int y f(x, y) dy}{f(x)}.$$

Where $f(y, x)$ is a joint density of x and y, $f(y|x)$ is a conditional density of Y given x and $f(x)$ is a marginal density of X. We used proposed transformed weighted kernel density estimate in estimating the regression function.

1.8 Failure Rate Function Estimation

Let X represents time to failure of a subject with density $f(x)$ and distribution $F(x)$. The univariate failure rate function is given by

$$h(x) = -\frac{d}{dx}\log(1 - F(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{\overline{F}(x)}, \quad x \geq 0. \tag{1.4}$$

where $\overline{F}(x) > 0$ is a survival function and is given by

$$\overline{F}(x) = P(X > x) = 1 - \int_{-\infty}^{x} f(u)du.$$

When the random variables X come from nonexperimental, nonreplicated and non-random categories, the transformed weighted kernel estimators are used.

CHAPTER 2

Literature Review

The nonparametric density estimation was first introduced by Fix and Hodges [16] in the form of histogram in their unpublished manuscript in 1952. Rosenblat [30] proposed a naive kernel-based estimator where the kernel function was a simple uniform distribution and in 1962, Parzen [26] presented theoretical and mathematical framework for the kernel estimator including large sample theory and point-wise consistency. Some of the most important works over the last six decades were accomplished by many including Rudemo [31], Stone [35], Bowman [6], Silverman [34], Scott and Terrell [33], Jones and Marron [20], Wand and Jones [36], and Mnatsakanov and Sarkisian [22]. Rudemo [31], and Bowman [6] had independently developed least square cross validation (LSCV) technique to compute a data based bandwidth. Also, Scott and Terrell [33] presented biased cross validation (BCV) method of bandwidth selection. These two methods remain the most popular techniques in selecting data based bandwidth.There have been number of papers devoted to density estimation of non negative random variables. Many authors including Rao and Bagai [28], Comte and Catalot [14] briefly mentioned the necessity of transformation while estimating density of nonnegative random variables. Silverman [34] provides some adaptations of the existing methods when handling the nonnegative random variables. Marron and Ruppert [13], and Alberts and Karunamuni [4] used transformation to reduce the bias at the boundaries.

The study of suicidal data by Silverman [34] shows that the estimate of nonnegative random variables with the standard kernel function has positive values on the negative half of the real line [34] (page 18). However, if the area to the left of the origin is ignored then the density estimate would not integrate out to one. To avoid this problem, Rao and Bagai [28] used kernels with the support of positive half of the real line. Their use of exponential kernel has poor performance near the origin and in the tail area. The optimal bandwidth has slower convergence rate than the bandwidth of standard kernel. Also, their use of exponential kernel may encounter non integrability issue if $x < X_{(1)}$. Chen [11] and Scaille [32] proposed boundary bias free estimate by implementing asymmetric kernels with the support of positive half of the real line. The convergence rate of their estimates depends on the position of data point from the origin.

Failure rate function is widely used in life time data analysis. The failure rate function provides important information about the distributions of random failure times of objects in reliability and survival analysis. The pioneering work in nonparametric estimation of the univariate failure rate function can be found in Watson and Leadbetter [37], [38] and Ahmad and Lin [2]. Basu [8], Cox [15], and Puri and Rubin [27] have proposed a scalar-valued multivariate analog of univariate failure rate function which does not possess similar relationships between survival probability and the failure rate function which in otherwise well established in univariate case. Johnson and Kotz [19], and Marshal [21] defined bivariate failure rate function as vector-valued bivariate failure rate function which is in agreement with univariate case. The nonparametric kernel-type estimation of vector-valued bivariate failure rate function have been considered by Ahmad and Lin [1]. In univariate and bivariate estimation of failure rate functions, Watson and Leadbetter [37], [38] Ahmad and Lin [2], and Ahmad and Lin [1] used classical kernel density functions without transformation and hence their estimators give positive weight to the area where random variables do not exit. Fisher [17] introduced the concept of weighted distribution in the study of the effects of methods of ascertainment upon the estimation of frequencies and Rao [29] formalized weighted distribution in a unifying theory. As a special case of weighted distribution, Zelen [39] introduced weighted distribution to represent length-biased sampling in the context of cell kinetics and the early detection of disease. Patil [25] provided various examples such as encountered data analysis, equilibrium population analysis subject to harvesting and predation, meta-analysis incorporating publication bias and heterogeneity, clustering and extraneous variation using length biased distributions. Bhattacharyya, Franklin and Richardson [5] proposed kernel density estimate of length-biased distribution. Bhattacharyya estimate failed to be density and their estimate is erroneous near origin. Jones [12] proposed a new kernel based estimator for the length-biased distribution which is analogous to the kernel estimator of direct sampling case.

CHAPTER 3

Transformed Nonparametric Functions Estimation

This research focuses on nonnegative random variables. Since standard kernel density is inadequate in estimating functions that involves random variables which has support on positive half of the Euclidean space, we propose transform kernel functions estimations. We show that our estimators have better performance in tail area and have smaller mean integrated square errors by simulation. Proofs of theorems are presented in the appendix. New notations that are used in this research are introduced in this section.

3.1 Transformed Kernel Density Estimation (TKDE)

In situations where random variables come from life time distribution or from a finite support, transformation of random variables is one possibility of solving spill over effect in standard kernel density and distribution functions estimations. Let $X_1, X_2, ..., X_n$ be nonnegative random variables with density function $f(x)$ and cumulative distribution function $F(x)$. We want to estimate $f(x)$ and $F(x)$. Let $Y = \phi(X)$ denote a known transformation of X such that $Y \in \mathbb{R}$ and the new random variables Y may have a density that can be more easily estimated using the standard kernel. Then one would invert the density estimate of Y to the density estimate of original random variables X. The new estimator is called the transformed kernel density estimator (TKDE). The TKDE is simply based on the standard statistical distribution theory:

$$f(x) = g(t(x))|t'(x)|$$

. Our proposed estimator of $f(x)$ is

$$\hat{f}(x) = \frac{|\phi'(x)|}{na} \sum_{i=1}^{n} K\left(\frac{\phi(x) - \phi(X_i)}{a}\right) \tag{3.1}$$

where a is bandwidth. The detailed method for finding the optimal bandwidth is discussed later. The kernel function is a symmetric probability density function about the origin and it has the following properties

1. $\int K(u)du = 1$,
2. $\int |u|K(u)du = 0$,
3. $\int K^2(u)du < \infty, \; \int u^2K^2(u)du < \infty$

In similar fashion, we propose an estimate of $F(x)$ as

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{K}\left(\frac{\phi(x) - \phi(X_i)}{a}\right). \tag{3.2}$$

The following theorem summarizes the properties of $\hat{f}(x)$.

Theorem 3.1 *1. $E(\hat{f}(x)) \to f(x)$ as $n \to \infty$ such that $na \to \infty$ and for all x, $f(x)$ is continuous.*

2. $E\{\hat{f}(x) - f(x)\}^2 \to 0$ as $n \to \infty$ such that $na \to \infty$ for all continuity points of $f(x)$.

3. The optimal choice of the bandwidth is $a_{opt} = \left\{ \frac{R(K)\int |\phi'(x)| f(x)dx}{S(G^2,(\phi')^8)n} \right\}^{\frac{1}{5}}$,

where $R(K) = \int K^2(u)du$ and $S(G^2, (\phi')^8) = \int \frac{G^2(x)}{(\phi'(x))^8} dx$ with

$$G(x) = f(x)\left\{3(\phi''(x))^2 - \phi'(x)\phi'''(x)\right\} - 3f'(x)\phi'(x)\phi''(x) + f''(x)\{\phi'(x)\}^2$$

Corollary 3.1 *For the log transformation* $\phi(x) = \ln x$, $a_{opt} = \left\{ \frac{R(K)(\int \frac{f(x)}{|x|})dx}{\Psi(f,f',f")n} \right\}^{\frac{1}{5}}$, *where* $\Psi(f, f', f'') = \int \left(f(x)dx + 3|x|f'(x)dx + x^2 f''(x)\right)^2 dx$.

Proof: Let $\phi(x) = \ln x$ then $\phi'(x) = \frac{1}{x}, \phi''(x) = -\frac{1}{x^2}$ and $\phi'''(x) = \frac{2}{x^3}$

Thus,

$$G(x) = f(x) \left\{ \frac{3}{x^4} - \frac{2}{x^4} \right\} - 3f'(x)\frac{1}{x}\left(-\frac{1}{x^2}\right) + f''(x)\frac{1}{x^2}$$

together with $|\phi'(x)|^8 = \frac{1}{x^8}$ gives us

$$\begin{aligned} S(G^2, |\phi'|^8) \quad &= \int \left(\frac{\left(\frac{f(x)}{x^4} + 3\frac{f'(x)}{x^3} + \frac{f''(x)}{x^2}\right)^2}{\frac{1}{x^8}} \right) dx \\ &= \int \left\{ f(x) + 3|x|f'(x) + x^2 f''(x) \right\}^2 dx \\ &= \Psi(f, f', f'') \end{aligned}$$

Therefore

$$E\left\{\hat{f}(x) - f(x)\right\}^2 \approx \frac{R(K)}{na} \int \frac{f(x)}{|x|} dx + \frac{a^4}{4} \Psi(f, f', f''))$$

Therefore, the result follows $a_{opt} = \left\{ \frac{R(K)\left(\int \frac{f(x)}{|x|}\right)dx}{\Psi(f,f',f'')n} \right\}^{\frac{1}{5}}$.

The expression for optimal bandwidth involves an unknown density which is to be estimated.

The most crucial part in the density estimation is a selection of optimal bandwidth. Any attempt to decrease either the bias or the variance with respect to smoothing parameter a will result in an increase of the other. We adapted some techniques such as Unbiased Cross Validation (UCV) and Biased Cross Validation (BCV) methods to evaluate the optimal bandwidths.

3.2 Unbiased Cross Validation (UCV)

The usual criteria to assess the accuracy of density estimate is the integrated square error (ISE). The idea of the cross validation(UCV) arises from expanding the integrated square error of $\hat{f}(x)$.

$$\begin{aligned}\mathrm{ISE}(\hat{f}(x)) &= \int \left(\hat{f}(x) - f(x)\right)^2 dx \\ &= \int \hat{f}^2(x)dx - 2\int f(x)\hat{f}(x)dx + \int f^2(x)dx\end{aligned}$$

The last term is independent of bandwidth a. Therefore, minimizing $ISE(\hat{f}(x))$ is the same as minimizing the first two terms of above expression. The UCV is the procedure of obtaining unbiased estimate of the $ISE(\hat{f}(x))$.

$$\begin{aligned}ISE(\hat{f}(x)) &= \int \left(\hat{f}(x) - f(x)\right)^2 dx = \int \hat{f}^2(x)dx - 2\int \hat{f}(x)dF_n \\ &\approx \int \hat{f}^2(x)dx - 2\int \hat{f}(x)dF_n \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\int \frac{|\phi'(x)|^2}{n^2a^2}K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)K\left(\frac{\phi(x)-\phi(X_j)}{a}\right)dx \\ &\quad -2\int \frac{|\phi'(x)|}{na}\sum_{i=1}^{n}K\left(\frac{(\phi(x)-\phi(X_i))}{a}\right)dF_n(x)\end{aligned}$$

$$\begin{aligned}\text{Let } I_1 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\int \frac{|\phi'(x)|^2}{n^2a^2}K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)K\left(\frac{\phi(x)-\phi(X_j)}{a}\right)dx \\ \text{and } I_2 &= -2\int \frac{|\phi'(x)|}{na}\sum_{i=1}^{n}K\left(\frac{(\phi(x)-\phi(X_i))}{a}\right)dF_n(x)\end{aligned}$$

Finding I_1 depends upon the choice of a kernel function and a type of transformation. For example if $\phi(x) = \log x$ and standard normal for the kernel function, I_1 can be simplified as:

$$I_1 = \sum_{i=1}^{n}\sum_{j=1}^{n}\int \frac{|\phi'(x)|^2}{n^2a^2}K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)K\left(\frac{\phi(x)-\phi(X_j)}{a}\right)dx$$

Let $\phi(x) = \log x = y$ then $x = e^y$ and $\phi'(x) = \frac{1}{x}$ and a standard normal kernel.

Thus,

$$\begin{aligned}
I_1 &= \frac{1}{n^2a^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int \frac{1}{x^2}K\left(\frac{\log x - \log X_i}{a}\right)K\left(\frac{\log x - \log X_j}{a}\right)dx \\
&= \frac{1}{n^2a^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int e^{-y}K\left(\frac{y - \log X_i}{a}\right)K\left(\frac{y - \log X_j}{a}\right)dy \\
&= \frac{1}{n^2a^2}\frac{1}{2\pi}\sum_{i=1}^{n}\sum_{j=1}^{n}\int e^{-y}e^{-\left(\frac{(y-\log X_i)^2}{2a^2}\right)-\left(\frac{(y-\log X_j)^2}{2a^2}\right)}dy
\end{aligned}$$

Completing square in the exponent,

$$-\left\{y + \frac{1}{2a^2}\left[2y^2 - 2y(\log x_i + logx_j) + \log x_i^2 + logx_j^2\right]\right\}$$

$$= -\frac{1}{a^2}\left\{y^2 - y(\log x_i + \log x_j - a^2) + \frac{c_{ij}}{2}\right\}$$

$$= -\frac{1}{a^2}\left(y - (\log x_i + \log x_j - a^2)\right)^2 + \frac{1}{4a^2}(\log x_i + \log x_j - a^2)^2 - \frac{c_{ij}}{2a^2}$$

where $\quad c_{ij} = \log(x_i)^2 + \log(x_j)^2$

Hence,

$$I_1 = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2\sqrt{\pi}an^2}exp\left\{\frac{1}{4a^2}(logx_i + logx_j - a^2)^2 - \frac{c_{ij}}{2a^2}\right\}$$

Also,

$$I_2 = -\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-1}(X_i) = -\frac{2}{n^2}\sum_i\sum_j\frac{\phi'(X_i)}{a}K\left(\frac{\phi(X_i) - \phi(X_j)}{a}\right)$$

which reduces in the log and normal kernel function case to

$$I_2 = -\frac{2}{n^2}\sum_i\sum_j\frac{1}{X_ia\sqrt{2\pi}}exp^{\frac{-\left(log(X_i)-log(X_j)\right)^2}{2a^2}}$$

Then $\widehat{ISE} = I_1 + I_2$

3.3 Bias Cross Validation (BCV)

The biased cross validation method of obtaining optimal bandwidth a_{opt} is based on minimizing Asymptotic Mean Integrated Square Error(AMISE) of $\hat{f}(x)$.

$$\text{AMISE}\hat{f}(x) = \frac{R(k)}{na}\int|\phi'(x)|f(x)dx + \frac{a^4}{4}\mu_2^2(K)S(G^2, |\phi'|^8)$$

So,

$$\hat{a}_{opt} = \left\{ \frac{R(K)\int |\phi'(x)| f(x)dx}{\mu_2(K) S(G^2, (\phi')^8) n} \right\}^{\frac{1}{5}}$$

$\hat{a}_{opt}$ is obtained by estimating $\int |\phi'(x)| dF(x)$ and $S(G^2, (\phi')^8)$.

Then $\int |\phi'(x)| f(x)dx$ is estimated by $\int |\phi'(x)| dF_n(x) = \frac{1}{n} \sum_{i=1}^{n} |\phi'(X_i)|$.

Thus,

$$a_{opt} = \left\{ \frac{R(K)\frac{1}{n} \sum_{i=1}^{n} |\phi'(X_i)|}{S(G^2, (\phi')^8) n} \right\}^{\frac{1}{5}}.$$

For example when $\phi(x) = \ln x$, we have

$$S(G^2, |\phi'|^8) = \int \left\{ f(x) + 3|x| f'(x) + x^2 f''(x) \right\}^2 dx$$

To estimate a_{opt}, we estimate $S(G^2, |\phi'|^8)$ by plugging in the estimates of $f(x)$, $f'(x)$ and $f''(x)$ in $S(G^2, |\phi'|^8)$ and put that estimate of $S(G^2, |\phi'|^8)$ in the expression of a_{opt}.

Theorem 3.2 *The mean square error of $\hat{F}(x)$ is given by*

$$MSE(\hat{F}(x)) = \frac{F(x)(1 - F(x))}{n} - \frac{a}{n}\tau(\mathbb{K}, F, \phi) + \left[\frac{a^2}{2} (F \circ \phi^{-1})''(\phi(x)) \int z^2 K(z) dz \right]^2$$

Corollary 3.2 *The optimal bandwidth of $\hat{F}(x)$ is given by*

$$a^* = \left\{ \frac{\tau(\mathbb{K}, F, \phi)}{n \left[(F \circ \phi^{-1})''(\phi(x)) \int z^2 K(z) dz \right]^2} \right\}^{\frac{1}{3}}.$$

Where,

$$\tau(\mathbb{K}, F, \phi) = 2(F \circ \phi^{-1})'\phi(x) \int u \mathbb{K}(u) K(u) du$$

Proof: Differentiating $MSE(\hat{F}(x))$ with respect to a and setting it to zero immediately gives the optimal bandwidth that minimizes the MSE.

3.4 Regression Function Estimation

The transformed nonparametric density estimation can be applied in estimating regression function when non negative random variables are regressed to a response variable y.

Let Y_i and X_i be response variable and non negative explanatory variables respectively. The relation of Y to X is described by

$$y_i = m(x_i) + \epsilon_i, \quad i = 1, 2, 3, ..., n, \epsilon_i \overset{i.i.d}{\sim} N(0, \sigma^2).$$

We are interested in estimating the regression $m(x)$ with out any assumption on ϵ_i. The regression $m(x)$ can be expressed as

$$m(x) = E[Y|X = x] = \frac{\int yf(x, y)dy}{f(x)}$$

where $f(x, y)$ is a joint density of x and y. Then the regression estimate is

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)}{\sum_{i=1}^{n} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)} \tag{3.3}$$

The mean square error of the estimate is given by the following equation. Its derivation is provided in the appendix.

$$\begin{aligned} &E\left(\hat{m}(x) - m(x)\right)^2 \\ &\quad = \frac{R(K)|\phi'(x)|}{nam^2(x)f_1(x)}\left[\psi_w(\phi(x)) + m^2(x) - 2m(x)\eta_w(\phi(x))\right] \\ &\quad + \left[\frac{a^2\mu_2(K)}{2}\left(2(f_1 \circ \phi^{-1})'(\phi(x))(m \circ \phi^{-1})'(\phi(x)) \right.\right. \\ &\qquad \left.\left. + f_1(x)(m \circ \phi^{-1})''(\phi(x))\right)/f_1(x)\right]^2 \end{aligned} \tag{3.4}$$

Thus, IMSE is obtained by integrating equation 3.4 with respect to x, we assume that all integrals exist.

