Recovery of the Regression Functions and Distributions in some Incomplete Models

Broti Garai

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
Garai, Broti, "Recovery of the Regression Functions and Distributions in some Incomplete Models" (2016). Graduate Theses, Dissertations, and Problem Reports. 5648.
https://researchrepository.wvu.edu/etd/5648

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Recovery of the Regression Functions and Distributions in some Incomplete Models.

Broti Garai

Dissertation submitted to the Eberly College of Arts and Sciences at West Virginia University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computational Statistics

Robert Mnatsakanov, Ph.D., Chair
Michael Mays, Ph.D.
E. James Harner, Ph.D.
Krzysztof Chris Ciesielski, Ph.D.
Erin Leatherman, Ph.D.

Department of Statistics
Morgantown, WV
2016

Keywords: Moment-determinate functions; Laplace transform inversion; Nonparametric estimation

Copyright 2016 Broti Garai
Abstract

Recovery of the regression functions and distributions in some indirect models

Broti Garai

In this research, we propose formulas for recovering the regression function based on the product moments as well as recovering the distributions and derivative function in some indirect models. The upper bounds for the uniform rate of approximations are also derived. For regression functions two cases where the support of underlying functions is bounded and unbounded from above are studied. Based on the proposed approximations, new types of nonparametric estimates of the cumulative distribution, the density functions and the derivative functions in multiplicative-censoring model, as well as the approximations of conditional variance, are constructed. The suggested approach is also applied in the demixing problem and the constrained deconvolution problem, as well as for recovery of a distribution for unknown finite support. The absolute and mean squared errors of corresponding estimates are investigated as well. A simulation study justifies the consistency of the proposed estimates.
Acknowledgement

I would like to express my special appreciation and thanks to Dr. Robert Mnatsakanov, I am extremely fortunate for having him as my advisor. I would like to thank him for all the encouragement during my research and for allowing me to grow as a researcher. It was he who provided proper direction and guidance for the completion of the dissertation.

I would also like to thank my committee members, Dr. Michael Mays, Dr. E. James Harner, Dr. Chris Ciesielski and Dr. Erin Leatherman for their thoughtful input, cooperation, expertise and patience throughout the entire phase of this research.

I would also like to express my sincere gratitude to Dr. Robert Mnatsakanov, Dr. Mark Culp, Dr. Arun Ross and the Department of Statistics for proving me financial support throughout my Masters and Ph.D. I want to especially thank Mrs. Betty Wearden and the Department of Statistics for the Stanley Wearden Memorial Graduate Endowment Scholarship. I would also like to extend my thanks to Dr. Huey Minn Lee who gave me the opportunity to instruct a course and Barbara Dailey for all her efforts in solving our problem with Mathematica.

And talking about Morgantown: I have couple of special friends to thank, Dibyadhuti, Madhurima, Eepsa, Mimi, Susmita, Arnab, Sreeparna, Abhik, little Adi and little Ashmit who were a good support for me and important for all the fun I had in Morgantown. I also want to specially thank Dibyadhuti for giving his useful feedback and corrections during my dissertation writing.

Lastly, I would like to thank my family for all their love and encouragement. And most of all to my loving, supportive, and patient husband Prithish whose constant support and encouragement from the beginning to the final stages of this Ph.D. is so appreciated.
# Contents

1 Introduction  
  1.1 Regression Setup ........................................ 3  
  1.2 Multiplicative-censoring model ........................... 5  
  1.3 Other Incomplete Models .................................. 6  
  1.4 Recovery of Density and Distribution Function .......... 7  
  1.5 Organization of Dissertation ............................. 9  

2 Approximation of Regression Function .................. 11  
  2.1 Distribution of $X$ is known ............................ 11  
      2.1.1 Model 1: $X \sim U(0,1)$ ......................... 12  
      2.1.2 Model 2: $X \sim f, f$ is known ................ 18  
  2.2 Distribution of $X$ is unknown .......................... 26  

3 Estimation of Regression Functions .................... 33  
  3.1 Estimation of $r(x), x \in [0,1]$: Distribution of $X$ known . 34  
  3.2 Estimation of $r(x), x \in \mathbb{R}_+$: Distribution of $X$ unknown .. 36  

4 Recovery of Derivative Function with Application .... 42  
  4.1 Approximating the derivative function .................. 42  
  4.2 Multiplicative-censoring model ........................ 45  

5 Reconstruction of Distributions in other Incomplete Models 56  
  5.1 Demixing Problem ...................................... 57  
      5.1.1 Binomial Mixture ................................... 57  
      5.1.2 Negative Binomial Mixture ........................ 59  
      5.1.3 Exponential Mixture .............................. 61  
  5.2 Constrained Deconvolution Problem .................... 64  
  5.3 Recovery of Distribution with Unknown Support ....... 67  

6 Conclusion and Future Work .............................. 71  

Bibliography .............................................. 74
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Approximated regression curves (blue dots) (a): $r_{\alpha,1}$ with $\alpha = 30$; and (b): $\tilde{r}_{\alpha,1}$ with $\alpha = 15$. In both plots the true regression function is $r(x) = (2x + 3)/12$ (orange curve).</td>
</tr>
<tr>
<td>2.2</td>
<td>(a) $\sigma^2_{\alpha,1}(x)$ (blue dots) when $\alpha = 30$ and (b) $\tilde{\sigma}^2_{\alpha,1}(x)$ (blue dots) when $\alpha = 15$. In both plots, The true conditional variance $\sigma^2(x)$ is displayed as well (orange).</td>
</tr>
<tr>
<td>2.3</td>
<td>Errors of approximation in (a) $\sup$-norm and (b) $L_2$-norm for $r_{\alpha,1}$, when $\alpha \in [30, 230]$.</td>
</tr>
<tr>
<td>2.4</td>
<td>Approximated regression functions (blue dots) (a) $r_{\alpha,2}$ with $\alpha = 40$ and (b) $\tilde{r}_{\alpha,2}$ with $\alpha = 20$. The true regression $r(x) = 2(1 - x)/5$ (orange curve).</td>
</tr>
<tr>
<td>2.5</td>
<td>Plots of the approximated conditional variances (blue dots) in (a): $\sigma^2_{\alpha,2}(x)$ and (b): $\tilde{\sigma}^2_{\alpha,2}(x)$ when $\alpha = \alpha_2 = 30$.</td>
</tr>
<tr>
<td>2.6</td>
<td>The error of approximations in (a) $\sup$-norm and (b) $L_2$-norm for $\alpha$ ranging from 30 to 230.</td>
</tr>
<tr>
<td>2.7</td>
<td>Approximated curve (blue) (a) $r_{\alpha,2}$ and (b) $\tilde{r}_{\alpha,2}$ along with the true regression $r(x) = \frac{T_1 x}{T_2}$, when $\alpha = \alpha_2 = 60$.</td>
</tr>
<tr>
<td>2.8</td>
<td>(a) $\sigma^2_{\alpha,2}(x)$ and (b) $\tilde{\sigma}^2_{\alpha,2}(x)$ (both blue) represents the approximated conditional variance, when $\alpha = \alpha_2 = 60$ and $T = 3$.</td>
</tr>
<tr>
<td>2.9</td>
<td>The errors of approximations in $r_{\alpha,2}$ in (a) $\sup$-norm and (b) $L_2$-norm between when $\alpha \in [30, 230]$.</td>
</tr>
<tr>
<td>2.10</td>
<td>Approximated (a) $r_{\alpha,b}$ and (b) $\tilde{r}_{\alpha,b}$ (blue) of the true regression $r(x) = (2x^2 + 8)/(x^2 + 2)$ (orange) when $\alpha = \alpha_2 = 40$, and $b = 1.15$.</td>
</tr>
<tr>
<td>2.11</td>
<td>(a) $\sigma^2_{\alpha,b}(x)$ and (b) $\tilde{\sigma}^2_{\alpha,b}(x)$ represents the approximated conditional variance, when $\alpha = \alpha_2 = 40$, and $b = 1.15$.</td>
</tr>
<tr>
<td>2.12</td>
<td>The (a) $\sup$-norm distance and (b) $L_2$-norm distance for $r_{\alpha,b}$, when $\alpha \in [30, 230]$ and $b \in [1.1, 2.4]$.</td>
</tr>
<tr>
<td>2.13</td>
<td>The approximated (a) $r_{\alpha,b}$, (b) $\tilde{r}_{\alpha,b}$ and the true regression function $r(x) = \frac{1}{x^2} + \frac{\theta}{12}(1 - 2e^{-\lambda_1 x})$ (orange) are displayed. Here $\alpha = \alpha_2 = 30$ and $b = 1.2$.</td>
</tr>
</tbody>
</table>
2.14 (a) $\sigma_{\alpha,b}^2(x)$ and (b) $\tilde{\sigma}_{\alpha,b}^2(x)$ represents the approximated conditional variance, when $\alpha = \alpha_2 = 30$, and $b = 1.2$.  

2.15 The (a) $sup$-norm distance and (b) $L_2$-norm distance when $\alpha$ ranges from 30 to 230 and $b$ ranges from 1.1 to 2.5.  

3.1 $\hat{r}_{\alpha,1}$ (blue) and $r(x) = \frac{(2x+3)}{12}$ (orange) when $\alpha = 30$.  

3.2 $\hat{r}_{\alpha,2}$ (blue) and $r(x) = \frac{b(1-x)}{b+c}$ (orange) when $\alpha = 30$.  

3.3 Estimated $\hat{r}_{\alpha,b}$ (blue) of the true regression $r(x) = \frac{1}{\alpha}y + \frac{b}{2\alpha^2}(1 - 2e^{-\lambda x})$ (orange) when $\alpha = 30$ and $b = 1.2$.  

3.4 Estimated $\hat{r}_{\alpha}$ (red), $\hat{r}_b$ (blue) and $\tilde{r}_{\alpha,b}$ (green) of the true regression $r(x) = (x - 1.5)^2$ (black) when $\alpha = 100$ and $b = 1.52$.  

3.5 Estimated $\hat{r}_{\alpha}$ (red), $\hat{r}_b$ (blue) and $\tilde{r}_{\alpha,b}$ (green) of the true regression $r(x) = \frac{1}{x^2}$ (black) when $\alpha = 100$ and $b = 3.515$.  

4.1 Approximations (the dotted curves) (a): $g'_\alpha$ when $\alpha = 50$; and (b): $\tilde{g}'_\alpha$ when $\alpha = 50$. In both plots the true $g'(y) = 60y(1-y)(1-2y)$.  

4.2 Estimated $\hat{F}_{\alpha,n}$ (dotted curve) (a): when $n = 500$, $\alpha = 35$; and (b): when $n = 1000$, $\alpha = 30$. In both plots the true $F$ (continuous curve) represents the cdf of a $Beta(3,3)$ distribution.  

4.3 Estimated $\hat{f}_{\alpha,n}$ (dotted curve) (a): when $n = 1000$, $\alpha = 25$; and (b): when $n = 1500$, $\alpha = 30$. Here $X \sim Beta(3,3)$.  

4.4 Estimated $\hat{f}'(x)$ (dotted curve) (a): when $n = 3000$, $\alpha = 25$; and (b): when $n = 5000$, $\alpha = 30$. Here $X \sim Beta(3,3)$.  

4.5 Estimated $\hat{F}_{\alpha,n}$ (dotted curve) when (a): $n = 2000$, $\alpha = 25$; and (b): $n = 3000$, $\alpha = 30$. In both plots the true $F$ (continuous curve) represents cdf of a $X \sim (1/3)Beta(4,6) + (2/3)Beta(3,4)$ distribution.  

4.6 Estimated $\hat{f}_{\alpha,n}$ (dotted curve) when (a): $n = 5000$, $\alpha = 20$; and (b): $n = 4000$, $\alpha = 25$. Here $X \sim (1/3)Beta(4,6) + (2/3)Beta(3,4)$.  

4.7 Estimated $\hat{f}'(x)$ (dotted curve) when (a): $n = 5000$, $\alpha = 20$; and (b): $n = 4000$, $\alpha = 25$. Here $X \sim (1/3)Beta(4,6) + (2/3)Beta(3,4)$.  

5.1 Estimated pdf (a): $f'_{\alpha,T}(x)$ and cdf (b): $F_{\alpha,T}(x)$ for the binomial mixture when $X \sim Beta(3,2)$  

5.2 Approximated (a) pdf and (b) cdf for negative binomial mixture (black). In both plots the red line denotes the true pdf $f(x) = 2x$ and cdf $f(x) = x^2$, when $\alpha = 30$.  

5.3 Approximated $a: f_\alpha(x)$ and $b: F_\alpha(x)$ (black) for $f(x) = 2(x+1)^{-3}$ and $F(x) = 1 - (x+1)^{-1}$ (red) when $\alpha = 30$.  

5.4 Approximated $a: f_\alpha(x)$ and $b: F_\alpha(x)$ (black) for $f(x) = 4(x+1)^{-5}$ and $F(x) = 1 - (x+1)^{-4}$ (red) when $\alpha = 30$.  

vi
5.5 Approximated (a) pdf and (b) cdf for Constrained devolution (black). In both plots the red line denotes the true pdf $f(x) = e^{-x}$ and cdf $F(x) = 1 - e^{-x}$. ................................................................. 67

5.6 Approximated $F_{\alpha,m}$ when $\alpha = 30$ and $m = 500$ corresponding to example 14. ................................................................. 68

5.7 Approximated (a) $f_{\alpha,m}$ and (b) $F_{\alpha,m}$ of $f$ and $F$ for unknown $T$
when $\mu_{j,F} = \frac{9}{(j+3)^2}$, $f(x) = -2x^2 \log(x^3)$ and $F(x) = x^3 - x^3 \log(x^3)$. 69

5.8 Approximated (a) $f_{\alpha,m}$ and (b) $F_{\alpha,m}$ of $f$ and $F$ for unknown $T$
when $\mu_{j,F} = \frac{1}{(j+1)^2}$, $f(x) = -2 \log(x)$ and $F(x) = x - x \log(x)$. ... 70
List of Tables

2.1 The error of approximations of $r_{\alpha,2}$, $\tilde{r}_{\alpha,2}$, and $r_n$ in sup-norm and $L_2$-norm for $\alpha = \alpha_2 = n = 30$. ............................................ 24

3.1 Error in approximation for $\hat{r}_{\alpha,1}$, when $X \sim U(0, 1)$ and $Y \sim F(y) = y^{1/2}$. 35

3.2 Error in approximation for $\hat{r}_{\alpha,2}$ when $f(x, y) = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}x^{a-1}y^{b-1}(1-x-y)^{c-1}$. .................................................. 36

3.3 The errors of approximations for $\hat{r}_{\alpha,b}$, when $h(x, y) = g(x, 1, 1)g(y, 4, 1)/2 + g(x, 3, 1)g(y, 2, 1)/2$. .................................................. 38

3.4 The errors of approximations for NW estimate, Koul and Song estimate and $\hat{r}_{\alpha,b}$ in $L_2$-error, when $X \sim Log-Normal(0, 1)$ and $r(x) = (x-1.5)^2$. .................................................. 40

3.5 The errors of approximations for NW estimate, Koul and Song estimate and $\hat{r}_{\alpha,b}$ in $L_2$-error, when $X \sim Log-Normal(0, 1)$ and $r(x) = \frac{1}{x}$. 41

4.1 The errors of approximations $\hat{F}_{\alpha,n}$ in $L_2$-distance. Here $X \sim Beta(3, 3)$. 51

4.2 The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha,n}$ in $L_2$-distance. Here $X \sim Beta(3, 3)$. .................................................. 52

4.3 The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha,n}$ in $L_2$-distance. Here $X \sim Beta(3, 3)$. .................................................. 52

4.4 The errors of approximations for $\hat{F}_{\alpha,n}$ in $L_2$-distance. Here $X \sim (1/3)Beta(4, 6) + (2/3)Beta(3, 4)$. .................................................. 53

4.5 The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha,n}$ in $L_2$-distance. .................................................. 54

4.6 The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha,n}$ in $L_2$-distance. .................................................. 54

5.1 The errors of approximations for $f_{\alpha,T}^*$ and $F_{\alpha,T}^*$ in $L_2$-distance, when $Y|X = x \sim Bin(m, x)$. .................................................. 58
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm and $L_2$-norm, when $Y</td>
<td>X = x \sim NegBin(r, \phi(x))$.</td>
</tr>
<tr>
<td>5.3</td>
<td>The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm and $L_2$-norm, when $Y</td>
<td>X = x \sim Exp(\phi(x))$ and $F(x) = 1 - (x + 1)^{-2}$.</td>
</tr>
<tr>
<td>5.4</td>
<td>The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm and $L_2$-norm, when $Y</td>
<td>X = x \sim Exp(\phi(x))$ and $F(x) = 1 - (x + 1)^{-4}$.</td>
</tr>
<tr>
<td>5.5</td>
<td>The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm.</td>
<td>66</td>
</tr>
<tr>
<td>5.6</td>
<td>The errors of approximations for $F_{\alpha,m}$ in $sup$-norm.</td>
<td>68</td>
</tr>
<tr>
<td>5.7</td>
<td>The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm.</td>
<td>69</td>
</tr>
<tr>
<td>5.8</td>
<td>The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm.</td>
<td>70</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The moment problem is one of the most famous classical problem in mathematics and applied statistics. The goal of this project is to determine regression functions and distributions based on their sequence of moments. There are three well known moment problems, including the Hausdorff moments problem, the Stieltjes moment problem, and the Hamburger moment problem. We say the Hamburger moment problem has a unique solution if there exists only one distribution $F$ on the real line $(-\infty, \infty)$ such that the given sequence $m_{j,F} = \int_{-\infty}^{\infty} t^j dF(t)$, $j = 0, 1, \ldots$, identifies $F$ uniquely. If the support of $F$ is compact, say, $[0, 1]$ then the problem is known as the Hausdorff moment problem, whereas if the support of $F$ is $(0, \infty)$, we have a Stieltjes moment problem [40].

