Asymptotic Properties of the Deconvolution Kernel Density Estimate based on 2-Dependent Error Structure with Applications to Remaining Useful Life Problems in Reliability Theory

by

GEOFFREY H SCHUETTE

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON
May 2018

In rememberance of everyone who touched my life:

No man or woman is an island.

To exist just for yourself is meaningless.

You can achieve the most satisfaction

when you feel related to some greater purpose in life,

something greater than yourself.

– Denis Waitley

To the man I hope to be:

There goes the only man I ever respected.

He's what every boy thinks he's going to be when he grows up

and wishes he had been when he's an old man.

– Robert Ryan, The Tall Men(1955)

Something we should always remember:

Intelligence plus character

that is the goal of true education.

– Martin Luther King Jr.

Acknowledgments

I would like to thank my thesis advisor, Dr. Shan Sun-Mitchell, for the constant support and belief in me throughout my graduate career. Her guidance has led me to be a better mathematician and individual over the past five years. To the faculty and staff in the Mathematics Department at The University of Texas at Arlington(UTA), especially my committee: Dr. Paul Componation, Dr. Andrzej Korzeniowski, Dr. Suvra Pal, and Dr. Jonghyun Yun, thank you for all the support and encouragement you have given me and other students over the years.

I would like to acknowledge Dr. Paul Componation and Dr. Susan Ferreira in the Industrial, Manufacturing, & Systems Engineering Department at UTA for the oppurtunity to do research under the NASA Systems Engineering Constortium. I would aslo like to the National Science Foundation(NSF) East Asia and Pacific Summer Institute(EAPSI) program, the Ministry of Science and Technology(MOST) Taiwan, National Taiwan University, and Dr. Ming-Yen Cheng for the amazing summer research experience. Finally, I would like to thank the Science, Mathematics, and Research for Transformation Scholarship for Service Program and the TRADOC Analysis Center, Fort Leavenworth, for providing me a scholarship and employement opportuinities.

To my parents, family, friends, and God, thank you for the love and support you have given me all of my life. I would not be the man I am today without you, and I will be forever grateful that my life has been blessed to have you all in my life.

April, 2018

Abstract

Asymptotic Properties of the Deconvolution Kernel Density Estimate based on 2-Dependent Error Structure with Applications to Remaining Useful Life Problems in Reliability Theory

Geoffrey H Schuette , Ph.D.

The University of Texas at Arlington, 2018

Supervising Professor: Shan Sun-Mitchell

This thesis is motivated from an engineering question, which led us to the deconvolution problem with a dependent error structure. We establish a deconvolution kernel density estimator by adapting the methods of kernel density estimates and Fourier Transforms. In this approach, the contaminated data with additive random errors are assumed dependent and satisfying smooth or super smooth conditions. Under both smooth and supper smooth conditions, we derived:

1. optimal rates of convergence in terms of mean integrated squared error for deconvolution kernel density estimator;
2. the limiting distribution of the estimator.

Table of Contents

Acknowledgments iv

Abstract v

Chapter Page

1. Motivation 1

1.1 Introduction 1

1.2 Engineering Motivation 2

1.3 Density Estimation 5

1.3.1 Kernel Density Estimator 6

1.3.2 Sampling 7

2. Deconvolution Estimator 8

3. Optimal Rates of Convergence for the Deconvolution Kernel Desnity Estimator with 2-Dependent Error Structure 13

3.1 MISE Upper Bound 13

3.2 Cases of MISE for Smooth and Supersmooth Assumptions 20

3.2.1 Case 1: Ordinary Smooth Distribution 21

3.2.2 Case 2: Supersmooth Distribution 23

3.3 Theorems and Corollaries 26

4. Limiting Distrubtion of the Deconvolution Kernel Density Estimator . . . 27

4.1 Preliminaries 27

4.2 Limiting Distribution for Deconvolution Kernel Density Estimator . . 30

4.2.1 Case 1: Smooth Error Distribution 30

4.2.2 Case 2: Super Smooth Error Distribution 32

4.2.3 Sample Variance 35
References 38
Biographical Statement 41

Chapter 1

Motivation

1.1 Introduction

As systems in engineering continue to become larger and more complex in nature, real time condition monitoring of such systems continues to be an important area of study in engineering and for the long term health of the systems. Real time condition monitoring is the process of collecting real-time sensor information from a functioning device in order to infer or reason about the health of the device. Some examples included:

- Vibration analysis and Diagnostics
- Model-based voltage and current systems
- Ultrasound testing (Material Thickness/Flaw Testing)
- Infrared thermography

Degradation models are often used to model condition monitoring of the systems. In the Bayesian approach, prior distributions are assumed, and posterior distributions are then found from the assumptions of the prior and the model. After the posterior distribution is found, we can find the probability via the appropriate threshold failure time and access the health of the system. However, when the assumption of the prior distribution is not accurate, then it is important to be able to estimate the distribution from the samples generated by the signal time. Therefore, the focus of this research proposal is to fit the unknown distribution with a smooth empirical distribution function based on kernel function and sequence of bandwidths.

An improved model will lead to a more accurate distribution and a truer probability to the failure threshold.

1.2 Engineering Motivation

A common technique in engineering to estimate when a device will fail is to use a Bayesian approach with a stochastic model. The researchers will assume the parameters in the model follow known distributions and use the bayesian approach to find the probability of when the device will reach a failure thresehold. Below is some problem defintions common to this area:

Definition 1.1. *Problem Definitions*

- *Degradation Signal is the real condition of the device as detected through conditioning monitoring.*
- *Residual Life is the remaining wait time till the device fails.*
- *Residual Life Distribution: For a nonnegative random variable, X, the residual life distribution, denoted by $R_X(t)$, is defined by:*

$$R_X(t) = P(X > x + t | X > x)$$

For the purpose of this dissertation, we will assume the degradation model is

$$S(t_i) = \Theta t_i + \epsilon(t_i) \qquad \text{for } i = 1, 2, \dots. \tag{1.1}$$

Let D denote the failure threshold, and let the vector $\vec{S} = (S(t_1) = s_1, S(t_2) = s_2, \dots, S(t_n) = s_n)$ be the observed degradations signals up to time t_n. We want to find $R_S(t)$ given that degradation signal follows the model above, i.e.:

$$\begin{aligned} R_S(t) &= \mathrm{P}\left(S(t + t_n) \geq D | \vec{S}\right) \\ &= \mathrm{P}\left(\Theta(t + t_n) + \epsilon(t + t_n) \geq D | \vec{S}\right) \end{aligned}$$

To calculate this probability, we need the conditional probability density function(pdf) of $\Theta|\vec{S}$. To find this pdf, the following relationship is commonly used:

$$p\left(\theta|\vec{S}\right) \propto \pi(\theta) f\left(\vec{S}|\theta\right) \tag{1.2}$$

In the Bayesian framework, p is called the posterior distribution, π is called the prior distribution of Θ, and f is the likelihood fucntion. The two sides of the equations are proportional because the left hand side doesn't have the normalizing constant needed to make the right hand side a probability density functions.

In [18], Gabraeel, et. al. explore a degradation models with symmetric priors. They proposed the following degradation model for $i = 1, 2, \ldots$:

$$S(t_i) = \phi + \theta \exp\left(\beta t_i + \epsilon(t_i) - \frac{\sigma^2}{2}\right) \tag{1.3}$$

where ϕ is a known constant, θ is a lognormal random variable, where $\ln\theta$ has mean μ_0 and variance σ_0^2, and β is a normal random variable with mean μ_1 and variance σ_1^2.

First, the authors assume the term $\epsilon(t_i)$ is a random error term that follows a normal distribution with mean zero and variance σ^2. This model describes how the devices being monitored degrade over time. For convenience, they took the natural log of both sides to introduce the logged signal at time t_i, denoted L_i, which yielded the following model:

$$L_i = \ln(S(t_i) - \phi) = \theta' + \beta t_i + \epsilon(t_i) \tag{1.4}$$

where $\theta' = \ln\theta - \frac{\sigma^2}{2}$ is a random variable with mean $\mu_0 - \frac{\sigma^2}{2}$ and variance σ_0^2. Using a Bayesian approach, the authors found that the posterior distribution given the observed data $L_1, \ldots, L_k$ with the prior distribution of θ' and β assumed to be normal is a bivariate normal distribution containing the parameters related to equation (1.3).

Using this posterior distribution, they find an approximation for the residual life c.d.f. for time $T \in (-\infty, \infty)$. Then the truncated conditional c.d.f. for failure time, T, with the constraint $T \geq 0$ can be derived easily.

Then in [5], Chakraborty et. al. explore a model that will allow for a skewed prior distribution, as the method developed above was for a normal distribution and symmetric priors. Hence, a linear degradation model is given by the following with signal time $S(t_i) = S_i$:

$$S_i = \theta t_i + \epsilon(t_i), \qquad i = 1, \ldots k, \tag{1.5}$$

where $\theta \sim \Gamma(\alpha, \beta)$ and iid $\epsilon(t_i) \sim N(0, \sigma^2)$. They find that the posterior distribution of θ given the signal times is:

$$f(\theta | S_1, S_2, \ldots, S_k) = \frac{\theta^{\alpha-1}}{c} \exp\left[-\frac{1}{2\sigma_1^2}(\theta - \mu_1)^2\right], \qquad \theta \in \mathbb{R}^+, \tag{1.6}$$

where

$$c = \int_0^\infty \theta^{\alpha-1} \exp\left[-\frac{1}{2\sigma_1^2}(\theta - \mu_1)^2\right] d\theta,$$

$$\text{where } \mu_1 = \frac{b}{2a},\ \sigma_1^2 = \frac{1}{2a},\ a = \frac{1}{2\sigma^2}\sum_{i=1}^k t_i^2,\ b = \frac{1}{\sigma^2}\sum_{i=1}^k S_i t_i - \frac{1}{\beta}$$

With the posterior distribution known, Chakraborty et al., found the distribution of the residual life, L_r, of the signal, which is given by:

$$\mathrm{P}(L_r \leq t | S_1, \ldots, S_k) = 1 - \int_0^\infty \frac{y^{\alpha-1}}{c} \exp\left[\frac{1}{2\sigma_1^2}\left(\frac{y}{t+t_k} - \mu_1\right)^2\right] \Phi\left(\frac{T-y}{\sigma}\right) dy \tag{1.7}$$

This residual life distribution requires numerical integration to calculate at any time t. The authors then give an alternative method using a simulation based approach to compute an empirical residual life distribution. They conducted the simulations of both the methods in [5] and [18] to compare the two methods with prior signals being both skewed and symmetric. They conclude that the method in the [5] is better suited to handle both the skewed prior as well as the symmetric priors.

1.3 Density Estimation

To our knowledge, the methods and models available in literature of real time condition monitoring of devices using degradation stochastic models all follow the Bayesian framework. The focus of this thesis started by stepping away from the Bayesian framework and assume the distribution of Θ is unknown and develop an asymptotic estimation of the unknown distribution from random samples. We begin by giving a definiton of density estiamtion, and a breif summary of common techniques used in density estimation.

Definition 1.2. *Density Estimation, as defined in [25] by Silverman*

Given a set of observed data points sampled from an unknown probability density function, density estimation is the construction of an estimate of a probability density function from the observed data.

There are many methods of density estimation ranging from parametric to nonparametric. In the parametric approach, one would assume a known distribution and estimate the parameters which distinguishes said distribution. An example would be to assume a distribution follows a normal density with parameters μ and σ^2, and then use the observed data to estimate μ and σ^2. In this thesis, we will be using the nonparametric approach, more specifically the kernel density estimator. Other density estiamtors include: histogram, naive estimator, nearest neighborhood method, variable kernel method, orthogonal series estimators, maximum penalized likelihood estimators, and general weight function estimators.

1.3.1 Kernel Density Estimator

Definition 1.3. *Kernel Density Estimator*

With observed values $X_1, \ldots, X_n$ *from unknown pdf* f*, let* $K(x)$ *be a function which satisfies the following property:*

$$\int_{-\infty}^{\infty} K(x)dx = 1.$$

Then the kernel density estimator $\hat{f}$ *of* f *based on kernel* K *is given by:*

$$\hat{f}(x;\lambda_n) = \frac{1}{n\lambda_n}\sum_{j=1}^{n} K\left(\frac{x - X_i}{\lambda_n}\right),$$

where λ_n *is the bandwidth such that* $\lambda_n \to 0$ *as* $n \to \infty$*.*

The function K is called a kernel function and λ_n is also called the window width or smoothing parameter. For more information on this standard kernel density estimator, you can find more details in [25] and [29], by Silverman and Wand and Jones respectively. In statistics, we want the mean squared error and the mean integrated square error of any estimator to tend to zero as our sample size gets large. To prove that this kernel density estimator is a good estimator in terms of mean square error, we need to restrict the type of kernel functions we consider. To do this, assume that K is a symmetric function which satisfies the following additional properties to the definition:

$$\int_{-\infty}^{\infty} tK(t)dt = 0 \qquad \text{and} \qquad \int_{-\infty}^{\infty} t^2K(t)dt = c \neq 0$$

Some common examples of Kernel functions that satisfy these properties are: Normal, Epanechnikov, and Tri-Cube Kernels functions.