3.5 Failure Rate Function Estimation

In this section, we propose estimates of univariate and bivariate failure rate functions. These estimates are based on our density estimate of nonnegative random variables. It is noted that bivariate case can be generalized to any multivariate case.

(A) **Univariate Failure Rate Function Estimation**

As mentioned earlier, univariate failure rate function is defined as

$$h(x) = -\frac{d}{dx}\log(1 - F(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{\overline{F}(x)}. \quad x \geq 0$$

Where $\overline{F}(x) > 0$ is a survival function and is given by

$$\overline{F}(x) = P(X > x) = 1 - \int_0^x f(u)du$$

In the literature, hazard rate is also known as conditional failure rate, instantaneous death rate, force of mortality etc. The failure rate $h(x)dx$ represents the instantaneous chance that subject fails in the time interval $(x, x + dx)$, given that it has survived of age x.

Our estimate of univariate failure rate function is

$$\hat{h}(x) = \frac{\hat{f}(x)}{\hat{\overline{F}}(x)} = \frac{\hat{f}(x)}{1 - \hat{F}(x)}, \quad x \geq 0$$

where

$$\hat{f}(x) = \frac{|\phi'(x)|}{na}\sum_{i=1}^{n} K\left(\frac{\phi(x) - \phi(X_i)}{a}\right)$$

and

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{K}\left(\frac{\phi(x) - \phi(X_i)}{a}\right).$$

Theorem 3.3 *Expected value of $\hat{h}(x)$ is*

$$E\hat{h}(x) = h(x)\left\{1 + \frac{a^2}{2}\left(\frac{\sqrt{S(G^2, |\phi'(x)|^8)}}{f(x)} - (F \circ \phi^{-1})''|\phi(x)| \int z^2K(z)dz\right)\right\}.$$

Corollary 3.3 *Bias of $\hat{h}(x)$ is*

$Bias(\hat{h}(x)) = \frac{h(x)a^2}{2}\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)} - (F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz\right\}.$

Theorem 3.4 *Var $\hat{h}(x)$ is given by*

$V\hat{h}(x) = \frac{h^2(x)|\phi'(x)|R(K)}{naf(x)}$

Corollary 3.4

$$\begin{aligned} MSE(\hat{h}(x)) &= \frac{h^2(x)|\phi'(x)|R(K)}{naf(x)} + \left\{\frac{h(x)a^2}{2}\left(\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)}\right.\right. \\ &\qquad \left.\left. -(F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz\right)\right\}^2 \qquad (3.5) \end{aligned}$$

Corollary 3.5 *The optimal bandwidth that minimizes the $MSE(\hat{h}(x))$ is given by*

$$a^*_{Opt} = \left[\frac{|\phi'(x)|R(K)}{nf(x)\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)} - (F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz\right\}}\right]^{\frac{1}{5}}.$$

Proof: Differentiating equation 3.5 w.r.t. a and setting it to 0, immediately gives the optimal bandwidth.

In practice, we generally deal with multivariate cases. In the following section, we propose estimate of vector-valued bivariate failure rate function and its properties.

(B) **Bivariate Failure Rate Function Estimation**

Let $\mathbf{X} = [X_{11}, X_{21}]', ..., [X_{1,n}, X_{2,n}]'$ be a bivariate nonnegative random vector with cumulative distribution function F and probability density function f. The bivariate vector-valued failure rate function is defined as

$$\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), h_2(\mathbf{x})]'$$

where

$$h_i(\mathbf{x}) = \frac{-\delta}{\delta x_i}\ln\overline{F}(\mathbf{x}) = \frac{g_i(\mathbf{x})}{\overline{F}(\mathbf{x})},\ i = 1, 2.$$

$$g_1(\mathbf{x}) = \int_{x_i}^{\infty} f(x_1, y_2)dy_2 \text{ and } g_2(\mathbf{x}) = \int_{x_i}^{\infty} f(y_1, x_2)dy_1.$$

Let $y_i = \phi_i(X_{i1})$ be the transformation $i = 1, 2$. We propose estimates of f, g and F respectively as follows.

$$\begin{aligned}
\hat{f}(x_1, x_2) &= \frac{|\phi'_{x_1}(x_1)\phi'_{x_2}(x_2)|}{na_1a_2} \sum_{i=1}^{n} K_1\left(\frac{\phi_1(x_1) - \phi_1(x_{1i})}{a_1}\right) \\
&\qquad \times K_2\left(\frac{\phi_2(x_2) - \phi_2(x_{2i})}{a_2}\right), i = 1, .., n. \\
\hat{g}_1(x_1, x_2) &= \int_{x_2} \hat{f}(x_1, y_2)dy_2 = \frac{|\phi'_{x_1}(x_1)|}{na_1} \sum_{i=1}^{n} K_1\left(\frac{\phi_1(x_1) - \phi_1(x_{1i})}{a_1}\right) \\
&\qquad \times \overline{\mathbb{K}}_2\left(\frac{\phi_2(x_2) - \phi_2(x_{2i})}{a_2}\right) \\
\hat{g}_2(x_1, x_2) &= \int_{x_1} \hat{f}(y_1, x_2)dy_1 = \frac{|\phi'_{x_2}(x_2)|}{na_2} \sum_{i=1}^{n} \overline{\mathbb{K}}_1\left(\frac{\phi_1(x_1) - \phi_1(x_{1i})}{a_1}\right) \\
&\qquad \times K_2\left(\frac{\phi_2(x_2) - \phi_2(x_{2i})}{a_2}\right) \\
\hat{\overline{F}}(x_1, x_2) &= \frac{1}{n} \sum_{i=1}^{n} \overline{\mathbb{K}}_1\left(\frac{\phi_1(x_1) - \phi_1(x_{1i})}{a_1}\right) \overline{\mathbb{K}}_2\left(\frac{\phi_2(x_2) - \phi_2(x_{2i})}{a_2}\right)
\end{aligned}$$

Finally, our estimator of vector-valued bivariate failure rate function is

$$\hat{\mathbf{h}}(\mathbf{x}) = [\hat{h}_1(\mathbf{x}), \hat{h}_2(\mathbf{x})]' \text{ with } \hat{h}_1(\mathbf{x}) = \frac{\hat{g}_1(x_1, x_2)}{\hat{\overline{F}}(x_1, x_2)} \text{ with } \hat{h}_2(\mathbf{x}) = \frac{\hat{g}_2(x_1, x_2)}{\hat{\overline{F}}(x_1, x_2)}.$$

The vector-valued bivariate failure rate function is also known as hazard gradient. Remaining part of this section provides asymptotic properties of $\hat{f}(x_1, x_2)$, $\hat{g}_1(x_1, x_2)$ and $g_2(x_1, x_2)$. Also, the consistency of $\hat{h}(\mathbf{x})$ is established in this section .

Theorem 3.5 *The expected value of $\hat{h}(x)$ is*

$$E(\hat{h}(\mathbf{x})) = \Bigg[\quad \Big\{h_1(\mathbf{x})\Big(1 + \frac{bias\hat{g}_1(\mathbf{x})}{g_1(\mathbf{x})} + \frac{bias\hat{\overline{F}}(\mathbf{x})}{\overline{F}(\mathbf{x})}\Big)\Big\}$$
$$, \Big\{h_2(\mathbf{x})\Big(1 + \frac{bias\hat{g}_2((\mathbf{x})}{g_2(\mathbf{x})} + \frac{bias\hat{\overline{F}}(\mathbf{x})}{\overline{F}(\mathbf{x})}\Big)\Big\}\Bigg]'$$

where

$$bias\hat{F}(\mathbf{x}) = \frac{1}{2}\Bigg[a_1^2 z_1^2 F((\phi_1^{-1})''(\phi(x_1)), (\phi_2^{-1})''(\phi(x_2))$$

$$+a_2^2 z_2^2 F((\phi_1^{-1})''(\phi(x_1)), (\phi_2^{-1})''(\phi(x_2))\Bigg] \times K_1(z_1)K_2(z_2)dz_1dz_2$$

$$bias\hat{g}_1(\mathbf{x}) = \frac{1}{2}\int_{x_2}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)$$

$$+a_2^2 z_2^2 f\left((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)\Bigg]dz_1dz_2dy$$

$$bias\hat{g}_2(\mathbf{x}) = \frac{1}{2}\int_{x_1}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)$$

$$+a_2^2 z_2^2 f\left((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)\Bigg]dz_1dz_2dy.$$

Theorem 3.6

$$\lim_{na\to+\infty} Cov[\hat{h}_i(\mathbf{x}_\alpha), \hat{h}_j(\mathbf{x}_\beta)]' = \frac{|\phi_i'(\mathbf{x}_\alpha)|R(k_i)}{g_i(\mathbf{x}_\alpha)} + h_i^2(\mathbf{x}_\alpha)\frac{[1-\overline{F}(\mathbf{x}_\alpha)]}{\overline{F}(\mathbf{x}_\alpha)} \quad for\ \ i=j; \quad \alpha=\beta,$$

$$= g_i(\mathbf{x}_\alpha)g_j(\mathbf{x}_\alpha)\frac{[1-\overline{F}(\mathbf{x}_\alpha)]}{\overline{F}^3(\mathbf{x}_\alpha)} \quad for\ \ i\neq j\ \ and\ \ \alpha=\beta$$

$$= \frac{g_i(\mathbf{x}_\alpha)g_j(\mathbf{x}_\beta)}{\overline{F}^2(\mathbf{x}_\alpha)\overline{F}^2(\mathbf{x}_\beta))}(\overline{F}(\mathbf{x}_{max(\alpha,\beta)}) - \overline{F}(\mathbf{x}_\alpha)\overline{F}(\mathbf{x}_\beta)) \quad for\ \ i,j=1,2; \quad \alpha\neq\beta$$

$$where \quad \mathbf{x}'_{max(\alpha,\beta} = (x_{\alpha max}, x_{\beta max}) = (max(x_{1\alpha}, x_{1\beta}), max(x_{2\alpha}, x_{2\beta})),$$

$$for\ \ \alpha\neq\beta=1,2.$$

Proof of this theorem follows immediately after using theorem 4.1 of Ahmad and Lin[1]. The mean square error is then calculated using bias and variance of $h(\mathbf{x})$.

3.6 Computational Study

In this section, we present simulated and a real data examples of transformed kernel density estimation. We computed mean integrated square errors to asses the estimates.

3.7 Simulations

For optimal bandwidth of density and distribution functions estimate, Monte Carlo simulations of 500 iterations was carried out in R and sample of sizes n= 10, 20, 40, 60, 100 were generated from log normal distribution with mean 0 and variance 1. The table below summarizes the optimal bandwidths using UCV and BCV along with their corresponding standard errors in the parenthesis. The simulation result shows that unbiased cross validation of bandwidth selection method does not work for log transformation. This is partly because there are a terms with high power in the denominator when we use Newton-Raphson method and those terms unusually get too big. However, that problems does not appear in biased cross validation.

Table 3.1: Optimal bandwidth using unbiased and biased cross validation for lognormal(n,0,1) and optimal bandwidth for its distribution.

n	h_{ucv} (se)	h_{bcv}(se)	h^*(se)
10	0.495(0.083)	0.592(0.203)	0.205(0.168)
20	0.486(0.0540)	0.540(0.175)	0.097(0.010)
40	0.481(0.035)	0.497(0.136)	0.047(0.006)
60	0.480 (0.029)	0.425(0.118)	0.033(0.004)
100	0.478(0.023)	0.394(0.099)	0.033(0.002)

The failure rate function of log normal data is estimated at 40th percentile. The table below summarizes the result. The simulation shows that standard error for this example is quite big but it is in decreasing order with larger sample size. We believe that the sample size for this type of estimation is too small. With the better computational resources, we can get better result with the larger sample size.

Table 3.2: Failure Rate function estimated at 40th percentile of lognormal(n,0,1) with standard error in parenthesis.

n	$h(se)$	h_{bcv}	h^*
10	0.345(0.19)	0.59	0.205
20	0.331(0.187)	0.540	0.097
40	0.323(0.177)	0.497	0.047
60	0.301 (0.162)	0.425	0.033
100	0.289(0.154)	0.394	0.033

We performed transformed kernel, standard kernel and parametric simple linear regression analysis on child's weights on a data set child's weight to child's height found on sas manual with bandwidths 0.033 and 3.24 for transformed and standard kernel cases respectively. Transformed kernel regression estimation performs almost as good as the parametric regression. the prediction result shows that standard kernel regression analysis slightly overestimates or underestimates than the transformed kernel regression analysis. The following table summarizes transformed nonparametric(tkde), standard nonparametric(kde) and parametric(ppred) prediction of child's weight on given height. The lower(plwr) and upper(pupr) 95% prediction intervals are provided in the last two columns.

Table 3.3: Regression function estimation of child's weight vs child's height data set found in sas manual.

height	weight(tkde)	weight(ppredict)	weight(kde)	plwr	pupr
51	51.17834	55.82363	65.50283	28.56741	83.07984
53	63.59846	63.62169	74.57251	37.28011	89.96327
55	81.14582	71.41975	81.52674	45.83965	96.99985
57	84.69134	79.21781	86.50457	54.23203	104.20358
60	92.81439	90.91490	94.27482	66.48142	115.34837
64	105.16550	106.51102	104.81542	82.14301	130.87903
67	113.13985	118.20811	111.30079	93.38269	143.03353
69	119.20833	126.00617	116.42506	100.64559	151.36675
72	136.09423	137.70326	127.25181	111.22252	164.18400

Random samples of sizes n= 200, 400, 600 and 800 were generated from Log normal(0,1) and gamma(2,1) densities and the corresponding mean integrated square errors were computed. We compare these results with varying kernel and standard kernel densities results found in table 1 of [22]. The following tables summarize the results.

Table 3.4: MISEs of varying, standard and transformed kernel densities for lognormal(n,0,1).

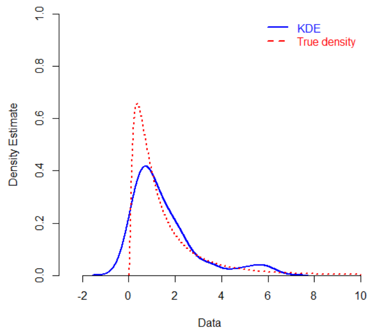
n	$\hat{f}_\alpha$	$\hat{f}^*_\alpha$	$\hat{f}_h$	$\hat{f}_{tkde}(StdErr)$	α_{cv}	α^*_{cv}	h_{cv}	h
200	0.0092	0.0066	0.0166	0.0053(0.0030)	14	10	0.30	0.375
400	0.0057	0.0043	0.0103	0.0034(0.0021)	18	14	0.25	0.335
600	0.0039	0.0030	0.0075	0.0023(0.0013)	22	16	0.22	0.317
800	0.0029	0.0022	0.0059	0.0019(0.0010)	24	18	0.19	0.263

Table 3.5: MISEs of varying, standard and transformed kernel densities for gamma(n,2,1).

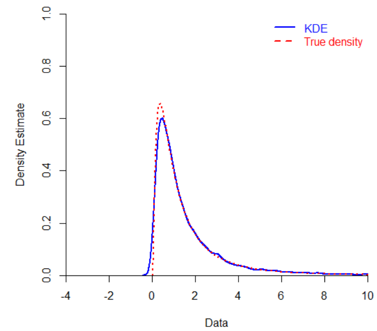
n	$\hat{f}_\alpha$	$\hat{f}^*_\alpha$	$\hat{f}_h$	$\hat{f}_{tkde}(StdErr)$	α_{cv}	α^*_{cv}	h_{cv}	h
200	0.0060	0.0046	0.0080	0.0041(0.0021)	11	7	0.14	0.29
400	0.0033	0.0026	0.0045	0.0023(0.0011)	14	9	0.11	0.27
600	0.0020	0.0016	0.0030	0.0016(0.0009)	17	11	0.10	0.25
800	0.0018	0.0015	0.0026	0.0013(0.0008)	19	12	0.09	0.22

The computational results show that TKDE performs as good as or better than varying kernel estimator in terms of integrated mean square error. It is always better than the standard kernel density estimation.

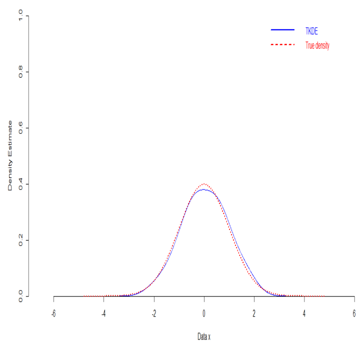
The figures 3.1.a and 3.1.b are the density estimates using standard kernel density of Log Normal data of sample sizes 40 and 10000 respectively. Clearly, the standard kernel assigns positive probabilities to the left of the origin. The figures 3.1.c and 3.1.d are the density estimates of log transformed data using standard kernel and the figures 3.1.e and 3.1.f can be considered as back transformed density estimates of 3.1.c and 3.1.d which are given by our proposed density estimate (TKDE).