There exist several methods that are used to recover distributions and probability density functions e.g., the Maximum Entropy method, the approach based on the Laplace transform inversion, and methods based on mixtures. The Laplace transform of a function $f(x)$ is defined as, $\mathcal{L}_f(s) = \int_0^\infty e^{-st} f(t) dt$, for $s > 0$. In order to recover the distribution function $F$ and its density $f$, the Laplace transform inversion is of-
ten implemented. Mnatsakanov ([28], [29] and [31]) approximated univariate and bivariate distributions given the sequence of moments or a finite number of Laplace transformed functions. In these papers it was shown that the analytical forms of the approximations are easily evaluated depending on the exponential (transformed) moments of an underlying distribution. It is also seen that for a large number of moments \( \alpha \), the approximated pdf \( f_\alpha \) is stable. It should be noted that this method is applicable for estimation of \( f_\alpha \) in situations when only the estimated moments of the target distributions are available.

Consider the maximum entropy setup: denote \( f(x) \) as the density function of \( X \), supported by \( S \subset \mathbb{R} \) and assume the moment values of \( m(j) = \int_0^1 t^j f(t)dt, \ j = 0,1,2,\cdots \), are given. The approximations of \( f \) belong to the class of functions \( \mathcal{F} = \{ f_M : f_M(x) = c e^{-\left(\sum_{j=1}^{M} \lambda_j x^j\right)} \text{ for all } x \in S \} \), where \( c \) and \( \lambda_j, \ j = 0,1,2,\cdots \), are determined so that the integral of \( f_M(x) \) over \( S \) is 1. Maximizing the differential entropy \( H(f) = -\int_0^1 f(x) \log f(x)dx \) with respect to \( f \in \mathcal{F} \) yields the maximum entropy approximation. The maximum entropy method was used in [15], [35], [41] and [42] to reconstruct the unknown density from its moments. Gzyl and Tagliani [15] investigated Stieltjes and Hausdorff moment problems in the maximum entropy setup. They calculated the condition number in entropy, endowing both Stieltjes moment problem existence conditions and Hausdorff moment problem, determinacy conditions by a geometric meaning. Inverardi et. al. [35] used low number of fractional moments to minimize the entropy of the approximate densities. Tagliani [41] and [42] considered the classical Hausdorff moment problem and reduced Hamburger moment problem to recover the density from fractional moments under the existence of maximum entropy set up.
The method of moments can be employed in various statistical problems. Several applications of moment problems are considered in this dissertation. One such application is recovering the regression functions when the moments of the distributions are available. Another implementation studied is reconstructing the derivative function. Rebuilding the mixing distribution of a mixture model using the moments is also considered. In addition we approximated the distribution and density function in some incomplete models. All the proposed approximations are implemented using several examples and comparisons are made with well-known methodology.

1.1 Regression Setup

Often, two components of a dataset are observed, the response vector \( Y = (y_1, y_2, \ldots, y_n) \) and the predictive data matrix \( X = (x_1, x_2, \ldots, x_n) \), where each \( x_i \) is a \( p \)-dimensional vector. The goal of many statistical problems is to express the response \( Y \) as a function of the predictive variables in \( X \), i.e., \( Y = r(x) + \epsilon \), where \( \epsilon \) is the error associated with the model. Optimizing the unknown parameters of a given regression model is fairly straightforward when both \( Y \) and \( X \) are observed directly. One such example is the linear regression model that represent response \( Y \) as an additive function of \( X \) as: \( Y = \alpha + \beta X + \epsilon \), where \( \alpha \) and \( \beta \) are unknown parameters of the model. A very popular regression model is the simple linear regression model or ordinary least squares regression model, when \( Y \in \mathbb{R} [11] \).

Suppose instead of observing the variables \((Y, X)\), we only have information about the averages/product moments of the form: \( m(k, j) = E(Y^kX^j) \), with \( k = 0, 1, 2 \) and \( j = 1, 2, \ldots, \alpha \). The representation of the response \( Y \) as a functional form \( r(x) \) becomes a complex problem as no common regression techniques can be
implemented with product moments. The main objective of this research is to re-
cover the regression function $r(x)$ given the information contained in the sequence of
product moments $m(k, j)$, $k = 0, 1, 2$ and $j = 1, 2, \ldots, \alpha$.

A nonparametric estimator of the regression function $r(x) \equiv E(Y \mid X = x)$ with
product (empirical) moments of $Y$ and $X$ is used in this dissertation. The most
common choice of nonparametric regression is kernel based methods such as the
Nadaraya-Watson (NW) estimator. The NW estimator is given by

$$
\hat{r}_h(x) = \frac{\sum_{i=1}^{n} K_h(X_i - x)Y_i}{\sum_{i=1}^{n} K_h(X_i - x)}
$$

(1.1)

where $K$ is a symmetric kernel density function and $K_h(.) = K(. / h) / h$. One of the
most commonly used kernels is the Gaussian kernel $K(t) = (\sqrt{2\pi})^{-1} e^{-t^2/2}$. Mnat-
sakanov and Sarkisian [32] (M-S) proposed a new nonparametric density estimator
on $\mathbb{R}_+$ based on a sequence of varying kernels. The derived estimator of the pdf $f(x)$
for $x \in \mathbb{R}_+$ is

$$
\hat{f}_\alpha(x) = \frac{1}{n} \sum_{i=1}^{n} K_\alpha(x, X_i),
$$

where

$$
K_\alpha(x, t) = \frac{1}{t \Gamma(\alpha)} \left( \frac{\alpha x}{t} \right)^{\alpha - 1} e^{-\alpha x / t}.
$$

(1.2)

A varying kernel estimator of $r(x)$ was proposed by Koul and Song [23]. The esti-
mates are particularly designed for positive random variables. The M-S kernel esti-
mator of $r(x)$ is defined as,

$$
\hat{r}_\alpha(x) = \frac{\sum_{i=1}^{n} K_\alpha(x, X_i)Y_i}{\sum_{i=1}^{n} K_\alpha(x, X_i)},
$$

with the same $K_\alpha(x, t)$ defined in (1.2).

Under certain smooth conditions of the underlying regression function $r(x)$, and
by requiring its moment-determinacy, Hagwood [17] approximated the conditional expectation $E(Y|X = x) = r(x)$ using the sequence of moments $m(1,j)$ with $j = 0, 1, \ldots, 2n$. When the support of $r$ is finite, say $[0,1]$, the following approximation was introduced as

$$r_n(x) = \frac{1}{f(x)} \sum_{j=0}^{2n} \nu_{[n(M_n+w)/M_n],j} \frac{n_j+1}{M_n^{j+1}} m(1,j), \quad 0 \leq x \leq 1,$$

(1.3)

where $\nu_{i,j}$ is the $(i,j)$th entry of the inverted Vandermonde matrix $V(-n, -n+1, \ldots, n)$ [3] and [4]. The marginal density function of $X$ is $f(x) = F'(x)$ and $M_n$ is a sequence of positive real numbers such that $M_n/n \to 0$. It was mentioned that the reconstruction of $r(x)$ based on (1.3) provides an improved approximation compared to normal approximation.

Moreover, it is well known that the conditional variance $\sigma^2(x)$ affects the association function $r(x)$. Therefore the knowledge of $\sigma^2(x)$ is necessary to fully understand the measure of association between response $Y$ and predictor $X$. Thus another related problem considered in this dissertation is the retrieval of the conditional variance $\sigma^2(x)$ from the product moments $m(k,j)$ [27].

### 1.2 Multiplicative-censoring model

We address the problem of recovering the derivative of a density function $g$ of a random variable $Y$, in the case where only the sequence of moments $E[Y^j], j \in \mathbb{N}_+$, is given. The application in multiplicative-censoring model is outlined as well. Suppose we observe $Y = X U$, where $X$ and $U$ both are unobserved independent random variables, with $X \sim F$(unknown) and $U \sim Uniform(0,1)$. In the multiplicative-
censoring model our aim is to estimate the cdf $F$, the pdf $f$ and the derivative function $f'$ given $n$ independent copies of $Y$ [27]. In particular, we derived the consistent nonparametric estimates of the associated distribution, density and and its derivative function. Abbaszadeh et. al. [1] studied the problem of estimating the density $f$ and its derivative function $f'$. The authors used an approach based on wavelets and explored the rates of convergence of the estimates using simulated data. Anderson and Hansen [2] proposed a different approach for estimating the density function using the linear inverse problem. Their approach is based on singular value decomposition of the desired density.

1.3 Other Incomplete Models

An interesting problem is the recovery of the distribution for a mixture model. In the mixture distribution we assume $Y|X = x$ follows a known distribution $g(y) = P(Y = y)$ with a random parameter $X \sim f(x)$. Therefore another objective of the work is to recover the mixing distributions in several mixture models, e.g., Binomial, Negative Binomial and Exponential. A simulation study illustrates the performance of our model in recovering the mixing distributions.

The deconvolution problem can be formulated as follows: assume one observes $n$ copies of $Z = X + Y$, where $X$ and $Y$ are independent and the distribution of $X$ is unknown. Our aim is to recover the distribution of $X$. In this dissertation we will follow a special case of the deconvolution problem (known as constrained deconvolution) when $Z = aX + \beta Y$ where $X$ and $Y$ follows the same unknown distribution $F$, and $(\alpha, \beta)$ are unknown parameters. Belomestny [5] constructed a non-parametric estimator of $f = F'$ using a Fourier transform inversion. He also investigated the
asymptotic behavior of the mean integrated square error. For this model we suggest two approximations: the cdf $F$ and the pdf $f$, based on the inversion of the scaled Laplace transformation.

Lastly, the information regarding the support of a distribution may be incomplete in a sense that, $X$ follows some distribution which lies within $l_1$ and $l_2$. Information of either or both of $l_1$ and $l_2$ may be missing. Estimation of the missing support via the method of moments is demonstrated in this dissertation.

1.4 Recovery of Density and Distribution Function

In this dissertation it is assumed that the regression function defined on $\mathbb{R}_+$ is moment-determinate ($M$-determinate). The conditions under which a function is $M$-determinate or $M$-indeterminate is a classical mathematical problem and has been studied in many works, see, for example, [21], [24], [40] and [38]. Additionally, there are several techniques that provide the reconstruction of the probability density function from the sequence of its moments.

Ha and Provost [36] showed that under certain conditions the density function can be approximated by $f_d(x) = \psi(x) \sum_{\ell=0}^{d} \xi_\ell x^\ell$, where $\psi(x)$ is the initial density approximation of $f(x)$. The values of $\xi_\ell$, $\ell = 0, 1, \cdots, d$ are determined by equating the first $d$ moments obtained from $f_d(x)$ to those of $X$, i.e. via

$$
\begin{bmatrix}
\xi_0 \\
\xi_1 \\
\vdots \\
\xi_d
\end{bmatrix} = \begin{bmatrix}
m(0) & m(1) & \cdots & m(d) \\
m(1) & m(2) & \cdots & m(d+1) \\
\vdots \\
m(d) & m(d+1) & \cdots & m(2d)
\end{bmatrix}^{-1} \begin{bmatrix}
m(1) \\
\mu_X(1) \\
\vdots \\
\mu_X(d)
\end{bmatrix},
$$

where
where $\mu_{X}(j) \equiv E(X^{j})$ and $m(j) \equiv \int_{a}^{b} x^{j} \psi(x) dx$.

In order to recover and estimate the distribution of the summands of the Poisson distribution, Mnatsakanov and Sarkisian [30] suggested a uniform upper bound for the MR-approximation of the cdf $F$. In the framework of moment problems (Hausdorff, Stieltjes, and Hamburger), Mnatsakanov [28], [29], [31], and [33] proposed a method for recovering the cumulative distribution function as well as the corresponding pdf. The reconstructions of the cdf and pdf are based on the sequence of $\phi$-transformed moments $\{m_{\phi}(j)\}_{j=0}^{\alpha}$ of $F$ up to a given order $\alpha$. The approximated cdf and pdf can be written in a unified form as follows: for each $\alpha \in \mathbb{N}$, $F_{\alpha, \phi} := \mathcal{K}_{\alpha}^{-1} m_{\phi}$ and $f_{\alpha, \phi} := \mathcal{B}_{\alpha}^{-1} m_{\phi}$ we have

$$
(\mathcal{K}_{\alpha}^{-1} m_{\phi})(x) = \sum_{k=0}^{[\alpha \phi(x)]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{\alpha}{k} (-1)^{j-k} m_{\phi}(j) \quad (1.4)
$$

and

$$
(\mathcal{B}_{\alpha}^{-1} m_{\phi})(x) = \frac{\Gamma(\alpha + 2)|\phi'(x)|}{\Gamma([\alpha \phi(x)] + 1)} \sum_{j=0}^{[\alpha \phi(x)]} (-1)^{j} m_{\phi}(j + [\alpha \phi(x)]) \frac{(\alpha - [\alpha \phi(x)] - j)!}{j!(\alpha - [\alpha \phi(x)] - j)!}, \quad (1.5)
$$

respectively. Here, $m_{\phi}(j) = \int \phi(t)^{j} dF(t)$, and depending on the form of support of $F$, the function $\phi$ is defined. For example,

$$
\phi(x) = \begin{cases} 
\frac{x}{T} & \text{if } 0 \leq x \leq T \\
 b^{-x} & \text{if } x \in \mathbb{R}_{+}.
\end{cases}
$$

As it is shown in [28] and [29], the recovered constructions $F_{\alpha, \phi}$ and $f_{\alpha, \phi}$ provide a stable approximation of $F$ and $f$. 
It is worth mentioning that the methods 1.5 and 1.3 are applied when reconstructing the pdf, while the method 1.4 can be used for recovering the $M$-determinate function as well (see Mnatsakanov [33] and Example 7 in Section 4.1). This fact enables us to apply construction (1.5) to recover the regression function, the conditional variance, and the density functions as well. Construction (1.4) is used for the approximation and estimation of the cumulative distribution functions.

### 1.5 Organization of Dissertation

The organization of this dissertation is as follows. In Chapter 1 we introduced the methods and problems that will be discussed in depth in this dissertation. In Chapter 2 two cases of recovering regression functions are investigated: (a) when the distribution of a predictive variable $X$ is known and (b) when the distribution of $X$ is unknown. The proposed construction of the regression function is based on the knowledge of the sequence of product moments of the response $Y$ and the covariate variables $X$ up to order $\alpha$. In Chapter 3 the regression functions are estimated for both case (a) and (b). The comparisons of proposed estimates of the regression function with kernel-type constructions are conducted via a simulation study, and plots of the recovered functions are presented for several examples.

In Chapter 4 we discuss the recovery of the derivative function and the distributions in case of the multiplicative-censoring model, constrained deconvolution problem, and when the upper bound of support of $F$ is unknown. In Section 4.1 we also approximate the derivative function $g'$ of a continuous density function $g$ and address the problem of estimating the distribution function $F$, the corresponding density $f$ and the derivative of the density $f'$ in the framework of the multiplicative-censoring
The recovery of density and distribution functions in the demixing problem is explained in Chapter 5, Section 5.1. Namely in the framework of binomial, exponential, and negative binomial mixture models, the problem of nonparametric estimation and approximation of the mixing distribution and its density function is studied. The approximation of the pdf and cdf of $X$ for the constrained deconvolution problem is considered in Section 5.2. In Section 5.3 we consider several examples to recover the distribution and density functions when the upper bound is unknown.

In the chapters that follow, simulation studies and comparisons with other constructions are conducted. Graphical illustrations and tables with the values of the errors of recovered functions are also provided. The upper bounds for the uniform rate of convergence for the approximated regression and the derivative functions are derived. Moreover, we investigate the asymptotic behavior of the distances between the approximated and true functions. Chapter 6 outlines some concluding remarks about this research and some prominent future research directions.
Chapter 2

Approximation of Regression Function

The motivation behind the problem of recovering the regression function \( r(x) \) is provided in Section 1.1. In order to achieve this task two separate cases are investigated in this Chapter: (a) the distribution of \( X \) is known in Section 2.1 and (b) the distribution of \( X \) is unknown in Section 2.2. The upper bounds for the uniform rate of convergence for the regression function are derived. Moreover, we investigate the asymptotic behavior of the distances between the approximated and the true regression functions for all cases [27].

2.1 Distribution of \( X \) is known

In this section we only concentrate on the approximation of the regression function \( r(x) \) for a known distribution of \( X \). In addition, we assume that the distribution function \( F(X) \) maintains a finite support, i.e, \( supp\{F\} = [0,T]; \ 0 \leq T < \infty \). For the
sake of simplicity, we first consider $T = 1$. Section 2.1.1 describes such a process with $X$ following the uniform distribution on $(0, 1)$, i.e., $X \sim U(0, 1)$ (see Example 1). Theoretical explanations of the uniform rate of convergence and the asymptotic behavior are also provided in this section. Two more general cases are investigated in Section 2.1.2: (1) $X \sim Beta(a, b)$ with known $a$ and $b$ (see Example 2) and (2) $X \sim U(0, T)$ with finite $T$ (see Example 3).

2.1.1 Model 1: $X \sim U(0, 1)$

Let us assume $X \sim U(0, 1)$ with some joint probability density function of $(X, Y)$ given by $h : [0, 1]^2 \rightarrow \mathbb{R}_+$. Denote the product moments $m_k$ with $k = \{0, 1, 2\}$ as,

$$m_k = m(k, j) = E(Y^k X^j) = \int \int y^k x^j h(y, x) dy dx, \quad (2.1)$$

for all $j = 0, 1, \ldots, \alpha$. The first approximation of the regression function $r(x) = E(Y|X = x)$ is constructed similar to (1.5) with $\phi(x) = x$. That is,

$$r_{\alpha,1}(x) = (\mathcal{B}^{-1}_\alpha m_1)(x), \quad (2.2)$$

where

$$(\mathcal{B}^{-1}_\alpha m_1)(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x] + 1)} \sum_{j=0}^{\alpha-\lfloor \alpha x \rfloor} \frac{(-1)^j m(1, j + \lfloor \alpha x \rfloor)}{j! (\alpha - \lfloor \alpha x \rfloor - j)!}, \quad 0 \leq x \leq 1. \quad (2.3)$$

The next task is to seek the circumstances under which the approximated regression function $r_{\alpha,1}(x)$ is bounded from above and to find the rate of convergence of $r_{\alpha,1}(x)$. With that in mind, let us consider the following statement regarding the upper bound of $r_{\alpha,1}$. 

12
Theorem 1.

(i) If $r$ is continuous on $[0, 1]$, then $r_{\alpha,1} \xrightarrow{u} r$, and

$$
\| r_{\alpha,1} - r \| \leq \Delta(r; \delta) + \frac{2 \| r \|}{\delta^2(\alpha+2)},
$$

where $0 < \delta < 1$, $\Delta(r; \delta) = \sup_{|u-x| \leq \delta} | r(u) - r(x) |$ represents the module of continuity of $r$, and $\| r \|$ is the sup-norm of $r$.