1.3.2 Sampling

As a reminder, we are modeling the degradation signals using the following stochastic model:

$$S(t_i) = \Theta t_i + \epsilon(t_i) \qquad \text{for } i = 1, 2, \ldots.$$

Our focus is to create a kernel density estimator for the distrubtion of Θ. From the model, we can see Θ represents the slope of the stochastic signal. Thus it is reasonable to assume the point-slope formula would be an appropriate estimate for Θ. Assume we have the following observations from the degradtion model: $(t_i, S(t_i))$ $i = 1, 2, \ldots, n$. Now we define the following estiamtor:

$$\widetilde{\theta}_i = \frac{S(t_{i+1}) - S(t_i)}{t_{i+1} - t_i} \qquad \text{for } i = 1, 2 \ldots, n-1 \tag{1.8}$$

Using the assumption that the signals $S(t_i)$ follow the model above, we see that

$$\begin{aligned}
\widetilde{\theta}_i &= \frac{S(t_{i+1}) - S(t_i)}{t_{i+1} - t_i} \qquad i = 0, \ldots, n \\
&= \frac{\theta t_{i+1} + \epsilon(t_{i+1}) - \theta t_i - \epsilon(t_i)}{t_{i+1} - t_i} \\
&= \theta + \frac{\epsilon(t_{i+1}) - \epsilon(t_i)}{t_{i+1} - t_i} \\
&= \theta + X_i, \text{ where}
\end{aligned}$$

$X_i = \dfrac{\epsilon(t_{i+1}) - \epsilon(t_i)}{t_{i+1} - t_i}$. Without loss of generality, let $t_{i+1} - t_i = 1$ for all i. Our model now how the following form: $\widetilde{\Theta} = \theta + X$. We can see the density function of $\widetilde{\Theta}$ is a convolution of the density function of Θ and the density function of X. To create a density estimiation for Θ from the distribution of $\widetilde{\Theta}$, we will explore the idea of deconvolution in the next chapter.

Chapter 2

Deconvolution Estimator

We wish to estimate the distribution of Θ, when $\widetilde{\Theta} = \Theta + X$. From the observed signals $\vec{S}$, we can find the associated $\widetilde{\theta}_i$'s. As is standard in engineering models, we will assume $\epsilon(t_i)$'s are independent and identically distributed for all i and assume that Θ is independent of all $\epsilon(t_i)$'s. We can see that $\widetilde{\theta}_i$'s are identically distributed but not independent since the observations are dependent of the errors from previous signals. Since Θ and X are independent, we know that the density of $\widetilde{\Theta}$ is a convolution of the density of Θ and the density of X, which we can write as:

$$f_{\widetilde{\Theta}}(y) = \int \pi(z) f_X(z-y)dz$$

where $f_{\widetilde{\Theta}}(x)$, $\pi(x)$, and $f_X(x)$ denote the distribution of $\widetilde{\Theta}$, Θ and X respectively. Using properties of Fourier Transforms, we get $\phi_{\widetilde{\Theta}}(u) = \phi_\pi(u)\phi_X(u)$. As long as $|\phi_X(u)| > 0$ for all u, we have $\phi_\pi(u) = \dfrac{\phi_{\widetilde{\Theta}}(u)}{\phi_X(u)}$, and assuming the inverse characteristic equation exists, we can write

$$\pi(x) = \phi^{-1}\left(\frac{\phi_{\widetilde{\Theta}(u)}}{\phi_X(u)}\right) = \frac{1}{2\pi}\int e^{-iux}\frac{\phi_{\widetilde{\Theta}(u)}}{\phi_X(u)}du \tag{2.1}$$

We wish to estimate $\pi(x)$. Using eqaution (2.1), we can create a kernel estimator of $f_{\widetilde{\Theta}}$ and find the corresponding characteristic equation to create an estimate of $\pi(x)$. Let $K(x)$ be a bounded even probability density as proposed in Stefanski and Carroll (1990), and let $\hat{f}$ be an ordinary kernel density estimator of $f_{\widetilde{\Theta}}$ based on kernel K, i.e.,

$$\hat{f}(x;\lambda_n) = \frac{1}{n\lambda_n}\sum_{j=1}^{n} K\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)$$

where λ_n is the bandwidth. Then the characteristic equation of $\hat{f}$ is:

$$\begin{aligned}
\phi_{\hat{f}}(u) &= \mathrm{E}\left[e^{iuX^*}\right] \\
&= \int e^{iu\tilde{x}} \hat{f}(\tilde{x}; \lambda_n) d\tilde{x} \\
&= \int e^{iu\tilde{x}} \frac{1}{n\lambda_n} \sum_{j=1}^{n} K\left(\frac{\tilde{x} - \widetilde{\theta}_j}{\lambda_n}\right) d\tilde{x} \\
&= \frac{1}{n\lambda_n} \sum_{j=1}^{n} \int e^{iu\tilde{x}} K\left(\frac{\tilde{x} - \widetilde{\theta}_j}{\lambda_n}\right) d\tilde{x}
\end{aligned}$$

Now, we will do a change of variables. Let $v = \dfrac{\tilde{x} - \widetilde{\theta}_j}{\lambda_n}$, which means $\tilde{x} = \widetilde{\theta}_j + \lambda_n v$ and $dv = \dfrac{d\tilde{x}}{\lambda_n}$. So,

$$\begin{aligned}
\phi_{\hat{f}}(u) &= \frac{1}{n\lambda_n} \sum_{j=1}^{n} \int e^{iu(\widetilde{\theta}_j + \lambda_n v)} K(v)\, \lambda_n dv \\
&= \frac{1}{n} \sum_{j=1}^{n} \int e^{iu\widetilde{\theta}_j} e^{iu\lambda_n v} K(v) dv \\
&= \frac{1}{n} \sum_{j=1}^{n} e^{iu\widetilde{\theta}_j} \int e^{iu\lambda_n v} K(v) dv \\
&= \frac{1}{n} \sum_{j=1}^{n} e^{iu\widetilde{\theta}_j} \phi_K(u\lambda_n)
\end{aligned}$$

Let $\widehat{\phi}(u) = \frac{1}{n} \sum_{j=1}^{n} e^{iu\widetilde{\theta}_j}$. Then $\phi_{\hat{f}}(u) = \widehat{\phi}(u)\phi_K(u\lambda_n)$. So an appropriate estimator of $\pi(x)$, denoted by $\widehat{\pi}(x; \lambda_n)$, is

$$
\begin{aligned}
\widehat{\pi}(x;\lambda_n) &= \phi^{-1}\left(\frac{\phi_{\hat{f}}(u)}{\phi_X(u)}\right) \\
&= \frac{1}{2\pi}\int e^{-iux}\frac{\phi_{\hat{f}}(u)}{\phi_X(u)}du \\
&= \frac{1}{2\pi}\int e^{-iux}\frac{\widehat{\phi}(u)\phi_K(u\lambda_n)}{\phi_X(u)}du \\
&= \frac{1}{2\pi}\int e^{-iux}\frac{\frac{1}{n}\sum_{j=1}^{n} e^{iu\widetilde{\theta}_j}\phi_K(u\lambda_n)}{\phi_X(u)}du \\
&= \frac{1}{2n\pi}\int\sum_{j=1}^{n} e^{iu\widetilde{\theta}_j}e^{-iux}\frac{\phi_K(u\lambda_n)}{\phi_X(u)}du \\
&= \frac{1}{2n\pi}\sum_{j=1}^{n}\int e^{-iu(x-\widetilde{\theta}_j)}\frac{\phi_K(u\lambda_n)}{\phi_X(u)}du
\end{aligned}
$$

Now with the change of variable: $y = u\lambda_n$. We have:

$$
\begin{aligned}
\widehat{\pi}(x;\lambda_n) &= \frac{1}{2n\pi}\sum_{j=1}^{n}\int e^{-iy\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)}\frac{\phi_K(y)}{\phi_X\left(\frac{y}{\lambda_n}\right)}\frac{dy}{\lambda_n} \\
&= \frac{1}{n\lambda_n}\sum_{j=1}^{n}\frac{1}{2\pi}\int e^{-iy\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)}\frac{\phi_K(y)}{\phi_X\left(\frac{y}{\lambda_n}\right)}dy
\end{aligned}
$$

Let

$$
K^*(t) = \frac{1}{2\pi}\int e^{-iyt}\frac{\phi_K(y)}{\phi_X\left(\frac{y}{\lambda_n}\right)}dy \tag{2.2}
$$

Then

$$
\widehat{\pi}(x;\lambda_n) = \frac{1}{n\lambda_n}\sum_{j=1}^{n} K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right) \tag{2.3}
$$

Equation (2.3) is called the Deconvolution kernel density estimator. This approach uses fourier transforms and are predominatly used for statistical inference in errors-in-variables problems. At this point, we will give a brief history of work that has been done with this approach. Carroll and Hall [3] and Stefanski and Carroll [27] constructed a consistent deconvolution kernel desnity estimator. Stefanski [26] and Fan [15] & [17] derived the general asympototic properties of the deconvolution kernel density estimator. In [16], Fan shows the limiting distribution of this estimator is asmyptotically normal.

Li and Vuong [21], Lin and Carrol [22], Hall and Ma [19], Delaigle *et al.* [7], and Stefanski and McIntrye [28] explore measurement error problems and replicated data. Diggle and Hall [14] and Neumann [24] study the deconvolution problem when samples of error data are available. Butucea and Matias [1], Butucea *et al.* [2], and Kneip *et al.* [20] focused on problems where the error term in the model follows a supersmooth distribution known up to a scale parameter, and Meister [23] assumed the error term follows a normal distribution. In Carroll *et al.* [4], the authors provided an account of the general methodology for the deconvolution problem. Delaigle and Hall [12] explore choices for the smoothing parameters choice in error-in-variables problems using SIMEX, and in [13], they remove the assumption of the distribution of the error term is known and prove properties of the deconvolution estimator in such scenario. Delaigle and Gijbels in [9], [10], and [11] explore different bandwidth selections for the deconvolution problem.

For the remainder of this thesis, we will follow the frame work Fan used in [15] and [16]. We will need the following defintions:

Definition 2.1.

- *A random variable, X, is said to have an ordinary smooth distribution of order β if the characteristic function of X, denoted $\phi_X(t)$, satisfies:*

$$d_0|t|^{-\beta} \leq |\phi_X(t)| \leq d_1|t|^{-\beta} \qquad \text{as } t \to \infty,$$

for some positive constants d_0, d_1, and β. Examples include gamma, double exponential and symmetric gamma distributions.

- *A random variable, X, is said to have a supersmooth distribution of order β if the characteristic function of X, denoted $\phi_X(t)$, satisfies:*

$$d_0|t|^{-\beta_0}\exp\left(\frac{-|t|^\beta}{\gamma}\right) \le |\phi_X(t)| \le d_1|t|^{-\beta_1}\exp\left(\frac{-|t|^\beta}{\gamma}\right) \qquad \text{as } t \to \infty,$$

for some positive constants d_0, d_1, γ and β and constants β_0 and β_1 . Examples include normal, mixture normal, and Cauchy.

As we know the convergence of the estimators are directly related to the smoothness of the error term, it is appropriate to consider both cases in our deconvolution kernel density estimator. The final tool we will need before we move onto the next chapter is Plancherel's Theorem:

Theorem 2.2.

$$\int_{-\infty}^{\infty} |E(t)|^2 dt = \int_{-\infty}^{\infty} |E_v|^2 dv$$

where $E(t) = \int_{-\infty}^{\infty} E_v e^{-2\pi i v t} dv$, i.e. the integral of the squared modulus of a function is equal to the integral of the squared modulus of its spectrum.

Note: E(t) and E$_v$ are Fourier transform pairs.