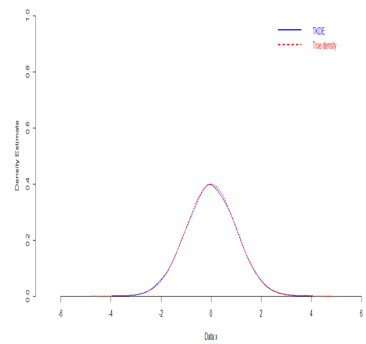
(a) n=40, h=0.50



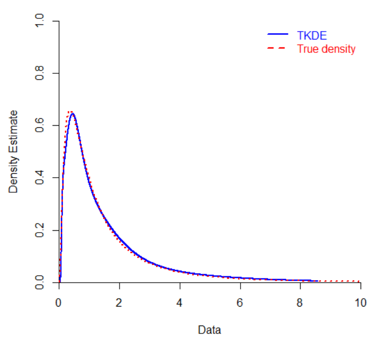
(b) n=10000, h=0.15



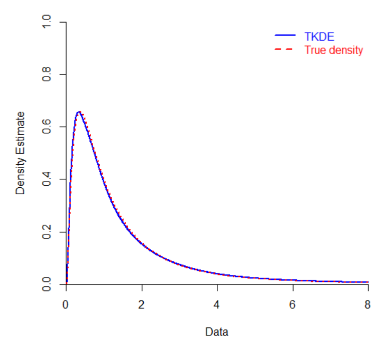
(c) n=40, h=0.45



(d) n=10000, h=0.20



(e) n=40, h=0.39

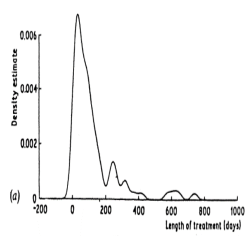


(f) n=10000, h=0.27

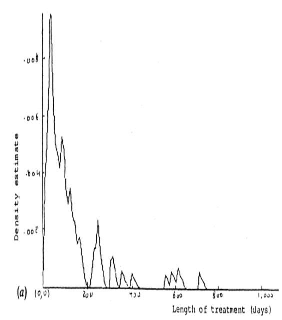
Figure 3.1: The figures (a) and (b) are density estimates of Log Normal data using standard kernel. Figure (c) and (d) are density estimate of log transformed data using standard kernel and Figure (e) and (f) are density estimates using proposed TKDE

3.8 Real Data Examples

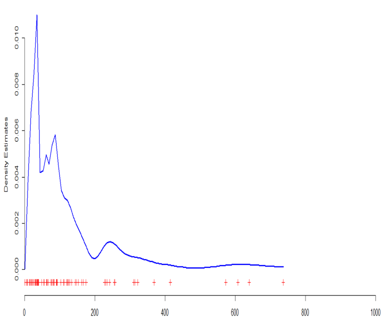
The figure 3.2.a is taken from Silverman[34](page 18) which shows the density estimate of suicidal data using standard kernel density estimate whereas figure 3.2.b is taken from Rao and Bagai [28] and is also the density estimate of the suicide study data using exponential kernel. The graph 3.2.c is the density estimate of the suicide study data by using TKDE. The TKDE performs better in the tail area and near zero.



(a) suicide data, n= 86, h=20



(b) suicide data, n= 86, h=20



(c) suicide data, n= 86, h=0.10

Figure 3.2: The figures (a), (b), and (c) are kernel density estimates of suicide data found in Silverman[34] page 18 using standard , exponential and transformed kernel(3.1) densities respectively. The figures (a) and (b) are taken from Silverman[34] and Rao and Bagai[28] respectively.

CHAPTER 4

Transformed Nonparametric Weighted Functions Estimation

The new transformed weighted kernel estimators are proposed along with their properties. The short mathematical results are presented in this section and the results requiring rigorous proofs are provided in the appendix. New notations that are used in this research are introduced in this section.

4.1 Transformed Weighted Kernel Density Estimation(TWKDE)

We propose transformed density estimate as:

Let $W_i = \phi(Y_i)$ then the TWKDE of X_i is

$$\hat{f}(x) = \frac{|\phi'(x)|}{a}\left(\sum \frac{1}{|w(\phi(Y_i))|}\right)^{-1} \sum \frac{1}{|w(\phi(Y_i))|} K\left(\frac{\phi(x) - \phi(Y_i)}{a}\right) \tag{4.1}$$

Similarly, An estimate of $F(x)$ is

$$\hat{F}(x) = \left(\sum \frac{1}{|w(\phi(Y_i))|}\right)^{-1} \sum \frac{1}{|w(\phi(Y_i))|} \mathbb{K}\left(\frac{\phi(x) - \phi(Y_i)}{a}\right) \tag{4.2}$$

Where $\mathbb{K}(u)$ is corresponding distribution function of $K(u)$.

We write

$$f(x) = \frac{g(x)\theta_w}{|w(\phi(x))|} \quad where \;\; \theta_w = E_f(|w(\phi(X))|)$$

Then, we can write the cdf as

$$\begin{aligned} F(x) &= \theta_w \int_0^{\phi(x)} (|w(u)|)^{-1} dG(u) \\ &= |E_g \left[\frac{I(\phi(Y) \le \phi(x))}{|w(\phi(Y))|}\right] / E_g \left[|w(\phi(Y))|^{-1}\right]. \end{aligned}$$

The empirical distribution of $F(x)$ is given by

$$F_n(x) = \left[\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\right]^{-1}\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}I(\phi(Y_i) \leq \phi(x)). \tag{4.3}$$

The kernel estimate of a density function f is defined as

$$\hat{f}(x) = \frac{d}{dx}\int \frac{1}{a}K\left(\frac{\phi(x)-\phi(w)}{a}\right)F_n((w))dw. \tag{4.4}$$

Using eqn 4.3 in 4.4 gives

$$\left[\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\right]^{-1}\frac{d}{dx}\int \frac{1}{a}K\left(\frac{\phi(x)-\phi(w)}{a}\right)$$
$$\times\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}I(\phi(Y_i) \leq \phi(x))dw \tag{4.5}$$

The derivative is equal to

$$\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\frac{d}{dx}\int_{\phi(Y_i)}^{\infty}\frac{1}{a}K\left(\frac{\phi(x)-\phi(w)}{a}\right)dw$$
$$= \sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\frac{|\phi'(x)|}{a^2}\int_{\phi(Y_i)}^{\infty}K'\left(\frac{\phi(x)-\phi(w)}{a}\right)dw$$
$$= -\frac{|\phi'(x)|}{a^2}\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\int_{\frac{\phi(x)-\phi(Y_i)}{a}}^{\infty}K'(w)dw$$
$$= \frac{|\phi'(x)|}{a}\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)$$

Hence

$$\hat{f}(x) = \frac{|\phi'(x)|}{a}\left[\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\right]^{-1}\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right).$$

Then the estimate of $F(x)$ is obtained by using $\int_{0}^{\phi(x)}\hat{f}(w)dw$.

In the following section, we study the asymptotic properties of $\hat{f}(x)$ and $\hat{F}(x)$.

4.2 Asymptotic MSE and IMSE of the Estimates

In this section, we derive mean square errors, integrated mean square errors and optimal bandwidths of weighted density and distribution functions. The mean square error of the weighted density estimate is given by

$$E(\hat{f}(x) - f(x))^2 = \frac{f(x)\theta_w|\phi'(x)|R(K)}{na|w(\phi(x))|} + \frac{a^4}{4}\mu_2^2(K)S(G^2, |\phi'(x)|^8). \tag{4.6}$$

The optimal bandwidth a_{opt} is given by

$$a_{opt} = \left\{ \frac{f(x)\theta_w|\phi'(x)|R(K)}{n|w(\phi(x))|\mu_2^2(K)S(G^2, |\phi'(x)|^8)} \right\}^{\frac{1}{5}}. \tag{4.7}$$

The integrated mean square error is obtained by integrating equation 4.6 with respect x provided that all integrals exist.

$$\begin{aligned} E(\hat{F}(x) - F(x))^2 = \frac{\theta_w}{n|w(\phi(x))|} \Bigg[& (F(x)(1 - F(x)) - a\tau(\mathbb{K}, F, \phi)) \\ & + |w(\phi(x))|F^2(x)\theta_w\eta(\theta_w, \nu_w) - \frac{2F^3(x)}{\theta_w^2} \Bigg] \\ & + \frac{a^4\mu_2^2(K)((F \circ \phi^{-1})''(\phi(x)))^2}{4} \end{aligned} \tag{4.8}$$

The optimal bandwidth a^*_{opt} is given by

$$a^*_{opt} = \left\{ \frac{\theta_w\tau(\mathbb{K}, F, \phi)}{n|w(\phi(x))|\mu_2^2(K)((F \circ \phi^{-1})''(\phi(x)))^2} \right\}^{\frac{1}{3}} \tag{4.9}$$

also

$$\begin{aligned} \int E(\hat{F}(x) - F(x))^2 = \frac{\theta_w}{n} \int \Bigg[& \frac{F(x)}{|w(\phi(x))|} - \frac{a\tau(\mathbb{K}, F, \phi)}{|w(\phi(x))|} \\ & - F^2(x) \left(\frac{1}{|w(\phi(x))|} - \theta_w\eta(\theta_w, \nu_w) + \frac{2F(x)}{|w(\phi(x))|\theta_w^2} \right) \\ & + \frac{a^4\mu_2^2(K)}{4}((F \circ \phi^{-1})''(\phi(x)))^2 \Bigg] dx \end{aligned} \tag{4.10}$$

Next, we discuss large sample properties of the estimates.

4.3 Large Sample Properties of the Estimates

it is clear that $\hat{f}(x) \to f(x)$ as $n \to \infty$ in probability at every continuity point x of f.

Theorem 4.1 *(I) If $na \to \infty$ and $na^5 \to 0$ as $n \to \infty$, if f'' exists and is bounded, then $\sqrt{na}(\hat{f}(x) - f(x)$ is asymptotically normal with mean 0 and variance $\sigma^2 = \frac{|\phi'(x)|f(x)R(K)}{|w(\phi(x))|}$.*
(II) If $na^2 \to \infty$ as $n \to \infty$ and if f is uniformly continuous, and if $\int e^{-itu} \frac{\left[|\phi'(x)|K\left(\frac{\phi(x)-\phi(y)}{a}\right)\right]}{|w(\phi(x))|} dx$ is absolutely integrable in t, then $sup_x|\hat{f}(x) - f(x)| \to 0$ in probability as $n \to \infty$.
(III) If for any $\epsilon > 0$, $\sum_{i=1}^{n} e^{-\epsilon na^2} < \infty$, if f is uniformly continuous and if $\frac{\left[|\phi'(x)|K\left(\frac{\phi(x)-\phi(y)}{a}\right)\right]}{|w(\phi(x))|}$ is a function of bounded variation,then $sup_x|\hat{f}(x) - f(x)| \to 0$ with probability one as $n \to \infty$.

The proof of this theorem is provided in the appendix. Next, we summarize the large sample properties of $\hat{F}(x)$ in the following theorem without proof. The proof follow in similar fashion as above, so is omitted.

Theorem 4.2 *(I) If $na^4 \to 0$ as $n \to \infty$, if f' exits and is bounded, then $\sqrt{n}(\hat{F}(x) - F(x))$ is asymptotically normal with mean 0 and variance*

$$\frac{1}{\theta_w n|w(\phi(x))|}[F(x)(1 - F(x) - a\tau(\mathbb{K}, F, \phi)]$$

(II) If F is uniformly continuous, then $\sup_x |\hat{F}(x) - F(x)| \to 0$ with probability one as $n \to \infty$.

4.4 Regression Estimation for Transformed Weighted Data

This section illustrates regression estimate of weighted data, its mean square error and large sample properties.

Since

$$f(x, y) = \frac{\theta_w g(x, y)}{|w(\phi(x), \phi(y))|},$$

thus

$$F(x, y) = \theta_w \int_{-\infty}^{\phi(x)} \int_{-\infty}^{\phi(y)} (|w(r, s))^{-1} dG(r, s)$$

$$= \frac{E_g\left\{I(\phi(U) \leq \phi(x), \phi(V) \leq \phi(y)/|w(\phi(U), \phi(V)|\right\}}{E_g(|w(\phi(U), \phi(V)))^{-1}}.$$

Let $(U_1, V_1), ..., (U_n, V_n)$ be a random sample from G, then the empirical estimate of $F(x, y)$ is given by

$$F_n(x, y) = \frac{\sum_{i=1}^{n}(|w(\phi(U_i), \phi(V_i)|)^{-1} I(\phi(U_i) \leq \phi(x), \phi(V_i) \leq \phi(y))}{\sum_{i=1}^{n}(|w(\phi(U_i), \phi(V_i)|)^{-1}}.$$

Then the joint kernel density estimate is given by

$$\begin{aligned} \hat{f}(x, y) \quad &= \quad \left[\sum_{i=1}^{n} \frac{|w(\phi(U_i), \phi(V_i))|a^2}{|\phi'(x)\phi'(y)|}\right]^{-1} \sum_{i=1}^{n}(|w(\phi(U_i), \phi(V_i)|)^{-1} \\ &\qquad \times K^{(2)}\left(\frac{\phi(x) - \phi(U_i)}{a}, \frac{\phi(y) - \phi(V_i)}{a}\right). \end{aligned} \tag{4.11}$$

Where $K^{(2)}$ is a known density which is bounded and $\|.\|K^{(2)}(.,.) \to 0$ as $\|(u, v)'\| \to \infty$. The regression $m(x) = E(\phi(Y)|\phi(X = x)) = \int \phi(y)f(x, y)/f_1(x)$. Then the regression estimate is

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} V_i K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)/|w(\phi(U_i), \phi(V_i))|}{\sum_{i=1}^{n} K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)/|w(\phi(U_i), \phi(V_i))|}. \tag{4.12}$$

Theorem 4.3 *The mean square error of the regression estimate is given by the following equation*

$$\begin{aligned} &E\left(\hat{m}(x) - m(x)\right)^2 \\ &\qquad = \frac{R(K)|\phi'(x)|}{nam^2(x)f_1(x)}\left[\psi_w(\phi(x)) + m^2(x)\nu_w(\phi(x)) - 2m(x)\eta_w(\phi(x))\right] \\ &\qquad + \left[\frac{a^2\mu_2(K)}{2}\left(2(f_1 \circ \phi^{-1})'(\phi(x))(m \circ \phi^{-1})'(\phi(x))\right.\right. \\ &\qquad\qquad \left.\left. + f_1(x)(m \circ \phi^{-1})''(\phi(x))\right)/f_1(x)\right]^2 \end{aligned} \tag{4.13}$$

Thus, IMSE is obtained by integrating (4.13) with respect to x, we assume that all integrals exist.

Theorem 4.4 *(I) If* $na \to \infty$ *and* $na^5 \to 0$ *as* $n \to \infty$, *if* f'' *and* h'' *exist and are bounded, then* $\sqrt{na}(\hat{m}(x) - m(x)$ *is asymptotically normal with mean* 0 *and variance* $\sigma^2 = R(K)|\phi'(x)| \left[\psi_w(\phi(x)) + m^2(x)\nu_w(\phi(x)) - 2m(x)\nu_w(\phi(x))\right]/m^2(x)f_1(x)$.

(II) If $na^2 \to \infty$ *as* $n \to \infty$ *and if* f *and* h *are uniformly continuous, and if* $\int e^{-itu} \frac{|\phi'(x)|K(u)}{|w(\phi(u),\phi(v))|}du$ *is absolutely integrable in* u,*if* $\inf_{a \le x \le b} h(x) = \alpha > 0$ *and if* $0 \le \alpha \le v \le b < \infty$, *then* $\sup_{v<a\le x\le b<\infty} |\hat{m}(x) - m(x)| \to 0$ *in probability as* $n \to \infty$. *(III) If for any* $\epsilon > 0$, $\sum_{i=1}^{n} e^{-\epsilon na^2} < \infty$, *if* f_1 *and* k *are uniformly continuous and if* $\frac{\left[|\phi'(x)|K\left(\frac{\phi(x)-\phi(y)}{a}\right)\right]}{|w(\phi(x))|}$ *is a function of bounded variation (in u)and ,if* $\inf_{a\le x\le b} h(x) = \alpha > 0$ *and if* $0 \le \alpha \le v \le b < \infty$, *then* $\sup_{v<a\le x\le b<\infty} |\hat{m}(x) - m(x)| \to 0$ *in probability as* $n \to \infty$.

,then $\sup_x |\hat{f}(x) - f(x)| \to 0$ *with probability one as* $n \to \infty$.

4.5 Univariate Failure Rate Function of Weighted Data

The estimate of failure rate function for weighted data is obtained by plugging the estimates $\hat{f}(x)$ and $\hat{F}(x)$ in $\hat{h}(x) = \frac{\hat{f}(x)}{\hat{\bar{F}}(x)}, x > 0$ and $\hat{\bar{F}}(x) > 0$. Note that expected value of failure rate function in biased and unbiased cases are identical.The following theorems provide variance and mean square error of the failure rate function estimate of weighted data.

Theorem 4.5 $Var\hat{h}(x)$ *is*
$Var(\hat{h}(x) = \frac{h^2(x)|\phi'(x)|R(K)\theta_w}{naf(x)|w(\phi(x))|}$

Theorem 4.6 $MSE\hat{h}(x)$ *is*

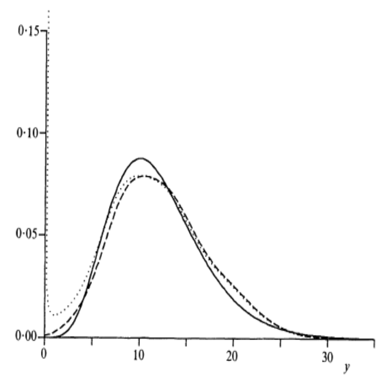
$$MSE(\hat{h}(x) = \left\{ \frac{h(x)a^2}{2} \left(\frac{\sqrt{S(G^2, |\phi'(x)|^8)}}{f(x)} - (F \circ \phi^{-1})''|\phi(x)| \times \int z^2 K(z)dz \right) \right\}^2 + \frac{h^2(x)|\phi'(x)|R(K)\theta_w}{naf(x)|w(\phi(x))|}$$

4.6 Computational Study

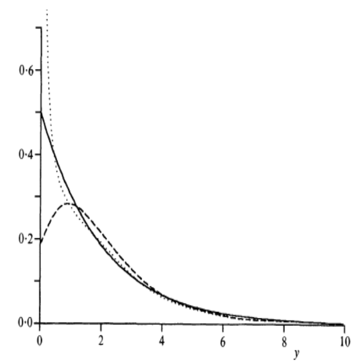
In this section, we present some simulated examples. The graphical results are compared with Jones [12] in which standard normal kernel was used in estimating the density of length biased data. Note that the length biased distribution is a special case of weighted distribution. The length biased data arises when the probability of an observation to be included in the sample is proportional to its length.

4.7 Simulations

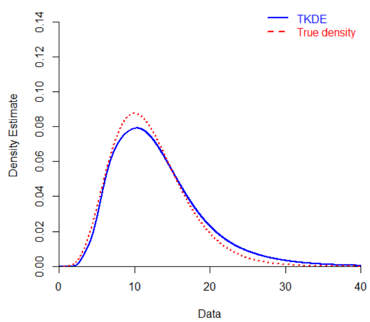
Random samples of size n= 200 are generated from chi-square density with 12 and 2 degrees of freedom. These distributions are chosen to make comparison with Jones [12]. Patil [25] shows that length biasing χ_p^2 results in χ_{p+2}^2 distribution. Figures 4.1a and 4.1b are taken from Jones [12]. Figures 4.1c, and 4.1d are produced by TWKDE. In figure d, we clearly see that Jones [12] truncated the density estimate about Y axis which they also acknowledged. This problem of spill over effect has taken care by the proposed transformed weighted kernel density estimate. The density estimate using standard kernel will be worse if the data come from compact interval. The TWKDE handles appropriately when data comes from finite support. As in the transformed kernel density estimate, the TWKDE will have lower bias and so smaller mean integrated square error. The detail simulation studies with various cases will be immediate future study.