(ii) If $r''$ is continuous on $[0, 1]$, then

$$
r_{\alpha,1}(x) - r(x) = \frac{1}{\alpha+2} \left[ (1-2x)r'(x) + \frac{1}{2}(1-x)r''(x) \right] + o \left( \frac{1}{\alpha} \right), \text{ as } \alpha \to \infty.
$$

Proof. (i) The approximated regression function

$$
r_{\alpha,1}(x) = \frac{\Gamma(\alpha+2)}{\Gamma(\lfloor \alpha x \rfloor + 1)} \frac{\alpha^{-\lfloor \alpha x \rfloor}}{j! (\alpha - [\alpha x] - j)!} \sum_{j=0}^{\infty} (-1)^j E(YX^{j+[\alpha x]})
$$

(2.4)

can be rewritten as follows:

$$
r_{\alpha,1}(x) = \frac{\Gamma(\alpha+2)}{\Gamma(\lfloor \alpha x \rfloor + 1)} \frac{\alpha^{-\lfloor \alpha x \rfloor}}{j! (\alpha - [\alpha x] - j)!} \int_0^1 E(YX^{j+[\alpha x]} | X = u) f(u)du
$$

$$
= \frac{\Gamma(\alpha+2)}{\Gamma(\lfloor \alpha x \rfloor + 1)} \frac{\alpha^{-\lfloor \alpha x \rfloor}}{j! (\alpha - [\alpha x] - j)!} \int_0^1 u^{j+[\alpha x]} E(Y | X = u) f(u)du.
$$

Changing the order of integration and summation and considering $X \sim U(0, 1)$, we arrive at,

$$
r_{\alpha,1}(x) = \int_0^1 E(Y | X = u) \frac{\Gamma(\alpha+2)}{\Gamma(\lfloor \alpha x \rfloor + 1)} \frac{\alpha^{-\lfloor \alpha x \rfloor}}{j! (\alpha - [\alpha x] - j)!} u^{\lfloor \alpha x \rfloor} du
$$

$$
= \int_0^1 \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) r(u) du.
$$

(2.5)
Splitting the range of integration in equation (2.5) into two parts yields,

\[ ||r_{\alpha,1} - r|| = \sup_{0 \leq x \leq 1} \left( \int_{|u-x| \leq \delta} |r(u) - r(x)| \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) du + \int_{|u-x| > \delta} |r(u) - r(x)| \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) du. \]

Denoting \( \Delta(r, \delta) = \sup_{|u-x| \leq \delta} |r(u) - r(x)| \) we obtain,

\[ ||r_{\alpha,1} - r|| \leq \Delta(r, \delta) + R_{\alpha} \text{ with } 0 \leq \delta \leq 1. \] (2.6)

Using Chebychev’s inequality we get the upper bound of \( R_{\alpha} \) as,

\[ R_{\alpha} \leq 2 \||r|| \sup_{0 \leq x \leq 1} \int_{|u-x| > \delta} \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) du \leq \frac{2 \||r||}{\delta^2(\alpha + 2)}. \] (2.7)

Substituting (2.7) in (2.6) completes the proof.

\( (ii) \) From (2.5) we can derive \( r_{\alpha,1}(x) \to r(x) \) uniformly (Feller [13]). We also know that \( \beta(\cdot, [\alpha x] + 1, \alpha - [\alpha x] + 1) \) forms a \( \delta \)-sequence as \( \alpha \to \infty \) with mean,

\[ \theta_{\alpha} = \frac{([\alpha x] + 1)}{(\alpha + 2)} \]

and variance

\[ \sigma_{\alpha}^2 = \frac{([\alpha x] + 1)(\alpha - [\alpha x] + 1)}{(\alpha + 2)^2(\alpha + 1)} \]

such that, \( \theta_{\alpha} - x = \frac{1 - 2x}{(\alpha + 2)^2} + \Delta_{1,\alpha}(x) \) and \( \sigma_{\alpha}^2 - x = \frac{x(1-x)}{(\alpha + 2)^2} + \Delta_{2,\alpha}(x) \), where \( \Delta_{k,\alpha} \leq \frac{2}{(\alpha + 2)^2} \) (Johnson et al. [18], Chen [7]). The Taylor series expansion of \( r \) at \( x \) and the continuity property of \( r'' \) proves the theorem. ■

Theorem 1 provides the upper bound of \( r_{\alpha,1}(x) \) defined in (2.3) and also the order of convergence to the true regression function \( r(x) \). In Corollary 1, we explore some
generalization to Theorem 1 under certain specifications.

**Corollary 1.** If \( \alpha_2 = 2\alpha \), \( \tilde{r}_{\alpha,1}(x) = 2r_{\alpha_2,1}(x) - r_{\alpha,1}(x) \), and \( r'' \) is a continuous function, then

\[
\tilde{r}_{\alpha,1}(x) - r(x) = \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ r'(x)(1 - 2x) + \frac{1}{2} r''(x)x(1 - x) \right] \\
+ (2\Delta_{1,\alpha_2}(x) - \Delta_{1,\alpha}(x)) r'(x) + \frac{1}{2} (2\Delta_{2,\alpha_2}(x) - \Delta_{2,\alpha}(x)) r''(x) + O\left(\frac{1}{\alpha^2}\right).
\]

**Proof.** The difference between \( \tilde{r}_{\alpha,1}(x) \) and \( r(x) \) can be expressed as

\[
\tilde{r}_{\alpha,1}(x) - r(x) = 2r_{\alpha_2,1}(x) - r_{\alpha,1}(x) - r(x) \\
= 2(r_{\alpha_2,1}(x) - r(x)) - (r_{\alpha,1}(x) - r(x)). \tag{2.8}
\]

Therefore (2.8) combined with Theorem 1 (ii) yields

\[
\tilde{r}_{\alpha,1}(x) - r(x) = 2 \left\{ \left[ \frac{1 - 2x}{\alpha_2 + 2} + \Delta_{1,\alpha_2}(x) \right] r'(x) + \frac{1}{2} \left[ \frac{x(1 - x)}{\alpha_2 + 2} \right] r''(x) \right\} \\
- \left\{ \left[ \frac{1 - 2x}{\alpha + 2} + \Delta_{1,\alpha}(x) \right] r'(x) + \frac{1}{2} \left[ \frac{x(1 - x)}{\alpha + 2} \right] r''(x) \right\} \\
+ 2O\left(\frac{1}{(\alpha_2 + 2)^2}\right) - O\left(\frac{1}{(\alpha + 2)^2}\right) \\
= \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ r'(x)(1 - 2x) + \frac{1}{2} r''(x)x(1 - x) \right] \\
+ (2\Delta_{1,\alpha_2}(x) - \Delta_{1,\alpha}(x)) r'(x) + \frac{1}{2} (2\Delta_{2,\alpha_2}(x) - \Delta_{2,\alpha}(x)) r''(x) \\
+ O\left(\frac{1}{\alpha^2}\right).
\]

\[\blacksquare\]

**Remark 1.** Analogous to the conditional expectation, two approximations of the conditional variance \( \sigma^2(x) = E(Y^2|X = x) - (E(Y|X = x))^2 \) can be formulated as
well:

\[ \sigma^2_{\alpha,1}(x) = (\mathcal{B}_\alpha^{-1} m_2)(x) - \left( (\mathcal{B}_\alpha^{-1} m_1)(x) \right)^2 \]

and

\[ \tilde{\sigma}^2_{\alpha,1}(x) = 2\sigma^2_{\alpha_2,1}(x) - \sigma^2_{\alpha,1}(x) \quad \text{with} \quad \alpha_2 = 2\alpha. \]

The proofs of the statements are very similar to the proofs of Theorem 1 and Corollary 1, and are therefore omitted from this dissertation.

So far we have established the approximation of \( r(x) \), some boundary conditions and also the order of convergence of the approximation \( r_{\alpha,1}(x) \). Implementation of all of these is shown in Example 1.

**Example 1.** Let us specify the marginal distributions of \( X \) and \( Y \) as follows:

\[ X \sim U(0, 1) \quad \text{and} \quad Y \sim F(y) = y^{1/2}. \]

Define the joint density of \( X \) and \( Y \) according to the Farlie-Gumbel-Morgenstern copulas with \( \theta = \frac{1}{2} \) [9] as,

\[ h(x,y) = \frac{3}{4} y^{-1/2} - \frac{1}{2} - \frac{1}{2} xy^{-1/2} + x, \quad \text{for} \quad 0 < x, y < 1. \]

The corresponding regression function is \( r(x) = (2x+3)/12 \), for \( x \in [0, 1] \), and the product moments

\[ m(1, j) = (5j+8)/(12(j+1)(j+2)), \quad j = 0, 1, \ldots. \]

The approximated regression functions \( r_{\alpha,1} \) and \( \tilde{r}_{\alpha,1} \) are constructed utilizing (2.2) and Corollary 1, respectively for \( \alpha = \alpha_2 = 30 \). In Figure 2.1, the curves of \( r_{\alpha,1} \) and \( \tilde{r}_{\alpha,1} \) along with the true regression function \( r \) are displayed. It is evident from
Figure 2.1: Approximated regression curves (blue dots) (a): $r_{\alpha,1}$ with $\alpha = 30$; and (b): $\tilde{r}_{\alpha,1}$ with $\alpha = 15$. In both plots the true regression function is $r(x) = (2x + 3)/12$ (orange curve).

Let us apply (1.5) for approximating the conditional variance $\sigma^2(x) = E(Y^2 | X = x) - (E(Y | X = x))^2$. In this example we have,

$$E(Y^2 | X = x) = (28 + 40x)/210$$

and $m(2, j) = \frac{68j + 96}{210(j + 1)(j + 2)}$.

The figure that the approximations $r_{\alpha,1}$ and $\tilde{r}_{\alpha,1}$ closely resembles the true regression function $r(x)$. From the figure it is also evident that $\tilde{r}_{\alpha,1}$ performs better than $r_{\alpha,1}$.

Let us apply (1.5) for approximating the conditional variance $\sigma^2(x) = E(Y^2 | X = x) - (E(Y | X = x))^2$. In this example we have,

$$E(Y^2 | X = x) = (28 + 40x)/210$$

and $m(2, j) = \frac{68j + 96}{210(j + 1)(j + 2)}$.

Figure 2.2 displays the approximated conditional variances for $\alpha = \alpha_2 = 30$, along with the true $\sigma^2(x)$, and from the figures it is clear that $\tilde{\sigma}^2_{\alpha,1}(x)$ converges faster than $\sigma^2_{\alpha,1}(x)$.

Figure 2.2: (a) $\sigma^2_{\alpha,1}(x)$ (blue dots) when $\alpha = 30$ and (b) $\tilde{\sigma}^2_{\alpha,1}(x)$ (blue dots) when $\alpha = 15$ in both plots. The true conditional variance $\sigma^2(x)$ is displayed as well (orange).
Next we study the asymptotic behavior of the errors for approximate $r_{\alpha,1}$ for different $\alpha$ values. Let us denote the distances $d_1(\cdot,\cdot)$ and $d_2(\cdot,\cdot)$ based on $sup$-norm and $L_2$-norm, respectively as,

$$d_1(f,g) = \sup_x |f(x) - g(x)| \quad and \quad d_2(f,g) = \left( \int_0^\infty (f(x) - g(x))^2 \, dx \right)^{1/2}.$$  \hfill (2.9)

In order to study the deviation of approximate $r_{\alpha,1}$ from the exact regression function $r$, we have constructed the approximated regressions $r_{\alpha,1}$ for $\alpha$ ranging from 30 to 230. The $sup$-norm and $L_2$-norm distances between $r_{\alpha,1}$ and $r$ are calculated and displayed in Figure 2.3. Within this range of $\alpha$ the minimum $sup$-norm and $L_2$-norm distances are found to be 0.00128205 and 0.00837484 respectively, at $\alpha = 128$.

![Figure 2.3: Errors of approximation in (a) $sup$-norm and (b) $L_2$-norm for $r_{\alpha,1}$, when $\alpha \in [30,230]$.](image)

**2.1.2 Model 2: $X \sim f$, $f$ is known**

Let us assume that $X$ follows a known continuous pdf $f$, with cdf $F$, and has finite support, $supp\{F\} = [0,1]$. Given the sequence $m_1$, consider the approximation:

$$r_{\alpha,2}(x) = \frac{1}{f(x)} (B_\alpha^{-1}m_1)(x), \quad 0 \leq x \leq 1,$$  \hfill (2.10)
where $B_\alpha^{-1}m_1$ is defined in (2.3). Theorem 2 discusses some properties of $r_{\alpha,2}(x)$.

**Theorem 2.**

(i) If $r$ is continuous, then $r_{\alpha,2} \to r$, as $\alpha \to \infty$.

(ii) If $f$ is bounded from above and $\inf f(x) = \beta > 0$, then

$$
\| r_{\alpha,2} - r \| \leq \frac{\| r \| \Delta(f, \delta) + \| f \| \Delta(r, \delta)}{\beta} + \frac{2 \| r \| \| f \|}{\delta^2 \beta(\alpha+2)}.
$$

**Proof.** (i) It is known that $r(x)$ is continuous on $[0, 1]$ and

$$
r_{\alpha,2}(x) = \frac{\Gamma(\alpha+2)}{\Gamma(\lceil \alpha \rceil + 1)} \int_0^1 u^{\{\alpha\}} \sum_{j=0}^{\alpha-\{\alpha\}} \frac{(-u)^j}{j! (\alpha-\{\alpha\}-j)!} E(Y|X = u) \frac{f(u)}{f(x)} du.
$$

(2.11)

Changing the order of integration and summation of (2.11) we get

$$
r_{\alpha,2}(x) = \int_0^1 \beta(u, \lceil \alpha \rceil + 1, \alpha - \{\alpha\} + 1) r(u) \frac{f(u)}{f(x)} du.
$$

(2.12)

Therefore from (2.12) we get $r_{\alpha,2}(x) \to r(x)$ uniformly (Feller [13]).

(ii) Following the steps of Theorem 1, we split the above integration into two parts

$$
\| r_{\alpha,2} - r \| = \sup_{0 \leq x \leq 1} \frac{1}{f(x)} \int_{|u-x| \leq \delta} | \psi(u) - \psi(x) | r(u, [\alpha] + 1, \alpha - [\alpha] + 1) du \\
+ \sup_{0 \leq x \leq 1} \frac{1}{f(x)} \int_{|u-x| > \delta} | \psi(u) - \psi(x) | \beta(u, [\alpha] + 1, \alpha - [\alpha] + 1) du
$$

(2.13)

where $\psi(u) = r(u) f(u)$. In addition let us denote

$$
\Delta(\psi, \delta) = \sup_{|u-x| \leq \delta} | \psi(u) - \psi(x) |.
$$

(2.14)
Therefore from (2.13) and (2.15) we have,

$$\Delta(\psi, \delta) \leq || r || \Delta(f, \delta) + || f || \Delta(r, \delta)$$  \hspace{1cm} (2.15)

and

$$|| r_{\alpha,2} - r || \leq \frac{\Delta(\psi, \delta)}{\beta} + R_\alpha \text{ where } 0 \leq \delta \leq 1.$$

Using Chebychev’s inequality we finally get,

$$R_\alpha \leq \frac{2 \cdot || \psi ||}{\beta} \sup_{0 \leq x \leq 1, |u-x|>\delta} \beta(u, [\alpha x] + 1, \alpha - [\alpha x] + 1) du \leq \frac{2 \cdot || r || \cdot || f ||}{\delta^2 \beta (\alpha + 2)}$$

which completes the proof.  \hspace{1cm} ■

**Corollary 2.** Suppose \( \psi = r f, \alpha_2 = 2 \alpha, \tilde{r}_{\alpha,2}(x) = 2r_{\alpha_2,2}(x) - r_{\alpha,2}(x), \) and \( \psi'' \) is continuous, then

$$\tilde{r}_{\alpha,2}(x) - r(x) = \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ \frac{1}{f(x)} \psi'(x)(1 - 2x) + \frac{1}{2f(x)} \psi''(x)x(1-x) \right]$$

$$+ \frac{1}{f(x)} (2\Delta_{1,\alpha_2}(x) - \Delta_{1,\alpha}(x)) \psi'(x) + \frac{1}{f(x)} \Delta_{2,\alpha_2}(x) \psi''(x)$$

$$- \frac{1}{2f(x)} \Delta_{2,\alpha}(x) \psi''(x) + O \left( \frac{1}{\alpha^2} \right),$$

as \( \alpha \to \infty. \)

**Proof.** Let us rewrite \( \tilde{r}_{\alpha,2}(x) - r(x) \) as follows:

$$\tilde{r}_{\alpha,2}(x) - r(x) = 2(r_{\alpha_2,2}(x) - r(x)) - (r_{\alpha,2}(x) - r(x)).$$  \hspace{1cm} (2.16)
Similar to Corollary 1, (3.3) yields

\[
\tilde{r}_{\alpha,2}(x) - r(x) = 2 \left\{ \frac{1}{f(x)} \left[ \frac{1 - 2x}{\alpha_2 + 2} + \Delta_1,\alpha_2(x) \right] \psi'(x) + \frac{1}{2f(x)} \psi''(x) \left[ \frac{x(1-x)}{\alpha_2 + 2} \right] \right\} \\
- \left\{ \frac{1}{f(x)} \left[ \frac{1 - 2x}{\alpha + 2} + \Delta_1,\alpha(x) \right] \psi'(x) + \frac{1}{2f(x)} \psi''(x) \left[ \frac{x(1-x)}{\alpha + 2} \right] \right\} \\
+ 2O \left( \frac{1}{(\alpha_2 + 2)^2} \right) - O \left( \frac{1}{(\alpha + 2)^2} \right) \\
= \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ \frac{1}{f(x)} \psi'(x)(1 - 2x) + \frac{1}{2f(x)} \psi''(x)x(1 - x) \right] \\
+ \frac{1}{f(x)} \left( 2\Delta_1,\alpha_2(x) - \Delta_1,\alpha(x) \right) \psi'(x) + \frac{1}{f(x)} \Delta_2,\alpha_2(x) \psi''(x) \\
- \frac{1}{2f(x)} \Delta_2,\alpha(x) \psi''(x) + O \left( \frac{1}{\alpha_2^2} \right).
\]

The above subsection demonstrates the approximation of \( r(x) \) using \( r_{\alpha,2}(x) \), and the upper bound for \( r_{\alpha,2}(x) \). The approximation \( r_{\alpha,2}(x) \) and the distances are explained in Examples 2 and 3.