Chapter 3

Optimal Rates of Convergence for the Deconvolution Kernel Desnity Estimator with 2-Dependent Error Structure

3.1 MISE Upper Bound

In this chapter, we will create an upper bound of the mean integrated square error of the deconvolution estimator constructed in chapter 2, and find optimal rates of λ_n which in the worst case scenario will ensure us the mean integrated square error of the deconvolution estimator will go to zero as n goes to infinity. The mean squared error of an arbitary estimator $\hat{\theta}$ is defined by $\mathrm{MSE}(\hat{\theta}) = \mathrm{Var}(\hat{\theta}) + (\mathrm{E}[\hat{\theta}] - \theta)^2$. So,

$$\begin{aligned} \mathrm{MISE}\{\widehat{\pi}(\cdot;\lambda_n)\} &= \int_0^T \mathrm{MSE}\{\widehat{\pi}(x,\lambda_n)\}dx \\ &= \int_0^T \mathrm{Var}\{\widehat{\pi}(x;\lambda_n)\} + (\mathrm{bias}(\widehat{\pi}(x;\lambda_n)))^2 dx \end{aligned} \tag{3.1}$$

where $\mathrm{bias}(\hat{f}(x;\lambda_n)) = \mathrm{E}[\hat{f}(x:\lambda_n)] - f(x)$. First, we will find $\mathrm{E}[\widehat{\pi}(x;\lambda_n)]$ by using the property of expectation, $\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Y]]$. So,

$$\begin{aligned} \mathrm{E}[\widehat{\pi}(x;\lambda_n)|\Theta = z] &= \mathrm{E}\left[\frac{1}{n\lambda_n}\sum_{j=1}^{n} K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\middle|\Theta = z\right] \\ &= \frac{1}{n\lambda_n}\sum_{j=1}^{n} \mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\middle|\Theta = z\right] \end{aligned}$$

since $\widetilde{\theta}_j$'s are identically distributed, the equation above becomes:

$$
\begin{aligned}
&= \frac{1}{\lambda_n} \mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right) \middle| \theta = z \right] \\
&= \frac{1}{\lambda_n} \int K^*\left(\frac{x-w}{\lambda_n}\right) f_{\widetilde{\Theta}|\Theta}(w|z)dw \\
&= \frac{1}{\lambda_n} \int \left(\frac{1}{2\pi} \int e^{-iy\left(\frac{x-w}{\lambda_n}\right)} \frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)} dy \right) f_{\widetilde{\Theta}|\Theta}(w|z)dw \\
&= \frac{1}{2\pi\lambda_n} \int\int e^{-iy\left(\frac{x-w}{\lambda_n}\right)} \frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)} f_{\widetilde{\Theta}|\Theta}(w|z)dydw
\end{aligned}
$$

using Fubini's Theorem, we now get:

$$
\begin{aligned}
&= \frac{1}{2\pi\lambda_n} \int e^{-\frac{iyx}{\lambda_n}} \frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)} \int e^{iw\left(\frac{y}{\lambda_n}\right)} f_{\widetilde{\theta}_1|\theta}(w|z)dwdy \\
&= \frac{1}{2\pi\lambda_n} \int e^{-\frac{iyx}{\lambda_n}} \frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)} \phi_{\widetilde{\Theta}|\Theta}\left(\frac{y}{\lambda_n}\middle| z\right) dy
\end{aligned}
$$

Since $\widetilde{\Theta} = \Theta + X_1$ and $\Theta = z$, we know that $\phi_{\widetilde{\Theta}|\Theta}(u|z) = \phi_{X_1}(u)e^{iuz}$. Thus, we have:

$$
\begin{aligned}
\mathrm{E}[\widehat{\pi}(x;\lambda_n)|\Theta = z] &= \frac{1}{2\pi\lambda_n} \int e^{\frac{-iyx}{\lambda_n}} \frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)} \phi_{X_1}\left(\frac{y}{\lambda_n}\right) e^{iz\left(\frac{y}{\lambda_n}\right)} dy \\
&= \frac{1}{2\pi\lambda_n} \int e^{iy\left(\frac{z-x}{\lambda_n}\right)} \phi_K(y)dy \\
&= \frac{1}{2\pi\lambda_n} \int e^{-iy\left(\frac{x-z}{\lambda_n}\right)} \phi_K(y)dy \\
&= \frac{1}{\lambda_n} K\left(\frac{x-z}{\lambda_n}\right) \qquad (3.2)
\end{aligned}
$$

Since $\mathrm{E}[\mathrm{E}[\widehat{\pi}(x;\lambda_n)|\Theta]] = \frac{1}{\lambda_n}\mathrm{E}\left[K\left(\frac{x-\Theta}{\lambda_n}\right)\right]$, we know $\mathrm{E}[\widehat{\pi}(x;\lambda_n)] = \frac{1}{\lambda_n}\mathrm{E}\left[K\left(\frac{x-\Theta}{\lambda_n}\right)\right]$. Thus, we know that the expected value of $\widehat{\pi}(x)$ is the same as the ordinary kernel density.

With this result, we are now ready to investigate the bias term. Using a result from [29] by Wand and Jones, we can show:

$$\begin{aligned}
\text{bias}(\widehat{\pi}(x;\lambda_n)) &= \frac{\lambda_n^2}{2}\pi''(x)\int y^2K(y)dy + \mathcal{O}(\lambda_n^2) \\
(\text{bias}(\widehat{\pi}(x;\lambda_n)))^2 &= \frac{\lambda_n^4}{4}(\pi''(x))^2\left(\int y^2K(y)dy\right)^2 + \frac{\lambda_n^2}{2}\pi''(x)(\mathcal{O}(\lambda_n^2))\int y^2K(y)dy + \mathcal{O}(\lambda_n^4) \\
&= \frac{\lambda_n^4}{4}(\pi''(x))^2\left(\int y^2K(y)dy\right)^2 + \mathcal{O}(\lambda_n^4) \\
\int(\text{bias}(\widehat{\pi}(x;\lambda_n)))^2dx &= \int\left(\frac{\lambda_n^4}{4}(\pi''(x))^2\left(\int y^2K(y)dy\right)^2 + \mathcal{O}(\lambda_n^4)\right)dx \\
&= \frac{\lambda_n^4}{4}\left(\int y^2K(y)dy\right)^2\int(\pi''(x))^2dx + \int\mathcal{O}(\lambda_n^4)dx
\end{aligned}$$

Thus, we have:

(3.3)
$$\int(\text{bias}(\widehat{\pi}(x;\lambda_n)))^2dx = \frac{\lambda_n^4}{4}\left(\int y^2K(x)dx\right)^2\int(\pi''(x))^2dx + \int\mathcal{O}(\lambda_n^4)dx = \mathcal{O}(\lambda_n^4)$$

Next, we need to find $\text{Var}\{\widehat{\pi}(x;\lambda_n)\}$ and integrate.

$$\begin{aligned}
\text{Var}\{\widehat{\pi}(x;\lambda_n)\} &= \text{Var}\left\{\frac{1}{n\lambda_n}\sum_{j=1}^{n}K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right\} \\
&= \frac{1}{n^2\lambda_n^2}\text{Var}\left\{\sum_{j=1}^{n}K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right\} \\
&= \frac{1}{n^2\lambda_n^2}\left[\sum_{j=1}^{n}\text{Var}\left\{K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right\} + 2\sum_{j<k}\text{Cov}\left\{K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right), K^*\left(\frac{x-\widetilde{\theta}_k}{\lambda_n}\right)\right\}\right]
\end{aligned}$$

(3.4)
$$= \frac{1}{n\lambda_n^2}\text{Var}\left\{K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right\} + \frac{2}{n\lambda_n^2}\text{Cov}\left\{K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right), K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right\},$$

since $\widetilde{\theta}_j$'s are identically distributed, and for j and k satisfying $|k-j|>1$, we know $\widetilde{\theta}_j$ is independent of $\widetilde{\theta}_k$. Thus, we have

$$\text{Cov}\left\{K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right), K^*\left(\frac{x-\widetilde{\theta}_k}{\lambda_n}\right)\right\} = 0 \qquad \text{for } j \text{ and } k \text{ satisfying } |k-j|>1.$$

Now, we will explore the integral $\frac{1}{n\lambda_n^2}\text{Var}\left\{K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right\}$ in equation (3.4):

$$\int \frac{1}{n\lambda_n^2}\text{Var}\left\{K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right\}dx = \int \frac{1}{n\lambda_n^2}\text{E}\left[\left(K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right)^2\right] - \frac{1}{n\lambda_n^2}\left(\text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\text{E}\left[\int\left(K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right)^2 dx\right] - \frac{1}{n\lambda_n^2}\int\left(\text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2 dx \tag{3.5}$$

Next, we will show the second term in equation (3.5) is as follows:

$$\frac{1}{n\lambda_n^2}\int\left(\text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2 dx = \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt$$

Proof.

$$\frac{1}{n\lambda_n^2}\int\left(\text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2 dx = \frac{1}{n\lambda_n^2}\int\left(\text{E}\left[K^*\left(\frac{x-(\Theta+X_1)}{\lambda_n}\right)\right]\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int\int K^*\left(\frac{x-(u+v)}{\lambda_n}\right) f_{\theta;X_1}(u,v)dudv\right)^2 dx$$

Using θ independent of X_1

$$= \frac{1}{n\lambda_n^2}\int\left(\int\int K^*\left(\frac{x-(u+v)}{\lambda_n}\right)\pi(u)f_{X_1}(v)dudv\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int\int \frac{1}{2\pi}\int e^{-iy(\frac{x-(u+v)}{\lambda_n})}\frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)}dy\pi(u)f_{X_1}(v)dudv\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int\int \frac{1}{2\pi}\int e^{iy(\frac{u+v-x}{\lambda_n})}\frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)}\pi(u)f_{X_1}(v)dydudv\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int\int \frac{1}{2\pi}\int e^{iy(\frac{u+v-x}{\lambda_n})}\frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)}\pi(u)f_{X_1}(v)dudvdy\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int \frac{1}{2\pi}e^{iy(\frac{-x}{\lambda_n})}\frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)}\int e^{iy(\frac{u}{\lambda_n})}\pi(u)du\int e^{iy(\frac{v}{\lambda_n})}f_{X_1}(v)dvdy\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int \frac{1}{2\pi}e^{iy(\frac{-x}{\lambda_n})}\frac{\phi_K(y)}{\phi_{X_1}\left(\frac{y}{\lambda_n}\right)}\phi_\theta\left(\frac{y}{\lambda_n}\right)\phi_{X_1}\left(\frac{y}{\lambda_n}\right)dy\right)^2 dx$$

$$= \frac{1}{n\lambda_n^2}\int\left(\int \frac{1}{2\pi}e^{iy(\frac{-x}{\lambda_n})}\phi_K(y)\phi_\theta\left(\frac{y}{\lambda_n}\right)dy\right)^2 dx$$

Let $t = \frac{y}{\lambda_n}$ and $w = \frac{x}{2\pi}$, then $dt = \frac{1}{\lambda_n} dy$ and $dw = \frac{1}{2\pi} dx$, and we get:

$$\frac{1}{n\lambda_n^2} \int \left(\mathrm{E}\left[K^* \left(\frac{x - \widetilde{\theta}_1}{\lambda_n} \right) \right] \right)^2 dx = \frac{1}{n\lambda_n^2} \int \left(\int \frac{1}{2\pi} e^{-2\pi wti} \phi_K(t\lambda_n) \phi_\theta(t) \lambda_n dt \right)^2 2\pi dw$$
$$= \frac{1}{2\pi n} \int \left(\int e^{-2\pi wti} \phi_K(t\lambda_n) \phi_\theta(t) dt \right)^2 dw$$

Using Plancherel's Theorem:

$$= \frac{1}{2\pi n} \int |\phi_K(t\lambda_n) \phi_\theta(t)|^2 dt$$
$$= \frac{1}{2\pi n} \int \phi_K^2(t\lambda_n) |\phi_\theta(t)|^2 dt$$

Thus,

$$\frac{1}{n\lambda_n^2} \int \left(\mathrm{E}\left[K^* \left(\frac{x - \widetilde{\theta}_1}{\lambda_n} \right) \right] \right)^2 dx = \frac{1}{2\pi n} \int \phi_K^2(\lambda_n t) |\phi_\theta(t)|^2 dt$$