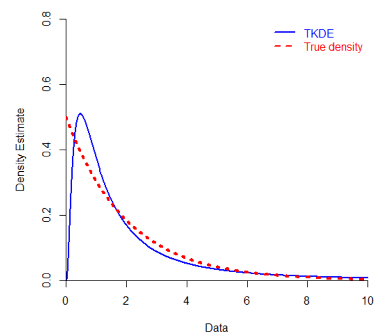
(a) χ^2_{12} with h=1.88, n=200



(b) χ^2_2 with h=0.72, n=200



(c) χ^2_{12} with h=0.22, n=200

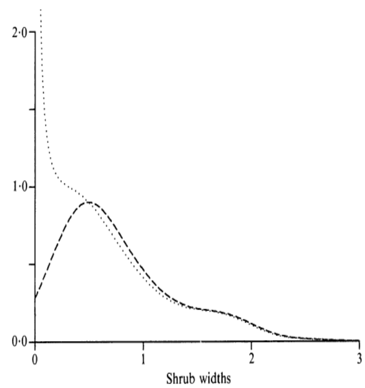


(d) χ^2_2 with h=0.38, n=200

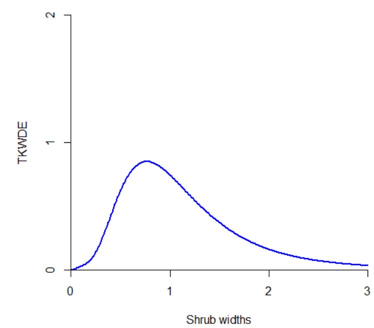
Figure 4.1: The figures (a) and (b) are taken from Jones[12]. The figures (c) and (d) are produced by TWKDE.

4.8 Real Data Example

The density estimate of the data on the widths of n=46 shrubs obtained by line transect sampling found in Table 3 of Muttlak and McDonald (1990) is shown below. The figure 4.2a is taken from Jones (1990) and figure 4.2b is obtained by using the proposed TWKDE



(a) n=46, h= 0.23



(b) h= 0.48

Figure 4.2: The density estimate of the data on the widths of n=46 shrubs obtained by line transect sampling found in Table 3 of Muttlak and McDonald(1990). The figure a is taken from Jones(1990) and figure b is obtained by using the proposed TWKDE.

CONCLUSIONS

The standard kernel density estimation performs well in many situations when the parametric models does not. The parametric models can not detect the multi modes if they are present but the kernel density accurately identifies not only the modes but also the nature of the modes over time. This is the reason that many researchers recommended using the kernel density estimate to estimate the density of random variables. But dealing with non negative random variables or the variables from compact intervals with standard kernel density estimate technique encounters boundary bias problem. Simply, the standard kernel density estimation technique is inadequate to estimate the density of the non-negative random variables. By comparing our technique with other currently existing ones, the non-parametric transformed kernel density estimate is recommended to estimate the density of non-negative random variables. It does not suffer from spill over effect and boundary biased problem. It has better performance in tail area. The simulation study shows that, for smaller sample sizes, it has smaller integrated mean square than the varying kernel density estimation.

TKWDE appropriately takes care of spill over effect when dealing random variables from weighted distribution. It does not suffers from any boundary effect problem. We are free to choose a kernel function from large class of densities.

Various aspects of transformed kernel case is under review. Of course, finding the right choice of transformation is a topic for future study along with better simulations work specially in the area of bandwidth selection.

CHAPTER 5

Future Study

5.1 Transformed Multivariate Kernel Density Estimation

The purpose of this study is to generalize the univariate transformed kernel density estimate to the multivariate setting. The generalization is carried out using the product of univariate kernel functions provided in Cacoullos[10].

Suppose that $\mathbf{X_1}, ..., \mathbf{X_n}$ is i.i.d. q-vector where $\mathbf{X_i} \in \mathbb{R}_+^q$, for $q > 1$ having a common pdf $f(\mathbf{X}) = \mathbf{f}(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_q})$. Let $\mathbf{Y} = \phi(\mathbf{X})$ be known transformation such that the density of new random vector can be easily estimated by using standard kernel density estimate. The density of the original random vector is then obtained by back transforming the density of the new random vector. We propose an estimate of $f(\mathbf{X})$ as

$$\hat{f}(\mathbf{X}) = \frac{||\phi'(\mathbf{X})||}{n} \sum_{j=1}^{n} \left\{ \prod_{i=1}^{q} \frac{1}{h_i} K\left(\frac{\phi(\mathbf{x_i}) - \phi(\mathbf{X_{ij}})}{h_i} \right) \right\}. \tag{5.1}$$

Where h_i are bandwidths and $K(.)$ is known multivariate probability density with the following properties:

$$\sup_{\mathbf{y} \in q} |K(\mathbf{y})| < \infty \tag{5.2}$$

$$\int |K(\mathbf{y})| \mathbf{dy} = 1 \tag{5.3}$$

$$\lim_{|\mathbf{y}| \to \infty} |y|^p K(\mathbf{y}) = 0 \tag{5.4}$$

Similarly, the estimate of the distribution function is

$$\hat{F}(\mathbf{X}) = \frac{||\phi'(\mathbf{X})||}{n} \sum_{j=1}^{n} \left\{ \prod_{i=1}^{q} \mathbb{K}\left(\frac{\phi(\mathbf{x_i}) - \phi(\mathbf{X_{ij}})}{h_i} \right) \right\}. \tag{5.5}$$

We will investigate the asymptotic properties and bandwidth selection of our estimate in future study.

5.2 Transformed Multivariate Weighted Kernel Density Estimation

In this section, we generalized transformed weighted kernel density estimate to its multivariate version. Weighted distributions arise in area such as sociological, economical, missing data or damaged data. Let $\mathbf{X}$ be a q dimensional random vector with common probability density $f(\mathbf{X})$. Suppose random vector $\mathbf{X}$ is not observable but we observe another random vector $\mathbf{Y}$ with distribution G and density g which is related to f as

$$g(\mathbf{y}) = \frac{w(\mathbf{y})f(\mathbf{y})}{\theta_w},$$

where $\theta_w = \int w(\mathbf{x})f(\mathbf{x})d\mathbf{x} < \infty$.

Note that weighted distribution makes sense only for nonnegative data. We want to estimate natural multivariate density f of random vector $\mathbf{X}$ when we observe random vector $\mathbf{Y}$ from weighted distribution G. We propose the generalization of transformed weighted kernel density estimate as follows: Let $\mathbf{W_i} = \phi(\mathbf{Y_i})$

$$\hat{f}(\mathbf{X}) = \frac{||\phi'(\mathbf{x})||}{a^q}\left(\sum_{j=1}^{n}\prod_{i=1}^{q}\frac{1}{|w(\phi(\mathbf{Y_{ij}}))|}\right)^{-1}\sum\prod_{i=1}^{q}\frac{1}{|w(\phi(\mathbf{Y_{ij}}))|}K\left(\frac{\phi(\mathbf{y_i})-\phi(\mathbf{Y_{ij}})}{a_i}\right) \tag{5.6}$$

Similarly, An estimate of $F(x)$ is

$$\hat{F}(\mathbf{X}) = \left(\sum_{j=1}^{n}\prod_{i=1}^{q}\frac{1}{|w(\phi(\mathbf{Y_{ij}}))|}\right)^{-1}\sum_{j=1}^{n}\prod_{i=1}^{q}\frac{1}{|w(\phi(\mathbf{Y_{ij}}))|}\mathbb{K}\left(\frac{\phi(\mathbf{y_i})-\phi(\mathbf{Y_{ij}})}{a_i}\right) \tag{5.7}$$

where $\mathbb{K}(\mathbf{u})$ is corresponding distribution function of $K(\mathbf{u})$.

The asymptotic properties, bandwidth selection and applications of the propose estimators are left for the future study.

5.3 Transformed Kernel Mean Residual Life Estimation

Suppose that a subject or a component survived of age t, the remaining life time after t is random. The expected value of this random remaining life time is called the mean residual life(MRL) and it has great interest in many areas including survival analysis, reliability analysis and actuarial science. The MRL is an important criterion for finding an optimal burn-in time for a component.

Let X be a life time random variable with survival function function $\overline{F}(X) > 0$. The residual life random variable at age t is given by $X_t = X - t \backslash X > t$. Then the MRL is define as

$$\mu(t) = E(X - t \backslash X > t) = \frac{\left[\int\limits_t^{\infty} \overline{F}(x)dx\right]}{\overline{F}(t)}$$

where $\mu(0) = \mu = E(X)$.

which can be also written as

$$\mu(t) = \frac{\left(\int\limits_t^{\infty} xf(x)dx\right)}{\overline{F}(t)} - t$$

We like to estimate $\mu(t)$ using transformed kernel density and distribution functions by plugging them in the above equation. The properties and applications are left for the future study.

CHAPTER 6

Appendix

THEOREM 3.1:

(i) $E(\hat{f}(x) \to f(x)$ as $n \to \infty$ such that $na \to \infty$ and for all x, $f(x)$ is continuous.

(ii) $E\{\hat{f}(x) - f(x)\}^2 \to 0$ as $n \to \infty$ such that $na \to \infty$ for all continuity points of f.

(iii) The optimal choice of a is $a_{opt} = \left\{ \dfrac{R(K) \int |\phi'(x)| f(x) dx}{S(G^2, (\phi')^8) n} \right\}^{\frac{1}{5}}$

Proof: (i) Note that

$$\begin{aligned} E(\hat{f}(x)) &= \frac{\phi'(x)}{a} EK\left(\frac{\phi(x) - \phi(x_1)}{a}\right) \\ &= \phi'(x) \int_0^\infty \frac{1}{a} K\left(\frac{\phi(x) - \phi(y)}{a}\right) f(y) dy \\ &= |\phi'(x)| \int K(u) f \circ \phi^{-1}(\phi(x) - au)(\phi^{-1})'(\phi(x) - au) du \end{aligned}$$

But $(\phi^{-1})'$ is continuous and $(\phi^{-1})' \to 0$ as $x \to \infty$. It is bounded and hence, we have

$$E(\hat{f}(x)) \to ||\phi'|(x) f(x) \left[(\phi^{-1})'(\phi(x)\right] \int_{\mathbb{R}} K(u) du = f(x)$$

at every continuity point x of f(x).

(ii) Since the kernel is symmetric about the origin, we only need to look at the Var($\hat{f}$).

So,

$$\begin{aligned} \text{Var}(\hat{f}(x) &= \frac{\phi'^2(x)}{n} Var\left\{\frac{1}{a} K\left(\frac{(\phi(x) - \phi(x_i)}{a}\right)\right\} \\ &= \frac{\phi'^2(x)}{na^2} EK^2\left(\frac{(\phi(x) - \phi(x_1)}{a}\right) \end{aligned}$$

But as in (i), we have,

$$
\begin{aligned}
&\frac{\phi'^2(x)}{na^2} EK^2\left(\frac{(\phi(x)-\phi(x_1)}{a}\right)\\
&\quad= \frac{\phi'^2(x)}{na}\int K^2(u)(f\circ\phi^{-1})(\phi(x)-au)\frac{1}{|\phi'(\phi^{-1}(\phi(x)-au)|}du\\
&\quad\approx \frac{|\phi'(x)|}{na} f(x)\int K^2(u)du\\
&\quad= \frac{|\phi'(x)|}{na} f(x)R(K)
\end{aligned}
$$

(iii) We shall define the argument used in (i) and (ii).

Let $P_n(x,u) = f\circ\phi^{-1}(\phi(x)-au)|(\phi^{-1})'(\phi(x)-au)|$.

$$
\begin{aligned}
= &\left\{f(x) - au(f\circ\phi^{-1})'(\phi(x)) + \frac{a^2u^2}{2}(f\circ\phi^{-1})''(\phi(x)) - ...\right\}\\
&\quad\times\left\{(\phi^{-1})'\phi(x) - au(\phi^{-1})''\phi(x) + \frac{a^2u^2}{2}(\phi^{-1})'''\phi(x) - ...\right\}
\end{aligned}
$$

Let $\phi(x) = lnx = y \Rightarrow \phi^{-1}(y) = e^y = x$ then $(\phi^{-1})'(\phi(x)) = (\phi^{-1})'(y) = (\phi^{-1})''(y) = ... = (\phi^{-1})''''(y) = e^y = e^{lnx} = x$

Hence

$$
\begin{aligned}
= &\left\{f(x) - au(f\circ\phi^{-1})'(\phi(x)) + \frac{a^2u^2}{2}(f\circ\phi^{-1})''(\phi(x)) - ...\right\}\\
&\quad\times\left\{(\phi^{-1})'\phi(x) - au(\phi^{-1})''\phi(x) + \frac{a^2u^2}{2}(\phi^{-1})'''\phi(x) - ...\right\}\\
= &\frac{1}{ax}\int K(u)\left[f(x) - auf'(x)x + \frac{a^2u^2}{2}\{f''(x)x^2 + f'(x)x\}\right]\\
&\qquad\times\left\{x - aux + \frac{a^2u^2}{2}\right\}du\\
\approx &\left[f(x) + \frac{a^2}{2}\mu_2(K)f(x) + \mu_2(K)a^2xf'(x) + \frac{a^2}{2}\mu_2(K)x^2f''(x)\right.\\
&\qquad\left.+\frac{a^2}{2}\mu_2(K)xf'(x)\right].
\end{aligned}
$$

Hence, The bias is

$$
\frac{a^2}{2}\mu_2(K)\left[f(x) + 3xf'(x) + x^2f''(x)\right]. \tag{6.1}
$$

Also,

$$
\begin{aligned}
MSE(\hat{f}(x)) \qquad &= E(\hat{f}(x)-f(x))^2 = var(\hat{f}(x)) + \Big[bias\hat{f}(x)\Big]^2 \\
&= \Big[E\hat{f}^2(x) - \Big(E\hat{f}(x)\Big)^2\Big] + \Big[E\hat{f}(x)-f(x)\Big]^2 \\
&= \frac{(\phi'(x))^2}{na}\left\{\int K^2(u)P_n(x,u)du - a\left[\int K(u)P_n(x,u)du\right]^2\right\} \\
&\qquad + \left\{\int K(u)|\phi'(x)|P_n(x,u)du - f(x)\right\}^2
\end{aligned}
$$

Substitution $p_n(x,u)$ in the above expression will give

$$MSE(\hat{f}(x)) \approx \frac{R(K)}{na}|\phi'(x)|f(x) + \frac{a^4}{4}\frac{\mu_2^2(K)G^2(x)}{(|\phi'|^8(x))}.$$

Where

$$G(x) = f(x)\left\{3(\phi''(x))^2 - \phi'(x)\phi'''(x)\right\} - 3f'(x)\phi'(x)\phi''(x) + f''(x)\{\phi'(x)\}^2.$$

Then,

$$MISE\hat{f}(x) \approx \frac{R(K)}{na}\int|\phi'(x)|f(x)dx + \frac{a^4}{4}\mu_2^2(K)\int S(G^2,|\phi'|^8(x)) \tag{6.2}$$

where $S(G^2,(\phi')^8) = \frac{G^2(x)}{(\phi'(x))^8}dx$, $\Big[Bias\hat{f}(x)\Big]^2 = \frac{a^4}{4}\mu_2^2(K)S(G^2,|\phi'|^8)$ and $Var\hat{f}(x) = \frac{R(K)}{na}|\phi'(x)|f(x)dx$.

Differentiating (6.2) with respect to a and solving for a by setting it to 0 , gives the desired result.

Theorem 3.2 The mean square error of $\hat{F}(x)$ is given by

$$\text{MSE}(\hat{F}(x)) = \frac{F(x)(1-F(x))}{n} - \frac{a}{n}\tau(\mathbb{K},F,\phi) + \left[\frac{a^2}{2}(F\circ\phi^{-1})''(\phi(x))\int z^2K(z)dz\right]^2$$

Proof:

$$MSE\hat{F}(x) = E(\hat{F}(x)-F(x))^2 = V(\hat{F}_n(x)) + (Bias(\hat{F}_n(x))^2$$

Let I=$Bias(\hat{F}(x))$ and II=$V(\hat{F}(x))$

$$I = \int\limits_y \mathbb{K}\left(\frac{\phi(x)-\phi(y)}{a}\right)f(y)dy - F(x)$$

First, consider only $\int\limits_{y} \mathbb{K}\left(\frac{(\phi(x)-\phi(y)}{a}\right) f(y)dy.$

$$
\begin{aligned}
&= F(y)\mathbb{K}\left(\frac{(\phi(x)-\phi(y))}{a}\right)\Big|_{-\infty}^{\infty} + \int F(y)K\left(\frac{\phi(x)-\phi(y)}{a}\right)\frac{\phi'(y)}{a}dy \\
&= 0 + \int F(y)K\left(\frac{\phi(x)-\phi(y)}{a}\right)\frac{\phi'(y)}{a}dy \\
&= \int F(y)K\left(\frac{\phi(x)-\phi(y)}{a}\right)\frac{\phi'(y)}{a}dy
\end{aligned}
$$

Let $z \ = \ \dfrac{\phi(x)-\phi(y)}{a}$ then $\phi(y) = \phi(x) - az.$

$y = \phi^{-1}(\phi(x) - az)$ and so $|\phi'(y)dy| = |-adz| = adz,$

Then

$$
\int F(y)K\left(\frac{\phi(x)-\phi(y)}{a}\right)|\frac{1}{a}\phi(y)|dy = \int F(\phi^{-1}(\phi(x)-az))K(z)dz
$$

By using Taylor series expansion,

$$
\begin{aligned}
&= \int [F\circ\phi^{-1}(\phi(x)) - az(F\circ\phi^{-1})'(\phi(x)) + \frac{a^2z^2}{2}(F\circ\phi^{-1})''(\phi(x))]K(z)dz \\
&= F\circ\phi^{-1}(\phi(x))\int K(z)dz - a(F\circ\phi^{-1})'(\phi(x))\int zK(z)dz \\
&\qquad + \frac{a^2}{2}(F\circ\phi^{-1})''(\phi(x))\int z^2K(z)dz \\
&= F(x) + \frac{a^2}{2}(F\circ\phi^{-1})''(\phi(x))\int z^2K(z)dz,
\end{aligned}
$$

So,

$$
\begin{aligned}
I &= \frac{a^2}{2}(F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz\\
I^2 &= \left[\frac{a^2}{2}(F\circ\phi^{-1})''(\phi(x))\int z^2K(z)dz\right]^2\\
II &= V\hat{F}(x) = V\left\{\frac{1}{n}\sum_{i=1}^{n}\mathbb{K}\left(\frac{(\phi(x)-\phi(X_i)}{a}\right)\right\} = \frac{1}{n}V\mathbb{K}\left(\frac{\phi(x)-\phi(y)}{a}\right)\\
&= \frac{1}{n}\left\{E\mathbb{K}^2\left(\frac{\phi(x)-\phi(y)}{a}\right)-\left(E\mathbb{K}\left(\frac{\phi(x)-\phi(y)}{a}\right)\right)^2\right\}\\
&= \frac{1}{n}\left\{\int\mathbb{K}^2\left(\frac{\phi(x)-\phi(y)}{a}\right)dF(y)-F^2(x)+o\left(\frac{a}{n}\right)\right\}
\end{aligned}
$$

Now, consider $\int\mathbb{K}^2\left(\frac{\phi(x)-\phi(y)}{a}\right)dF(y)$.