**Example 2.** Let us consider a Bivariate Dirichlet density,

\[
h(x, y) = \frac{\Gamma(a + b + c)}{\Gamma(a)\Gamma(b)\Gamma(c)} x^{a-1} y^{b-1} (1 - x - y)^{c-1}
\]

with parameters \((a, b, c) > 0, 0 < x, y < 1\) and \( x + y < 1 \).

The true regression function for this example is \( r(t) = \frac{a(1-t)}{a + c} \), for \( t \in [0, 1] \) and the product moment is

\[
m(1, j) = \frac{(a)_j (b)_1}{(a + b + c)_{1+j}},
\]

where \((a)_k = \Gamma(a + k) / \Gamma(a)\).
The construction of the approximations \( r_{\alpha,2} \) and \( \tilde{r}_{\alpha,2} \) of \( r \) follows from (2.4) and Corollary 2 for \( \alpha = 160, \ a = 1/3, \ b = 1 \) and \( c = 1/2 \). The curves \( r_{\alpha,2} \) and \( \tilde{r}_{\alpha,2} \) along with the true regression function \( r \) are displayed in Figure 2.4. From the figure it is clear that \( \tilde{r}_{\alpha,2} \) converges faster than \( r_{\alpha,2} \).

![Figure 2.4: Approximated regression functions (blue dots) (a) \( r_{\alpha,2} \) with \( \alpha = 40 \) and (b) \( \tilde{r}_{\alpha,2} \) with \( \alpha = 20 \). The true regression \( r(x) = 2(1 - x)/5 \) (orange curve).](image)

The approximation of the conditional variance: \( \sigma^2(x) = E(Y^2|X = x) - (E(Y|X = x))^2 \) with \( E(Y^2|X = x) = \frac{b(b+1)(1-x)^2}{(b+c)(b+c+1)} \), is also considered. Define the approximated conditional variance \( \sigma^2_{\alpha,2}(x) \) as

\[
\sigma^2_{\alpha,2}(x) = \frac{1}{f(x)} (\mathcal{B}_\alpha^{-1} m_2)(x) - \left( \frac{1}{f(x)} (\mathcal{B}_\alpha^{-1} m_1)(x) \right)^2
\]

and its modified version as follows:

\[
\tilde{\sigma}^2_{\alpha,2}(x) = 2\sigma^2_{\alpha,2}(x) - \sigma^2_{\alpha,2}(x).
\]

Note that

\[
m(2, j) = \frac{(a)_j(b)_2}{(a+b+c)_{2+j}}.
\]

Figure 2.5 displays the approximated and true conditional variances for \( \alpha = \alpha_2 = 22 \).
30, $a = 1/3$, $b = 1$ and $c = 1/2$. From the plot it is clear that $\tilde{\sigma}^2_{\alpha,2}(x)$ approximate $\sigma^2(x)$ better than $\sigma^2_{\alpha,2}(x)$ when $\alpha = \alpha_2 = 30$.

Figure 2.5: Plots of the approximated conditional variances (blue dots) in (a): $\sigma^2_{\alpha,2}(x)$ and (b): $\tilde{\sigma}^2_{\alpha,2}(x)$ when $\alpha = \alpha_2 = 30$.

Similar to Example 1 we observe the deviation of approximate $r_{\alpha,2}$ from the exact regression function $r$ for different $\alpha$. The minimum sup-norm and $L_2$-norm distance between $r_{\alpha,2}$ and $r$ are calculated for $\alpha$ ranging from 30 to 230. Figure 2.6 displays sup-norm and $L_2$-norm distance for $r_{\alpha,2}$ and $r$. For $\alpha = 228$ and the minimum sup-norm and $L_2$-norm are 0.00138935 and 0.0117026 respectively.

Figure 2.6: The error of approximations in (a) sup-norm and (b) $L_2$-norm for $\alpha$ ranging from 30 to 230.

The sup-norm and $L_2$-norm distances for the approximations of $r_{\alpha,2}, \tilde{r}_{\alpha,2}$, and $r_n$ defined in (1.3) are presented in Table 2.1. Note that the values of the considered
approximations are evaluated at the Chebychev’s nodes as presented in Hagwood [17].

Table 2.1: The error of approximations of \( r_{\alpha,2}, \tilde{r}_{\alpha,2}, \) and \( r_n \) in \( sup \)-norm and \( L_2 \)-norm for \( \alpha = \alpha_2 = n = 30. \)

<table>
<thead>
<tr>
<th>Approximation</th>
<th>( sup )-norm</th>
<th>( L_2 )-norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{\alpha,2} ) with ( \alpha = 30 )</td>
<td>0.0140185</td>
<td>0.0460327</td>
</tr>
<tr>
<td>( \tilde{r}_{\alpha,2} ) with ( \alpha = 15 )</td>
<td>0.00551802</td>
<td>0.017884</td>
</tr>
<tr>
<td>( r_n ) with ( n = 30 )</td>
<td>0.0510259</td>
<td>0.0521238</td>
</tr>
</tbody>
</table>

From Table 2.1 we see that constructions \( r_{\alpha,2} \) and \( \tilde{r}_{\alpha,2} \) if compared with \( r_n \), proposed in [17], show more accurate approximations when \( \alpha = 30. \)

**Example 3.** Suppose \( X \sim U[0,T] \) and \( Y|X = x \sim U[0,x] \) such that the joint distribution of \( X \) and \( Y \) is

\[
h(x,y) = \frac{1}{xT},
\]

where \( 0 < y < x < T. \) The regression function for this given distribution is \( r(x) = \frac{y^2}{2} \) and the product moment is

\[
m(1,j) = \frac{T}{2(j+2)}.
\]

Similar to Example 2, the approximated regression functions \( r_{\alpha,2} \) and \( \tilde{r}_{\alpha,2} \) for \( r \) are constructed using (2.10) and Corollary 2. The approximated curves \( r_{\alpha,2} \) and \( \tilde{r}_{\alpha,2} \) for \( r \) are displayed in Figure 2.7 for \( \alpha = 60 \) and \( T = 3. \) From this figure we can say that \( r_{\alpha,2} \) and \( \tilde{r}_{\alpha,2} \) follow the same general trend as the true regression function \( r(x). \)

To approximate the conditional variance \( \sigma^2(x) \) where \( E(Y^2|X = x) = \frac{T^2}{3} \), we
define the product moment
\[
m(2, j) = \frac{T^2}{3(j + 3)}.
\]

Similar to Example 2 we calculate the approximated conditional variance \(\sigma^2_{\alpha,2}(x)\) and \(\tilde{\sigma}^2_{\alpha,2}(x)\). The graphical representation of \(\sigma^2_{\alpha,2}(x)\) and \(\tilde{\sigma}^2_{\alpha,2}(x)\) is displayed in Figure 2.8 and we can see that the approximate \(\tilde{\sigma}^2_{\alpha,2}(x)\) is much better than \(\sigma^2_{\alpha,2}(x)\).

Figure 2.8: (a) \(\sigma^2_{\alpha,2}(x)\) and (b) \(\tilde{\sigma}^2_{\alpha,2}(x)\) (both blue) represents the approximated conditional variance, when \(\alpha = \alpha_2 = 60\) and \(T = 3\).

To study the deviation of approximate for different \(\alpha\) we calculate the sup-norm and \(L_2\)-norm between \(r_{\alpha,2}\) and \(r\), for \(\alpha\) ranging from 30 to 230. The minimum sup-norm and \(L_2\)-norm are 0.00487769 and 0.0326301, respectively for \(\alpha = 202\). The distances are displayed in Figure 2.9.
2.2 Distribution of $X$ is unknown

In this subsection we consider

Model 3: $X \sim F$ with unknown cdf $F$ and suppose $supp\{F\} = \mathbb{R}_+$. To approximate the conditional mean $r(x)$ and conditional variance $\sigma^2(x)$, let us introduce the following sequences of moments $m_k = \{m(k,j), j = 0, 1, \ldots, \alpha\}$ with

$$m(k,j) = E(Y^k \phi(X)^j),$$

(2.17)

where $\phi(x) = b^{-x}$ for some $b > 1$, and $k = 0, 1, 2$. The case with a finite support $[0, T], \ T < \infty$, can also be handled as well by applying (2.17) and (2.18) below with $\phi(x) = x/T, \ 0 \leq x \leq T$.

In this subsection we will evaluate the values of all proposed approximations at $x_j = \log\left(\frac{\alpha}{\alpha - j + 1}\right)/\log(b)$ for $j = 1, \ldots, \alpha$.

Assuming that the two moment sequences $m_0$ and $m_1$ are given, we suggest the approximation of $r(x)$ to be

$$r_{\alpha,b}(x) = \frac{((\mathcal{B}^{-1}_\alpha m_1) \circ \phi)(x)}{((\mathcal{B}^{-1}_\alpha m_0) \circ \phi)(x)}, \ x \in \mathbb{R}_+.$$ 

(2.18)

Figure 2.9: The errors of approximations in $r_{\alpha,2}$ in (a) $sup$-norm and (b) $L_2$-norm between when $\alpha \in [30, 230]$. 

26
Similarly to Corollary 1 and 2, the modified version of \( r(x) \) is

\[
\tilde{r}_{\alpha,b}(x) = 2r_{\alpha_2,b}(x) - r_{\alpha,b}(x)
\]  

(2.19)

Here, in (2.18) both the numerator and denominator can be written as follows:

\[
\left( (\mathcal{B}_\alpha^{-1} m_k \circ \phi) \right)(x) = C_\alpha(x) \sum_{j=0}^{\alpha-\lfloor \alpha \phi(x) \rfloor} (-1)^j m(k, j + \lfloor \alpha \phi(x) \rfloor) \frac{j!}{(\alpha - \lfloor \alpha \phi(x) \rfloor - j)!}
\]  

(2.20)

where \( C_\alpha(x) = \frac{\lfloor \alpha \phi(x) \rfloor \log(b)}{\alpha} \frac{\Gamma(\alpha + 2)}{\Gamma(\lfloor \alpha \phi(x) \rfloor + 1)} \) and \( k = 0, 1 \). The expression (2.20) with \( k = 2 \) will be used for recovering the conditional variance.

In this section we discuss the approximations \( r_{\alpha,b}(x) \) and \( \tilde{r}_{\alpha,b}(x) \) of regression function \( r(x) \) when the distribution of \( X \) is unknown. We will implement these approximations in Examples 4 and 5, and we will also study the deviation of the approximations from the true one.

**Example 4.** Let us consider that the pdf of \( X \) is a mixture of two gamma densities \( g(\cdot, 1, 1) \) and \( g(\cdot, 3, 1) \) with equal weights, i.e,

\[
f(x) = g(x, 1, 1)/2 + g(x, 3, 1)/2,
\]

while the joint pdf of \( X \) and \( Y \) is

\[
h(x,y) = g(x,1,1)g(y,4,1)/2 + g(x,3,1)g(y,2,1)/2.
\]

For this example the true regression function is \( r(x) = (2x^2 + 8)/(x^2 + 2) \) for \( x > 0 \) and the product moments \( m_0 \) and \( m_1 \) are
\[
m(0, j) = \frac{1}{2(j \ln b + 1)} + \frac{1}{2(j \ln b + 1)^3}, \quad \text{and} \quad m(1, j) = \frac{2}{j \ln b + 1} + \frac{1}{(j \ln b + 1)^3},
\]

Figure 2.10: Approximated (a) \( r_{\alpha, b} \) and (b) \( \tilde{r}_{\alpha, b} \) (blue) of the true regression \( r(x) = (2x^2 + 8)/(x^2 + 2) \) (orange) when \( \alpha = \alpha_2 = 40 \), and \( b = 1.15 \).

The approximated regression curves \( r_{\alpha, b} \) and \( \tilde{r}_{\alpha, b} \) of \( r \) are constructed using (2.18) and (2.19). Figure 2.10 illustrates the behavior of \( r_{\alpha, b} \) and \( \tilde{r}_{\alpha, b} \) for parameters \( \alpha = \alpha_2 = 40 \) and \( b = 1.15 \). From the plot it is clear that curves \( r_{\alpha, b} \) and \( \tilde{r}_{\alpha, b} \) are similar to the true regression function \( r(x) \).

The approximations of the conditional variance \( \sigma^2(x) \) with \( E(Y^2|X = x) = \frac{6x^2 + 40}{x^2 + 2} \) are defined as

\[
\sigma_{\alpha, b}^2(x) = \frac{((\mathcal{B}_\alpha^{-1}m_2) \circ \phi)(x)}{((\mathcal{B}_\alpha^{-1}m_0) \circ \phi)(x)} - \left(\frac{((\mathcal{B}_\alpha^{-1}m_1) \circ \phi)(x)}{((\mathcal{B}_\alpha^{-1}m_0) \circ \phi)(x)}\right)^2
\]

and

\[
\tilde{\sigma}_{\alpha, b}^2(x) = 2\sigma_{\alpha_2, b}^2(x) - \sigma_{\alpha, b}^2(x),
\]

with \( m(2, j) = \frac{10}{1 + j \ln b} + \frac{3}{(1 + j \ln b)^3} \).
Figure 2.11: (a) $\sigma_{\alpha,b}^2(x)$ and (b) $\tilde{\sigma}_{\alpha,b}^2(x)$ represents the approximated conditional variance, when $\alpha = \alpha_2 = 40$, and $b = 1.15$.

Figure 2.11 shows that the approximated conditional variances $\sigma_{\alpha,b}^2$ and $\tilde{\sigma}_{\alpha,b}^2$ almost coincide with the true variance $\sigma^2(x)$ when $\alpha = 40$, $\alpha_2 = 40$, and $b = 1.15$.

Figure 2.12: The (a) sup-norm distance and (b) $L_2$-norm distance for $r_{\alpha,b}$, when $\alpha \in [30, 230]$ and $b \in [1.1, 2.4]$.

Two parameters $\alpha$ and $b$ are optimized in this example using (2.9). The sup-norm and $L_2$-norm distances are calculated between $r_{\alpha,b}$ and $r$ for the pair $(\alpha, b)$ for different choices of parameters $\alpha \in [30, 230]$ and $b \in [1.1, 2.5]$. Figure 2.12 displays the surface of the two distance measures for varying $\alpha$ and $b$. The minimum value of sup-norm is 0.0140471 for $\alpha = 200$ and $b = 1.4$ and the minimum value of $L_2$-norm is 0.0631728 for $\alpha = 230$ and $b = 2.5$. 
Example 5. The marginal distributions of $X$ and $Y$ considered for this example are exponential with

$$G(x) = 1 - e^{-\lambda_1 x} \text{ and } F(y) = 1 - e^{-\lambda_2 y} \text{ for } \lambda_1, \lambda_2 > 0.$$  

The joint cdf of $X$ and $Y$ is defined according to the copula function $C_{\theta}(u, v) = uv + \theta u(1-u)v(1-v)$ with parameter $\theta$ (see [9]). The conditional expectation of $Y$ given $X = x$ has the following form:  

$$r(x) = \frac{1}{\lambda_2} + \frac{\theta \lambda_1}{2\lambda_2} (1 - 2e^{-\lambda_1 x}).$$  

It is easy to see that the sequences $m_0$ and $m_1$ are

$$m(0, j) = \frac{\lambda_1}{\lambda_1 + j \ln(b)}, \text{ and }$$

$$m(1, j) = \frac{\lambda_1}{\lambda_2 (\lambda_1 + j \ln(b))} + \frac{\theta \lambda_1}{2\lambda_2 (\lambda_1 + j \ln(b))} - \frac{\theta \lambda_1}{\lambda_2 (2\lambda_1 + j \ln(b))}.$$

Similar to Example 4 we construct the approximations $r_{\alpha, b}$ and $\tilde{r}_{\alpha_2, b}$ for the regression function $r(x)$. A comparison between approximated regression functions $r_{\alpha, b}$ and $\tilde{r}_{\alpha_2, b}$ with $r$ for $\alpha = \alpha_2 = 30$, $\theta = -0.9$, $\lambda_1 = 0.5$, $\lambda_2 = 0.7$, $b = 1.2$ is displayed in Figure 2.13. It is clear that the approximated regression functions $r_{\alpha, b}$ and $\tilde{r}_{\alpha_2, b}$ estimates $r(x)$ well.
To approximate the conditional variance in Example 5, we have $E(Y^2|X = x) = \frac{2}{\lambda_2^2} + \frac{3\theta}{2\lambda_2^2}(1 - 2e^{-\lambda_1x})$ and

$$m(2, j) = \frac{2\lambda_1}{\lambda_2^2(\lambda_1 + j\ln b)} + \frac{3\theta\lambda_1}{2\lambda_2^2(\lambda_1 + j\ln b)} - \frac{3\theta\lambda_1}{\lambda_2^2(2\lambda_1 + j\ln b)}.$$ 

The approximated conditional variances $\sigma^2_{\alpha,b}(x)$ and $\tilde{\sigma}^2_{\alpha,b}(x)$ closely resemble the true conditional variance $\sigma^2(x)$ as shown in Figure 2.14.

Figure 2.14: (a) $\sigma^2_{\alpha,b}(x)$ and (b) $\tilde{\sigma}^2_{\alpha,b}(x)$ represents the approximated conditional variance, when $\alpha = \alpha_2 = 30$, and $b = 1.2$.

Figure 2.15: The (a) sup-norm distance and (b) $L_2$-norm distance when $\alpha$ ranges from 30 to 230 and $b$ ranges from 1.1 to 2.5

Analogous to Example 4 we calculate the distances between the approximate and
true regression function for different \((\alpha, b)\) pairs. Figure 2.15 displays the surface of the distance functions. The minimum \(sup\)-norm in this case is 0.00796570 for \(\alpha = 230\) and \(b = 2.0\) and minimum \(L_2\)-norm is 0.0575141 for \(\alpha = 230\) and \(b = 2.5\).