□

Similarly, one can show using Plancherel's Theorem and the substitution $t = \dfrac{x - \widetilde{\theta}_1}{\lambda_n}$:

$$\frac{1}{n\lambda_n^2} \int \left(K^* \left(\frac{x - \widetilde{\theta}_1}{\lambda_n} \right) \right)^2 dx = \frac{1}{2\pi\lambda_n n} \int \frac{\phi_K^2(t)}{\left| \phi_{X_1}\left(\frac{t}{\lambda_n} \right) \right|^2} dt \tag{3.6}$$

Notice that the right hand side does not depend on $\widetilde{\theta}_1$. So equation (3.5) now becomes:

$$\int \frac{1}{n\lambda_n^2} \mathrm{Var} \left\{ K^* \left(\frac{x - \widetilde{\theta}_1}{\lambda_n} \right) \right\} dx = \frac{1}{2\pi\lambda_n n} \int \frac{\phi_K^2(t)}{|\phi_{X_1}\left(\frac{t}{\lambda_n} \right)|^2} dt - \frac{1}{2\pi n} \int \phi_K^2(\lambda_n t) |\phi_\theta(t)|^2 dt$$

One can show by definition of X_1 that $\phi_{X_1}(u) = \phi_{\epsilon_1}(u) \phi_{\epsilon_0}(-u)$ which gives us:

$$= \frac{1}{2\pi\lambda_n n} \int \frac{\phi_K^2(t)}{\left| \phi_{\epsilon_1}\left(\frac{t}{\lambda_n} \right) \phi_{\epsilon_0}\left(\frac{-t}{\lambda_n} \right) \right|^2} dt - \frac{1}{2\pi n} \int \phi_K^2(\lambda_n t) |\phi_\theta(t)|^2 dt$$

Now, we will work on the integral of the covariate term of equation (3.4):

$$\text{Cov}\left\{K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right), K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right\} = \text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right] - \text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right]$$
$$= \text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right] - \left(\text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2$$

since $\widetilde{\theta}_i$ are identically distributed.

$$\leq \text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right]$$

First we will define the following notation: let $d\vec{z} = dzdadbdc$. So,

$$\frac{2}{n\lambda_n^2}\int \text{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right]dx = \frac{2}{n\lambda_n^2}\int \text{E}\left[K^*\left(\frac{x-(\theta+\epsilon_1-\epsilon_0)}{\lambda_n}\right)K^*\left(\frac{x-(\theta+\epsilon_2-\epsilon_1)}{\lambda_n}\right)\right]dx$$
$$= \frac{2}{n\lambda_n^2}\int\cdots\int K^*\left(\frac{x-(z+b-a)}{\lambda_n}\right)K^*\left(\frac{x-(z+c-b)}{\lambda_n}\right)f_{\Theta;\epsilon_0;\epsilon_1;\epsilon_2}(z_1,z_2,z_3,z_4)d\vec{z}dx$$
$$= \frac{2}{n\lambda_n^2}\int\cdots\int K^*\left(\frac{x-(z+b-a)}{\lambda_n}\right)K^*\left(\frac{x-(z+c-b)}{\lambda_n}\right)\pi(z)f_{\epsilon_0}(a)f_{\epsilon_1}(b)f_{\epsilon_2}(c)d\vec{z}dx$$
$$= \frac{2}{n\lambda_n^2}\int\cdots\int \frac{1}{2\pi}\int e^{-iy_1\left(\frac{x-(z+b-a)}{\lambda_n}\right)}\frac{\phi_K(y_1)}{\phi_{X_1}\left(\frac{y_1}{\lambda_n}\right)}dy_1\frac{1}{2\pi}\int e^{-iy_2\left(\frac{x-(z+c-b)}{\lambda_n}\right)}\frac{\phi_K(y_2)}{\phi_{X_1}\left(\frac{y_2}{\lambda_n}\right)}dy_2\pi(z)f_{\epsilon_0}(a)f_{\epsilon_1}(b)f_{\epsilon_2}(c)d\vec{z}dx$$
$$= \frac{1}{2\pi^2 n\lambda_n^2}\int\cdots\int e^{iy_1\left(\frac{z+b-a-x}{\lambda_n}\right)}\frac{\phi_K(y_1)}{\phi_{X_1}\left(\frac{y_1}{\lambda_n}\right)}e^{iy_2\left(\frac{z+c-b-x}{\lambda_n}\right)}\frac{\phi_K(y_2)}{\phi_{X_1}\left(\frac{y_2}{\lambda_n}\right)}\pi(z)f_{\epsilon_0}(a)f_{\epsilon_1}(b)f_{\epsilon_2}(c)dy_1dy_2d\vec{z}dx$$
$$= \frac{1}{2\pi^2 n\lambda_n^2}\int\cdots\int e^{iy_1\left(\frac{z+b-a-x}{\lambda_n}\right)}\frac{\phi_K(y_1)}{\phi_{X_1}\left(\frac{y_1}{\lambda_n}\right)}e^{iy_2\left(\frac{z+c-b-x}{\lambda_n}\right)}\frac{\phi_K(y_2)}{\phi_{X_1}\left(\frac{y_2}{\lambda_n}\right)}\pi(z)f_{\epsilon_0}(a)f_{\epsilon_1}(b)f_{\epsilon_2}(c)d\vec{z}dy_2dx$$

Now we can seperate the different integrals of a, b, c, and z and write them as characteristic equations:

$$= \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{X_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{X_1}\left(\frac{y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_0}\left(\frac{-y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right)\phi_{\epsilon_2}\left(\frac{y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx$$

Usine $\phi_{X_1}(u) = \phi_{\epsilon_1}(u)\phi_{\epsilon_0}(-u)$ we have,

$$= \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_0}\left(\frac{-y_1}{\lambda_n}\right)\phi_{X_1}\left(\frac{y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_0}\left(\frac{-y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right)\phi_{\epsilon_2}\left(\frac{y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx$$
$$= \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_2}{\lambda_n}\right)\phi_{\epsilon_0}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right)\phi_{\epsilon_2}\left(\frac{y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx$$

But the characteristic equations of the ϵ_i's are the same since they are i.i.d. So,

$$= \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx$$

Thus, we have:

$$\frac{2}{n\lambda_n^2}\int \mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)K^*\left(\frac{x-\widetilde{\theta}_2}{\lambda_n}\right)\right]dx =$$

(3.7)

$$\frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx$$

Now let $Y_1 = e^{i\epsilon_1\left(\frac{y_1}{\lambda_n}\right)}$ and $Y_2 = e^{i\epsilon_1\left(\frac{-y_2}{\lambda_n}\right)}$, then:

$$\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right) = \mathrm{E}\left[e^{i\epsilon_1\left(\frac{y_1-y_2}{\lambda_n}\right)}\right] = \mathrm{E}\left[Y_1Y_2\right]$$

using definition of covariance, we get:

$$\begin{aligned}\mathrm{E}\left[Y_1Y_2\right] &= \rho_{Y_1,Y_2}\sigma_{Y_1}\sigma_{Y_2} + \mathrm{E}\left[Y_1\right]\mathrm{E}\left[Y_2\right]\\ &= \sigma_{Y_1}\sigma_{Y_2} + \mathrm{E}\left[Y_1\right]\mathrm{E}\left[Y_2\right]\end{aligned}$$

since Y_1 and Y_2 are functions of ϵ_1, we know by definition:

$$= \sigma_{Y_1}\sigma_{Y_2} + \phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)$$

Using this in equation (3.7), we get:

$$\begin{aligned}&\frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1-y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx =\\
&= \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\left[\sigma_{Y_1}\sigma_{Y_2} + \phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)\right]e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx\\
&= \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\sigma_{Y_1}\sigma_{Y_2}e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx\\
&\quad + \frac{1}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx\\
&= \frac{\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)}\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx\\
&\quad + \frac{1}{2\pi^2 n\lambda_n^2}\iiint \phi_K(y_1)\phi_K(y_2)\phi_\Theta\left(\frac{y_1+y_2}{\lambda_n}\right)e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx\end{aligned}$$

But since $\pi(x)$ is bounded, there exists a c such that $\phi_\Theta \leq c$

$$\leq \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)} e^{ix(\frac{y_1+y_2}{\lambda_n})}dy_1dy_2dx + \frac{c}{2\pi^2 n\lambda_n^2}\iiint \phi_K(y_1)\phi_K(y_2)e^{ix(\frac{y_1+y_2}{\lambda_n})}dy_1dy_2dx$$

$$= \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)} e^{ix(\frac{y_1+y_2}{\lambda_n})}dy_1dy_2dx + \frac{c}{2\pi^2 n\lambda_n^2}\int\left(\int \phi_K(y_1)e^{ix(\frac{y_1}{\lambda_n})}dy_1\right)^2 dx$$

Using Plancherel's Theorem,

$$= \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)} e^{ix(\frac{y_1+y_2}{\lambda_n})}dy_1dy_2dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1$$

So we have established the following bound on the variance term of the MISE for $\widehat{\pi}(x;\lambda_n)$:

$$\int \text{Var}\{\widehat{\pi}(x;\lambda_n)\}dx \leq \frac{1}{2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|\phi_{\epsilon_1}\left(\frac{t}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-t}{\lambda_n}\right)\right|^2}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt$$

(3.8)

$$+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\phi_{\epsilon_1}\left(\frac{y_1}{\lambda_n}\right)\phi_{\epsilon_1}\left(\frac{-y_2}{\lambda_n}\right)} e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1$$

With equation (3.3) and (3.8), we have established an upper bound for the MISE given in equation (3.1).

3.2 Cases of MISE for Smooth and Supersmooth Assumptions

In this section, we will consider two different cases: ϵ_1 follows an ordinary smooth distribution and a super smooth distribution. We will show the upper bound given in equation (3.8) converges to zero with the appropriate selection of bandwidth. This will show the MISE given in equation (3.1) of our deconvolution kernel denstiy estimator goes to zero as n gets large.

3.2.1 Case 1: Ordinary Smooth Distribution

We will assume that ϵ_1 follows an ordinary smooth distribution. This means $\phi_{\epsilon_1}(t)$ satisfies:

$$d_0|t|^{-\beta} \le |\phi_{\epsilon_1}(t)| \le d_1|t|^{-\beta} \qquad \text{as } t \to \infty, \tag{3.9}$$

for some positive constants d_0, d_1, and β. Using this assumption, we will find the upper bound of the variance term of the MISE for $\widehat{\pi}(x;\lambda_n)$ with the assumption that ϵ_1 follows an ordinary smooth distribution and get:

$$\begin{aligned}
\int \mathrm{Var}\{\widehat{\pi}(x;\lambda_n)\}dx \le{}& \frac{1}{2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|d_0\left|\frac{t}{\lambda_n}\right|^{-\beta} d_0\left|\frac{-t}{\lambda_n}\right|^{-\beta}\right|^2}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{d_0\left|\frac{y_1}{\lambda_n}\right|^{-\beta} d_0\left|\frac{-y_2}{\lambda_n}\right|^{-\beta}} e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
={}& \frac{1}{2d_0^2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|\frac{t}{\lambda_n}\right|^{-4\beta}}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{\left|\frac{y_1}{\lambda_n}\right|^{-\beta}\left|\frac{y_2}{\lambda_n}\right|^{-\beta}} e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
={}& \frac{1}{2d_0^2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|\frac{t}{\lambda_n}\right|^{-4\beta}}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^2}\int\left(\int \frac{\phi_K(y_1)}{\left|\frac{y_1}{\lambda_n}\right|^{-\beta}} e^{ix\left(\frac{y_1}{\lambda_n}\right)}dy_1\right)^2 dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
={}& \frac{1}{2d_0^2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|\frac{t}{\lambda_n}\right|^{-4\beta}}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^2}\int \frac{\phi_K^2(t)}{\left|\frac{t}{\lambda_n}\right|^{-2\beta}}dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1
\end{aligned}$$

the line above is a result of using Plancherel's Theorm. Therefore, we now have:

$$\begin{aligned}
\int \operatorname{Var}\{\widehat{\pi}(x;\lambda_n)\}dx &\leq \frac{1}{2d_0^2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|\frac{t}{\lambda_n}\right|^{-4\beta}}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&\quad + \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^2}\int \frac{\phi_K^2(t)}{\left|\frac{t}{\lambda_n}\right|^{-2\beta}}dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
&= \frac{1}{2d_0^2\pi\lambda_n^{1+4\beta}n}\int \phi_K^2(t)t^{4\beta}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&\quad + \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^{2(1+\beta)}}\int \phi_K^2(t)t^{2\beta}dt + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
&= \frac{1}{n\lambda_n^{1+4\beta}}\mathcal{O}(1) - \frac{1}{n}\mathcal{O}(1) + \frac{1}{n\lambda_n^{2(1+\beta)}}\mathcal{O}(1) + \frac{1}{n\lambda_n^2}\mathcal{O}(1)
\end{aligned}$$

$$\leq \begin{cases} \left[\frac{2}{n\lambda_n^{1+4\beta}} - \frac{1}{n} + \frac{1}{n\lambda_n^2}\right]\mathcal{O}(1) & \text{for } \beta \leq \frac{1}{2} \\ \left[\frac{2}{n\lambda_n^{2(1+\beta)}} - \frac{1}{n} + \frac{1}{n\lambda_n^2}\right]\mathcal{O}(1) & \text{for } \beta > \frac{1}{2} \end{cases} \tag{3.10}$$