Let $u=\left(\frac{(\phi(x)-\phi(y)}{a}\right)$ then $\phi(y)=\phi(x)-au\Rightarrow\phi'(y)dy=-adu$.

$$
\begin{aligned}
&= \frac{1}{n}\left\{F(y)\mathbb{K}\frac{\phi(x)-\phi(y)}{a}\Big|_{-\infty}^{\infty}+2\int(F\circ\phi^{-1})(\phi(x)-au)\mathbb{K}(u)K(u)du\right\}\\
&= \frac{1}{n}\left\{0+2\int(F\circ\phi^{-1})(\phi(x)-au)\mathbb{K}(u)K(u)du\right\}\\
&= \frac{1}{n}\left\{2\int\left(F(x)-au(F\circ\phi^{-1})'\phi(x)+\frac{a^2u^2}{2}\left(F\circ\phi^{-1}\right)''\phi(x)\right)\right.\\
&\qquad\left.\times\mathbb{K}(u)K(u)du\right\}\\
&= \frac{2}{n}F(x)\int\mathbb{K}(u)K(u)du-\frac{2a}{n}(F\circ\phi^{-1})'\phi(x)\int u\mathbb{K}(u)K(u)du\\
&\qquad+\frac{a^2}{2n}\left(F\circ\phi^{-1}\right)''\phi(x)\int u^2\mathbb{K}(u)K(u)du\\
&= \frac{2F(x)}{n}\left(\frac{1}{2}\right)-\frac{a}{n}\tau(\mathbb{K},f,\phi)\\
&= \frac{F(x)}{n}-\frac{a}{n}\tau(\mathbb{K},F,\phi)
\end{aligned}
$$

where

$$\begin{aligned}
\tau(\mathbb{K}, F, \phi) &= 2(F \circ \phi^{-1})'\phi(x) \int u\mathbb{K}(u)K(u)du \\
V\hat{F}(x) &= \frac{F(x)(1-F(x)}{n} - \frac{a}{n}\tau(\mathbb{K}, F, \phi)
\end{aligned}$$

$$\text{Thus,} \qquad MSE(\hat{F}(x)) = \frac{F(x)(1-F(x))}{n} - \frac{a}{n}\tau(\mathbb{K}, F, \phi) + \left[\frac{a^2}{2}(F \circ \phi^{-1})''(\phi(x)) \int z^2 K(z)dz\right]^2$$

Regression Function Estimation

Calculating mean square of the regression estimate.

Note that,

$$E\left(\frac{|\phi'(x)|}{na}\sum_{i=1}^{n} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right) = f_1(x) + \frac{a^2}{2}\mu_2(K)\sqrt{S(G^2, |\phi'|^8)} \tag{6.3}$$

Also,

$$\begin{aligned}
&E\left\{\sum_{i=1}^{n} \frac{|\phi'(x)\phi(Y_i)|}{a} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\} \\
&= \int\int K(z)\phi(y)(f \circ \phi^{-1})(\phi(x) - az, \phi(y))dzdy \\
&= \int (m \circ \phi^{-1})(\phi(x) - az)(f_1 \circ \phi^{-1})(\phi(x) - az)K(z)dz \\
&\approx \int \left\{m(x) - az(m \circ \phi^{-1})'(\phi(x)) + \frac{a^2z^2}{2}(m \circ \phi^{-1})''(\phi(x))\right\} \\
&\qquad \times \left\{f_1(x) - az(f_1 \circ \phi^{-1})'(\phi(x) + \frac{a^2z^2}{2}(f \circ \phi^{-1})''(\phi(x))\right\} K(z)dz \\
&\approx \Bigg[m(x)f_1(x) + \frac{a^2\mu_2(K)}{2}\Bigg(2(f_1 \circ \phi^{-1})'(\phi(x))(m \circ \phi^{-1})'(\phi(x)) \\
&\qquad + f_1(x)(m \circ \phi^{-1})''(\phi(x)) + (f_1 \circ \phi^{-1})''(\phi(x))m(x)\Bigg)\Bigg].
\end{aligned} \tag{6.4}$$

Hence

$$
\begin{aligned}
E(\hat{m}(x)) = m(x) + \frac{a^2\mu_2(K)}{2}\Bigg(& 2(f_1 \circ \phi^{-1})'(\phi(x))(m \circ \phi^{-1})'(\phi(x)) \\
& + f_1(x)(m \circ \phi^{-1})''(\phi(x)) \\
& + (f_1 \circ \phi^{-1})''(\phi(x))m(x)\Bigg)/f_1(x). \qquad (6.5)
\end{aligned}
$$

Next, we find the variance of the estimate.

First,

$$
\begin{aligned}
& V\left\{\sum_{i=1}^{n} \frac{|\phi'(x)\phi(Y_i)|}{na} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\} \\
&= \frac{(\phi'(x))^2}{na^2} V\left\{|\phi(Y_1)|K\left(\frac{\phi(x)-\phi(X_1)}{a}\right)\right\} \\
&\approx \frac{(\phi'(x))^2}{na^2} E\left\{|\phi(Y_1)|K\left(\frac{\phi(x)-\phi(X_1)}{a}\right)\right\}^2 \\
&= \frac{(\phi'(x))^2}{na^2} \int\int (\phi(y))^2 K^2\left(\frac{\phi(x)-\phi(x_1)}{a}\right) f(x,y)dxdy \\
&= \frac{|\phi'(x)|}{na} \int\left\{\int (\phi(y^*))^2 f\left(\phi(y^*)\backslash(\phi(x)-a\phi(x^*)\right)dy^*\right\} \\
&= \qquad \times f(x-ax^*)K^2(x^*)dx^* \\
&\approx \frac{|\phi'(x)|}{na} R(K) f_1(x)\psi_w(\phi(x)). \qquad (6.6)
\end{aligned}
$$

where $\psi_w(\phi(x)) = E\left\{\phi^2(Y)\backslash\phi(X=x)\right\}$. from earlier result, we have

$$
\begin{aligned}
& V\left\{\sum_{i=1}^{n} \frac{|\phi'(x)}{na} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\} \\
&= \frac{|\phi'(x)|R(K)f_1(x)}{na}. \qquad (6.7)
\end{aligned}
$$

Also,

$$
\begin{aligned}
& Cov\left\{\frac{|\phi'(x)|}{na}\sum_{i=1}^{n} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right), \frac{|\phi'(x)\phi(Y_i)|}{na}\sum_{i=1}^{n} K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\} \\
&\approx \frac{|\phi'(x)|\theta_w}{na} R(K) f_1(x)\eta_w(x). \qquad (6.8)
\end{aligned}
$$

where $\eta_w(\phi(x)) = E\{\phi(Y)\backslash\phi(X=x)\}$. Hence, putting (6.5),(6.6),(6.7) and (6.8) into V(U/V) and simplifying, we get that

$$V\hat{m}(x) = \frac{R(K)|\phi'(x)|\left[\psi_w(\phi(x)) + m^2(x) - 2m(x)\eta_w(\phi(x))\right]}{nam^2(x)f_1(x)}. \tag{6.9}$$

Hence

$$\begin{aligned} &E\left(\hat{m}(x) - m(x)\right)^2 \\ &\quad = \frac{R(K)|\phi'(x)|}{nam^2(x)f_1(x)}\left[\psi_w(\phi(x)) + m^2(x) - 2m(x)\eta_w(\phi(x))\right] \\ &\quad + \Bigg[\frac{a^2\mu_2(K)}{2}\Bigg(2(f_1\circ\phi^{-1})'(\phi(x))(m\circ\phi^{-1})'(\phi(x)) \\ &\qquad + f_1(x)(m\circ\phi^{-1})''(\phi(x))\Bigg)/f_1(x)\Bigg]^2 \end{aligned} \tag{6.10}$$

Thus, IMSE is obtained by integrating (6.10) with respect to x, we assume that all integrals exist.

Univariate Failure Rate Function

Theorem 3.3 Expected value of $\hat{h}(x)$ is

$$E\hat{h}(x) = h(x)\left(1 + \frac{a^2}{2}\left\{\frac{\sqrt{S(G^2, |\phi'(x)|^8)}}{f(x)} - (F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz\right\}\right)$$

Proof:

$$\begin{aligned} E\hat{h}(x) &= \frac{E\hat{f}(x)}{E\hat{\overline{F}}(x)} = \frac{f(x) + bias\hat{f}(x)}{\overline{F}(x) + bias\hat{\overline{F}}(x)} = \frac{f(x)}{\overline{F}(x)}\frac{1 + bias\frac{\hat{f}(x)}{f(x)}}{1 - bias\frac{\hat{F}(x)}{\overline{F}(x)}} \\ &= h(x)\left[1 + \frac{bias\hat{f}(x)}{f(x)}\right]\left[1 + \frac{bias\hat{F}(x)}{\overline{F}(x)}\right] \\ &= h(x)\left[1 + \frac{bias\hat{f}(x)}{f(x)} + \frac{bias\hat{F}(x)}{\overline{F}(x)} + bias\frac{\hat{f}(x)}{f(x)}bias\frac{\hat{F}(x)}{\overline{F}(x)}\right] \\ &= h(x)\left[1 + \frac{a^2}{2}\frac{\sqrt{S(G^2, |\phi'(x)|^8)}}{f(x)} + \frac{a^2(F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz)}{2\overline{F}(x)}\right] \\ &\quad + h(x)\left[\frac{a^2(F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz)}{2\overline{F}(x)} * \frac{a^2}{2}\frac{\sqrt{S(G^2, |\phi'(x)|^8)}}{f(x)}\right] \end{aligned}$$

Ignoring a^4 term, we have

$$\approx \quad h(x)\left[1+\frac{a^2}{2}\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)} + \frac{(F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz}{2\overline{F}(x)}\right\}\right]$$

Thus,

$$\begin{aligned}
E\hat{h}(x) &= [h(x)+h(x)(I_1+I_2)] \\
bias(\hat{h}(x)) &= \frac{a^2}{2}\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)} + \frac{(F\circ\phi^{-1})''|\phi(x)|\int z^2K(z)dz}{2\overline{F}(x)}\right\}
\end{aligned}$$

Theorem 3.4 $V\hat{h}(x) = \frac{h^2(x)|\phi'(x)|R(K)}{naf(x)}$

Proof:

$$V(\hat{h}(x)) = V\left(\frac{\hat{f}(x)}{\hat{\overline{F}}(x)}\right) = \frac{[E\hat{f}(x)]^2}{[E\hat{\overline{F}}(x)]^2}\left[\frac{V\hat{f}(x)}{[E\hat{f}(x)]^2} + \frac{V\hat{\overline{F}}(x)}{[E\hat{\overline{F}}(x)]^2} - \frac{2cov(\hat{f}(x),\hat{\overline{F}}(x))}{E\hat{f}(x)E\hat{\overline{F}}(x)}\right]$$

First, by theorem 3,

$$\begin{aligned}
\frac{[E\hat{f}(x)]^2}{[E\hat{\overline{F}}(x)]^2} &= [h(x)+h(x)(I_1+I_2)]^2 \\
&= h^2(x)\left[1+\frac{a^2}{2}\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)} + \frac{(F\circ\phi^{-1})''\phi(x)R(K)}{2\overline{F}(x)}\right\}\right]^2 \\
&\approx h^2(x)\left[1+a^2\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)} + \frac{(F\circ\phi^{-1})''\phi(x)R(K)}{2\overline{F}(x)}\right\}\right]
\end{aligned}$$

Second,

$$\begin{aligned}
\frac{V\hat{f}(x)}{[E\hat{f}(x)]^2} &= \frac{\frac{|\phi'(x)|f(x)\int K^2(u)du}{na}}{[f(x)+bias\hat{f}(x)]^2} = \frac{|\phi'(x)|f(x)R(K)}{na\left[f(x)+\frac{a^2}{2}\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)}\right]^2} \\
&= \frac{|\phi'(x)|R(K)}{naf(x)\left[1+\frac{a^2\sqrt{S(G^2,|\phi'(x)|^8)}}{2f(x)}\right]^2} \\
&\approx \frac{|\phi'(x)|R(K)}{naf(x)}
\end{aligned}$$

Third,

$$\begin{aligned}
\frac{V\hat{\overline{F}}(x)}{[E\hat{\overline{F}}(x)]^2} &= \frac{\frac{F(x)(1-F(x))}{n} - \frac{a}{n}\tau(\mathbb{K},F,\phi)}{[\overline{F}(x) - bias\hat{F}(x)]^2} \\
&= \frac{\frac{F(x)(1-F(x)}{n} - \frac{a}{n}\tau(\mathbb{K},F,\phi)}{\left[\overline{F}(x) - \frac{a^2}{2}(F\circ\phi^{-1})''(\phi(x))\int z^2K(z)dz\right]^2} \\
&\approx \frac{F(x)(1-F(x))}{n\overline{F}(x)} - \frac{a\tau(\mathbb{K},F,\phi)}{n\overline{F}(x)}
\end{aligned}$$

Fourth,

$$\frac{2cov(\hat{f}(x),\hat{\overline{F}}(x))}{E\hat{f}(x)E\hat{\overline{F}}(x)} = 2\left[\frac{E[\hat{f}(x)\hat{\overline{F}}(x)] - E\hat{f}(x)E\hat{\overline{F}}(x)}{E\hat{f}(x)E\hat{\overline{F}}(x)}\right] = 2\left[\frac{E[\hat{f}(x)\hat{\overline{F}}(x)]}{E\hat{f}(x)E\hat{\overline{F}}(x)} - 1\right]$$

Now,

$$\begin{aligned}
& E[\hat{f}(x)\hat{\overline{F}}(x)] \\
=\ & E\left\{\frac{\phi'(x)}{na}\sum_{i=1}^{n}K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\}\left\{\frac{1}{n}\sum_{j=1}^{n}\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_j)}{a}\right)\right\} \\
=\ & \frac{|\phi'(x)|}{n^2a}\sum_{i}E\left\{K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_j)}{a}\right)\right\} \\
=\ & \frac{|\phi'(x)|}{n^2a}\sum_{i}E\left\{K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\} \\
& \quad + \frac{|\phi'(x)|}{n^2a}\sum_{i\neq j}\sum EK\left(\frac{\phi(x)-\phi(X_i)}{a}\right)E\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_j)}{a}\right)
\end{aligned}$$

Let $I_1 = \frac{|\phi'(x)|}{n^2a}E\left\{K\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\}$

and

$I_2 = \frac{|\phi'(x)|}{n^2a}EK\left(\frac{\phi(x)-\phi(X_i)}{a}\right)E\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_j)}{a}\right)$.

Working on I_1 first,

$$\begin{aligned}
I_1 &= \frac{|\phi'(x)|}{n^2 a} E\left\{ K\left(\frac{\phi(x)-\phi(X_i)}{a}\right) \overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_i)}{a}\right)\right\} \\
&= \int \frac{|\phi'(x)|}{na} K\left(\frac{\phi(x)-\Phi(y)}{a}\right) \overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(y)}{a}\right) f(y)dy
\end{aligned}$$

Let

$$z = \left(\frac{\phi(x)-\phi(y)}{a}\right) \Rightarrow \phi(y) = \phi(x) - az \Rightarrow y = \phi^{-1}(\phi(x)-az).$$

Then,

$$|\phi'(y)dy| = |-adz| \Rightarrow dy = \frac{adz}{|\phi'(y)|} = \frac{adz}{|\phi'[\phi^{-1}(\phi(x)-az)]|}$$

So,

$$\begin{aligned}
I_1 &= \frac{\phi'(x)}{na} \int K(z)\overline{\mathbb{K}}(z) f(\phi^{-1}(\phi(x)-az)) \frac{adz}{\phi'[\phi^{-1}(\phi(x)-az)]} \\
&= \frac{\phi'(x)}{n} \int K(z)\overline{\mathbb{K}}(z) \frac{f\circ\phi^{-1}(\phi(x)-az)}{\phi'[\phi^{-1}(\phi(x)-az)]} dz
\end{aligned}$$

By Taylor series expansion,

$\phi^{-1}(\phi(x)-az) \approx x - (\phi^{-1})'(x)\phi(x)(-az) + \phi^{-1''}\phi(x)\frac{a^2z^2}{2} \approx x$

So, $\phi'[\phi^{-1}(\phi(x)-az)] \approx \phi'(x)$

Thus,

$$\begin{aligned}
I_1 &= \frac{\phi'(x)}{n} \int K(z)\overline{\mathbb{K}}(z) \frac{f\circ\phi^{-1}(\phi(x)-az)}{\phi'(x)} dz \\
&= \frac{1}{n} \int K(z)\overline{\mathbb{K}}(z) \eta(\phi(x)-az) dz
\end{aligned}$$

where

$$\eta(\phi(x)-az) = f\circ\phi^{-1}(\phi(x)-az),$$

Again, by Tayler series expansion,

$$\begin{aligned}
I_1 &= \frac{1}{n}\int K(z)\overline{\mathbb{K}}(z)[\eta(\phi(x)) - az\eta'\phi(x) + \frac{a^2z^2}{2}\eta''(\phi(x)]dz \\
&\approx \frac{\eta(\phi(x))}{n}\int K(z)\overline{\mathbb{K}}(z)dz - \frac{a}{n}\eta'(\phi(x))\int zK(z)\overline{\mathbb{K}}(z)dz \\
&= \frac{\eta(\phi(x))}{n}\left[\int K(z)dz - \int \mathbb{K}(z)d\mathbb{K}(z)\right] - \frac{a}{n}\eta'(\phi(x))\int zK(z)\overline{\mathbb{K}}(z)dz \\
&= \frac{\eta(\phi(x))}{n}(1-\frac{1}{2}) - \frac{a}{n}\eta'(\phi(x))\int zK(z)\overline{\mathbb{K}}(z)dz \\
&= \frac{\eta(\phi(x))}{2n} - S(K) \\
\text{where } S(K) &= \frac{a}{n}\eta'(\phi(x))\int zK(z)\overline{\mathbb{K}}(z)dz \text{ is a known quantity.}
\end{aligned}$$