We see that the approximates \(\tilde{r}_{\alpha,m}\) and \(\tilde{\sigma}_{\alpha,m}^2\), \(m = 1, 2\), introduced in Corollaries 1 and 2, as well as \(\tilde{r}_{\alpha,b}\) and \(\tilde{\sigma}_{\alpha,b}^2\) perform better than \(r_{\alpha,m}\), \(r_{\alpha,b}\) and \(\sigma_{\alpha,m}^2\), \(\sigma_{\alpha,b}^2\). Also from each of the above examples, we see that the error rate decreases with the increase in values of \(\alpha\).
Chapter 3

Estimation of Regression Functions

Chapter 1 reviews the applications of the classical moment problem. One such application of approximating the regression functions is presented in Chapter 2. This chapter will examine another aspect of estimating the regression functions when the empirical moments are given. Here, the sequence of actual moments \( m_k = \{m(k,j); k = \{0, 1\}, j = 0, 1, \ldots, \alpha\} \) are estimated using the empirical moments:

\[
\hat{m}(k,j) = \frac{1}{n} \sum_{i=1}^{n} (Y_i^k X_i^j), \tag{3.1}
\]

for \( k = 0, 1 \) and \( j = 0, 1, \ldots, \alpha \). In order to establish the entire process of regression function estimation, first we need to find \( \hat{m}(k,j) \). Given some distribution function, random numbers \( (Y_l, X_l) \) for \( l \in 1, \ldots, n \) are generated. The empirical moments \( \hat{m}(k,j) \) are constructed using the random numbers which are utilized ultimately for the estimation process of the regression function. The examples considered here to demonstrate the estimation process are Examples 1, 2, and 4 from Chapter 1 and the example from [23] where, \( X \sim LN(0, 1) \), \( \epsilon \sim N(0, 0.5) \), and \( X \) and \( \epsilon \) are inde-
dependent random variables with $Y = r(X) + \varepsilon$. The results from the last example are compared with the estimate provided in [23].

### 3.1 Estimation of $r(x), x \in [0, 1]$: Distribution of $X$ known

This section will demonstrate the estimation of the regression function $r(x)$ for a known distribution of the variable $X$. The estimated regression function is proposed analogously to the approximated regression function described in (2.10), with the replacement of actual product moments with the empirical moments. Therefore the estimator $\hat{r}_{\alpha,k}(x)$ of $r(x)$ for $k = 1, 2$ is defined as

$$
\hat{r}_{\alpha,k}(x) = \frac{1}{f(x)} \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha] + 1)} \sum_{j=0}^{\alpha-[\alpha]} (-1)^j \hat{M}(1, j + [\alpha]) \frac{\alpha - [\alpha]}{j!(\alpha - [\alpha] - j)!}.
$$

(3.2)

A comparative study of the estimates $\hat{r}_{\alpha,k}(x)$ with the true regression function $r(x)$ is presented in this section. The average $L_2$-error used for this study is defined as

$$
d_{\alpha,N} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{\alpha} \sum_{j=1}^{\alpha} \left( \hat{r}_{\alpha,k} \left( \frac{j}{\alpha} \right) - r \left( \frac{j}{\alpha} \right) \right)^2 \right)^{1/2},
$$

where $k$ takes values 1, 2 and $N = 50$ is the number of replications of the experiment.

**Example 1 (continued)** Considering the same setup of Example 1, we have generated $n = 5000$ independent random samples of $(X,Y)$ from the joint distribution $h(x,y)$ to estimate the actual regression function $r(x) = \frac{(2x+3)}{12}$. Using $\hat{M}_1$ instead of the product moments $m(1, j)$ in (2.1) provides the estimate $\hat{r}_{\alpha,1}$. Figure 3.1 shows the estimated regression function for $\alpha = 30$ and the $L_2$-distances computed for different
choices of $\alpha$ and sample sizes $n$ are shown in Table 3.1.

![Figure 3.1: $\hat{r}_{\alpha, 1}$ (blue) and $r(x) = \frac{(2x+3)}{12}$ (orange) when $\alpha = 30.$](image)

Table 3.1: Error in approximation for $\hat{r}_{\alpha, 1}$, when $X \sim U(0, 1)$ and $Y \sim F(y) = y^{1/2}$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n = 500$</th>
<th>$n = 1000$</th>
<th>$n = 2000$</th>
<th>$n = 5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.031258</td>
<td>0.025103</td>
<td>0.014944</td>
<td>0.011085</td>
</tr>
<tr>
<td>25</td>
<td>0.036825</td>
<td>0.025222</td>
<td>0.018560</td>
<td>0.011736</td>
</tr>
<tr>
<td>30</td>
<td>0.031768</td>
<td>0.025759</td>
<td>0.021272</td>
<td>0.010548</td>
</tr>
</tbody>
</table>

It is easily noticeable that the $L_2$-error is minimal at the bigger sample size, and is minimum for $\alpha = 30$ and $n = 5000$. Therefore with the increase in sample size, the estimated regression model performs better.

**Example 2 (continued)** Consider $h(x, y) = \frac{\Gamma(a + b + c)}{\Gamma(a)\Gamma(b)\Gamma(c)}x^{a-1}y^{b-1}(1 - x - y)^{c-1}$ where $0 < x, y < 1$, $x + y < 1$ and $a = 1$, $b = 1/3$ and $c = 0.5$, as in Example 2. We have simulated $n = 5000$ joint random samples $(X, Y)$ from $h(x, y)$ to calculate $\hat{m}_1$. Substituting the value of $\hat{m}_1$ in (3.2) for $\alpha = 30$ yields $\hat{r}_{\alpha, 2}$. The estimated regression line is demonstrated in Figure 3.2 along with the actual regression line. The $L_2$-error in Table 3.2 shows that the performance of our model gets better with increase in sample size and the error is minimum when $\alpha$ is 30.
Figure 3.2: $\hat{r}_{\alpha,2}$ (blue) and $r(x) = \frac{b(1-x)}{b+c}$ (orange) when $\alpha = 30$.

It is also evident from both Tables 3.1 and 3.2 that the estimated regression lines $\hat{r}_{\alpha,1}$ and $\hat{r}_{\alpha,2}$ behave better with a larger random sample ($n = 5000$) and $\alpha = 30$ for both examples.

### Table 3.2: Error in approximation for $\hat{r}_{\alpha,2}$ when $f(x, y) = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}x^{a-1}y^{b-1}(1-x-y)^{c-1}$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>n = 500</th>
<th>n = 1000</th>
<th>n = 2000</th>
<th>n = 5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.028496</td>
<td>0.020238</td>
<td>0.015335</td>
<td>0.010049</td>
</tr>
<tr>
<td>60</td>
<td>0.038779</td>
<td>0.027045</td>
<td>0.018739</td>
<td>0.011445</td>
</tr>
<tr>
<td>100</td>
<td>0.043254</td>
<td>0.032562</td>
<td>0.022266</td>
<td>0.011401</td>
</tr>
</tbody>
</table>

### 3.2 Estimation of $r(x), x \in \mathbb{R}_+$: Distribution of $X$ unknown

In this section we discuss the estimation of the regression function $r(x)$ when the pdf of $X$ is unknown. Unlike the previous section where simple substitution provided the estimated regression function, here we estimate the regression function using the
following expression
\[ \hat{r}_{\alpha,b}(x) = \frac{\left( (\mathcal{B}_{\alpha}^{-1}\hat{m}_1) \circ \phi \right)(x)}{\left( (\mathcal{B}_{\alpha}^{-1}\hat{m}_0) \circ \phi \right)(x)} \] (3.3)

where,
\[
\left( (\mathcal{B}_{\alpha}^{-1}\hat{m}_k) \circ \phi \right)(x) = C_{\alpha}(x) \sum_{j=0}^{\alpha-\lceil \alpha\phi(x) \rceil} \frac{(-1)^j \hat{m}(k,j+\lceil \alpha\phi(x) \rceil)}{j! (\alpha-\lceil \alpha\phi(x) \rceil-j)!},
\]
\[
C_{\alpha}(x) = \frac{[\alpha\phi(x)] \log(b)}{\alpha} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha\phi(x)+1)},
\]
\[
\hat{m}(k,j) = \frac{1}{n} \sum_{i=1}^{n} (Y_i^k [\phi(X_i)]^j), \quad \text{and}
\]
\[
\phi(x) = b^{-x} \quad \text{for some } b > 1, \ k = 0, 1 \text{ and } x > 0.
\]

The performance of the aforementioned estimate in (3.3) is studied using the following examples.

**Example 4 (continued)** In order to make the comparison between the estimated and true regression function, we generate 500 random samples \((X,Y)\) from \(h(x,y) = \left( g(x,1,1)g(y,4,1) + g(x,3,1)g(y,2,1) \right)/2\). These random samples are utilized to calculate the empirical moments \(\hat{m}_0\) and \(\hat{m}_1\). Next we evaluate the true regression curve \(r(x) = \frac{1}{\lambda y} + \frac{\theta}{2 \alpha y} (1-2e^{-\lambda x})\) and also the estimated \(\hat{r}_{\alpha,b}(x)\) curve according to (3.3) at \(x = \frac{\log\left( \frac{\alpha-j+1}{\alpha-j} \right)}{\log(b)}\), where \(j = 1, 2, \ldots, \alpha\). The plot for \(\hat{r}_{\alpha,b}(x)\) is presented in the Figure 3.3 for \(\alpha = 30\) and \(b = 1.21\). A comparative result is presented in Table 3.3 between the regression curves using the average \(L_2\)-error \(d_{\alpha,b,n}\), which is defined as:

\[
d_{\alpha,b,n} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{\alpha} \sum_{j=1}^{\alpha} \left( \hat{r}_{\alpha,b}(j) - r\left(\frac{j}{\alpha}\right) \right)^2 \right)^{1/2}, \quad (3.4)
\]

where \(N = 50\) is the number of replications. From Table 3.3 we can conclude that
generally, when the sample size increases, the $L_2$-distance is smaller and our model performs better. It is also clear that for this example the distance is minimum when $\alpha = 25$ and $b = 1.65$ respectively.

Figure 3.3: Estimated $\hat{r}_{\alpha,b}$ (blue) of the true regression $r(x) = \frac{1}{\lambda x} + \frac{\theta}{2\lambda_x} (1 - 2e^{-\lambda x})$ (orange) when $\alpha = 30$ and $b = 1.21$.

Table 3.3: The errors of approximations for $\hat{r}_{\alpha,b}$, when $h(x,y) = g(x,1,1)g(y,4,1)/2 + g(x,3,1)g(y,2,1)/2$.

<table>
<thead>
<tr>
<th>$d_{\alpha,n}$</th>
<th>n = 500</th>
<th>n = 1000</th>
<th>n = 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>b=1.21</td>
<td>b=1.40</td>
<td>b=1.65</td>
</tr>
<tr>
<td>20</td>
<td>0.1805</td>
<td>0.1502</td>
<td>0.1696</td>
</tr>
<tr>
<td>25</td>
<td>0.1963</td>
<td>0.1554</td>
<td>0.1746</td>
</tr>
<tr>
<td>30</td>
<td>0.2215</td>
<td>0.1624</td>
<td>0.1799</td>
</tr>
<tr>
<td>$\frac{\hat{r}_\alpha(x)}{x^2}$</td>
<td>$\frac{\sum_{i=1}^{n} K_{\alpha}(x,X_i)Y_i}{\sum_{i=1}^{n} K_{\alpha}(x,X_i)}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example 6. Now we compare our proposed construction with Koul and Song [23]. Let us consider $X \sim Log-Normal(0,1)$, $\varepsilon \sim Normal(0,0.5)$, and $X$ and $\varepsilon$ are independent random variables with $Y = r(X) + \varepsilon$. Two cases are considered in [23] (a) $r(x) = (x - 1.5)^2$ and (b) $r(x) = \frac{1}{x^2}$. The estimated regression proposed in Koul and Song is

\[ \hat{r}_\alpha(x) = \frac{\sum_{i=1}^{n} K_{\alpha}(x,X_i)Y_i}{\sum_{i=1}^{n} K_{\alpha}(x,X_i)} \]
where

\[ K_\alpha(x,t) = \frac{1}{t^{\Gamma(\alpha)/\alpha}} \left( \frac{\alpha x}{t} \right)^\alpha e^{-\frac{\alpha x}{t}}. \]

and \( \alpha = n^{2/5}. \)

Figure 3.4: Estimated \( \hat{r}_\alpha \) (red), \( \hat{r}_h \) (blue) and \( \hat{r}_{\alpha,b} \) (green) of the true regression \( r(x) = (x - 1.5)^2 \) (black) when \( \alpha = 100 \) and \( b = 1.52. \)

The construction of the estimated regression functions are same as [23]. For case (a) we took 1024 equally spaced values from (0.1, 1) and sample size \( n = 100 \), and for case (b) we took 1024 equally spaced values from (0, 3) and \( n = 200 \). We also extend our comparison with the N–W kernel as in (1.1), where they used the standard normal kernel for their construction. Figure 3.4 represents case (a) for \( \alpha = 100 \), \( b = 1.52 \), \( n = 200 \) and \( h = n^{-1/5} \) and Figure 3.5 represents case (b) for \( \alpha = 100 \), \( b = 3.515 \), \( n = 100 \) and \( h = 0.1 \ast n^{-1/5} \). From both figures we can say that our estimate is competitive with respect to the other two.
Figure 3.5: Estimated $\hat{r}_\alpha$ (red), $\hat{r}_h$ (blue) and $\hat{r}_{\alpha,b}$ (green) of the true regression $r(x) = \frac{1}{x^2}$ (black) when $\alpha = 100$ and $b = 3.515$.

Table 3.4: The errors of approximations for NW estimate, Koul and Song estimate and $\hat{r}_{\alpha,b}$ in $L_2$-error, when $X \sim \text{Log-Normal}(0, 1)$ and $r(x) = (x - 1.5)^2$.

<table>
<thead>
<tr>
<th>Methods</th>
<th>$d_{\alpha,n}$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
<th>$n = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{r}_h$, $h = n^{-1/5}$</td>
<td>0.2265</td>
<td>0.1922</td>
<td>0.1096</td>
<td></td>
</tr>
<tr>
<td>$\hat{r}_\alpha$, $\alpha = n^{2/5}$</td>
<td>0.2668</td>
<td>0.1783</td>
<td>0.1324</td>
<td></td>
</tr>
<tr>
<td>Methods $\hat{r}_{\alpha,b}$, $\alpha = \alpha$, $b = \beta$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$n = 100$</td>
<td>$n = 200$</td>
<td>$n = 500$</td>
<td></td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}$, $\alpha = 80$, $b = 1.35$</td>
<td>0.1078</td>
<td>0.1700</td>
<td>0.1742</td>
<td>0.1140</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}$, $\alpha = 100$, $b = 1.35$</td>
<td>0.1610</td>
<td>0.1608</td>
<td>0.1674</td>
<td>0.1021</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}$, $\alpha = 120$, $b = 1.35$</td>
<td>0.1561</td>
<td>0.1579</td>
<td>0.1655</td>
<td>0.0959</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}$, $\alpha = 200$, $b = 1.35$</td>
<td>0.1509</td>
<td>0.1266</td>
<td>0.0771</td>
<td>0.0938</td>
</tr>
</tbody>
</table>

Similarly to the previous example, we calculate the $L_2$-error for the estimated regression $\hat{r}_{\alpha,b}(x)$, $\hat{r}_h(x)$ and $\hat{r}_\alpha(x)$. From Tables 3.4 and 3.5 we can see that our approximation $\hat{r}_{\alpha,b}(x)$ is more accurate than $\hat{r}_h(x)$ and $\hat{r}_\alpha(x)$ for nearly all $\alpha$ and $n$. From both tables it is also clear that the performance of $\hat{r}_{\alpha,b}(x)$ generally gets better with an increase of $\alpha$ and $n$. 

40
Table 3.5: The errors of approximations for NW estimate, Koul and Song estimate and $\hat{r}_{\alpha,b}$ in $L_2$-error, when $X \sim Log-Normal(0,1)$ and $r(x) = \frac{1}{x^2}$.

<table>
<thead>
<tr>
<th>Methods</th>
<th>$d_{\alpha,n}$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
<th>$n = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{r}_{h}, h = 0.1n^{-1/5}$</td>
<td>16.924</td>
<td>14.541</td>
<td>14.157</td>
<td></td>
</tr>
<tr>
<td>$\hat{r}_{\alpha}, \alpha = n^{2/5}$</td>
<td>5.6743</td>
<td>4.6548</td>
<td>2.9380</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>$n = 100$</td>
<td>$n = 200$</td>
<td>$n = 500$</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>$b=2.318$</td>
<td>$b=3.515$</td>
<td>$b=3.964$</td>
<td>$b=2.318$</td>
<td>$b=3.515$</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}, \alpha = 80$</td>
<td>5.9009</td>
<td>4.9314</td>
<td>4.8393</td>
<td>5.8166</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}, \alpha = 100$</td>
<td>5.6478</td>
<td>4.1789</td>
<td>3.929</td>
<td>5.6737</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}, \alpha = 120$</td>
<td>4.5386</td>
<td>4.1047</td>
<td>3.8650</td>
<td>4.6650</td>
</tr>
<tr>
<td>$\hat{r}_{\alpha,b}, \alpha = 200$</td>
<td>3.4042</td>
<td>3.2329</td>
<td>3.3124</td>
<td>3.6431</td>
</tr>
</tbody>
</table>
Chapter 4

Recovery of Derivative Function with Application

In this chapter our aim is to recover the derivative function when the moment sequence $E[Y^j], j \in \mathbb{N}_+$ is given. Here we assume that $Y \sim g$ and we want to recover the derivative function $g'$, where $G(y) = P(Y \leq y)$. We also consider the problem of recovering the distribution function, the density, and the derivative function in the framework of the multiplicative-censoring model [27].