We want equation (3.10) to go to zero for all values of β. We will let $\lambda_n = n^{-\frac{a}{2}}\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + n^{-b}\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta)$, where a and b are chosen so that $0 < a < \frac{1}{1+4\beta}$ and $0 < b < \frac{1}{2(1+\beta)}$, and we will focus on both the first and third terms in equation (3.10), for thier respective values of β. We will show they go to zero with the choice of a and b. For $\beta \in \left(0, \frac{1}{2}\right]$, we have:

$$\begin{aligned}
\frac{2}{n\lambda_n^{1+4\beta}} &= \frac{2}{n\left(n^{-\frac{a}{2}}\right)^{1+4\beta}} \qquad \text{and} \qquad & \frac{1}{n\lambda_n^2} &= \frac{1}{n\left(n^{-\frac{a}{2}}\right)^2} \\
&= 2n^{-\left(1-\frac{a(1+4\beta)}{2}\right)} & &= n^{-(1-a)}
\end{aligned}$$

It is easy to see by the choice of a that $1 - \frac{a(1+4\beta)}{2} > 0$ and $1 - a > 0$. Let $\omega_1 = 1 - \frac{a(1+4\beta)}{2}$ and $\omega_2 = 1 - a$. Then:

$$\begin{aligned}
\int \operatorname{Var}\{\widehat{\pi}(x;\lambda_n)\}dx &\leq k\left[\frac{2}{n\lambda_n^{1+4\beta}} - \frac{1}{n} + \frac{1}{n\lambda_n^2}\right] \\
&= o(n^{-\omega_1}) - \mathcal{O}(n^{-1}) + o(n^{-\omega_2}) = o(n^{-\omega_a}),
\end{aligned}$$

where $-\omega_a = \max\{-\omega_1, -1, \omega_2\}$.

Now to consider the case: $\beta \in \left[\frac{1}{2}, \infty\right)$:

$$\frac{2}{n\lambda_n^{2(1+\beta)}} = \frac{2}{n\left(n^{-b}\right)^{2(1+\beta)}} = 2n^{-(1-2b(1+\beta))} \qquad \text{and} \qquad \frac{1}{n\lambda_n^2} = \frac{1}{n\left(n^{-b}\right)^2} = n^{-(1-2b)}$$

It is easy to see by the choice of b that $1 - 2b(1+\beta) > 0$ and $1 - 2b > 0$. Let $\omega_3 = 1 - 2b(1+\beta)$ and $\omega_4 = 1 - 2b$. Then:

$$\begin{aligned}\int \mathrm{Var}\{\widehat{\pi}(x;\lambda_n)\}dx &\leq k\left[\frac{2}{n\lambda_n^{2(1+\beta)}} - \frac{1}{n} + \frac{1}{n\lambda_n^2}\right] \\ &= o(n^{-\omega_3}) - O(n^{-1}) + o(n^{-\omega_4}) = o(n^{-\omega_b}),\end{aligned}$$

where $-\omega_b = \max\{-\omega_3, -1, \omega_4\}$. If $\lambda_n = n^{-\frac{a}{2}}\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + n^{-b}\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta)$, where a and b are chosen so that $0 < a < \frac{1}{1+4\beta}$ and $0 < b < \frac{1}{2(1+\beta)}$. Then we have that:

$$\int \mathrm{Var}\{\widehat{\pi}(x;\lambda_n)\}dx \leq o(n^{-\omega_a})\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + o(n^{-\omega_b})\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta) \tag{3.11}$$

Finally when ϵ_1 follows a smooth distribution and for $\lambda_n = n^{-\frac{a}{2}}\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + n^{-b}\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta)$, where a and b are chosen so that $0 < a < \frac{1}{1+4\beta}$ and $0 < b < \frac{1}{2(1+\beta)}$, we have from equation (3.3) and (3.11):

$$\mathrm{MISE}\{\widehat{\pi}(\cdot;\lambda_n)\} = o(n^{-\omega_a})\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + o(n^{-\omega_b})\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta) + o(\lambda_n^4)$$

3.2.2 Case 2: Supersmooth Distribution

We will assume that ϵ_1 follows a super smooth distribution. This means $\phi_{\epsilon_1}(t)$ satisfies:

$$d_0|t|^{\beta_0}\exp\left(\frac{-|t|^\beta}{\gamma}\right) \leq |\phi_{\epsilon_1}(t)| \leq d_1|t|^{\beta_1}\exp\left(\frac{-|t|^\beta}{\gamma}\right) \qquad \text{as } t \to \infty, \tag{3.12}$$

for some positive constants d_0, d_1, γ and β and constants β_0 and β_1. Using this assumption, we will find the upper bound of the variance term of the MISE for $\widehat{\pi}(x;\lambda_n)$ with the assumption that ϵ_1 follows an supersmooth distribution and get:

$$\begin{aligned}
\int \operatorname{Var}\{\widehat{\pi}(x;\lambda_n)\}dx \leq \frac{1}{2\pi\lambda_n n}\int & \frac{\phi_K^2(t)}{\left|d_0\left|\frac{t}{\lambda_n}\right|^{\beta_0}\exp\left(\frac{-|t|^\beta}{\gamma\lambda_n^\beta}\right)d_0\left|\frac{-t}{\lambda_n}\right|^{\beta_0}\exp\left(\frac{-|-t|^\beta}{\gamma\lambda_n^\beta}\right)\right|^2}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\iiint \frac{\phi_K(y_1)\phi_K(y_2)}{d_0\left|\frac{y_1}{\lambda_n}\right|^{\beta_0}\exp\left(\frac{-|y_1|^\beta}{\gamma\lambda_n^\beta}\right)d_0\left|\frac{-y_2}{\lambda_n}\right|^{\beta_0}\exp\left(\frac{-|-y_2|^\beta}{\gamma\lambda_n^\beta}\right)}e^{ix\left(\frac{y_1+y_2}{\lambda_n}\right)}dy_1dy_2dx \\
&+ \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
= \frac{1}{2\pi\lambda_n n}\int & \frac{\phi_K^2(t)}{\left|d_0^2\left|\frac{t}{\lambda_n}\right|^{2\beta_0}\exp\left(\frac{-2|t|^\beta}{\gamma\lambda_n^\beta}\right)\right|^2}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&+ \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\int\left(\int \frac{\phi_K(y_1)}{d_0\left|\frac{y_1}{\lambda_n}\right|^{\beta_0}\exp\left(\frac{-|y_1|^\beta}{\gamma\lambda_n^\beta}\right)}e^{ix\left(\frac{y_1}{\lambda_n}\right)}dy_1\right)^2 dx + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1
\end{aligned}$$

Using Plancherel's Theorem, we have:

$$\begin{aligned}
&= \frac{1}{2\pi\lambda_n n}\int \frac{\phi_K^2(t)}{\left|d_0^4\left|\frac{t}{\lambda_n}\right|^{4\beta_0}\exp\left(\frac{-4|t|^\beta}{\gamma\lambda_n^\beta}\right)\right|}dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&\quad + \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2\pi^2 n\lambda_n^2}\int \frac{\phi_K^2(t)}{d_0^2\left|\frac{t}{\lambda_n}\right|^{2\beta_0}\exp\left(\frac{-2|t|^\beta}{\gamma\lambda_n^\beta}\right)}dt + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1 \\
&= \frac{1}{2d_0^4\pi\lambda_n^{1-4\beta_0}n}\int \phi_K^2(t)\exp\left(\frac{4|t|^\beta}{\gamma\lambda_n^\beta}\right)|t|^{-4\beta_0}\,dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt \\
&\quad + \frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^{2-2\beta_0}}\int \phi_K^2(t)\exp\left(\frac{2|t|^\beta}{\gamma\lambda_n^\beta}\right)|t|^{-2\beta_0}\,dt + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1
\end{aligned}$$

But we can bound $\exp\left(c|t|^\beta\right)$, where c is considered a positive constant, by $\exp\left(c|T|^\beta\right)$. Thus we have that:

$$\int \mathrm{Var}\{\widehat{\pi}(x;\lambda_n)\}dx \le \frac{1}{2d_0^4\pi\lambda_n^{1-4\beta_0}n}\exp\left(\frac{4|T|^\beta}{\gamma\lambda_n^\beta}\right)\int \phi_K^2(t)\,|t|^{-4\beta_0}\,dt - \frac{1}{2\pi n}\int \phi_K^2(\lambda_n t)|\phi_\theta(t)|^2 dt$$
$$+\frac{c\sigma_{Y_1}\sigma_{Y_2}}{2d_0^2\pi^2 n\lambda_n^{2-2\beta_0}}\exp\left(\frac{2|T|^\beta}{\gamma\lambda_n^\beta}\right)\int \phi_K^2(t)\,|t|^{-2\beta_0}\,dt + \frac{c}{2\pi^2 n\lambda_n^2}\int \phi_K^2(y_1)dy_1$$
$$= \frac{1}{\lambda_n^{a_1}n}\mathcal{O}\left(\exp\left(\frac{4|T|^\beta}{\gamma\lambda_n^\beta}\right)\right) - \frac{1}{n}\mathcal{O}(1)$$
$$+\frac{1}{n\lambda_n^{a_2}}\mathcal{O}\left(\exp\left(\frac{2|T|^\beta}{\gamma\lambda_n^\beta}\right)\right) + \frac{1}{n\lambda_n^2}\mathcal{O}(1)$$

where $$a_1 = \begin{cases} 1 & \text{if } \beta_0 \ge 0 \\ 1-4\beta_0 & \text{if } \beta_0 < 0 \end{cases}$$

and $$a_2 = \begin{cases} 2 & \text{if } \beta_0 \ge 0 \\ 2-2\beta_0 & \text{if } \beta_0 < 0 \end{cases}$$

Now, we will choose our bandwidth to be $\lambda_n = T\left(\frac{a_0}{\gamma}\right)^{\frac{1}{\beta}}(\log n)^{\frac{-1}{\beta}}$, where $a_0 > 4$ and is chosen to be fixed. So we can now change the big O notation to little o notation.

$$\frac{1}{\lambda_n^{a_1}n}\mathcal{O}\left(\exp\left(\frac{4|T|^\beta}{\gamma\lambda_n^\beta}\right)\right) = \frac{1}{n\left(T\left(\frac{a_0}{\gamma}\right)^{\frac{1}{\beta}}(\log n)^{\frac{-1}{\beta}}\right)^{a_1}}\exp\left(\frac{4|T|^\beta}{\gamma\left(T\left(\frac{a_0}{\gamma}\right)^{\frac{1}{\beta}}(\log n)^{\frac{-1}{\beta}}\right)^\beta}\right)$$
$$= \frac{k'(\log n)^{\frac{a_1}{\beta}}}{n}\exp\left(\frac{4}{a_0}\log n\right)$$
$$= \frac{k'(\log n)^{\frac{a_1}{\beta}}}{n}\exp\left(\log n^{\frac{4}{a_0}}\right)$$
$$= \frac{k'(\log n)^{\frac{a_1}{\beta}}}{n^{1-\frac{4}{a_0}}}$$
$$= o(n^{-\omega_1}), \qquad \text{where } \omega_1 < 1-\frac{4}{a_0}$$

Similarly, we can see that

$$\frac{1}{n\lambda_n^{a_2}}\mathcal{O}\left(\exp\left(\frac{2|T|^\beta}{\gamma\lambda_n^\beta}\right)\right) = o(n^{-\omega_2}) \qquad \& \qquad \frac{1}{n\lambda_n^2}\mathcal{O}(1) = o(n^{-\omega_3}) \qquad \text{where } \omega_2 < 1-\frac{2}{a_0} \text{and } \omega_3 < 1$$

Thus we have that:

$$\int \text{Var}\{\widehat{\pi}(x;\lambda_n)\}dx \le o(n^{-\omega_1}) - o(n^{-1}) + o(n^{-\omega_2}) + o(n^{-\omega_3})$$

$$= o(n^{-\omega_1}), \qquad \text{since } \max_{i=1,2,3}\{\max\{-\omega_i\}, -1\} = \max\{-\omega_1\} \tag{3.13}$$

Finally when ϵ_1 follows a super smooth distribution and for $\lambda_n = T\left(\frac{a_0}{\gamma}\right)^{\frac{1}{\beta}}(\log n)^{\frac{-1}{\beta}}$, we have from equation (3.3) and (3.13):

$$\text{MISE}\{\widehat{\pi}(\cdot;\lambda_n)\} = o(n^{-\omega_1}) + o(\lambda_n^4) = o(n^{-\omega_1})$$

3.3 Theorems and Corollaries

Theorem 3.1.