Now,

$$I_2 = \frac{\phi'(x)}{n^2a}\sum\sum_{i\neq j} EK\left(\frac{\phi(x)-\phi(X_i)}{a}\right)E\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(X_j)}{a}\right)$$

If $i \neq j$ then X_i and X_j are independent. Thus,

$$\begin{aligned}
I_2 &= \frac{|\phi'(x)|}{n^2a}n(n-1)\left\{EK\left(\frac{\phi(x)-\phi(y)}{a}\right)E\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(y)}{a}\right)\right\} \\
&= \frac{|\phi'(x)|}{na}(n-1)\left\{EK\left(\frac{\phi(x)-\phi(y)}{a}\right)E\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(y)}{a}\right)\right\}
\end{aligned}$$

Using earlier results, we have

$$\begin{aligned}
&E\frac{\phi'(x)}{a}K\left(\frac{\phi(x)-\phi(y)}{a}\right) \approx f(x) \text{ and} \\
&E\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(y)}{a}\right) \\
&= 1 - E\mathbb{K}\left(\frac{\phi(x)-\phi(y)}{a}\right) \approx 1 - F(x) = \overline{F}(x)
\end{aligned}$$

Thus,

$E[\hat{f}(x)\hat{\overline{F}}(x)] = \frac{\eta(\phi(x))}{2n} - S(K) + \frac{n-1}{n}f(x)(1-F(x)),$

and

$$
\begin{aligned}
\frac{2cov(\hat{f}(x),\hat{\overline{F}}(x))}{E\hat{f}(x)E\hat{\overline{F}}(x)} &= 2\left[\frac{E[\hat{f}(x)\hat{\overline{F}}(x)]}{E\hat{f}(x)E\hat{\overline{F}}(x)}-1\right] \\
\approx\ 2\left[\frac{\frac{\eta(\phi(x))}{2n}-S(K)+\frac{n-1}{n}f(x)(1-F(x))}{f(x)\overline{F}(x)}-1\right] \\
=\ 2\left[\frac{\eta(\phi(x))}{2nf(x)\overline{F}(x)}-\frac{S(K)}{f(x)\overline{F}(x)}+1-\frac{1}{n}-1\right] \\
\approx\ \frac{\eta(\phi(x))}{n\overline{F}(x)f(x)}-\frac{2S(K)}{\overline{F}(x)f(x)}-\frac{1}{n}
\end{aligned}
$$

Combining all the results, we have

$$
\begin{aligned}
V\hat{h}(x) \quad &\approx h^2(x)\left[1+a^2\left\{\frac{\sqrt{S(G^2,|\phi'(x)|^8)}}{f(x)}+\frac{(F\circ\phi^{-1})''\phi(x)R(K)}{2\overline{F}(x)}\right\}\right]\times \\
&\left\{\frac{|\phi'(x)|R(K)}{naf(x)}+\frac{F(x)}{n}+\frac{a\tau(\mathbb{K},F,\phi)}{n\overline{F}(x)}-\frac{\eta(\phi(x))}{n\overline{F}(x)f(x)}\right. \\
&\left.-\frac{2S(K)}{\overline{F}(x)f(x)}-\frac{1}{n}\right\} \\
&\approx h^2(x)\left\{\frac{|\phi'(x)|R(K)}{naf(x)}+\frac{F(x)}{n}+\frac{a\tau(\mathbb{K},F,\phi)}{n\overline{F}(x)}-\frac{\eta(\phi(x))}{n\overline{F}(x)f(x)}\right. \\
&\left.-\frac{2S(K)}{\overline{F}(x)f(x)}-\frac{1}{n}\right\} \\
&\approx \frac{h^2(x)|\phi'(x)|R(K)}{naf(x)}
\end{aligned}
$$

Theorem 3.5

$$
\begin{aligned}
E\hat{h}(\mathbf{x}) = \Bigg[\quad &\left\{h_1(\mathbf{x})\Big(1+\frac{bias\hat{g}_1(\mathbf{x})}{g_1(\mathbf{x})}+\frac{bias\hat{\overline{F}}(\mathbf{x})}{\overline{F}(\mathbf{x})}\Big)\right\} \\
&,\left\{h_2(\mathbf{x})\Big(1+\frac{bias\hat{g}_2((\mathbf{x})}{g_2(\mathbf{x})}+\frac{bias\hat{\overline{F}}(\mathbf{x})}{\overline{F}(\mathbf{x})}\Big)\right\}\Bigg]'
\end{aligned}
$$

where

$$bias\hat{F}(\mathbf{x}) = \frac{1}{2}\Bigg[a_1^2 z_1^2 F((\phi_1^{-1})''(\phi(x_1)), (\phi_2^{-1})''(\phi(x_2))$$

$$+a_2^2 z_2^2 F((\phi_1^{-1})''(\phi(x_1)), (\phi_2^{-1})''(\phi(x_2))\Bigg] \times K_1(z_1)K_2(z_2)dz_1dz_2$$

$$bias\hat{g}_1(\mathbf{x}) = \frac{1}{2}\int_{x_2}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)$$

$$+a_2^2 z_2^2 f\left((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)\Bigg]dz_1dz_2dy$$

$$bias\hat{g}_2(\mathbf{x}) = \frac{1}{2}\int_{x_1}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)$$

$$+a_2^2 z_2^2 f\left((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)\Bigg]dz_1dz_2dy$$

Proof: We derive for $E\hat{h}_1(\mathbf{x})$, $E\hat{h}_2(\mathbf{x})$, follows exactly.

$E\hat{h}_1(\mathbf{x}) = \frac{E\hat{g}_1(\mathbf{x})}{E\hat{\hat{F}}(\mathbf{x})}$

First, $E\hat{g}_1(\mathbf{x}) = E\int_{x_2}^{\infty}\hat{f}(x_1,y)dy = \int_{x_2}^{\infty}\left[E\hat{f}(x_1,y)\right]dy$

Let's consider

$$E\hat{f}(x_1,y) = \int\int\frac{|\phi'_{x_1}(x_1)\phi'_y(y)|}{a_1a_2}K_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right)\times$$
$$K_2\left(\frac{\phi_2(y)-\phi_2(v)}{a_2}\right)f(u,v)dudv$$

Using Taylor series expansion, we get

$$\hat{f}(x_1, y) = f(x_1, y) + \frac{1}{2}\int\int\Bigg[\quad a_1^2 z_1^2 f\left(((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)$$

$$+a_2^2 z_2^2 f\left(((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)\Bigg] dz_1 dz_2$$

$$E\hat{g}_1((\mathbf{x})) = \int_{x_2}^{\infty} E\hat{f}(x_1, y) dy$$

$$= g_1(x_1, x_2) + \frac{1}{2}\int_{x_2}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left(((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)$$

$$+a_2^2 z_2^2 f\left(((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)\Bigg] dz_1 dz_2 dy$$

$$bias\hat{g}_1(\mathbf{x}) = \frac{1}{2}\int_{x_2}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left(((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)$$

$$+a_2^2 z_2^2 f\left(((\phi_1^{-1})''(\phi_1(x_1)), (\phi_y^{-1})''(\phi_2(y))\right)\Bigg] dz_1 dz_2 dy$$

Similarly,

$$\hat{f}(y, x_2) = f(y, x_2) + \frac{1}{2}\int\int\Bigg[\quad a_1^2 z_1^2 f\left(((\phi_1^{-1})''(\phi_2(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)$$

$$+a_2^2 z_2^2 f\left(((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)\Bigg] dz_1 dz_2$$

$$E\hat{g}_2((\mathbf{x})) = \int_{x_1}^{\infty} E\hat{f}(y, x_2) dy$$

$$= g_2(x_1, x_2) + \frac{1}{2}\int_{x_1}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left(((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)$$

$$+a_2^2 z_2^2 f\left(((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)\Bigg] dz_1 dz_2 dy$$

$$bias\hat{g}_2(\mathbf{x}) = \frac{1}{2}\int_{x_1}^{\infty}\int\int\Bigg[\quad a_1^2 z_1^2 f\left((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)$$

$$+a_2^2 z_2^2 f\left(((\phi_1^{-1})''(\phi_1(y)), (\phi_2^{-1})''(\phi_2(x_2))\right)\Bigg] dz_1 dz_2 dy$$

now,

$$
\begin{aligned}
E\hat{F}(\mathbf{x}) &= \int\int \mathbb{K}_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right)\mathbb{K}_2\left(\frac{\phi_2(x_2)-\phi_2(v)}{a_2}\right)f(u,v)dudv \\
&= \int\Bigg[\quad F_u(u,v)\mathbb{K}_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right)\Big|_{-\infty}^{\infty} + \int F_u(u,v) \\
&\qquad \times K_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right)\frac{|\phi'(u)|}{a_2}du\Bigg] \times \mathbb{K}_2\left(\frac{\phi_2(x_2)-\phi_2(v)}{a_2}\right)dv \\
&= \int F_u(u,v)\mathbb{K}_2\left(\frac{\phi_2(x_2)-\phi_2(v)}{a_2}\right)dv \int K_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right) \\
&\qquad \times \frac{|\phi'(u)|}{a_1}du \\
&= \int\Bigg[\quad F_{uv}(u,v)\mathbb{K}_2\left(\frac{\phi_2(x_2)-\phi_2(v)}{a_2}\right)\Big|_{-\infty}^{\infty} + \int F_{uv}(u,v) \\
&\qquad \times K_2\left(\frac{\phi_2(x_2)-\phi_2(v)}{a_2}\right)\frac{|\phi'(v)|}{a_2}dv\Bigg] \times K_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right) \\
&\qquad \times \frac{|\phi'(u)|}{a_1}du \\
&= \int\int F_{uv}(u,v)K_1\left(\frac{\phi_1(x_1)-\phi_1(u)}{a_1}\right)K_2\left(\frac{\phi_2(x_2)-\phi_2(v)}{a_2}\right)\frac{|\phi'(u)|}{a_1} \\
&\qquad \times \frac{|\phi'(v)|}{a_2}dudv
\end{aligned}
$$

Using Taylor series expansion,

$$
\begin{aligned}
&\int\int F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)))K_1(z_1)K_2(z_2)dz_1dz_2 \\
&\approx \int\int\Bigg[\quad F(x_1,x_2)+\frac{1}{2}\Bigg[\quad a_1^2z_1^2F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)) \\
&\qquad +a_2^2z_2^2F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)\Bigg]\Bigg]K_1(z_1)K_2(z_2)dz_1dz_2 \\
&= F(x_1,x_2)+\frac{1}{2}\Bigg[\quad a_1^2z_1^2F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)) \\
&\qquad +a_2^2z_2^2F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)\Bigg]K_1(z_1)K_2(z_2)dz_1dz_2
\end{aligned}
$$

Thus, bias$\hat{F}(\mathbf{x})$ is given by

$$\frac{1}{2}\left[a_1^2 z_1^2 F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)+a_2^2 z_2^2 F((\phi_1^{-1})''(\phi(x_1)),(\phi_2^{-1})''(\phi(x_2)\right]$$
$$\times K_1(z_1)K_2(z_2)dz_1dz_2$$

TWKDE

Asymptotic MSE and IMSE of the Estimates

We study MSE and IMSE by evaluating the mean and variance of the estimates. First,

Let

$$\begin{aligned}\hat{f}(x) &= \frac{|\phi'(x)|}{a}\left[\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}\right]^{-1}\sum_{i=1}^{n}(|w(\phi(Y_i)|)^{-1}K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\\ &= \frac{U_{n,a}(x)}{V_n} \qquad (6.11)\end{aligned}$$

We know,

$$E(U/V)\approx E(U)/E(V),\quad \text{and}$$
$$V(U/V)\approx\left[\frac{E(U)}{E(V)}\right]^2\left(\frac{V(U)}{(E(U))^2}+\frac{V(V)}{(E(V))^2}-2\frac{Cov(U,V)}{E(U)E(V)}\right)$$

So,

$$E\hat{f}(x)\approx\frac{EU_{n,a}(x)}{EV_n}=\theta_w EU_{n,a}(x)$$

But

$$\begin{aligned}EU_{n,a}(x) &= E\left\{\frac{|\phi'(x)|}{a}(|w(\phi(Y_1)|)^{-1}K\left(\frac{\phi(x)-\phi(Y_1)}{a}\right)\right\}\\ &= \int\frac{|\phi'(x)|}{a}(|w(\phi(Y_1)|)^{-1}K\left(\frac{\phi(x)-\phi(y)}{a}\right)g(y)dy\\ &= \int\theta_w^{-1}\frac{|\phi'(x)|}{a}K\left(\frac{\phi(x)-\phi(y)}{a}\right)f(y)dy\end{aligned}$$

Thus,

$$EU_{n,a}(x)\approx\theta_w^{-1}\left\{f(x)+\frac{a^4}{4}\mu_2^2(K)S(G^2,|\phi'(x)|^8)\right\}. \qquad (6.12)$$

where $S(G^2, |\phi'|^8(x))$ is defined in theorem 3.1.

Hence

$$E\hat{f}(x) \approx f(x) + \frac{a^2}{2}\mu_2(K)\sqrt{S(G^2, |\phi'(x)|^8)}. \tag{6.13}$$

Where $\mu_2(K) = \int u^2 K(u)du$ and we use the fact that $\int uK(u)du = 0$. Next,

$$\begin{aligned}
V(U_{n,a}(x)) &\approx \frac{\phi'(x)|^2}{na^2} E\left\{(w(\phi(Y_1)))^{-2}K^2\left(\frac{\phi(x)-\phi(Y_1)}{a}\right)\right\} \\
&= \frac{\phi'(x)|^2}{na^2}\int (w(\phi(y)))^{-2}K^2\left(\frac{\phi(x)-\phi(y)}{a}\right)g(y)dy \\
&= \frac{\phi'(x)|^2}{\theta_w na^2}\int |(w(\phi(y)))|^{-1}K^2\left(\frac{\phi(x)-\phi(y)}{a}\right)f(y)dy \\
&= \frac{\phi'(x)|}{\theta_w na}\int |(w(\phi(x)-au)|^{-1}K^2(u)f(\phi(x)-au)du \\
&\approx \frac{\phi'(x)|(f(x)}{\theta_w na(w(\phi(x)))}\int K^2(u)du
\end{aligned}$$

Hence

$$V(U_{n,a}(x)) = \frac{|\phi'(x)|(f(x)R(K)}{\theta_w na(w(\phi(x)))}. \tag{6.14}$$

Also

$$\begin{aligned}
\frac{V(U_{n,a}(x))}{(E(U_{n,a}(x)))^2} &= \frac{\left\{\frac{|\phi'(x)|(f(x)R(K)}{\theta_w na(w(\phi(x)))}\right\}}{\left\{\theta_w^{-1}\left\{f(x) - \int\frac{a^2u^2K(u)(f\circ\phi^{-1})''(\phi(x))du}{2}\right\}\right\}^2} \\
&\approx \frac{\theta_w|\phi'(x)|R(K)}{f(x)naw(\phi(x))} \qquad (6.15)
\end{aligned}$$

Where $R(K) = \int K^2(u)du$. Next

$$nV(V_n) = V[(w\phi(Y_1))^{-1}] = \int (w(\phi(y))^{-2}g(y)dy - \theta_w^{-2} = \theta_w^{-1}\int (w(y))^{-1}f(y)dy - \theta_w^{-2}.$$

Thus

$$nV(V_n) = \theta_w^{-1}\left(\int (w(y))^{-1}f(y)dy - \theta_w^{-1}\right) = \eta(\theta_w, \nu_w)(suppose) \tag{6.16}$$

Hence

$$\frac{V(V_n)}{(E(V_n))} = \frac{\theta_w^2 \eta(\theta_w, \nu_w)}{n} \tag{6.17}$$

Finally,

$$\begin{aligned}
nCov(U_{n,a}(x), V_n) &= n\left\{E(U_{n,a}(x), V_n) - E(U_{n,a}(x))E(V_n)\right\} \\
&= E\left\{\frac{|\phi'(x)|}{a}(|w(\phi(Y_1)|)^{-2}K\left(\frac{\phi(x)-\phi(Y_1)}{a}\right)\right\} \\
&\qquad -\theta_w^{-1}\left\{f(x) - \int \frac{a^2u^2K(u)(f\circ\phi^{-1})''(\phi(x))du}{2}\right\}\theta_w^{-1} \\
&\approx \frac{|\phi'(x)|}{a}\int(|w(\phi(y)|)^{-2}K\left(\frac{\phi(x)-\phi(y)}{a}\right)g(y)dy \\
&= \frac{|\phi'(x)|}{a\theta_w}\int(|w(\phi(y)|)^{-1}K\left(\frac{\phi(x)-\phi(y)}{a}\right)f(y)dy.
\end{aligned}$$

Hence

$$Cov(U_{n,a}(x), V_n) \approx \frac{|\phi'(x)|f(x)}{n\theta_w|w(\phi(x))|} \tag{6.18}$$

Combining equations 6.17, 6.19, 6.20 and 6.18 into V(U/V) and simplifying gives

$$V(\hat{f}(x)) = \frac{f(x)\theta_w|\phi'(x)|R(K)}{na|w(\phi(x))|} + o\left(\frac{1}{na}\right). \tag{6.19}$$

Hence from equations 6.17 and 6.19, we have

$$E(\hat{f}(x) - f(x))^2 = \frac{f(x)\theta_w|\phi'(x)|R(K)}{na|w(\phi(x))|} + \frac{a^4}{4}\mu_2^2(K)S(G^2, |\phi'(x)|^8). \tag{6.20}$$

The optimal bandwidth a_{opt} is given by

$$a_{opt} = \left\{\frac{f(x)\theta_w|\phi'(x)|R(K)}{n|w(\phi(x))|\mu_2^2(K)S(G^2, |\phi'(x)|^8)}\right\}^{\frac{1}{5}}. \tag{6.21}$$

In similar manner, we derive MSE and IMSE of $\hat{F}(x)$.

$$\begin{aligned}
\hat{F}(x) &= \left(\sum \frac{1}{|w(\phi(Y_i))|}\right)^{-1} \sum \frac{1}{|w(\phi(Y_i))|}\mathbb{K}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right) \\
E(\hat{F}(x) &= E(U_{n,a}(x)/V_n) \approx \theta_w E\left[\frac{1}{|w(\phi(Y_i))|}\mathbb{K}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\right] \\
&\approx F(x) + \frac{a^2\mu_2(K)(F\circ\phi^{-1})''(\phi(x))}{2}
\end{aligned}$$

Thus,

$$E(\hat{F}(x)) = F(x) + \frac{a^2\mu_2(K)(F\circ\phi^{-1})''(\phi(x))}{2} \tag{6.22}$$