4.1 Approximating the derivative function

Consider a random variable $Y$ that has the absolute continuous density $g$. For simplicity assume $\text{supp} \{g\} = [0, 1]$. In this section we recover $g'$ given the moments sequence of $g$:

$$
\mu_G(j) = E[Y^j], \quad j = 1, 2, \ldots.
$$
Let us denote

\[ \mu_-(j) = (-j) \mu_G(j-1), \quad j = 1, 2, \ldots. \]

To recover \( g' \) consider \( \mathcal{B}_\alpha^{-1} \mu_- \) where \( \mathcal{B}_\alpha^{-1} \) is defined in (1.5) and \( m_\phi \) is replaced by \( \mu_- \) and \( \phi(x) = x \). Consider

\[ g'_\alpha(y) := (\mathcal{B}_\alpha^{-1} \mu_-)(y), \quad 0 \leq y \leq 1. \tag{4.1} \]

**Theorem 3.**

(i) If \( g(1) = g'(1) = 0 \) and \( g'' \) is bounded and continuous on \([0, 1]\), then \( g^{(k)}_\alpha \xrightarrow{\alpha \to \infty} g^{(k)} \) as \( \alpha \to \infty \), \( k = 1, 2 \) and

\[ || g^{(k)}_\alpha - g^{(k)} || \leq \Delta(g^{(k)}, \delta) + \frac{2 || g^{(k)} ||}{\delta^2(\alpha + 2)}. \]

(ii) If \( g^{(3)} \) is continuous on \([0, 1]\), then

\[ g'_\alpha(y) - g'(y) = \frac{1}{\alpha + 2} \left( g^{(2)}(y)(1 - 2y) + \frac{1}{2} g^{(3)}(y)y(1 - y) \right) + O \left( \frac{1}{\alpha} \right), \quad \text{as} \quad \alpha \to \infty. \]

**Proof.** Considering only the case when \( k = 1 \), using integration by parts and the condition \( g(1) = 0 \), we can easily obtain

\[ \mu_-(j) = -j \int_0^1 t^{j-1} g(t) \, dt = \int_0^1 t^j g'(t) \, dt. \]
Now, substituting $\mu_{-}(j)$ in (4.1), we get

$$g'_{\alpha}(y) = \int_{0}^{1} \beta(u, [\alpha y] + 1, \alpha - [\alpha y] + 1)g'(u)du.$$

The rest of the proof is based on the properties of functions $\beta(\cdot, [\alpha y] + 1, \alpha - [\alpha y] + 1)$ and the continuity of $g'$, which repeats the steps from Theorem 1 in Chapter 2.

\[\Box\]

**Corollary 3.** If $\alpha_2 = 2\alpha$, $\tilde{g}'_{\alpha}(x) = 2g'_{\alpha_2}(x) - g'_{\alpha}(x)$ and $g^{(3)}$ is a continuous function, then

$$
\tilde{g}'_{\alpha}(x) = g'(x) + \frac{1}{(\alpha + 1)(\alpha + 2)} \left[ g^{(2)}(x)(1 - 2x) + \frac{1}{2} g^{(3)}(x) x(1 - x) \right] \\
+ (2\Delta_{1,\alpha_2}(x) - \Delta_{1,\alpha}(x)) g^{(2)}(x) + \frac{1}{2} (2\Delta_{2,\alpha_2}(x) - \Delta_{2,\alpha}(x)) g^{(3)}(x) \\
+ O\left(\frac{1}{\alpha^2}\right).
$$

**Proof.** The proof mimics the steps of Corollary 1.

\[\Box\]

**Example 7.** Assume $g$ represents the pdf of a Beta $(a, b)$ distribution with $a = b = 3$. For the given distribution we want to recover $g'_{\alpha}(y) = 60y(1 - y)(1 - 2y)$, where $\mu_{-}(j) = (-j)\text{Beta}(a + j - 1, b)/\text{Beta}(a, b)$. Figure 4.1 shows the approximated curves $g'_{\alpha}$ and $\tilde{g}'_{\alpha}$ and true $g'$ when both $\alpha = 50$ and $\alpha_2 = 50$. 

Figure 4.1: Approximations (the dotted curves) (a): $g'_{\alpha}$ when $\alpha = 50$; and (b): $\tilde{g}'_{\alpha}$ when $\alpha_2 = 50$. In both plots the true $g'(y) = 60y(1-y)(1-2y)$.

### 4.2 Multiplicative-censoring model

In this subsection we estimate the underlying density function $f$ in a model when we observe $Y \sim g(y)$ where

$$g(y) = \int_{y}^{1} \frac{f(x)}{x} \, dx, \quad 0 < y \leq 1.$$  

Let $X$ be an unobserved random variable from the distribution $f$. We are given $n$ independent copies of $Y = X U$, where the random variable $U$ is another unobserved random variable uniformly distributed on $(0, 1)$ that is independent from $X$. The aim is to estimate the pdf $f$ given the sample $Y_1, \ldots, Y_n$ from $g$. Note that in this model $f(x) = -xg'(x)$. Using this relationship one can recover $f$ via

$$f_{\alpha}(x) = -xg'_{\alpha}(x).$$

Hence, by denoting $\tilde{g}'_{\alpha,n} = \mathcal{R}^{-1}_{\alpha} \hat{\mu}_-$, where the empirical counterpart $\hat{\mu}_-(j) = (-j)\hat{\mu}_G(j-1)$ is used instead of $\mu_-$, we obtain the estimate of $f$:

$$\hat{f}_{\alpha,n}(x) = -x\tilde{g}'_{\alpha,n}(x).$$
Here

$$\hat{\mu}_G(j) = \frac{1}{n} \sum_{i=1}^{n} Y_i^j$$

In addition, to estimate $f'(x) = -g'(x) - xg''(x)$, consider

$$\hat{f}'_\alpha(x) = -\hat{g}'_\alpha(x) - x\hat{g}''_\alpha(x).$$

where $\hat{g}''_{\alpha,n} = \mathcal{B}_\alpha^{-1} \hat{\mu}_{2-}$ and the empirical counterpart $\hat{\mu}_{2-}(j) = j(j-1)\hat{\mu}_G(j-2)$.

**Remark 1.** The moments of $f$ and $g$ are related by

$$E[X^j] = E[Y^j] / E[U^j]$$

where $E[U^j] = 1/(j+1)$. Hence, to estimate $F$ and $f$ one can apply (1.4) and (1.5), where $φ(x) = x$ and $mφ$ is replaced by

$$\hat{\mu}_n(j) = (j + 1)\hat{\mu}_G(j).$$

The estimates of $F$ and $f$ can be written as follows:

$$\hat{F}_{\alpha,n}(x) := (\mathcal{K}_\alpha^{-1} \hat{\mu}_n)(x) \quad \text{and} \quad \hat{f}_{\alpha,n}(x) := (\mathcal{B}_\alpha^{-1} \hat{\mu}_n)(x).$$

To estimate $f'$ we suggest to use

$$\hat{f}'_{\alpha,n}(x) := (\mathcal{B}_\alpha^{-1} \hat{\mu}_n^*)(x),$$

where $\hat{\mu}_n^*(j) = -j^2\hat{\mu}_G(j-1)$.  

46
From a simulation study we found that the asymptotic behavior of $\hat{f}_{\alpha,n}$ and $\hat{f}_{\alpha,n}$ are very similar to each other. In the following statement the weak consistency of $\hat{F}_{\alpha,n}$ and $\hat{f}_{\alpha,n}^*$ are also provided.

**Theorem 4.** Assume $g(0)$ is finite. We have

(i) $\hat{F}_{\alpha,n}(x) \rightarrow_p F(x)$ as $\sqrt{\alpha}/n \rightarrow 0, \alpha, n \rightarrow \infty,$

and if $f$ is a continuous function, then

(ii) $\hat{f}_{\alpha,n}^*(x) \rightarrow_p f(x)$ as $\alpha^{3/2}/n \rightarrow 0, \alpha, n \rightarrow \infty$

uniformly on any interval $[\delta, \gamma]$ with $0 < \delta, \gamma < 1.$

**Proof.** (i). Since the operator $\mathcal{H}_\alpha^{-1}$ in (1.4) is linear, we can rewrite $\hat{F}_{\alpha,n}$ as the sum

$$\hat{F}_{\alpha,n}(x) = (\mathcal{H}_\alpha^{-1}\hat{\mu}_G)(x) + (\mathcal{H}_\alpha^{-1}\hat{\mu}_+)(x), \tag{4.2}$$

where $\hat{\mu}_+(j) = j\hat{\mu}_G(j).$ Now, taking into account that

$$\sum_{k=0}^{[\alpha x]} \binom{j}{k} (-1)^k = (-1)^{[\alpha x]} \binom{j-1}{[\alpha x]}$$

(see [22]) and changing the summation in the both terms of (4.2), after some simple algebra, we obtain the following representation

$$\hat{F}_{\alpha,n}(x) = \hat{G}_{\alpha,n}(x) - \hat{\eta}_{\alpha,n}(x).$$

Here,

$$\hat{G}_{\alpha,n}(x) = \frac{1}{n} \sum_{i=1}^{n} B_\alpha(Y_i, x) \quad \text{with} \quad B_\alpha(t,x) = \sum_{k=0}^{[\alpha x]} \binom{\alpha}{k} t^k (1-t)^{\alpha-k}$$

and

$$\hat{\eta}_{\alpha,n}(x) = \frac{1}{n} \sum_{i=1}^{n} Y_i \beta(Y_i, [\alpha x] + 1, \alpha - [\alpha x]) := \frac{1}{n} \sum_{i=1}^{n} \xi_i.$$
where $\beta(\cdot, c, d)$ is the density function of a Beta$(c, d)$ distribution.

In the multiplicative-censoring model $F(x) = G(x) - xg(x)$, it is sufficient to show that $\hat{G}_{\alpha,n}(x) \to G(x)$, and $\hat{\eta}_{\alpha,n}(x) \to -xg(x)$ in probability. To complete the proof, let us mention that the means of $\hat{G}_{\alpha,n}(x)$ and $\hat{\eta}_{\alpha,n}(x)$ converge uniformly to $G(x)$ and $xg(x)$, respectively. The later follows from the statements from ([33], Theorem 3.1) and (Feller [13], v. II, Ch. VII), respectively. For the variances we have

\[
\text{Var}[\hat{G}_{\alpha,n}(x)] = \frac{1}{n} \text{Var}[B_{\alpha}(Y, x)] \leq \frac{1}{n} E[B_{\alpha}^2(Y, x)] \leq \frac{1}{n}.
\]

and

\[
\text{Var}[\hat{\eta}_{\alpha,n}(x)] \leq \frac{1}{n} E[\xi_1^2] = \frac{1}{n} \int_0^1 y^2 \beta(y, [\alpha x] + 1, \alpha - [\alpha x]) g(y) dy \leq \frac{C_1 g(0) \sqrt{\alpha}}{n \sqrt{x(1-x)}}.
\]

In the last inequality of (4.3) we apply the following property of $\beta(\cdot, c, d)$:

\[
\beta(y, [\alpha x] + 1, \alpha - [\alpha x]) \leq \frac{C_1 \sqrt{\alpha}}{\sqrt{x(1-x)}}
\]

for some positive constant $C_1$ (see, Chen [8]).

(ii) Note that $g'$ is continuous and $E[\xi_{\alpha,n}'] = g'_{\alpha}$. According to Theorem 3 (i) we have $E[\xi_{\alpha,n}'] \to u g'$ as $\alpha, n \to \infty$. Now, let us prove that

\[
\text{Var}[\hat{\xi}_{\alpha,n}(x)] \to 0 \text{ as } \frac{\alpha^{3/2}}{n} \to 0.
\]

After substituting $\hat{\mu}_-$ into $R_{\alpha}^{-1} \hat{\mu}_-$, simple algebraic and mathematical manipulations yields:
\[
\hat{g}^{'}_{\alpha, n}(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha] + 1)} \cdot \sum_{j=0}^{\infty} \frac{(-1)^j (j + [\alpha]) \frac{1}{2} \sum_{i=1}^{n} y_i^{j+[\alpha]-1}}{j! (\alpha - [\alpha] - j)!}
\]

\[
= \frac{\alpha + 1}{n} \sum_{i=1}^{n} \left( \frac{Y_i}{[\alpha]} - \frac{1 - Y_i}{[\alpha]} \right) \beta(Y_i, [\alpha], \alpha - [\alpha])
\]

\[
= \frac{\alpha + 1}{\alpha, \alpha - [\alpha]} \cdot \frac{1}{n} \sum_{i=1}^{n} \left( Y_i - \frac{[\alpha]}{\alpha} \right) \beta(Y_i, [\alpha], \alpha - [\alpha]).
\]

To prove (4.5), note that according to (4.4) the second moment of the summands in the last equation can be bounded above as follows:

\[
\int \left( y - \frac{[\alpha]}{\alpha} \right)^2 \beta^2(y, [\alpha], \alpha - [\alpha]) g(y) \, dy \leq \frac{C_1 g(0) \sqrt{\alpha}}{\sqrt{x(1-x)}} \int \left( y - \frac{[\alpha]}{\alpha} \right)^2 \beta(y, [\alpha], \alpha - [\alpha]) \, dy
\]

\[
= \frac{C_1 g(0) \sqrt{\alpha}}{\sqrt{x(1-x)}} \cdot \frac{[\alpha]}{\alpha} \frac{\beta(Y_i, [\alpha], \alpha - [\alpha])}{\alpha^2(\alpha + 1)}.
\]

In this section we constructed the estimates for \( F, f \) and \( f' \) in the framework of the multiplicative-censoring model. We also established some boundary conditions for \( \hat{F}_{\alpha, n} \) and \( \hat{f}_{\alpha, n}(x) \). Implementation of these results is shown in the next few examples.

**Example 7. (continued)** Assume \( X \sim Beta(3, 3) \) and \( U \sim U(0, 1) \) where \( X \) and \( U \) are independent. We generated \( n = 1000 \) random samples of \( X \) and \( U \). By taking \( Y_i = X_i U_i \), we estimated \( F, f \), and \( f' \) in the framework of the multiplicative-censoring model. For this example, \( f'(x) = 60x(1 - x)(1 - 2x) \).

In Figure 4.2 we plotted the estimated cdf of \( X \) using \( \hat{F}_{\alpha, n} \) when \( n = 500, \alpha = 35 \) and \( n = 1000, \alpha = 30 \). Figure 4.3 provides the estimated curves of \( \hat{f}_{\alpha, n}(x) \) when
\( n = 1000, \alpha = 25, \) and \( n = 2000, \alpha = 30, \) and Figure (4.4) shows the plot for \( \hat{f}_{\alpha,n}(x) \) when \( n = 3000, \alpha = 25, \) and \( n = 5000, \alpha = 30. \) Our simulation study shows that proposed nonparametric estimates of \( F, f, \) and \( f' \) are stable.

Figure 4.2: Estimated \( \hat{F}_{\alpha,n} \) (dotted curve) (a): when \( n = 500, \alpha = 35; \) and (b): when \( n = 1000, \alpha = 30. \) In both plots the true \( F \) (continuous curve) represents the cdf of a \( Beta(3,3) \) distribution.

Figure 4.3: Estimated \( \hat{f}_{\alpha,n} \) (dotted curve) (a): when \( n = 1000, \alpha = 25; \) and (b): when \( n = 1500, \alpha = 30. \) Here \( X \sim Beta(3,3). \)

In order to study the distance between the estimated functions and the true one, let us introduce the average \( L_2 \)-error for an arbitrary function \( g \) and its estimate \( \hat{g} \) as:

\[
d_{\alpha,N} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{\alpha} \sum_{j=1}^{\alpha} \left( \hat{g}_{\alpha,n}\left( \frac{i}{\alpha} \right) - g\left( \frac{i}{\alpha} \right) \right)^2 \right)^{1/2},
\]

(4.6)

where \( N \) is the number of replications.
The $L_2$-error and the standard deviation for the estimated $\hat{F}_{\alpha,n}$ of $F$ are calculated and displayed in Table 4.1. To calculate the distance we considered $N = 100$ replications. We can say that the distance between $\hat{F}_{\alpha,n}$ and $F$ decreases with the increase in $\alpha$ and $n$ and is minimum when $\alpha = 30$ and $n = 2048$.

Table 4.1: The errors of approximations $\hat{F}_{\alpha,n}$ in $L_2$-distance. Here $X \sim \text{Beta}(3,3)$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n = 256$</th>
<th>$n = 512$</th>
<th>$n = 1024$</th>
<th>$n = 2048$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha = 20$</td>
<td>0.0499 (0.0096)</td>
<td>0.0466 (0.0090)</td>
<td>0.0436 (0.0071)</td>
<td>0.0419 (0.0053)</td>
</tr>
<tr>
<td>$\alpha = 25$</td>
<td>0.0466 (0.0082)</td>
<td>0.0414 (0.0078)</td>
<td>0.0375 (0.0065)</td>
<td>0.0349 (0.0048)</td>
</tr>
<tr>
<td>$\alpha = 30$</td>
<td>0.0443 (0.0090)</td>
<td>0.0387 (0.0088)</td>
<td>0.0341 (0.0074)</td>
<td>0.0306 (0.0033)</td>
</tr>
</tbody>
</table>

The question of consistent estimation of the pdf $f$ have been studied by many authors, see, for instance, Abbaszadeh et al. [1] and references therein. Here, let us compare the $L_2$-error of $\hat{f}_{\alpha,n}$ and $\hat{f}'_{\alpha,n}$ with the errors corresponding to the estimators that were constructed by means of wavelets. To compare our estimate $\hat{f}_{\alpha,n}$ and $\hat{f}'_{\alpha,n}$ with the ones based on wavelets studied in Abbaszadeh et al. [1], we took $N = 100$. Tables 4.2 and 4.3 show the $L_2$-distances for the estimates. Those estimates based on the wavelets are borrowed from [20]. We see that our approximation yields much smaller errors than the approximation based on the wavelets.

51
Table 4.2: The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha,n}$ in $L_2$-distance. Here $X \sim \text{Beta}(3,3)$.

<table>
<thead>
<tr>
<th>Estimation Methods</th>
<th>$d_{\alpha,n}$ and (Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 256</td>
</tr>
<tr>
<td>Linear</td>
<td>12.7357 (0.9611)</td>
</tr>
<tr>
<td>Hard Thresholding</td>
<td>6.7080 (1.6041)</td>
</tr>
<tr>
<td>Smoothing</td>
<td>3.0102 (1.0102)</td>
</tr>
<tr>
<td>$\hat{f}_{\alpha,n}, \alpha = 20$</td>
<td>0.27015 (0.0813)</td>
</tr>
<tr>
<td>$\hat{f}_{\alpha,n}, \alpha = 25$</td>
<td>0.30762 (0.0845)</td>
</tr>
<tr>
<td>$\hat{f}_{\alpha,n}, \alpha = 30$</td>
<td>0.34365 (0.1055)</td>
</tr>
</tbody>
</table>

Table 4.3: The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}'_{\alpha,n}$ in $L_2$-distance. Here $X \sim \text{Beta}(3,3)$.