If $\epsilon(t_i)$ are i.i.d. and follow a smooth distribution and $\lambda_n(a,b) = n^{-\frac{a}{2}}\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + n^{-b}\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta)$, where a and b are chosen so that $0 < a < \frac{1}{1+4\beta}$ and $0 < b < \frac{1}{2(1+\beta)}$, then

$$MISE\{\widehat{\pi}_\Theta(\cdot;\lambda_n)\} = o(n^{-\omega_a})\mathbb{1}_{\left(0,\frac{1}{2}\right]}(\beta) + o(n^{-\omega_b})\mathbb{1}_{\left[\frac{1}{2},\infty\right)}(\beta) + o(\lambda_n^4),$$

where $\omega_a = \min\left\{1 - \frac{a(1+4\beta)}{2}, 1, 1-a\right\}$ and $\omega_b = \min\{1 - 2b(1+\beta), 1, 1-2b\}$

Theorem 3.2.

If $\epsilon(t_i)$ are i.i.d. and follow a supersmooth distribution and $\lambda_n(a_0) = T\left(\frac{a_0}{\gamma}\right)^{\frac{1}{\beta}}(\log n)^{\frac{-1}{\beta}}$, where a_0 is chosen so that $a_0 > 4$, then

$$MISE\{\widehat{\pi}_\Theta(\cdot;\lambda_n)\} = o(n^{-\omega_1}),$$

where $\omega_1 < 1 - \dfrac{4}{a_0}$

Corollary 3.3.

If $\epsilon(t_i)$ follow a standard Brownian motion, then $X_i = \epsilon(t_{i+1}) - \epsilon(t_i)$ are independent and by Theorem 2

$$MISE\{\widehat{\pi}(\cdot;\lambda_n)\} = o(n^{-\omega_1}), \qquad \text{where } \omega_1 < 1 - \frac{4}{a_0}$$

when $\lambda_n(a_0) = T\left(\frac{a_0}{\gamma}\right)^{\frac{1}{\beta}}(\log n)^{\frac{-1}{\beta}}$, where $a_0 > 4$.

Chapter 4

Limiting Distrubtion of the Deconvolution Kernel Density Estimator

We have shown the Deconvolution Kernel Density Estimator is a good estimator as the sample size gets large in terms of mean integrated squared error. Now we will now find the limiting distribution for the point wise evalution of the estimator. The standard technique for finding the limiting distribution is the Central Limit Theorem, but in our case, we cannot apply the Central Limit Theorem, since we do not have the independence assumption. In [16], Fan uses Lyapounov's condition and the triangular CLT on the standard deconvolution kernel density estimator to show the asymptotic normality of the estimator. As we cannot use the CLT, we will look to use a theorem that comes as a ramification of the CLT. In the following section, we provide the theorems and lemmas we will need for this section.

4.1 Preliminaries

The following theorem is found in [6] by Chung:

Theorem 4.1.

Suppose that $\{X_n\}$ *is a sequence of* m*-dependent, uniformly bounded random variables's and let*
$S_n = \sum_{i=1}^{n} X_i$ *such that*

$$\frac{\sigma(S_n)}{n^{1/3}} \to +\infty$$

as $n \to \infty$*. Then* $\dfrac{S_n - E(S_n)}{\sigma(S_n)}$ *converges to* $N(0,1)$*, where* σ *denotes the standard deviation of the random variable* S_n*.*

The sequence of $\{\widetilde{\theta}_n\}$ are 2-dependent and uniformly bounded, and hence $\left\{K^*\left(\frac{\widetilde{\theta}_n - x}{\lambda_n}\right)\right\}$ are also 2-depenent and uniformly bounded. To prove the asymptotic normality of $\hat{\pi}_\Theta$, we need to show that

$$\frac{\sigma\left(\sum_{i=1}^{n}\frac{1}{\lambda_n}K^*\left(\frac{x-\widetilde{\theta}_i}{\lambda_n}\right)\right)}{n^{1/3}} \to +\infty \tag{4.1}$$

Before we proceed with proofs of the two cases of error distributions, we will need the following lemmas from Fan [16]:

Lemma 4.2.

Suppose that $\{g_n\}$ is a sequence of Borel functions satisfying

$$g_n(y) \to g(y) \qquad \textit{and} \qquad \sup_n\{|g_n(y)|\} \le g^*(y),$$

where $g^*(y)$ satisfies

$$\int_{-\infty}^{\infty} g^*(y)dy < \infty \qquad \textit{and} \qquad \lim_{n\to\infty}|yg^*(y)| = 0.$$

If x is a point of continuity of a density $f(\cdot)$, then for any sequence $h_n \to 0$, we have

$$\lim_{n\to\infty}\frac{1}{h_n}\int_{-\infty}^{\infty} g_n\left(\frac{x-y}{h_n}\right)f(y)dy = f(x)\int_{-\infty}^{\infty} g(y)dy.$$

We will now use this lemma, and the following fact we proved earlier: $\frac{1}{\lambda_n}\mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right] = \frac{1}{\lambda_n}\mathrm{E}\left[K\left(\frac{x-\Theta}{\lambda_n}\right)\right]$ to show that $\frac{1}{\lambda_n}\mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right] \to \pi(x)$.

Proof.

$$\begin{aligned}\frac{1}{\lambda_n}\mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right] &= \frac{1}{\lambda_n}\mathrm{E}\left[K\left(\frac{x-\Theta}{\lambda_n}\right)\right] \\ &= \frac{1}{\lambda_n}\int_{-\infty}^{\infty}\pi(x)K\left(\frac{x-y}{\lambda_n}\right)dy\end{aligned}$$

using Lemma 4.2, we get as $n \to \infty$:

$$\to \pi(x) \int_{-\infty}^{\infty} K(y)dy$$

$$= \pi(x)$$

Hence, we have shown $\frac{1}{\lambda_n} \mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right] \to \pi(x)$. □

Since this limit exsists, we can say $\frac{1}{\lambda_n} \mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right]$ is bounded. This fact will be important for our proof. Next we will need another lemma for the super smooth case. We will write $\phi_\epsilon(t) = R_\epsilon(t) + iI_\epsilon(t)$, where $R_\epsilon(t)$ and $I_\epsilon(t)$ denote the real and imaginary parts of the characteristic function $\phi_\epsilon(t)$. From [16], Fan proves the following lemma:

Lemma 4.3.

If $\epsilon_i, i = 1, 2, \ldots, n$ follows a super smooth distribution and $I_\epsilon(t) = o(R_\epsilon(t))$ or $R_\epsilon(t) = o(I_\epsilon(t))$ as $t \to \infty$, then as $n \to \infty$

$$|K^*(y)| \geq cq(y) \exp\left(\frac{(1-b_n)^\beta}{\gamma h_n^\beta}\right) h_n^{\beta_0} b_n^4$$

uniformly over $y \in \left[0, \frac{\pi}{2}\right]$, where $b_n = h_n^{\frac{\beta}{10}}$, c is a positive constant, and

$$q(y) = \begin{cases} \cos y, & \text{if } I_\epsilon(t) = o(R_\epsilon(t)) \\ \sin y, & \text{if } R_\epsilon(t) = o(I_\epsilon(t)) \end{cases}$$

Finally, we will need two inequalities which will be written as lemmas:

Lemma 4.4.

For every real number $0 \geq r \geq 1$ and $x \geq -1$, then $(1+x)^r \leq 1 + rx$, which is a generalized version of Bernoulli's inequality.

Lemma 4.5.

For $x > 0$ and in a small neighborhood of zero and $\beta > 0$, we have $(1-x)^\beta \geq 1 - 2\beta x$.

The first inequality is fairly well known and can be proven either by induction or Taylors Formula. I will provide a proof of the second inequality. First we will write out Taylors formula for $(1-x)^\beta$ around the point 0: $(1-x)^\beta = 1-\beta x + R(x)$, where $R(x) = o(x)$. Since x is in a small positive neighborhood of zero, $R(x) = o(x)$ and $\beta > 0$, we know:

$$\begin{aligned} \beta x + R(x) &\geq 0 \\ 1 - \beta x + R(x) &\geq 1 - 2\beta x \\ (1-x)^\beta &\geq 1 - 2\beta x \end{aligned}$$

The above equation is a result of Taylors formula for $(1-x)^\beta$

Note: The second inequality is only true since x is in a small positive neighborhood of zero. If it wasn't, then the $R(x)$ wouldn't go to zero faster than the x term, and we wouldn't be able to write the first inequality. We are now ready to prove the two cases.

4.2 Limiting Distribution for Deconvolution Kernel Density Estimator

4.2.1 Case 1: Smooth Error Distribution

We will assume that $\epsilon_i, i = 1, 2, \ldots, n$ follows a smooth distribution with parameters: (d_0, d_1, β). Then X_i follows a super smooth distribution with parameters: $(d_0^2, d_1^2, 2\beta)$. We wish to show equation 4.1 is true with ϵ_i follow a smooth distribution. So for n sufficiently large:

$$\begin{aligned} \frac{\sigma\left(\sum_{i=1}^{n} \frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_i}{\lambda_n}\right)\right)}{n^{1/3}} &\geq n^{-\frac{1}{3}} \sigma\left(\frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right) \\ &= n^{-\frac{1}{3}} \left[\mathrm{E}\left[\left(\frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right)^2\right] - \left(\frac{1}{\lambda_n}\mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2\right] \end{aligned}$$

Since we have shown the second term is bounded, all we have to do is show the first term goes to ∞, and the proof is finished. We will focus on the first term.

$$\begin{aligned}
n^{-\frac{1}{3}}\mathrm{E}\left[\left(\frac{1}{\lambda_n}K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right)^2\right] &= \frac{1}{n^{\frac{1}{3}}\lambda_n^2}\int_{-\infty}^{\infty}\left[K^*\left(\frac{x-y}{\lambda_n}\right)\right]^2 f_{\widetilde{\theta}_j}(y)dy \\
&= \frac{1}{n^{\frac{1}{3}}\lambda_n^2}\int_{-\infty}^{\infty}\left[K^*\left(\frac{x-y}{\lambda_n}\right)\right]^2 f_{\widetilde{\theta}_j}(y)dy \\
&= \frac{f_{\widetilde{\theta}_j}(x)}{n^{\frac{1}{3}}\lambda_n^2}\int_{-\infty}^{\infty}\left[K^*\left(\frac{x-y}{\lambda_n}\right)\right]^2 dy
\end{aligned}$$

The equation above is derived by using Lemma 4.2 and now by using equation 3.6 we get:

$$\begin{aligned}
&= \frac{f_{\widetilde{\theta}_j}(x)}{2\pi\lambda_n n^{\frac{1}{3}}}\int \frac{\phi_K^2(t)}{\left|\phi_{X_1}\left(\frac{t}{\lambda_n}\right)\right|^2}dt \\
&\geq \frac{f_{\widetilde{\theta}_j}(x)}{2\pi\lambda_n n^{\frac{1}{3}}}\int \frac{\phi_K^2(t)}{\left|d_1^2\left|\frac{t}{\lambda_n}\right|^{-2\beta}\right|^2}dt \\
&= \frac{f_{\widetilde{\theta}_j}(x)}{2d_1^2\pi\lambda_n n^{\frac{1}{3}}}\int \frac{\phi_K^2(t)}{\left|\frac{\lambda_n}{t}\right|^{4\beta}}dt \\
&= \frac{f_{\widetilde{\theta}_j}(x)}{2d_1^2\pi\lambda_n^{1+4\beta} n^{\frac{1}{3}}}\int \phi_K^2(t)t^{4\beta}dt \\
&= \frac{kf_{\widetilde{\theta}_j}(x)}{\lambda_n^{1+4\beta} n^{\frac{1}{3}}},
\end{aligned}$$

where k is a positive constant. If we let $\lambda_n = an^{-b}$ where a, b are positive constants, then:

$$\begin{aligned}
\lim_{n\to\infty} n^{-\frac{1}{3}}\mathrm{E}\left[\left(\frac{1}{\lambda_n}K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right)^2\right] &\geq \lim_{n\to\infty} kf_{\widetilde{\theta}_j}(x)(an^{-b})^{-(1+4\beta)}n^{-\frac{1}{3}} \\
&= \lim_{n\to\infty} kf_{\widetilde{\theta}_j}(x)an^{b(1+4\beta)-\frac{1}{3}}
\end{aligned}$$

This last equation goes to infinity when $b(1+4\beta)-\frac{1}{3} > 0$. By choosing b such that $b > \frac{1}{3(1+4\beta)}$, we ensure the equation above tends to infinity as $n \to \infty$.