Next

$$\begin{aligned}
nV(U_{n,a}(x) &\approx E\left(|w(\phi(Y_1)))^{-2}\mathbb{K}\left(\frac{\phi(x)-\phi(Y_1)}{a}\right)\right) - F^2(x) \\
&= \int (|w(\phi(y)))^{-2}\mathbb{K}^2\left(\frac{\phi(x)-\phi(y)}{a}\right) g(y)dy - F^2(x) \\
&= \theta_w^{-1}\int (|w(\phi(y))|)^{-1}\mathbb{K}^2\left(\frac{\phi(x)-\phi(y)}{a}\right) f(y)dy - F^2(x) \\
&\approx \frac{1}{\theta_w|w(\phi(x))|}\left[F(x)(1-F(x) - a\tau(\mathbb{K},F,\phi)\right]
\end{aligned}$$

Hence

$$V(U_{n,a}(x)) = \frac{1}{n\theta_w|w(\phi(x))|}\left[F(x)(1-F(x)) - a\tau(\mathbb{K},F,\phi)\right] \tag{6.23}$$

Where

$$\tau(\mathbb{K},F,\phi) = 2(F\circ\phi^{-1})'(\phi(x))\int u\mathbb{K}(u)K(u)du$$

Next

$$\begin{aligned}
nCov((U_{n,a}(x), V_n) &= \{E(U_{n,a}(x), V_n) - E(U_{n,a}(x))E(V_n)\} \\
&\approx E\left\{(|w(\phi(Y_1)|)^{-2}\mathbb{K}\left(\frac{\phi(x)-\phi(Y_1)}{a}\right)\right\} \\
&= \int(|w(\phi(y)|)^{-2}\mathbb{K}\left(\frac{\phi(x)-\phi(y)}{a}\right)g(y)dy \\
&= \theta_w^{-1}\int(|w(\phi(y)|)^{-1}\mathbb{K}\left(\frac{\phi(x)-\phi(y)}{a}\right)f(y)dy \\
&\approx \frac{1}{\theta_w|w(\phi(x))|}\left(F(x)+\frac{a^2}{2}(F\circ\phi^{-1})''\mu_2(K)\right)
\end{aligned}$$

Thus

$$Cov((U_{n,a}(x), V_n) = \frac{1}{n\theta_w|w(\phi(x))|}\left(F(x)+\frac{a^2}{2}(F\circ\phi^{-1})''\mu_2(K)\right) \tag{6.24}$$

Combining all terms into V(U/V) and simplifying gives

$$\begin{aligned}
Var\hat{F}(x) = \frac{\theta_w}{n|w(\phi(x))|}\Bigg[&(F(x)(1-F(x)) - a\tau(\mathbb{K},F,\phi)) \\
&+|w(\phi(x))|F^2(x)\theta_w\eta(\theta_w,\nu_w) - \frac{2F^3(x)}{\theta_w^2}\Bigg]
\end{aligned} \tag{6.25}$$

Where $\tau(\mathbb{K},F,\phi), \eta(\theta_w,\nu_w)$ are as defined before. Hence

$$\begin{aligned}
E(\hat{F}(x)-F(x))^2 = \frac{\theta_w}{n|w(\phi(x))|}\Bigg[&(F(x)(1-F(x)) - a\tau(\mathbb{K},F,\phi)) \\
&+|w(\phi(x))|F^2(x)\theta_w\eta(\theta_w,\nu_w) - \frac{2F^3(x)}{\theta_w^2}\Bigg] \\
&+\frac{a^4\mu_2^2(K)((F\circ\phi^{-1})''(\phi(x)))^2}{4}
\end{aligned} \tag{6.26}$$

The optimal bandwidth a^*_{opt} is given by

$$a^*_{opt} = \left\{\frac{\theta_w\tau(\mathbb{K},F,\phi)}{n|w(\phi(x))|\mu_2^2(K)((F\circ\phi^{-1})''(\phi(x)))^2}\right\}^{\frac{1}{3}} \tag{6.27}$$

Also

$$\int E(\hat{F}(x) - F(x))^2 = \frac{\theta_w}{n} \int \Bigg[\frac{F(x)}{|w(\phi(x))|} - \frac{a\tau(\mathbb{K}, F, \phi)}{|w(\phi(x))|}$$
$$-F^2(x) \left(\frac{1}{|w(\phi(x))|} - \theta_w \eta(\theta_w, \nu_w) + \frac{2F(x)}{|w(\phi(x))|\theta_w^2} \right)$$
$$+ \frac{a^4 \mu_2^2(K)}{4} ((F \circ \phi^{-1})''(\phi(x)))^2 \Bigg] dx \tag{6.28}$$

Then IMSE is given by

$$\int E(\hat{f}(x) - f(x))^2 dx = \frac{\theta_w |\phi'(x)| R(K)}{na} \int |w(\phi(x)|^{-1} f(x) dx$$
$$+ \frac{a^4}{4} \mu_2^2(K) \int S(G^2, |\phi'(x)|^8). \tag{6.29}$$

THEOREM 4.1: (I) If $na \to \infty$ and $na^5 \to 0$ as $n \to \infty$, if f'' exists and is bounded, then $\sqrt{na}(\hat{f}(x) - f(x))$ is asymptotically normal with mean 0 and variance $\sigma^2 = \frac{|\phi'(x)| f(x) R(K)}{|w(\phi(x))|}$.
(II) If $na^2 \to \infty$ as $n \to \infty$ and if f is uniformly continuous, and if $\int e^{-itu \frac{\left[|\phi'(x)| K\left(\frac{\phi(x)-\phi(y)}{a}\right)\right]}{|w(\phi(x))|}} dx$ is absolutely integrable in t, then $sup_x |\hat{f}(x) - f(x)| \to 0$ in probability as $n \to \infty$.
(III) If for any $\epsilon > 0$, $\sum_{i=1}^{n} e^{-\epsilon n a^2} < \infty$, if f is uniformly continuous and if $\frac{\left[|\phi'(x)| K\left(\frac{\phi(x)-\phi(y)}{a}\right)\right]}{|w(\phi(x))|}$ is a function of bounded variation,then $sup_x |\hat{f}(x) - f(x)| \to 0$ with probability one as $n \to \infty$. To show uniform consistency and asymptotic normality, first we write the following decomposition:

$$\hat{f}(x) - f(x) = \theta_w \Bigg\{ E \frac{|\phi'(x)|}{a} \frac{1}{|w(\phi(Y_1)|} K\left(\frac{\phi(x) - \phi(Y_1)}{a} \right)$$
$$- \frac{g(x)}{w(\phi(x))} \Bigg\} + \theta_w \Bigg\{ \frac{|\phi'(x)|}{na} \sum_{i=1}^{n} \frac{1}{|w(\phi(Y_i))|} K\left(\frac{\phi(x) - \phi(Y_i)}{a} \right)$$
$$- E \frac{|\phi'(x)|}{a} \frac{1}{|w(\phi(Y_1))|} K\left(\frac{\phi(x) - \phi(Y_1)}{a} \right) \Bigg\}$$
$$- \theta_w \Bigg\{ \frac{|\phi'(x)|}{a} \sum_{i=1}^{n} \frac{1}{|w(\phi(Y_i))|} K\left(\frac{\phi(x) - \phi(Y_i)}{a} \right) \left(\frac{1}{|w(\phi(Y_i))|} \right)^{-1} \Bigg\}$$
$$\left\{ \frac{1}{n} \frac{1}{|w(\phi(Y_i))|} - \theta_w \right\} = I_{1n}(x) + I_{2n}(x) - I_{3n}(x) \tag{6.30}$$

Proof:(I) Under the given conditions, $\sqrt{na}\left(E\hat{f}(x) - f(x)\right) = o((na^5)^{1/2} = o(1)\Big)$.

Thus, $\sqrt{na}I_{1n}(x) = o(1)$ as $n \to \infty$. Also in $I_{3n}(x)$,$\frac{1}{n}\sum\frac{1}{|w\phi(y)|} - \theta_w = O_p(1)$ by the law of large number and since $\hat{f}(x) \to f(x)$ in probability for all x when f is continuous, it follows that $\sqrt{na}I_{3n}(x) = O_p(a^{1/2}) = o(1)$. Finally $\sqrt{na}I_{2n}(x)$ is asymptotically normal with 0 mean and variance given in equation 1.27 follows by standard argument provided by (cf. Parzen (1962)).

(II) We have $I_{1n} = \frac{|\phi'(x)|}{a}\int K\left(\frac{\phi(x)-\phi(y)}{a}\right)f(y)dy - f(x)$, again by arguments provided in Parzen(1962), we easily see that $sup_x|I_{1n}(x)| \to 0$ as $n \to \infty$, since f is uniformly continuous. Since K is bounded,

$sup_x|I_{3n}(x)| \leq \left(\frac{C\theta_w}{a}\right)\left\{\frac{1}{n}\sum\frac{1}{|w(\phi(Y_1))|} - \theta_w\right\} = o_p((na^2)^{1/2}) = o_p(1)$. Finally, by using Theorem 3A of Parzen (1962), It is obvious that $sup_x|I_{2n}(x)| \to 0$ in probability.

(III) By following the proof of Nadaraya(1965)(cf. Prakasa Rao(1983) p.37), it is obvious that $sup_x|I_{2n}(x)| \to 0$ in probability.

Next, we summarize the large sample properties of $\hat{F}(x)$ in the following theorem without proof. The proof follow in similar fashion as above, so is omitted.

THEOREM 4.3: The mean square error of the regression estimate is given by the following equation

$$\begin{aligned}
&E\left(\hat{m}(x) - m(x)\right)^2 \\
&\qquad = \frac{R(K)|\phi'(x)|}{nam^2(x)f_1(x)}\Bigg[\psi_w(\phi(x)) + m^2(x)\nu_w(\phi(x)) - 2m(x)\eta_w(\phi(x))\Bigg] \\
&\qquad + \Bigg[\frac{a^2\mu_2(K)}{2}\Bigg(2(f_1\circ\phi^{-1})'(\phi(x))(m\circ\phi^{-1})'(\phi(x)) \\
&\qquad\qquad + f_1(x)(m\circ\phi^{-1})''(\phi(x))\Bigg)/f_1(x)\Bigg]^2 \qquad (6.31)
\end{aligned}$$

First, we estimate $E\hat{m}(x)$. Note that,

$$
\begin{aligned}
& E\left\{\sum_{i=1}^{n} \frac{|\phi'(x)|}{a} K\left(\frac{\phi(x)-\phi(U_i)}{a}\right) /|w(\phi(U_i),\phi(V_i))\right\} \\
&= \frac{|\phi'(x)|}{a} E\left\{K\left(\frac{\phi(x)-\phi(U_1)}{a}\right) /|w(\phi(U_1),\phi(V_1))\right\} \\
&= \frac{|\phi'(x)|}{a} \int\int |w(\phi(u),\phi(v))|^{-1} K\left(\frac{\phi(x)-\phi(u)}{a}\right) g(u,v)dudv \\
&\approx \theta_w |\phi'(x)| \int\int K(z)(f\circ\phi^{-1})(\phi(x)-az,\phi(v))dz \\
&= \theta_w \left\{f_1(x) + \frac{a^4}{4}\mu_2(K)S(G^2,(\phi')^8)\right\}. \qquad (6.32)
\end{aligned}
$$

Also,

$$
\begin{aligned}
& E\left\{\sum_{i=1}^{n} \frac{|\phi'(x)\phi(V_i)|}{a} K\left(\frac{\phi(x)-\phi(U_i)}{a}\right) /|w(\phi(U_i),\phi(V_i))\right\} \\
&= \theta_w \int\int K(z)\phi(y)(f\circ\phi^{-1})(\phi(x)-az,\phi(y))dzdy \\
&= \theta_w \int (m\circ\phi^{-1})(\phi(x)-az)(f_1\circ\phi^{-1})(\phi(x)-az)K(z)dz \\
&\approx \theta_w \int \left\{m(x) - az(m\circ\phi^{-1})'(\phi(x)) + \frac{a^2z^2}{2}(m\circ\phi^{-1})''(\phi(x))\right\} \\
&\qquad \times \left\{f_1(x) - az(f_1\circ\phi^{-1})'(\phi(x) + \frac{a^2z^2}{2}(f\circ\phi^{-1})''(\phi(x))\right\} K(z)dz \\
&\approx \theta_w \Bigg[m(x)f_1(x) + \frac{a^2\mu_2(K)}{2}\Bigg(2(f_1\circ\phi^{-1})'(\phi(x))(m\circ\phi^{-1})'(\phi(x)) \\
&\qquad + f_1(x)(m\circ\phi^{-1})''(\phi(x)) + (f_1\circ\phi^{-1})''(\phi(x))m(x)\Bigg)\Bigg]. \qquad (6.33)
\end{aligned}
$$

Hence

$$
\begin{aligned}
E(\hat{m}(x)) = m(x) + \frac{a^2\mu_2(K)}{2}\Bigg(&2(f_1\circ\phi^{-1})'(\phi(x))(m\circ\phi^{-1})'(\phi(x)) \\
&+ f_1(x)(m\circ\phi^{-1})''(\phi(x)) \\
&+ (f_1\circ\phi^{-1})''(\phi(x))m(x)\Bigg)/f_1(x). \qquad (6.34)
\end{aligned}
$$

Next, we find the variance of the estimate.

$$
\begin{aligned}
& V\left\{\sum_{i=1}^{n}\frac{|\phi'(x)\phi(V_i)|}{na}K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)/|w(\phi(U_i),\phi(V_i))|\right\} \\
&= \frac{(\phi'(x))^2}{na^2}V\left\{|\phi(V_1)|K\left(\frac{\phi(x)-\phi(U_1)}{a}\right)/|w(\phi(U_1),\phi(V_1))|\right\} \\
&\approx \frac{(\phi'(x))^2}{na^2}E\left\{|\phi(V_1)|K\left(\frac{\phi(x)-\phi(U_1)}{a}\right)/|w(\phi(U_1),\phi(V_1))|\right\}^2 \\
& \frac{(\phi'(x))^2\theta_w}{na^2}\int\int(\phi(v))^2K^2\left(\frac{\phi(x)-\phi(u)}{a}\right)|w(\phi(u),\phi(v))|^{-1}f(u,v)dudv \\
& \frac{|\phi'(x)|\theta_w}{na}\int\left\{\int\frac{(\phi(v^*))^2}{|w(\phi(x),\phi(v^*))|}f(\phi(v^*))\backslash(\phi(x)-a\phi(u^*))dv^*\right\} \\
&\qquad \times f(x-a\phi(u^*))K^2(u^*)du^* \\
&= \frac{|\phi'(x)|\theta_w}{na}R(K)f_1(x)\psi_w(\phi(x)). \qquad (6.35)
\end{aligned}
$$

where $\psi_w(\phi(x)) = E\left\{\phi^2(Y)/|w(\phi(X),\phi(Y))\backslash\phi(X=x)|\right\}$. Next, following exactly the same steps above, one can show that

$$
\begin{aligned}
& V\left\{\sum_{i=1}^{n}\frac{|\phi'(x)}{na}K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)/|w(\phi(U_i),\phi(V_i))|\right\} \\
&= \frac{|\phi'(x)|\theta_w}{na}R(K)f_1(x)\nu_w(x). \qquad (6.36)
\end{aligned}
$$

where $\nu_w(\phi(x)) = E\left[1/|w(\phi(X),\phi(Y)\backslash\phi(X=x)|\right]$.

Also,

$$
\begin{aligned}
& Cov\left\{\frac{|\phi'(x)|}{na}\sum_{i=1}^{n}K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)/|w(\phi(U_i),\phi(V_i))|,\right. \\
&\qquad \left.\frac{|\phi'(x)|V_i}{na}\sum_{i=1}^{n}K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)/|w(\phi(U_i),\phi(V_i))|\right\} \\
&\approx \frac{|\phi'(x)|\theta_w}{na}R(K)f_1(x)\eta_w(x). \qquad (6.37)
\end{aligned}
$$

where $\eta_w(\phi(x)) = E\left\{\phi(Y)/|w(\phi(X),\phi(Y))\backslash\phi(X=x)|\right\}$. Hence, putting (3.37),(3.38),(3.39) and (3.40) into V(U/V) and simplifying, we get that

$$
V\hat{m}(x) = \frac{R(K)|\phi'(x)|\left[\psi_w(\phi(x)) + m^2(x)\nu_w(\phi(x)) - 2m(x)\eta_w(\phi(x))\right]}{nam^2(x)f_1(x)}. \qquad (6.38)
$$

Hence

$$
\begin{aligned}
&E\left(\hat{m}(x)-m(x)\right)^2 \\
&\qquad = \frac{R(K)|\phi'(x)|}{nam^2(x)f_1(x)}\left[\psi_w(\phi(x))+m^2(x)\nu_w(\phi(x))-2m(x)\eta_w(\phi(x))\right] \\
&\qquad + \left[\frac{a^2\mu_2(K)}{2}\Bigg(2(f_1\circ\phi^{-1})'(\phi(x))(m\circ\phi^{-1})'(\phi(x)) \right. \\
&\qquad\qquad \left. +f_1(x)(m\circ\phi^{-1})''(\phi(x))\Bigg)/f_1(x)\right]^2 \qquad (6.39)
\end{aligned}
$$