<table>
<thead>
<tr>
<th>Estimation Methods</th>
<th>$d_{\alpha,n}$ and (Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 256</td>
</tr>
<tr>
<td>Linear</td>
<td>40.4498 (3.3822)</td>
</tr>
<tr>
<td>Hard Thresholding</td>
<td>23.1572 (4.4096)</td>
</tr>
<tr>
<td>Smoothing</td>
<td>19.0204 (3.1333)</td>
</tr>
<tr>
<td>$\hat{f}'_{\alpha,n}, \alpha = 20$</td>
<td>3.3008 (0.9696)</td>
</tr>
<tr>
<td>$\hat{f}'_{\alpha,n}, \alpha = 25$</td>
<td>4.3253 (1.3721)</td>
</tr>
<tr>
<td>$\hat{f}'_{\alpha,n}, \alpha = 30$</td>
<td>5.5218 (1.5969)</td>
</tr>
</tbody>
</table>

**Example 8.** Let us consider $X \sim (1/3)\text{Beta}(4,6) + (2/3)\text{Beta}(3,4)$, $U \sim U(0,1)$ and $Y = XU$. Following the previous example we generate $n$ independent random samples of $X$ and $U$ to approximate $F(x)$, $f(x)$, and $f'(x) = 3(\text{Beta}(3,6) - \text{Beta}(4,5)) + 4(\text{Beta}(2,4) - \text{Beta}(3,3))$ for the multiplicative-censoring model.

The estimated cdf of $X$ using $\hat{F}_{\alpha,n}$ is shown in Figure 4.5 when $n = 2000$, $\alpha = 25$ and $n = 3000$, $\alpha = 30$. In Figure 4.6 we plotted the estimated pdf $\hat{f}_{\alpha,n}$ when $n = 5000$, $\alpha = 20$ and $n = 4000$, $\alpha = 25$ and the estimated derivative function $\hat{f}'_{\alpha,n}(x)$ is shown in Figure 5.6 when $n = 5000$, $\alpha = 20$, and $n = 4000$, $\alpha = 25$. From Figures 4.5, 4.6 and 4.7 we can say that the estimated distribution and density almost coincide with $F(x)$ and $f(x)$, and the estimated derivative functions follow the same path as
\[ f'(x). \]

Figure 4.5: Estimated \( \hat{F}_{\alpha,n} \) (dotted curve) when (a): \( n = 2000, \alpha = 25 \); and (b): \( n = 3000, \alpha = 30 \). In both plots the true \( F \) (continuous curve) represents cdf of a \( X \sim (1/3)Beta(4,6) + (2/3)Beta(3,4) \) distribution.

Figure 4.6: Estimated \( \hat{f}_{\alpha,n} \) (dotted curve) when : (a) \( n = 5000, \alpha = 20 \); and (b) \( n = 4000, \alpha = 25 \). Here \( X \sim (1/3)Beta(4,6) + (2/3)Beta(3,4) \).

Table 4.4: The errors of approximations for \( \hat{F}_{\alpha,n} \) in \( L_2 \)-distance. Here \( X \sim (1/3)Beta(4,6) + (2/3)Beta(3,4) \).

<table>
<thead>
<tr>
<th>( \hat{F}_{\alpha,n} )</th>
<th>( d_{\alpha,n} ) and (Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n = 256 )</td>
</tr>
<tr>
<td>( \alpha = 20 )</td>
<td>0.0458 (0.0125)</td>
</tr>
<tr>
<td>( \alpha = 25 )</td>
<td>0.0420 (0.0139)</td>
</tr>
<tr>
<td>( \alpha = 30 )</td>
<td>0.0419 (0.0158)</td>
</tr>
</tbody>
</table>
Figure 4.7: Estimated $\hat{f}'(x)$ (dotted curve) when, (a) $n = 5000$, $\alpha = 20$; and (b) $n = 4000$, $\alpha = 25$. Here $X \sim (1/3)Beta(4, 6) + (2/3)Beta(3, 4)$.

Table 4.5: The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha, n}$ in $L_2$-distance.

<table>
<thead>
<tr>
<th>Estimation Methods</th>
<th>$d_{\alpha, n}$ and (Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 256$</td>
</tr>
<tr>
<td>Linear</td>
<td>13.0786 (1.1155)</td>
</tr>
<tr>
<td>Hard Thresholding</td>
<td>9.0021 (1.8353)</td>
</tr>
<tr>
<td>Smoothing</td>
<td>3.0145 (1.0037)</td>
</tr>
<tr>
<td>$f_{\alpha, n}$, $\alpha = 20$</td>
<td>0.2378 (0.0477)</td>
</tr>
<tr>
<td>$f_{\alpha, n}$, $\alpha = 25$</td>
<td>0.3036 (0.1021)</td>
</tr>
<tr>
<td>$f_{\alpha, n}$, $\alpha = 30$</td>
<td>0.3375 (0.0923)</td>
</tr>
</tbody>
</table>

Table 4.6: The errors of approximations for Linear, Hard Thresholding, Smoothing wavelet approximates [1], and $\hat{f}_{\alpha, n}$ in $L_2$-distance.

<table>
<thead>
<tr>
<th>Estimation Methods</th>
<th>$d_{\alpha, n}$ and (Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 256$</td>
</tr>
<tr>
<td>Linear</td>
<td>64.1098 (7.1914)</td>
</tr>
<tr>
<td>Hard Thresholding</td>
<td>32.9540 (10.2683)</td>
</tr>
<tr>
<td>Smoothing</td>
<td>20.5346 (4.5341)</td>
</tr>
<tr>
<td>$f'_{\alpha, n}$, $\alpha = 20$</td>
<td>3.0135 (0.8901)</td>
</tr>
<tr>
<td>$f'_{\alpha, n}$, $\alpha = 25$</td>
<td>3.7812 (1.5173)</td>
</tr>
<tr>
<td>$f'_{\alpha, n}$, $\alpha = 30$</td>
<td>4.5196 (1.5099)</td>
</tr>
</tbody>
</table>

Similar to the previous example we calculate the errors for $\hat{F}_{\alpha, n}$, $\hat{f}_{\alpha, n}$ and $\hat{f}'_{\alpha, n}$.

Table 4.4 shows the error for $\hat{F}_{\alpha, n}$, it is clear that the error decreases with an increase in $\alpha$ and $n$, and is minimum for $\alpha = 30$ and $n = 2048$.
Next we compare the $L_2$-errors of $\hat{f}_{\alpha,n}$ and $\hat{f}'_{\alpha,n}$ with the estimates constructed by means of wavelets [1]. From Table 4.5 and 4.6 we can see that the error for our approximation decreases with the increase in $n$. It is also clear that $\hat{f}_{\alpha,n}$ and $\hat{f}'_{\alpha,n}$ yields much smaller errors than the approximation based on the wavelets.
Chapter 5

Reconstruction of Distributions in other Incomplete Models

An incomplete distribution function has no/missing information about a certain important aspect, e.g., (1) incomplete parameter set: $X$ follows Gaussian with unknown mean $\mu$, (2) missing support information: $X$ follows Uniform from 0 to unknown $T$, (3) partial distributional information: $Z = XY$ where the distribution of $X$ is unknown, etc. Classical moment problems can be incorporated for the recovery of the missing information in these incomplete models. In Section 5.1 we recover the distribution and density functions in the case of the mixture model, the recovery of the distribution and density function of $X$ for the constrained deconvolution problem are demonstrated in Section 5.2, and in Section 5.3 we recover the distribution when the upper bound is unknown. The performance of the proposed models are examined via several examples.
5.1 Demixing Problem

In this section our aim is to recover the mixture distribution of $Y$ and $X$ when the distribution of $Y|X = x$ is known. The general assumption in this section is that $X \sim F$ with unknown cdf $F$ and pdf $f$. Three separate distributional forms of $Y$ are considered here for recovering the distribution of $U$:

a) $Y|X = x \sim \text{Bin}(m,x)$, for some known $m \in \mathbb{N}_+$,

b) $Y|X = x \sim \text{NegBin}(r,\phi(x))$, for some $r \in \mathbb{N}_+$, and

c) $Y|X = x \sim \text{Exp}(\phi(x))$, with some known rate function $\phi(x)$, $x \in \mathbb{R}_+$.

5.1.1 Binomial Mixture

Assume $Y|X = x \sim \text{Bin}(m,x)$, where $m \in \mathbb{N}_+$ is a known integer and $X \sim F$, an unknown cdf $F$. Our objective is to estimate the cdf $F$ and pdf $f$ for the given $Y_1,\ldots,Y_n$ ($n$ independent copies of $Y$). The variable $Y$ is distributed according to the mixture

$$P(Y = y) = \int_0^1 \binom{m}{y} x^y (1-x)^{m-y} dF(x), \text{ for } y = 0,1,2\ldots,m.$$  \hspace{1cm} (5.1)

Let us also consider

$$\hat{\mu}_{j,F} = \frac{1}{m[j]} \frac{1}{n} \sum_{i=1}^{n} Y_i^{[j]}$$

with $y^{[j]} = y(y-1)\ldots(y-j+1)$. In order to estimate the cdf $F$, let us substitute $\phi(x) = x, T = 1$ and $\hat{\mu}_{j,F}$ instead of $m_{\phi}(j)$ in (1.4),

$$F_{\alpha,T}(x) = \sum_{k=0}^{[\alpha]} \frac{\alpha}{\alpha} \binom{\alpha}{j} (\frac{j}{k}) (-1)^{j-k} \hat{\mu}_{j,F}.$$  \hspace{1cm} (5.2)

57
The estimation of pdf $f = F'$ requires the following substitutions $\phi(x) = x, T = 1,$ and

$$\hat{\mu}_{j+\lfloor\alpha x\rfloor,F} = \frac{1}{m^{j+\lfloor\alpha x\rfloor}} \frac{1}{n} \sum_{i=1}^{n} Y_i^{j+\lfloor\alpha x\rfloor}$$

into (1.5) to obtain

$$f_{\alpha,T}^*(x) = \frac{\Gamma(\alpha + 2)}{\Gamma([\alpha x] + 1)} \sum_{j=0}^{\alpha-\lfloor\alpha x\rfloor} \frac{(-1)^j \hat{\mu}_{j+\lfloor\alpha x\rfloor,F}}{j!(\alpha - [\alpha x] - j)!}, \text{ for } x \in [0,1].$$

The following examples demonstrate the effectiveness of the proposed distributional estimate of $U$.

**Example 9.** Assume $X \sim Beta(3,2)$ and generate a random sample $Y_1, Y_2, \ldots, Y_n$ from

$$P(Y = y) = \binom{m}{y} \frac{\int_0^1 x^y(1-x)^{m-y}a^{-1}(1-x)^{b-1} du}{B(a,b)} = \frac{\binom{m}{y}B(a+y,b+m-y)}{B(a,b)},$$

where $(a,b) \in (2,3)$. Figure 5.1 shows the estimated pdf $f(x)$ and cdf $F(x)$, for

<table>
<thead>
<tr>
<th>$d_{\alpha,n}$</th>
<th>$\alpha = 25$</th>
<th>$\alpha = 28$</th>
<th>$\alpha = 32$</th>
<th>$\alpha = 35$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{\alpha,T}$, n=850</td>
<td>0.1657836</td>
<td>0.153522</td>
<td>0.1449409</td>
<td>0.1498753</td>
</tr>
<tr>
<td>$f_{\alpha,T}$, n=1750</td>
<td>0.1555064</td>
<td>0.1393623</td>
<td>0.1302612</td>
<td>0.1300159</td>
</tr>
<tr>
<td>$f_{\alpha,T}$, n=3000</td>
<td>0.155376</td>
<td>0.1396915</td>
<td>0.1284961</td>
<td>0.1247769</td>
</tr>
<tr>
<td>$F_{\alpha,T}$, n=900</td>
<td>0.02451202</td>
<td>0.02269515</td>
<td>0.02073172</td>
<td>0.01950324</td>
</tr>
<tr>
<td>$F_{\alpha,T}$, n=2000</td>
<td>0.02322071</td>
<td>0.0212571</td>
<td>0.01903979</td>
<td>0.01822334</td>
</tr>
<tr>
<td>$F_{\alpha,T}$, n=3000</td>
<td>0.02275302</td>
<td>0.02067746</td>
<td>0.01847673</td>
<td>0.01767947</td>
</tr>
</tbody>
</table>
the binomial mixture model. The $Y$ values are generated from the above distribution using the acceptance rejection sampling method [14] for $\alpha = 32$ and $m = 60$. We also calculated the error in approximation for $f$ and $F$ using (4.6), for $N = 50$. Table 5.1 illustrates that the estimate of the recovered distribution of $X$ is better with an increase in random sample size as well as an increase in the parameter $\alpha$, and for the given set of parameters the error is minimum when $\alpha = 35$ and $n = 3000$.

5.1.2 Negative Binomial Mixture

Let $Y|X = x \sim NegBin(r, \phi(x))$, where $X$ follows some unknown distribution $F$ similar to Section 5.1.1. Under this setup the probability density function of $Y$ comes out to be,

$$P(Y = y) = \int_0^1 \left( \frac{y + r - 1}{y} \right) x\phi(x)^y(1 - \phi(x))^{y}dF(x), \ y = 0, 1, 2, ..., r > 0$$
The approximation of the unknown cdf $F$ and the density function $f$ are considered in this section.

**Example 10.** Assume $\phi(x) = x$ and $F(x) = x^2$. The ordinary moments of $Y$ and $\phi$-transformed moments of $F$ are related to each other in this case by,

$$m_\phi(j) = \frac{E(Y^j)}{j!}. \quad (5.3)$$

Substituting the format of $\phi(x)$ and $F(x)$ into (5.3) we get,

$$m_\phi(j) = \int_0^\infty [\phi(x)]^j dF(x) = \frac{2}{j+2}.$$ 

Using the moment function we obtain the Moment-Recovered approximations of $F$ and $f$:

$$F_{\alpha}(x) = \sum_{k=0}^{[\alpha\phi(x)]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{j}{k} (-1)^{j-k} m_\phi(j), \quad \text{and} \quad (5.4)$$

$$f_{\alpha}(x) = \frac{\Gamma(\alpha + 2) |\phi'(x)|^{\alpha - [\alpha\phi(x)]} (-1)^j m_\phi(j + [\alpha\phi(x)])}{\Gamma([\alpha\phi(x)] + 1) \sum_{j=0}^{\alpha - [\alpha\phi(x)]} j!(\alpha - [\alpha\phi(x)] - j)!}. \quad (5.5)$$

The approximated cdf $F_{\alpha}(x)$ and pdf $f_{\alpha}(x)$ are plotted against their true values in Figure 5.4. Both the plots reveal a close resemblance between the approximated and their actual distributional counterparts. The sup-norm and $L_2$-norm deviation of the Moment-Recovered approximations are displayed in Table 5.2 which indicates a better resemblance to $f$ and $F$ for the values of $\alpha$ parameter at 20 and 28 respectively.
Figure 5.2: Approximated (a) pdf and (b) cdf for negative binomial mixture (black). In both plots the red line denotes the true pdf $f(x) = 2x$ and cdf $f(x) = x^2$, when $\alpha = 30$.

Table 5.2: The errors of approximations for $f_\alpha$ and $F_\alpha$ in sup-norm and $L_2$-norm, when $Y|X = x \sim NegBin(r, \phi(x))$.

<table>
<thead>
<tr>
<th>Distance</th>
<th>$\alpha = 20$</th>
<th>$\alpha = 24$</th>
<th>$\alpha = 28$</th>
<th>$\alpha = 32$</th>
<th>$\alpha = 35$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_\alpha$, sup-norm</td>
<td>0.09091</td>
<td>0.13077</td>
<td>0.10000</td>
<td>0.09412</td>
<td>0.33400</td>
</tr>
<tr>
<td>$F_\alpha$, sup-norm</td>
<td>0.03571</td>
<td>0.05327</td>
<td>0.03355</td>
<td>0.03433</td>
<td>0.03468</td>
</tr>
</tbody>
</table>

5.1.3 Exponential Mixture

Consider that a sample is chosen from an absolutely continuous distribution $G$ with corresponding density $g$ and that $\phi$ is an increasing/decreasing function in $U$. Let $Y|X = x \sim Exp(\phi(x))$, where $X \sim F$ (unknown). The marginal probability density function of $Y$ is,

$$g(y) = \int_0^\infty \frac{1}{\phi(x)} e^{-y/\phi(x)} dF(x)$$

(5.6)

for $x \in \mathbb{R}^+$. Our goal is to approximate the unknown cdf $F$ or its density function $f$, given a sample from $G$. Thus the Moment-Recovered approximation of $F$ and $f$ are
constructed as:

\[
F_{\alpha}(x) = 1 - \sum_{k=0}^{[\alpha \phi(x)]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{j}{k} (-1)^{j-k} m_{\phi}(j), \quad \text{and} \quad (5.7)
\]

\[
f_{\alpha}(x) = \frac{\Gamma(\alpha + 2)|\phi'(x)|}{\Gamma([\alpha \phi(x)] + 1)} \frac{\alpha^{-[\alpha \phi(x)]}}{j!} \frac{(-1)^j m_{\phi}(j + [\alpha \phi(x)])}{j!(\alpha - [\alpha \phi(x)] - j)!}. \quad (5.8)
\]

Examples 11 and 12 investigate the effects on the approximations for different choices of \(\phi(x)\) and \(F(x)\).

**Example 11.** Let us consider a specific case \(\phi(x) = (x + 1)^{-1}\) and \(F(x) = 1 - (x + 1)^{-2}\). Therefore for this example we have

\[
m_{\phi}(j) = \frac{E(Y^j)}{j!} = \int_0^\infty [\phi(x)]^j dF(x) = \frac{2}{j + 2}.
\]

The approximated density and distribution are plotted against their true density and distributions, respectively in Figure 5.4, which shows promising resemblance. From Table 5.3, it is clear that the distance between the approximated and actual pdf and cdf is minimum when \(\alpha\) is 32.

![Figure 5.3](a) ![Figure 5.3](b)

Figure 5.3: Approximated \(a: f_{\alpha}(x)\) and \(b: F_{\alpha}(x)\) (black) for \(f(x) = 2(x + 1)^{-3}\) and \(F(x) = 1 - (x + 1)^{-2}\) (red) when \(\alpha = 30\).
Table 5.3: The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm and $L_2$-norm, when $Y|X = x \sim Exp(\phi(x))$ and $F(x) = 1 - (x + 1)^{-2}$.