Theorem 4.6.

Let $\epsilon_i, i = 1, 2, \ldots, n$ follows a smooth distribution and $X_i = \epsilon_{i+1} - \epsilon_i$. Then as $n \to \infty$, we have

$$\frac{\widehat{\pi}(x;\lambda_n) - E[\widehat{\pi}(x;\lambda_n)]}{\sqrt{Var[\widehat{\pi}(x;\lambda_n)]}} \xrightarrow{D} N(0,1)$$

provided $\lambda_n = an^{-b}$ where a, b are positive constants and $b > \frac{1}{3(1+4\beta)}$.

4.2.2 Case 2: Super Smooth Error Distribution

We will assume that $\epsilon_i, i = 1, 2, \ldots, n$ follows a super smooth distribution with parameters: $(d_0, d_1, \beta_0, \beta_1, \beta, \gamma)$. Then X_i follows a super smooth distribution with parameters: $(d_0^2, d_1^2, 2\beta_0, 2\beta_1, \beta, \frac{\gamma}{2})$, and assume $I_{X_i}(t) = o(R_{X_i}(t))$ or $R_{X_i}(t) = o(I_{X_i}(t))$ as $t \to \infty$. We wish to show equation 4.1 is true with ϵ_i follow a supersmooth distribution. So for n sufficiently large:

$$\begin{aligned}
\frac{\sigma\left(\sum_{i=1}^{n} \frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_i}{\lambda_n}\right)\right)}{n^{1/3}} &\geq n^{-\frac{1}{3}} \sigma\left(\frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right) \\
&= n^{-\frac{1}{3}} \left[\mathrm{E}\left[\left(\frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right)^2\right] - \left(\frac{1}{\lambda_n}\mathrm{E}\left[K^*\left(\frac{x-\widetilde{\theta}_1}{\lambda_n}\right)\right]\right)^2\right]
\end{aligned}$$

Since we have shown the second term is bounded, all we have to do is show the first term goes to ∞, and the proof is finished. Now we will focus on the first term.

$$\begin{aligned}
n^{-\frac{1}{3}} \mathrm{E}\left[\left(\frac{1}{\lambda_n} K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right)^2\right] &= \frac{1}{n^{\frac{1}{3}}\lambda_n^2} \int_{-\infty}^{\infty} \left[K^*\left(\frac{x-y}{\lambda_n}\right)\right]^2 f_{\widetilde{\theta}_j}(y)dy \\
&= \frac{1}{n^{\frac{1}{3}}\lambda_n} \int_{-\infty}^{\infty} \left[K^*(y)\right]^2 f_{\widetilde{\theta}_j}(x-\lambda_n y)dy \\
&\geq \frac{1}{n^{\frac{1}{3}}\lambda_n} \int_{0}^{\frac{\pi}{2}} \left[cq(y)\exp\left(\frac{(1-b_n)^\beta}{\gamma\lambda_n^\beta}\right)\lambda_n^{2\beta_0} b_n^4\right]^2 f_{\widetilde{\theta}_j}(x-\lambda_n y)dy
\end{aligned}$$

The equaiton above is a result of using Lemma 4.3, and now by using the continuity of $f_{\widetilde{\theta}_j}$, we get:

$$\geq \frac{c_1 f_{\widetilde{\theta}_j}(x)}{n^{\frac{1}{3}}\lambda_n}\left[\exp\left(\frac{(1-b_n)^\beta}{\gamma\lambda_n^\beta}\right)\lambda_n^{2\beta_0}b_n^4\right]^2\int_0^{\frac{\pi}{2}}[q(y)]^2\,dy$$
$$\geq \frac{c_2 f_{\widetilde{\theta}_j}(x)}{n^{\frac{1}{3}}\lambda_n}\left[\exp\left(\frac{2(1-b_n)^\beta}{\gamma\lambda_n^\beta}\right)\lambda_n^{4\beta_0}b_n^8\right]$$

now using Lemma 4.5, we get:

$$\geq \frac{c_2 f_{\widetilde{\theta}_j}(x)}{n^{\frac{1}{3}}\lambda_n}\left[\exp\left(\frac{2(1-2\beta b_n)}{\gamma\lambda_n^\beta}\right)\lambda_n^{4\beta_0}b_n^8\right]$$
$$= \frac{c_2 f_{\widetilde{\theta}_j}(x)}{n^{\frac{1}{3}}\lambda_n}\exp\left(\frac{2\left(1-2\beta\lambda_n^{\frac{\beta}{10}}\right)}{\gamma\lambda_n^\beta}\right)\lambda_n^{4\beta_0}\lambda_n^{\frac{8\beta}{10}}$$
$$= \frac{c_2 f_{\widetilde{\theta}_j}(x)\lambda_n^{4\beta_0+\frac{8\beta}{10}-1}}{n^{\frac{1}{3}}}\exp\left(\frac{2\left(1-2\beta\lambda_n^{\frac{\beta}{10}}\right)}{\gamma\lambda_n^\beta}\right)$$

With this lower bound, we will choose $\lambda_n = a(\log n)^{\frac{-1}{\beta}}$ where a is a positive constant chosen such that $a < \left(\frac{6}{\gamma}\right)^{\frac{1}{\beta}}$, and we get:

$$\frac{c_2 f_{\widetilde{\theta}_j}(x)\lambda_n^{4\beta_0+\frac{8\beta}{10}-1}}{n^{\frac{1}{3}}}\exp\left(\frac{2\left(1-2\beta\lambda_n^{\frac{\beta}{10}}\right)}{\gamma\lambda_n^\beta}\right) = \frac{c_2 f_{\widetilde{\theta}_j}(x)\left(a(\log n)^{\frac{-1}{\beta}}\right)^{4\beta_0+\frac{8\beta}{10}-1}}{n^{\frac{1}{3}}}\exp\left(\frac{2\left(1-2\beta\left(a(\log n)^{\frac{-1}{\beta}}\right)^{\frac{\beta}{10}}\right)}{\gamma\left(a(\log n)^{\frac{-1}{\beta}}\right)^\beta}\right)$$
$$= \frac{c_3 f_{\widetilde{\theta}_j}(x)(\log n)^{\frac{-4\beta_0}{\beta}-\frac{8}{10}+\frac{1}{\beta}}}{n^{\frac{1}{3}}}\exp\left(\frac{2\left(1-2\beta a^{\frac{\beta}{10}}(\log n))^{-\frac{1}{10}}\right)}{\gamma a^\beta(\log n)^{-1}}\right)$$
$$= \frac{c_3 f_{\widetilde{\theta}_j}(x)(\log n)^{\frac{-4\beta_0}{\beta}-\frac{8}{10}+\frac{1}{\beta}}}{n^{\frac{1}{3}}}\exp\left(\frac{2}{\gamma a^\beta}\left(1-2\beta a^{\frac{\beta}{10}}(\log n))^{-\frac{1}{10}}\right)\log n\right)$$
$$= \frac{c_3 f_{\widetilde{\theta}_j}(x)(\log n)^{\frac{-4\beta_0}{\beta}-\frac{8}{10}+\frac{1}{\beta}}}{n^{\frac{1}{3}}}\exp\left(\log n^{\left(\frac{2}{\gamma a^\beta}\left(1-2\beta a^{\frac{\beta}{10}}(\log n))^{-\frac{1}{10}}\right)\right)}\right)$$
$$= \frac{c_3 f_{\widetilde{\theta}_j}(x)(\log n)^{\frac{-4\beta_0}{\beta}-\frac{8}{10}+\frac{1}{\beta}}}{n^{\frac{1}{3}}}\left(n^{\left(\frac{2}{\gamma a^\beta}\left(1-2\beta a^{\frac{\beta}{10}}(\log n))^{-\frac{1}{10}}\right)\right)}\right)$$
$$= c_3 f_{\widetilde{\theta}_j}(x)(\log n)^{\frac{-4\beta_0}{\beta}-\frac{8}{10}+\frac{1}{\beta}}\left(n^{\left(\frac{2}{\gamma a^\beta}\left(1-2\beta a^{\frac{\beta}{10}}(\log n))^{-\frac{1}{10}}\right)-\frac{1}{3}\right)}\right)$$

Since we have chosen a as we have, this leaves us with a limit of the form: $(\log n)^q n^{p-l(\log n)^{-k}}$ where p, l, k are all positive constants. Note: the choice of a above ensures us that p will be positive. So we have three cases to consider: when $q > 0$, $q < 0$, and $q = 0$. The solutions for $q > 0$ and $q = 0$ are essentially the same, and hence, we will only be showing the two cases:

- Case 1: When $q \geq 0$, define the following transformation: $n = \exp(m)$ then:

$$\begin{aligned}
\lim_{n\to\infty} (\log n)^q n^{p-l(\log n)^{-k}} &= \lim_{m\to\infty} m^q \exp\left(m(p - lm^{-k})\right) \\
&= \lim_{m\to\infty} m^q \exp\left(m\left(\frac{pm^k - l}{m^k}\right)\right) \\
&= \lim_{m\to\infty} m^q \exp\left(\frac{pm^k - l}{m^{k-1}}\right) = \infty
\end{aligned}$$

  since $q \geq 0$

- Case 2: When $q < 0$, define the following transformation: $n = \exp(m)$ then:

$$\begin{aligned}
\lim_{n\to\infty} (\log n)^q n^{p-l(\log n)^{-k}} &= \lim_{m\to\infty} m^q \exp\left(m\left(p - lm^{-k}\right)\right) \\
&= \lim_{m\to\infty} \frac{\exp\left(m\left(p - lm^{-k}\right)\right)}{m^{-q}}
\end{aligned}$$

  Note: $-q > 0$

$$\begin{aligned}
&= \lim_{m\to\infty} \frac{\exp(mp)\exp\left(-lm^{-k+1}\right)}{m^{-q}} \\
&= \lim_{m\to\infty} \frac{\exp(mp)}{m^{-q}\exp(lm^{-k})} \\
&> \lim_{m\to\infty} \frac{\exp(mp)}{3^{(lm^{-k})} m^{-q}} \\
&= \lim_{m\to\infty} \frac{\exp(mp)}{(1+2)^{(lm^{-k})} m^{-q}}
\end{aligned}$$

  Now by using Lemma 4.4, we get:

$$\begin{aligned}
&\geq \lim_{m\to\infty} \frac{\exp(mp)}{\left(1 + 2\left(lm^{-k}\right)\right) m^{-q}} \\
&= \lim_{m\to\infty} \frac{\exp(mp)}{m^{-q} + 2lm^{-(k+q)}} = \infty
\end{aligned}$$

  since $-q > 0$.

So we have shown that $(\log n)^q n^{p-l(\log n)^{-k}} \to \infty$ when p, l, k are all positive constants and q is a constant. With the choice of λ_n, we can now say:

$$n^{-\frac{1}{3}}\mathrm{E}\left[\left(\frac{1}{\lambda_n}K^*\left(\frac{x-\widetilde{\theta}_j}{\lambda_n}\right)\right)^2\right] \geq c_3 f_{\widetilde{\theta}_j}(x)\,(\log n)^{\frac{-2\beta_0}{\beta}-\frac{8}{10}+\frac{1}{\beta}}\left(n^{\left(\frac{2}{\gamma a^\beta}\left(1-2\beta a^{\frac{\beta}{10}}(\log n))^{-\frac{1}{10}}\right)-\frac{1}{3}\right)}\right)$$

$$\to \infty \qquad \text{as } n \to \infty$$

This shows

$$\frac{\sigma\left(\sum_{i=1}^{n}\frac{1}{\lambda_n}K^*\left(\frac{x-\widetilde{\theta}_i}{\lambda_n}\right)\right)}{n^{1/3}} \to \infty \qquad \text{as } n \to \infty$$

and by the theorem above, we have shown that $\dfrac{\widehat{\pi}(x;\lambda_n) - \mathrm{E}[\widehat{\pi}(x;\lambda_n)]}{\sqrt{\mathrm{Var}[\widehat{\pi}(x;\lambda_n)]}}$ converges in distribution to $N(0,1)$.