Large Sample Properties of the Regression Estimate

Writing $\hat{m}(x)=\hat{h}(x)/\hat{f}_1(x)$, it is obvious that following hold:

$$
\begin{aligned}
\hat{h}(x)-h(x) \quad = \quad & \theta_w\left\{|\phi'(x)|a^{-1}E\left(\frac{\phi(V_1)}{|w(\phi(U_1),\phi(V_1))|}K\left(\frac{\phi(x)-\phi(U_1)}{a}\right)\right)\right\} \\
& +\theta_w\left\{|phi'(x)|(na)^{-1}\sum_{i=1}^{n}\frac{\phi(V_i)}{|w(\phi(U_i),\phi(V_i))|}K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)\right. \\
& \left. -a^{-1}E\left(\frac{\phi(V_1)}{|w(\phi(U_1),\phi(V_1))|}K\left(\frac{\phi(x)-\phi(U_1)}{a}\right)\right)\right\} \\
& -\theta_w\left\{\sum_{i=1}^{n}\frac{|\phi'(x)\phi(V_i)|K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)}{|w(\phi(U_i),\phi(V_i))|}/a\sum_{i=1}^{n}\frac{1}{|w(\phi(U_i),\phi(V_i))|}\right\} \\
& \qquad \times\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|w(\phi(U_i),\phi(V_i))|}-\theta_w\right\} \\
& = I_{1n}(x)+I_{2n}(x)-I_{3n}(x). \qquad (6.40)
\end{aligned}
$$

where $h(x) = \int |\phi(y)| \frac{1}{|w(\phi(x),\phi(y))|} g(x,y)dy$, and

$$\begin{aligned}
\hat{f}_1(x) - f_1(x) &= \theta_w \left\{ |\phi'(x)| a^{-1} E\left(\frac{\phi(V_1)}{|w(\phi(U_1),\phi(V_1))|} K\left(\frac{\phi(x)-\phi(U_1)}{a}\right)\right) - f_1(x) \right\} \\
&\quad \theta_w \left\{ |phi'(x)|(na)^{-1} \sum_{i=1}^{n} \frac{1}{|w(\phi(U_i),\phi(V_i))|} K\left(\frac{\phi(x)-\phi(U_i)}{a}\right) \right. \\
&\quad \left. -a^{-1} E\left(\frac{1}{|w(\phi(U_1),\phi(V_1))|} K\left(\frac{\phi(x)-\phi(U_1)}{a}\right)\right) \right\} \\
&\quad -\theta_w \left\{ \sum_{i=1}^{n} \frac{|\phi'(x)| K\left(\frac{\phi(x)-\phi(U_i)}{a}\right)}{|w(\phi(U_i),\phi(V_i))|} / a \sum_{i=1}^{n} \frac{1}{|w(\phi(U_i),\phi(V_i))|} \right\} \\
&\quad \times \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{|w(\phi(U_i),\phi(V_i))|} - \theta_w \right\} \\
&\quad \times \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{|w(\phi(U_i),\phi(V_i))|} - \theta_w \right\} \\
&= J_{1n}(x) + J_{2n}(x) - J_{3n}(x). \qquad (6.41)
\end{aligned}$$

it is clear that $\hat{m}(x) \to m(x)$ in probability as $n \to \infty$ at every continuity of $m(.)$and that $na \to \infty$. Next, f_1 and m have bounded second derivatives, then $\sqrt(na) \left\{ \left(\hat{h}(x) - h(x)\right), \left(\hat{f}(x) - f(x)\right) \right\}'$ has the same limiting distribution as $\sqrt(na)\,(I_{2n}(x), J_{2n}(x))'$, which is bivariate normal with mean $(0,0)'$ and covariance matrix

$$\Sigma = |\phi'(x)|\theta_w R(K) f_1(x) \begin{bmatrix} \psi_w(\phi(x)) & \eta_w(\phi(x)) \\ \eta_w(\phi(x)) & \nu_w(\phi(x)) \end{bmatrix},$$

where η, ν ,and ψ are defined above. Hence using similar argument to that of Prakasa Rao(1983) p. 240-243, one can show that $\sqrt(na)\,(\hat{m}(x) - m(x))$ is asymptotically normal with mean 0 and variance

$\sigma^2 = R(K)|\phi'(x)| \left[\psi_w(\phi(x)) + m^2(x)\nu_w(\phi(x)) - 2m(x)\nu_w(\phi(x))\right]/m^2(x) f_1(x).$

To show uniform consistency(weak or strong), we use arguments as in Nadaraya((1970), if $\inf_{a \le x \le b} h(x) > 0$ and if $0 \le a \le V \le b < \infty$ with probability one, then

$$\begin{aligned}
P\left[\sup_{a \le x \le b} |\hat{m}(x) - m(x)| \ge \epsilon\right] &\le P\left[\sup_{a \le x \le b} |\hat{h}(x) - h(x)| \ge \epsilon^*\right] \\
&\quad + P\left[\sup_{a \le x \le b} |\hat{f}_1(x) - f_1(x)| \ge \epsilon^{**}\right]
\end{aligned}$$

where ϵ^* and ϵ^{**} are independent of n.Theorem 3.1 (II) and (III) gives the uniform convergence(weak and strong, respectively) of $\hat{f}_1(x)$ and applying similar argument yields analogous results for $\hat{h}(x)$ under the conditions: $\int e^{-itu} \frac{|\phi'(x)|K(u)}{|w(\phi(u),\phi(v))|} du$ is absolutely integrable in u and $K(u)/|w(\phi(u),\phi(v))$ is of bounded variation.

We summarize the result in the following theorem.

THEOREM 4.4: (I) If $na \to \infty$ and $na^5 \to 0$ as $n \to \infty$, if f'' and h'' exist and are bounded, then $\sqrt{na}(\hat{m}(x) - m(x)$ is asymptotically normal with mean 0 and variance $\sigma^2 = R(K)|\phi'(x)| [\psi_w(\phi(x)) + m^2(x)\nu_w(\phi(x)) - 2m(x)\nu_w(\phi(x))]/m^2(x)f_1(x)$. (II) If $na^2 \to \infty$ as $n \to \infty$ and if f and h are uniformly continuous, and if $\int e^{-itu} \frac{|\phi'(x)|K(u)}{|w(\phi(u),\phi(v))|} du$ is absolutely integrable in u,if $\inf_{a \leq x \leq b} h(x) = \alpha > 0$ and if $0 \leq \alpha \leq \phi(v) \leq b < \infty$, then $\sup_{v<a \leq x \leq b < \infty} |\hat{m}(x) - m(x)| \to 0$ in probability as $n \to \infty$.

(III) If for any $\epsilon > 0$, $\sum_{i=1}^{n} e^{-\epsilon na^2} < \infty$, if f_1 and k are uniformly continuous and if $\frac{\left[|\phi'(x)|K\left(\frac{\phi(x)-\phi(y)}{a}\right)\right]}{|w(\phi(x))|}$ is a function of bounded variation (in u) and ,if $\inf_{a \leq x \leq b} h(x) = \alpha > 0$ and if $0 \leq \alpha \leq v \leq b < \infty$, then $\sup_{v<a \leq x \leq b < \infty} |\hat{m}(x) - m(x)| \to 0$ in probability as $n \to \infty$.

,then $\sup_x |\hat{f}(x) - f(x)| \to 0$ with probability one as $n \to \infty$.

THEOREM 4.5: $Var\hat{h}(x)$ is

$Var(\hat{h}(x) = \frac{h^2(x)|\phi'(x)|R(K)\theta_w}{naf(x)|w(\phi(x))|}$

Proof:

$$V(\hat{h}(x)) = V\left(\frac{\hat{f}(x)}{\hat{\overline{F}}(x)}\right) = \frac{[E\hat{f}(x)]^2}{[E\hat{\overline{F}}(x)]^2}\left[\frac{V\hat{f}(x)}{[E\hat{f}(x)]^2} + \frac{V\hat{\overline{F}}(x)}{[E\hat{\overline{F}}(x)]^2} - \frac{2cov(\hat{f}(x),\hat{\overline{F}}(x))}{E\hat{f}(x)E\hat{\overline{F}}(x)}\right] \quad (6.42)$$

It is only remained to calculate $2cov(\hat{f}(x),\hat{\overline{F}})$ in the above expression.

$$\frac{2cov(\hat{f}(x),\hat{\overline{F}}(x))}{E\hat{f}(x)E\hat{\overline{F}}(x)} = 2\left[\frac{E[\hat{f}(x)\hat{\overline{F}}(x)] - E\hat{f}(x)E\hat{\overline{F}}(x)}{E\hat{f}(x)E\hat{\overline{F}}(x)}\right] = 2\left[\frac{E[\hat{f}(x)\hat{\overline{F}}(x)]}{E\hat{f}(x)E\hat{\overline{F}}(x)} - 1\right]$$

Now,

$$E\Bigg[\left\{\frac{|\phi'(x)|}{a}\left(\sum\frac{1}{|w(\phi(Y_i))|}\right)^{-1}\sum\frac{1}{|w(\phi(Y_i))|}K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\right\}$$
$$\left\{\left(\sum\frac{1}{|w(\phi(Y_i))|}\right)^{-1}\sum\frac{1}{|w(\phi(Y_i))|}\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\right\}\Bigg]$$

$$= E\Bigg[\frac{|\phi'(x)|}{a}\left(\sum_i\frac{1}{|w(\phi(Y_i))|}\right)^{-2}\sum_i\frac{1}{|w(\phi(Y_i))|^2}K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)$$
$$\times\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right) + E\Bigg[\frac{|\phi'(x)|}{a}\left(\sum\frac{1}{|w(\phi(Y_i))|}\right)^{-1}\sum\frac{1}{|w(\phi(Y_i))|}$$
$$\times K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\Bigg]E\Bigg[\left(\sum\frac{1}{|w(\phi(Y_i))|}\right)^{-1}\sum\frac{1}{|w(\phi(Y_i))|}$$
$$\times\sum\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\Bigg]$$

Let

$$I_1 = E\Bigg[\frac{|\phi'(x)|}{a}\left(\sum_i\frac{1}{|w(\phi(Y_i))|}\right)^{-2}\sum_i\frac{1}{|w(\phi(Y_i))|^2}K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)$$
$$\times\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)$$

and

$$I_2 = E\Bigg[\frac{|\phi'(x)|}{a}\left(\sum\frac{1}{|w(\phi(Y_i))|}\right)^{-1}\sum\frac{1}{|w(\phi(Y_i))|}$$
$$\times K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\Bigg]E\Bigg[\left(\sum\frac{1}{|w(\phi(Y_i))|}\right)^{-1}\sum\frac{1}{|w(\phi(Y_i))|}$$
$$\times\sum\overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\Bigg]$$

Then

$$I_1 = \int \frac{|\phi'(x)|}{a}\left(\sum_i \frac{1}{|w(\phi(Y_i))|}\right)^{-2} \sum_i \frac{1}{|w(\phi(Y_i))|^2} K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)$$
$$\times \overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right) g(y)dy$$
$$= \theta_w^{-1} \int \frac{|\phi'(x)|}{a}\left(\sum_i \frac{1}{|w(\phi(Y_i))|}\right)^{-1} \sum_i \frac{1}{|w(\phi(Y_i))|^2} K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)$$
$$\times \overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right) f(y)dy$$

using earlier results, it is straight forward to show that

$$I_1 = (\eta(\theta_w, \nu_w) + \theta_w^2)(\theta_w^{-1}|(w(\phi(x)|))^{-1}(\eta(\phi(x) - nS(K))$$

where $S(K)$ and $\eta(\phi(x)$ and $\eta(\theta_w, \nu_w)$ are defined as above.

Also

$$I_2 = E\Bigg[\frac{|\phi'(x)|}{a}\left(\sum \frac{1}{|w(\phi(Y_i))|}\right)^{-1} \sum \frac{1}{|w(\phi(Y_i))|}$$
$$\times K\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\Bigg] E\Bigg[\left(\sum \frac{1}{|w(\phi(Y_i))|}\right)^{-1} \sum \frac{1}{|w(\phi(Y_i))|}$$
$$\times \sum \overline{\mathbb{K}}\left(\frac{\phi(x)-\phi(Y_i)}{a}\right)\Bigg] \approx f(x)(1-F(x)$$

Substituting all results terms in equation the equation of $V(\hat{h}(x))$ then simplifying gives $var\hat{h}(x) = \frac{h^2(x)|\phi'(x)|R(K)\theta_w}{naf(x)|w(\phi(x))|}$.

THEOREM 4.6: $MSE\hat{h}(x)$ is

$$MSE(\hat{h}(x) = h(x)\Bigg\{\frac{h(x)a^2}{2}\Bigg(\frac{\sqrt{S(G^2, |\phi'(x)|^8)}}{f(x)} - (F \circ \phi^{-1})''|\phi(x)|$$
$$\times \int z^2 K(z)dz\Bigg)\Bigg\}^2 + \frac{h^2(x)|\phi'(x)|R(K)\theta_w}{naf(x)|w(\phi(x))|}$$

BIBLIOGRAPHY

[1] Ahmad, I. A. and Lin, P. E., Nonparametric estimation of a vector-valued bivariate failure rate, *The Annals of Statistics* **5** (1977) 1027-1038.

[2] Ahmad, I.A. and Lin, P.E., Strong convergence of an estimate of generalized failure rate function, *Bulletin of Mathematical Statistics*, **17** (1976), 77-84.

[3] Ahmad, I.A., On multivariate kernel estimation for samples from weighted distributions,*Statistics and Probability Letters*, **22** (1995), 121-129.

[4] Alberts, T. and Karunamuni, R., A locally adaptive transformation method of boundary correction in kernel density estimation, *J. Statist. Plann. Inference* ,**136** (2006) 2936-2960.

[5] Bhattacharyya, B.B., Kazempour and Richardson, G.D., Length biased density estimation of fibers, *Journal of Nonparametric Statistics* **1** (1991) 127-44.

[6] Bowman, A. W, An Alternative Method of Cross-Validation for the Smoothing of Density Estimates, *Biometrika* **71** (1984) 353-360.

[7] Barlow, R. E. and Proschan, F., Statistical Theory of Reliability and Life Testing, *Probability Models* (1976) Holt, Rinehart and Winston, Inc.

[8] Basu, A. P., Bivariate failure rate, *J. Amer. Statist. Assoc.* **66** (1971) 103-104.

[9] Lai, ,C.-D. and Xie M., Stochastic Ageing and Dependence for Reliability. *Springer, New York*, (2006).

[10] Cacoullos, T., Estimation of a multivariate density, *Ann. Inst. Statist. Math.* **18** (1966) 181- 190.

[11] Chen, S.X., Probability density function estimation using gamma kernels,*Ann. Inst. Stattist.Math.* **52** (2000) 471-480.

[12] Jones, M.C., Kernel density estimation for length biased data, *Biometrika.***78** (1991) 511-19.

[13] Marron, J.S. and Ruppert, D., Transformation to reduce bias in kernel density estimation,*J.R. Statist Soc.*, **56** (1994), 653-671.

[14] Comte, F. and Genon-Catalot, V., Convolution power kernels for density estimation, *Journal of Statistical Planning and Inference,* **142, 7,** (2012) 1698-1715.

[15] Cox, D. R., Regression models and life-tables,*J. Roy. Statist. Soc.* **34** (1972) 187-220.

[16] Fix, E. and Hodges, J.L., Discriminatory analysis. Nonparametric discrimination; consistency properties, **unpublished (1951)**.

[17] Fisher, R.A., The effects of methods of ascertainment upon the estimation of frequencies,*The Annals of Eugenics.* **6** (1934) 13-25.

[18] Everton, B.S. and Hothorn, H., A Handbook of Statistical Analyses Using R,*Florida: Chapman and Hall* (2009)

[19] Johnson, N. L. and Kotz, S., A vector-valued multivariate hazard rate,*J. Multivariate Anal.* **5** (1975) 53-66.

[20] Jones, M. C., Marron, J. S., and Sheather, S. J., A brief survey of bandwidth selection for density estimation,*Journal of the American Statistical Association* **91** (1996) 401407.

[21] Marshall, A. W., Some comments on the hazard gradient. *Stochastic Processes and Appl.* **3** (1975) 293-300.

[22] Mnatsakanov, R. and Sarkisian, K., Varying kernel density estimation on $\mathbb{R}_{+}$, *Statistics and Probability Letters* **82** (2012) 1337-1345.

[23] Muttlak, H.A. and McDonald, L.L., Ramked set sampling with size-biased probability of selection. *Biometric* **46** (1990) 435-445.

[24] Nadaraya, E. A., On nonparametric estimation of density function and regression curve, *Theor. Probability Appl.* **10** (1965) 186-190.

[25] Patil, G.P., Weighted Distributions,*Encyclopedia of Environmetrics***4** (2002) 2369-2377.

[26] Parzen, E., On estimation of probability density and mode, *Ann. Math. Statist.* **33** (1962) 1065- 1076.

[27] Puri, P. S. and Rubin, H., On a characterization of a family of distributions with constant multivariate failure rates,*Ann. Probability* **2** (1974) 738-740.

[28] Rao, Prakas,B.L.S. and Isha Bagai, Kernel Type Density Estimates for positive valued random variables,*Sankhya:The Indian Journal of Statistics* **57** (1995) 56-67.

[29] Rao, C.R., Weighted distributions arising out of methods of ascertainment, *A Celebration of Statistics* **Chapter 24** (1965) 543-569.

[30] Rosenblatt M., Remarks on some non-parametric estimates of a density function, *Ann. Math. Statist.,* **33** (1962), 1065-1076.

[31] Rudemo, M., Empirical Choice of Histogram and Kernel Density Estimators, *Scandinavian Journal of Statistics* **9** (1982) 65-78.

[32] Scaillet, O., Density estimation using inverse and reciprocal inverse Gaussian kernels, *Nonparametric Stat.* **16** (2004) 217-226.

[33] Scott, D. W., and Terrell, G. R., Biased and Unbiased Cross Validation in Density Estimation, *Journal of the American Statistical Association* **82** (1987) 1131-1146.

[34] Silverman, B. W., Density Estimation for Statistics and Data Analysis, *New York: Chapman and Hall* (1986).

[35] Stone, C.J., An asymptotically optimal window selection rule for kernel density estimates, *Ann. Statist* **12** (1984) 1285-1297.

[36] Wand, M. P. and Jones, M. C. 1995. Kernel Smoothing, *London: Chapman and Hall* (1995).

[37] Watson, G. S. and Leadbetter, M. R., Hazard analysis I, *Biometrika* **51** (1964a) 175-184.

[38] Watson, G. S. and Leadbetter, M. R., Hazard analysis II, *Sankhyd Ser. A* **26** (1964b) 101- 116.

[39] Zelen, M. Problems in cell kinetics and the early detection of disease, *Reliability and Biometry* **SIAM** (1974) 701706.

VITA

Thagendra Timsina

Candidate for the Degree of

Doctor of Philosophy

Dissertation: TRANSFORMED NONPARAMETRIC FUNCTIONS ESTIMATION

Major Field: Statistics

Biographical:

Personal Data: Born in Bayarwan, Nepal.

Education:

Receivd Associate Degree from Milwaukee Area Technical College, Milwaukee,WI, 2004, Computer Science

Received the B.S. degree from University of Wisconsin, Milwaukee, WI, 2006, Mathematics

Received the M.S. degree from University of Texas, Arlington, TX, 2008, Mathematics

Completed the requirements for the degree of Doctor of Philosophy with a major in statistics Oklahoma State University in 05, 2015.

Experience:

Teaching Assistant at Oklahoma State University from Aug 2010 - Dec 2014

Adjunct Math Faculty at Tarrant County College from Aug 2008 - Aug 2010

Teaching Assistant at University of Texas at Arlington from Jan 2007 - Aug 2008

Skills:

Proficient in C++, SQL, VB.NET, R, SAS, SAS EM, EXCEL, ACCESS

Hobby:

Traveling, Hiking, Trekking, Sports