<table>
<thead>
<tr>
<th>Distance</th>
<th>$\alpha = 20$</th>
<th>$\alpha = 24$</th>
<th>$\alpha = 28$</th>
<th>$\alpha = 32$</th>
<th>$\alpha = 35$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_\alpha, sup$-norm</td>
<td>0.08991</td>
<td>0.10403</td>
<td>0.07012</td>
<td>0.05882</td>
<td>0.10988</td>
</tr>
<tr>
<td>$F_\alpha, sup$-norm</td>
<td>0.03571</td>
<td>0.04798</td>
<td>0.02585</td>
<td>0.02307</td>
<td>0.03468</td>
</tr>
</tbody>
</table>

**Example 12.** Assume $F(x) = 1 - (x + 1)^{-4}$ and the similar form of $\phi(x)$ as Example 11. Then by substituting the $\phi$ and $F$ in (5.3) we get the moment function to be

$$m_\phi(j) = \frac{4}{j+4}.$$ 

Employing the above moment function we find the MR approximation of the pdf and cdf as in (5.8) and compare the results with the true cdf and pdf in Figure 5.4. A close affinity can be seen between the MR approximated distribution and actual distribution. The choice of $\alpha$ as 35 and 32 yields minimum distance between the approximated and true pdf $f$ and cdf $F$.

![Figure 5.4](image-url)

Figure 5.4: Approximated $a: f_\alpha(x)$ and $b: F_\alpha(x)$ (black) for $f(x) = 4(x + 1)^{-5}$ and $F(x) = 1 - (x + 1)^{-4}$ (red) when $\alpha = 30$.  

63
Table 5.4: The errors of approximations for $f_\alpha$ and $F_\alpha$ in $sup$-norm and $L_2$-norm, when $Y|X = x \sim Exp(\phi(x))$ and $F(x) = 1 - (x + 1)^{-4}$.

<table>
<thead>
<tr>
<th>Distance</th>
<th>$\alpha = 20$</th>
<th>$\alpha = 24$</th>
<th>$\alpha = 28$</th>
<th>$\alpha = 32$</th>
<th>$\alpha = 35$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_\alpha$, $sup$-norm</td>
<td>0.50000</td>
<td>0.58689</td>
<td>0.46129</td>
<td>0.36169</td>
<td>0.31914</td>
</tr>
<tr>
<td>$F_\alpha$, $sup$-norm</td>
<td>0.04636</td>
<td>0.092544</td>
<td>0.06060</td>
<td>0.03833</td>
<td>0.07837</td>
</tr>
</tbody>
</table>

### 5.2 Constrained Deconvolution Problem

Let us consider the following problem:

$$Z = a X + \beta Y,$$

where $Y$ is the error term, $(a, \beta)$ are known constants, $X$ and $Y$ are independent variables, and the distribution of $X$ is unknown. In this section we will consider the problem of recovering the distribution of $X$, when $Z$ is observed. In order to achieve this goal we will apply the Laplace transform inversion technique for estimation of the distribution of $X$.

Assume $0 < a < \beta$, $0 < \frac{a}{\beta} = \gamma < 1$ and $\frac{Z}{\beta} = \gamma X + Y$, where $\frac{Z}{\beta} \sim g(.)$ and $X \overset{d}{=} Y \sim f(.)$. The Laplace transformation of $\frac{Z}{\beta}$ can be expressed as

$$\mathcal{L}_g(s) = \int_0^\infty e^{-sx}g(x)dx$$

$$= E\left(e^{-s\gamma}\right) = E\left(e^{-s(\gamma X + Y)}\right)$$

$$= \mathcal{L}_f(\gamma s)\mathcal{L}_f(s).$$
Moreover the Laplace transformation of $Y$, i.e. $L_f(s)$ can be written as

$$L_f(s) = \frac{L_g(s)}{L_f(\gamma s)}. \quad (5.9)$$

Since $L_g(\gamma s) = L_f(\gamma s)L_f(\gamma^2 s)$, therefore $L_f(\gamma s)$ can be rewritten as

$$L_f(\gamma s) = \frac{L_g(\gamma s)}{L_f(\gamma^2 s)}. \quad (5.10)$$

Substituting $L_f(\gamma s)$ from (5.10) into (5.9) yields,

$$L_f(s) = \frac{L_g(s)}{L_g(\gamma s)}L_f(\gamma^2 s). \quad (5.11)$$

Again we know that,

$$L_g(\gamma^i s) = L_f(\gamma^{i+1} s)L_f(\gamma^{i+2} s)$$

for $i = 2, 3, \ldots$. Therefore we finally obtain

$$L_f(s) = \frac{L_g(s)}{L_g(\gamma s)} \frac{L_g(\gamma^2 s)}{L_g(\gamma^3 s)} \frac{L_g(\gamma^3 s)}{L_g(\gamma^4 s)} \ldots \frac{L_g(\gamma^{2k} s)}{L_g(\gamma^{2k+1} s)} = \prod_{k=0}^{\infty} \frac{L_g(\gamma^{2k} s)}{L_g(\gamma^{2k+1} s)}$$

Hence for convolution we consider

$$L_{f,N}(s) = \prod_{k=0}^{N} \frac{L_g(\gamma^{2k} s)}{L_g(\gamma^{2k+1} s)}$$

where $s = j \log(b)$. To approximate the cdf $F$ and the pdf $f$ we substitute $m_\phi(j)$ in
(1.4) and (1.5) with $\mathcal{L}_{f,N}(s)$ to get

$$F_\alpha(x) = 1 - \sum_{k=0}^{[\alpha b^{-x}]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} (-1)^{j-k} \mathcal{L}_{f,N}(j \log(b)) \quad (5.12)$$

and

$$f_\alpha(x) = \frac{\log(b)[\alpha b^{-x}]}{\alpha} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha b^{-x}+1)} \sum_{j=0}^{\alpha-\lfloor \alpha b^{-x} \rfloor} \frac{(-1)^{j} \mathcal{L}_{f,N}((j + \lfloor \alpha b^{-x} \rfloor) \log(b))}{j!(\alpha - \lfloor \alpha b^{-x} \rfloor - j)!},$$

for $x \in [0,1]$. The following example illustrates the recovery of the distribution of $X$
using the Laplace transform inversion.

**Example 13.** Assume $X \overset{d}{=} Y \sim \text{Exp}(1)$ and $\frac{Z}{\beta} = \gamma X + Y$, with $\gamma = \frac{\alpha}{\beta}$. The Laplace transform inversion for the distribution in the example is $\mathcal{L}_{f,N}(s) = \frac{s \gamma^{2N+2} + 1}{s+1}$ with $s = j \log(b)$. The approximated pdf $f_\alpha$ and cdf $F_\alpha$ are constructed from (5.13) and (5.12), respectively, for $\alpha = 32$, $\gamma = 0.1$, $b = 1.5$ and $N = 100$. In Figure 5.5, we can see the approximated cdf and pdf along with the true functions. Table 5.5 demonstrates the $sup$-norm distance between the approximated pdf $f_\alpha$ and cdf $F_\alpha$ from the pdf and cdf $f$ and $F$ under consideration. For the given set of parameter values, the minimum distance occurs at $\alpha = 35$ for both the pdf and the cdf.

<table>
<thead>
<tr>
<th>Distance</th>
<th>$\alpha = 20$</th>
<th>$\alpha = 24$</th>
<th>$\alpha = 28$</th>
<th>$\alpha = 32$</th>
<th>$\alpha = 35$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_\alpha$, $sup$-norm</td>
<td>0.16270</td>
<td>0.13692</td>
<td>0.11784</td>
<td>0.10352</td>
<td>0.09416</td>
</tr>
<tr>
<td>$F_\alpha$, $sup$-norm</td>
<td>0.09982</td>
<td>0.08322</td>
<td>0.07173</td>
<td>0.06181</td>
<td>0.04721</td>
</tr>
</tbody>
</table>
5.3 Recovery of Distribution with Unknown Support

The last implementation considered in this dissertation is the recovery of a distribution with unknown finite support. The structure of the problem is as follows. Suppose there is a random variable $X \sim F$ with $supp\{F\} = [0, T]$. Under this problem we consider $T$ to be unknown, where $0 < T < \infty$. The approximations of the cdf $F$ and the pdf $f$ can be calculated using the method proposed in [28] and [29] by approximating the support from their sequence of moments. For a given moment sequence of $F$:

$$
\mu_{j,F} = \int x^j dF(x).
$$

Using this moment sequence one can estimate $T$ using,

$$
T_m = (\mu_{m,F})^{\frac{1}{m}},
$$

where $m$ is a sufficiently large positive integer [25]. The Moment-Recovered approximations of $F$ and $f$ can be defined as follows:

$$
F_{\alpha,m}(x) = \sum_{k=0}^{[\frac{\alpha}{T_m}]} \sum_{j=k}^{\alpha} \binom{\alpha}{j} \binom{j}{k} \frac{(-1)^{j-k}}{T_m^j} \mu_{j,F}
$$
and
\[
f_{\alpha,m}(x) = \frac{\Gamma(\alpha + 2)|\frac{1}{T_m}|^{\alpha - \left\lfloor \frac{\alpha}{T_m} \right\rfloor}}{\Gamma\left(\left\lfloor \frac{\alpha}{T_m} \right\rfloor + 1\right)} \sum_{j=0}^{\left\lfloor \alpha - \left\lfloor \frac{\alpha}{T_m} \right\rfloor \right\rfloor} \frac{(-1)^j}{j!} \frac{\mu_{j+\left\lfloor \frac{\alpha}{T_m} \right\rfloor}}{\left(\alpha - \left\lfloor \frac{\alpha}{T_m} \right\rfloor - j\right)!} x^{T_m + \left\lfloor \frac{\alpha}{T_m} \right\rfloor}
\]
for \(0 \leq x \leq T_m\). Examples 14, 15, and 16 demonstrate the recovery of support for various distributions.

**Example 14.** Assume \(F(x) = x\), i.e., \(X \sim U(0,1)\) and \(T = 1\). In this case \(\mu_{j,F} = \frac{1}{(j+1)}, j \in \mathbb{N}, T_m = \left(\frac{1}{m+1}\right)^{1/m}\), and \(F(x) = \frac{x}{T_m}\) for \(0 \leq x \leq T_m\). To evaluate \(F_{\alpha,m}\) we take \(\alpha = 30\) and \(m = 500\). Figure 5.6 displays the curves of \(F_{\alpha,m}\) and \(F(x) = x\) on the interval \([0,T_m]\). The distance between the estimate and true function is minimum when \(\alpha = 32\), and is given in Table 5.6.

![Figure 5.6: Approximated \(F_{\alpha,m}\) when \(\alpha = 30\) and \(m = 500\) corresponding to example 14.](image)

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(F_{\alpha,m}, sup)-norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.03703</td>
</tr>
<tr>
<td>24</td>
<td>0.03038</td>
</tr>
<tr>
<td>28</td>
<td>0.02670</td>
</tr>
<tr>
<td>32</td>
<td>0.02293</td>
</tr>
<tr>
<td>35</td>
<td>0.03179</td>
</tr>
</tbody>
</table>

**Example 15.** Consider \(f(x) = -3x^2 \log(x^3)\) and \(F(x) = x^3 - x^3 \log(x^3)\) for \(0 \leq x \leq T_m\). Examples 14, 15, and 16 demonstrate the recovery of support for various distributions.

![Figure 5.6: Approximated \(F_{\alpha,m}\) when \(\alpha = 30\) and \(m = 500\) corresponding to example 14.](image)

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(F_{\alpha,m}, sup)-norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.03703</td>
</tr>
<tr>
<td>24</td>
<td>0.03038</td>
</tr>
<tr>
<td>28</td>
<td>0.02670</td>
</tr>
<tr>
<td>32</td>
<td>0.02293</td>
</tr>
<tr>
<td>35</td>
<td>0.03179</td>
</tr>
</tbody>
</table>

**Example 15.** Consider \(f(x) = -3x^2 \log(x^3)\) and \(F(x) = x^3 - x^3 \log(x^3)\) for \(0 \leq x \leq T_m\). Examples 14, 15, and 16 demonstrate the recovery of support for various distributions.

![Figure 5.6: Approximated \(F_{\alpha,m}\) when \(\alpha = 30\) and \(m = 500\) corresponding to example 14.](image)
Figure 5.7: Approximated (a): $f_{\alpha,m}$ and (b): $F_{\alpha,m}$ of $f$ and $F$ for unknown $T$ when $\mu_{j,F} = \frac{9}{(j+3)^2}$, $f(x) = -2x^2 \log(x^3)$ and $F(x) = x^3 - x^3 \log(x^3)$.

$x \leq T_m$. For the given example $\mu_{j,F} = \frac{9}{(j+3)^2}$, $j \in \mathbb{N}$, and $T_m = \left(\frac{9}{(m+3)^2}\right)^{1/m}$.

From Figure 5.7 it is clear that the approximated cdf is much better compared to the approximated pdf. Table 5.7 shows that for the given set of $\alpha$ values the distance between approximated $f_{\alpha,m}$ and $F_{\alpha,m}$ and true pdf and cdf is minimum at $\alpha=28$ and $\alpha=35$.

Table 5.7: The errors of approximations for $f_\alpha$ and $F_\alpha$ in sup-norm

<table>
<thead>
<tr>
<th>Distance</th>
<th>$\alpha = 20$</th>
<th>$\alpha = 24$</th>
<th>$\alpha = 28$</th>
<th>$\alpha = 32$</th>
<th>$\alpha = 35$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{\alpha,m}$, sup-norm</td>
<td>0.03402</td>
<td>0.04384</td>
<td>0.03176</td>
<td>0.03258</td>
<td>0.03514</td>
</tr>
<tr>
<td>$F_{\alpha,m}$, sup-norm</td>
<td>0.34386</td>
<td>0.32978</td>
<td>0.41081</td>
<td>0.34199</td>
<td>0.29815</td>
</tr>
</tbody>
</table>

Example 16. Let us assume $f(x) = -\log(x)$ and $F(x) = x - x \log(x)$ for $0 \leq x \leq T_m$, where $\mu_{j,F} = \frac{1}{(j+1)^2}$, $j \in \mathbb{N}$, and $T_m = \left(\frac{1}{(m+1)^2}\right)^{1/m}$. It can be seen clearly in Figure 5.8 that the true pdf and cdf are closely approximated by $f_\alpha$ and $F_\alpha$. Table 5.8 illustrates that both the approximations have minimum sup-norm distance at $\alpha = 35$. 69
Figure 5.8: Approximated (a) \( f_{\alpha,m} \) and (b) \( F_{\alpha,m} \) of \( f \) and \( F \) for unknown \( T \) when
\[ \mu_{j,F} = \frac{1}{(j+1)^2}, \]
\[ f(x) = -2\log(x) \] and \( F(x) = x - x\log(x) \).

Table 5.8: The errors of approximations for \( f_{\alpha} \) and \( F_{\alpha} \) in \textit{sup}-norm

<table>
<thead>
<tr>
<th></th>
<th>( \alpha = 20 )</th>
<th>( \alpha = 24 )</th>
<th>( \alpha = 28 )</th>
<th>( \alpha = 32 )</th>
<th>( \alpha = 35 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{\alpha,m} ), \textit{sup}-norm</td>
<td>0.11443</td>
<td>0.09381</td>
<td>0.07804</td>
<td>0.06554</td>
<td>0.05773</td>
</tr>
<tr>
<td>( F_{\alpha,m} ), \textit{sup}-norm</td>
<td>0.93495</td>
<td>0.76435</td>
<td>0.61865</td>
<td>0.49151</td>
<td>0.40575</td>
</tr>
</tbody>
</table>
Chapter 6

Conclusion and Future Work

To summarize, diverse applications of the moment problem are investigated in this dissertation. The motivation behind the proposed methodologies was the work by Mnatsakanov [28, 29], where recovery of the density and the distribution functions were studied in the framework of the Hausdorff moment problem. In this dissertation several new directions of the Moment-Recovered approximations are proposed, such as, recovery of regression functions, derivative functions, and distributions in several incomplete models. The performances of the recovered functions are validated through multiple examples. Moreover, our proposed techniques are shown to approximate the functions of interest better than several well-recognized methods.

In Chapters 2 and 3, the approximation and estimation of the regression function is conducted when both the distribution of the covariate $X$ is known (Section 2.1) and is unknown (Section 2.2). Comparison of our approximated constructions based on (1.5) with those given by (1.3) via the numerical evaluations shows the superiority of former method. Additionally, a comparative study of our estimated regression function with the estimate proposed by Koul and Song [23] is demonstrated. It is also
shown that the moment-determinate regression function and the conditional variance can be recovered with very high precision given a sequence of the product moments.

Chapter 4 demonstrates the reconstruction of derivative functions via a sequence of moments. Application of the proposed reconstruction for estimating $F$, $f$, and $f'$ in the framework of multiplicative-censoring model is discussed in Section 4.2. The simulation studies illustrated the close resemblance of the recovered $F$, $f$, and $f'$ with the original functions. In particular, the consistency in probability of proposed nonparametric estimates are established.

Applicability of the Moment-Recovered approach in other incomplete models is elaborated upon in Chapter 5. The problem of recovering mixing distributions and density functions in the binomial, negative binomial, and exponential mixture model is examined in Section 5.1. In addition, Section 5.2 provides the solution to the problem of recovering the unknown distribution in the constrained deconvolution problem. Finally in Section 5.3, we consider the problem of approximating the distribution function $F$ and corresponding density function $f$ when the support is unknown (the Hausdorff case). Thorough examination using distinct examples portrayed the efficiency of the proposed techniques.

As mentioned above, in Chapter 5 we discussed the problem of function approximation for the negative binomial mixture, exponential mixture, and deconvolution problem as well as recovery of distributions for unknown support. Future work revolves around the study of asymptotic properties of the estimates for these indirect models. The main aspect of this approach is to find the consistent estimate of the moments up to some order.

So far, most of the analysis is executed in *Mathematica*, as R cannot handle high
precision analysis. Recent development in R packages such as, \texttt{Rmpfr} and \texttt{gmp} makes it more convenient to handle problems requiring higher-precision numbers. Therefore, a future goal is to prepare an R package that can perform all the proposed Moment-Recovered implementations efficiently.

Kernel based regressions, smoothing splines, and wavelet-based methods are widely used in predictive analysis with real data problems. In Section 3.2, it is seen that the Moment-Recovered regression is competitive with respect to kernel based methods. Therefore, additional future research will focus on the application and examination of the proposed moment-based nonparametric regressions, this will also be performed with real data applications (where data is mostly incomplete).
Bibliography