Theorem 4.7.

Let $\epsilon_i, i = 1, 2, \ldots, n$ follows a super smooth distribution and $X_i = \epsilon_{i+1} - \epsilon_i$ and $I_{X_i}(t) = o(R_{X_i}(t))$ or $R_{X_i}(t) = o(I_{X_i}(t))$ as $t \to \infty$. Then as $n \to \infty$, we have

$$\frac{\widehat{\pi}(x;\lambda_n) - E[\widehat{\pi}(x;\lambda_n)]}{\sqrt{Var[\widehat{\pi}(x;\lambda_n)]}} \xrightarrow{D} N(0,1)$$

provided $\lambda_n = a(\log n)^{\frac{-1}{\beta}}$, for some a such that $0 < a < \left(\frac{6}{\gamma}\right)^{\frac{1}{\beta}}$.

4.2.3 Sample Variance

Ultimatley, the variance of the deconvolution estimator depends on the value of the unkown density function $\pi(x)$. We can use a consistent estimator to replace the value $\pi(x)$, but this requires extra computation. As is the case, we would rather use sample variance of this estimator for inferences. In order to use the sample variance, we must show the sample variance converges in probability to $\mathrm{Var}[\pi(x;\lambda_n)]$. In [16], Fan shows the sample variance converges to the variance for both the smooth and supersmooth case. We provide his results for both cases for just the deconvolution estimator and not all the derivatives here:

Corollary 4.8.

Under the assumptions the error term follows a smoot distribution, if $h_n = o\left(n^{-\frac{1}{2\beta+1}}\right)$, *then*

$$\sqrt{n}\left(\frac{\hat{f}_n(x) - f_X(x)}{s_n}\right) \to N(0,1) \tag{4.2}$$

where s_n *is given by either* $s_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{h_n}K^*\left(\frac{x-X_i}{h_n}\right)\right)^2$ *or the sample variace defined by:* $s_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{h_n}K^*\left(\frac{x-X_i}{h_n}\right) - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}K^*\left(\frac{x-X_i}{h_n}\right)\right)^2$ *provided that* $nh_n^{2\beta+1} \to \infty$ *and* $\hat{f}$ *is the standard deconvolution kernel density estimator.*

Corollary 4.9.

Under the assumptions the error term follows a super smoot distribution, if $h_n \sim \left(\frac{a\gamma}{2}\log n\right)^{-\frac{1}{\beta}}$ *for some* $a > 1$, *then*

$$\sqrt{n}\left(\frac{\hat{f}_n(x) - f_X(x)}{s_n}\right) \to N(0,1) \tag{4.3}$$

where s_n *is defined in 2.1.*

The difference between Fan's result and our result is again the added complexty of the variance term as a whole. As we have already shown the term X_i follows smooth and supersmooth distribtutions in the appropriate cases, we can apply his result for the sums of the variances, and all that is left to handle is the covariance term.

By applying Chebysev's Inequaltiy, we know that the variance term goes to zero. So we just need to show the Covariance term does as well:

$$\begin{aligned}
\frac{2}{n\lambda_n^2}\mathrm{Cov}\left\{K^*\left(\frac{\widetilde{\theta}_1 - x}{\lambda_n}\right), K^*\left(\frac{\widetilde{\theta}_2 - x}{\lambda_n}\right)\right\} &\le \frac{2}{n\lambda_n^2}\mathrm{E}\left[K^*\left(\frac{\widetilde{\theta}_1 - x}{\lambda_n}\right)K^*\left(\frac{\widetilde{\theta}_2 - x}{\lambda_n}\right)\right] \\
&= \frac{2}{n\lambda_n^2}\mathrm{E}\left[K^*\left(\frac{\theta + \epsilon_1 - \epsilon_0 - x}{\lambda_n}\right)K^*\left(\frac{\theta + \epsilon_1 - \epsilon_0 - x}{\lambda_n}\right)\right] \\
&= \frac{2}{n\lambda_n^2}\int\cdots\int K^*\left(\frac{z+b-a-x}{\lambda_n}\right)K^*\left(\frac{z+c-b-x}{\lambda_n}\right)f_{\Theta;\epsilon_0;\epsilon_1;\epsilon_2}(z,a,b,c)\,dzdadbdcdx \\
&= \frac{2}{n\lambda_n^2}\int\cdots\int K^*\left(\frac{z+b-a-x}{\lambda_n}\right)K^*\left(\frac{z+c-b-x}{\lambda_n}\right)\pi(z)f_{\epsilon_0}(a)f_{\epsilon_1}(b)f_{\epsilon_2}(c)\,dzdadbdc \\
&= \frac{2}{n\lambda_n^2}\int\cdots\int K^*\left(\frac{z+b-a-x}{\lambda_n}\right)f_{\epsilon_0}(a)da\,K^*\left(\frac{z+c-b-x}{\lambda_n}\right)f_{\epsilon_1}(b)f_{\epsilon_2}(c)\,dbdc\,\pi(z)dz
\end{aligned}$$

using Lemma 4.2, we get:

$$= \frac{2}{n\lambda_n} \iiint f_{\epsilon_0}(z+b-x) \int K^*\left(a\right) da K^*\left(\frac{z+c-b-x}{\lambda_n}\right) f_{\epsilon_1}(b) f_{\epsilon_2}(c) db dc \pi(z) dz$$

We can bount the intgral of $K^*(a)$. Let c_1 denote such an upper bound.

$$\leq \frac{2c_1}{n\lambda_n} \iiint f_{\epsilon_0}(z+b-x) K^*\left(\frac{z+c-b-x}{\lambda_n}\right) f_{\epsilon_1}(b) db f_{\epsilon_2}(c) dc \pi(z) dz$$

using 4.2 again, we get:

$$\leq \frac{2c_1}{n} \iint f_{\epsilon_1}(z+c-x) \int f_{\epsilon_0}(b\lambda_n + 2b - c) K^*\left(b\right) db f_{\epsilon_2}(c) dc \pi(z) dz$$

as the integrals above can all be bounded. Let c_2 be such an upper bound

$$\leq \frac{2c_2}{n}$$

So as $n \to \infty$, we see that $\frac{2c_2}{n} \to 0$. This shows that the covariance term goes to zero. This proves that the sample variance still converges to the variance of our estimator in the dependent case. Hence, we can replace the variance term in Theorems (4.6) and (4.7) by the sample variance formula in the corollaries above.

References

[1] C. Butucea and C. Matias, Minimax Estimation of the Noise Level and of the Deconvolution Density in a Semiparametric cCnvolution Model. , *Bernoulli* **11** (2005), 309-340.

[2] C. Butucea, C. Matias, and P. Pouet , Adaptivity in Convolution Models with Partially Known. , *Electron. J. Statist.* **2** (2008), 897-915.

[3] R. J. Carroll and P. Hall, Optimal Rates of Convergence for Deconvolving a Density., *J. Amer. Statist. Assoc.* **83** No. 404 (1988), 1184-1186.

[4] R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu, Measurement Error In Nonlinear Models, , *2nd Edn. Boca Raton: Chapman and Hall-CRC* (2006).

[5] S. Charkraborty, N. Gebraeel, M. A. Lawley, and H. Wan, Residual-Life Estimation for Components with Non-Symmetric Priors., *IIE Transcations* **41** (2009), 372-387.

[6] K. L. Chung, A Course in Probability Theory. *American Press* (2001).

[7] A. Delaigle, P. Hall, and A. Meister, On Deconvolution with Repeated Measurements., *Ann. Statist.* **36** No. 2 (2008), 665-685.

[8] A. Delaigle, An Alternate View of the Deconvolution Problem , *Statist. Sin.* **18** (2008), 1025-1045.

[9] A. Delaigle and I. Gijbels, Estimation of Intergrated Squared Density Derivatives from a Contaminated Sample , *J. R. Statist. Soc. B* **64** (2002), 869-886.

[10] A. Delaigle and I. Gijbels, Comparison of Data-Driven Bandwidth Selection Procedures in Deconvolution Kenerl Density Estimation. , *Computnl Statist. Data Anal.* **45** (2004), 249-267.

[11] A. Delaigle and I. Gijbels, Bootstrap Bandwidth Selection in Kernel Density Estimation from a Contaminated Sample, *I. Ann Inst Stat Math* **56** (2004), 19-47.

[12] A. Delaigle and P. Hall , Using SIMEX for Smoothing-Parameter Choice in Errors-in-Variables Problems. , *J. Am. Statist. Ass.* **103** (2008), 280-287.

[13] A. Delaigle and P. Hall, Methodology for Non-Parametric Deconvolution When the Error Distribution is Unkown. , *J. R. Statist. Soc. B* **78** Part 1, (2016), 231-252.

[14] P. J. Diggle and P. Hall , A Fourier Approach to Nonparametric Deconolution of a Density Estimate. , *J. R. Statist. Soc. B* **55** (1993), 523-531.

[15] J. Fan, On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems., *Ann. Statist.* **19** (1991), 1257-1272.

[16] J. Fan , Asymptotic Normality for Deconvolution Kernel Density Estimators. , *Sankhayā: The Indian Journal of Statistics, Series A* **53** No. 1 (1991), 97-110.

[17] J. Fan, Adaptively Local One-Dimensional Subproblems with Application to a Deconvolution Model. , *Ann. Statist.* **21** (1991), 600-610.

[18] N. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, Residual-Life Distributions from Componenet Degradation Signals: A Bayesian Approach., *IIE Transactions* **37** (2005), 543-557.

[19] P. Hall and Y. Ma, Semiparametric Estimators of Functional Measurement Error Models with Unkown Error. , *J. R. Statist. Soc. B* **69** (2007), 429-446.

[20] A. Kneip, L. Simar, and I. Van Keilegom, Boundary Estimation in the Presence of Measurement Error with Unkown Variance , *J. Econmetr.* to be published.

[21] T. Li and Q. Vuong, Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators. , *J. Multiv. Anal.* **65** (1998), 139-165.

[22] X. Lin and R. J. Carroll, Semiparametric Estimation in General Repeated Measures Problems. , *J. R. Statist. Soc. B* **68** (2006), 69-88.

[23] A. Meister, Density Estimation with Normal Measurement Error with Unkown Variance. , *Statist. Sin.* **16** (2006), 195-211.

[24] M. H. Neumann, On the Effect of Estimating the Error Density in Nonparametric Deconvolution. , *J. Nonparam. Statist.* **7** (1997), 307-330.

[25] B. W. Silverman, Density Estimation for Statistics and Data Analysis. ,*Chapman & Hall/CRC*(1998).

[26] L. A. Stefanski, Rates of COnvergence of some Estimators in a Class of Deconvolution Problems. , *Statist. Probab. Lett.* **9** (1990), 229-235.

[27] L. A. Stefanski and R. J. Carroll, Deconvolving Kernel Density Estimators., *Statist.* **21** No. 2 (1990), 169-184.

[28] L. A. Stefanski and J. McIntyre, Density Estimation with Replicate Heteroscedastic Measurements. , *Ann. Inst. Statist. Math* **63** (2011), 81-99

[29] M. P. Wand and M. C. Jones, Kernel Smoothing. ,*Chapman & Hall/CRC*(1994).

Biographical Statement

Geoffrey H Schuette was born in Midland, Texas in 1990 to Terry and Deniece Schuette. He has two half brothers and three half sisters. He grew up in Midland, Texas and graduated from Grady High School in Lenorah, Texas. He graduate fifth out of twenty students.

After graduating high school, Geoffrey attended Sul Ross State University to play college tennis. He studied mathematics and kinesiology. He received his Bachelor's in Science in Mathematics in 2013.

After graduating with his bachelor's, Geoffrey attened The University of Texas at Arlington to pursue his doctorate degree. During this time, he recieved his Masters in Applied Mathematics under the suprevision of Dr. Andrzej Korzeniowski, and for his dissertation, he worked under the supervision of Dr. Shan Sun-Mitchell. He earned his Ph.D. in Mathematics in May 2018.

Geoffrey's interests lie in density estimation techniques and nonparametric statistics, with particular interest in kernel density estimation and the Deconvolution probelm. He also has interests in Stochastic Analsysi and Alegbraic Statistics.