THE UNIVERSITY OF MANITOBA

SOME CONTRIBUTIONS TO NONPARAMETRIC ESTIMATION OF A PROBABILITY DENSITY FUNCTION AND ITS FUNCTIONALS

BY

LEO ODIWUOR ODONGO

A THESIS

SUBMITTED TO THE FACULTY OF

GRADUATE STUDIES IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Department of Statistics

WINNIPEG, MANITOBA

DECEMBER, 1992

ISBN 0-315-81670-8

Name ____________________

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

STATISTICS — SUBJECT TERM

0463 — SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Plant Physiology 0817
- Range Management 0777
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Physiology 0433
- Radiation 0821
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Chemotherapy 0992
- Dentistry 0567
- Education 0350
- Hospital Management 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology 0380
- Occupational Health and Therapy 0354
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Physical Therapy 0382
- Public Health 0573
- Radiology 0574
- Recreation 0575
- Speech Pathology 0460
- Toxicology 0383

Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Electronics and Electricity 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756
- Solid State 0611

Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984

Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Heat and Thermodynamics 0348
- Hydraulic 0545
- Industrial 0546
- Marine 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

SOME CONTRIBUTIONS TO NONPARAMETRIC ESTIMATION OF A PROBABILITY DENSITY FUNCTION AND ITS FUNCTIONALS

BY

LEO ODIWUOR ODONGO

A Thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

TABLE OF CONTENTS

Page

ABSTRACT

In this dissertation we consider four problems of nonparametric inference. The first problem deals with the estimation of the slope of a linear regression. We assume that the conditional density function of Y given X=x is f(y-α-βx) where the form of f is unknown and β is the slope of the regression of Y on X which is to be estimated. Without loss of generality we assume α to be zero. It is assumed that X is a bounded random variable. If the form of f were known, then the estimation problem could be treated by the classical maximum likelihood method. In the present study we first define an empirical likelihood equation based on the so-called kernel estimates of a probability density function and its derivatives and then propose a nonparametric estimate of β by combining a consistent estimate of β and the empirical likelihood equation. We derive the large sample distribution of our estimate by using the convergence properties of the kernel estimates of a probability density function and its derivatives in conjunction with the properties of U-statistics. It is found that the proposed estimate has the same large sample distribution as possessed by the maximum likelihood estimate.

In the second problem we study a naive estimate of the Hodges-Lehmann functional given by the integral of the square of an estimate of the probability density function. We establish the rate of strong convergence and the asymptotic normality of the estimate and derive an expression for the optimal smoothing parameter that minimizes the mean square error. We illustrate our results on simulated data to construct a 95% confidence interval for the Hodges-Lehmann functional when the underlying distributions are standard normal, standard Cauchy and standard logistic.

In the third problem we derive the mean square errors of the kernel estimates of the probability density function and its derivatives, the population distribution function, the quantiles and the mode. We obtain expressions for the smoothing parameter that minimize the mean square errors.

The fourth problem is a continuation of the third problem. In this problem we consider the computations of the mean square errors of the kernel estimates of the conditional probability density function, the regression function, the conditional quantile and the conditional mode. We derive expressions for the smoothing parameter that minimize the mean square errors.

ACKNOWLEDGEMENT

I would like to thank my supervisor Professor M. Samanta for his tireless effort in guiding me throughout the entire preparation of this dissertation. I also wish to thank the rest of the advisory committee, Professors J.F. Brewster, R.G. Woods and A. Thavaneswaran for many helpful suggestions that have led to an improved presentation of the dissertation. I would like to express my sincere thanks to Professors L.K. Chan and S.W. Cheng for providing me with all necessary facilities that made my graduate studies in the department of statistics both enjoyable and stimulating.

Financial support was offered to me by the Canadian Commonwealth Scholarship and Fellowship Committee to whom I owe a debt of gratitude. A study leave was kindly granted to me by Kenyatta University under their staff development scheme and to them I say thanks.

I am very thankful to my dear wife, Lily Atieno, for her encouragement and moral support, to my adorable daughters Rebecca Adhiambo, Lucy Achieng' and Brenda Akinyi for their patience during my absence from home, and to my parents for the sacrifices they made in financing my earlier education.

Finally, I wish to express my appreciation to the entire faculty of the department of Statistics at the University of Manitoba for their help in whatever form during my studies. Last but not least I would also like to mention my source of inspiration, Professor J.W. Odhiambo who introduced me to statistical thinking when I was a student at the University of Nairobi.

CHAPTER I

INTRODUCTION

1.1 A Brief Literature Review

The classical statistical methods are parametric. The assumption is that the sample observations come from a population having a probability density function (pdf) whose functional form is known but depends on a number of unknown parameters. In this case the problem is either to estimate the unknown parameters or to devise tests and derive confidence regions, etc, for the unknown parameters based on the sample. This assumption is rather strong because the assumed parametric model need not be the "actual" one if there is one, and the statistical methods developed for a particular parametric model could lead to erroneous conclusions when applied to a slightly perturbed model. The nonparametric statisical methods which do not assume a certain form of the statistical population mainly appeared in the literature during the second world war. One of the important topics within the nonparametric statistical methods is the nonparametric estimation of a probability density function using the kernel method as developed by Rosenblatt (1956). The idea of Rosenblatt was taken up more rigorously by Parzen (1962) to study the statistical properties of a general class of nonparametric estimates of a probability density function and the mode. Since then a number of researchers have investigated various aspects of nonparametric estimation of the probability density function and its functionals. Among researchers in this field we mention : Nadaraya (1964a, 1964b, 1965, 1974), Bhattacharya (1967), Schuster (1969, 1972), Samanta (1971, 1973, 1974a, 1989), Singh (1977a, 1979) and Karunamuni and Mehra (1990). Various methods which have been developed for density and curve estimation are found in monographs by Prakasa Rao (1983), Devroye and Györfi (1985), Silverman (1986), Devroye (1987), Müller (1988), Eubank (1988) and Nadaraya (1989).

1.2 Kernel Estimation of a Probability Density Function and Its Derivatives

Let $X_1, X_2, ..., X_n$ be independent random variables having a common probability density function (pdf) f and let ϕ be a real valued Borel measurable function such that the Lebesgue integral $\int_{-\infty}^{\infty}\phi(u)du=1$. Usually ϕ is taken to be a probability density function. An estimate $f_n(x)$ of the pdf f(x) is given by

$$f_n(x)=\frac{1}{na_n}\sum_{i=1}^{n}\phi\left(\frac{x-X_i}{a_n}\right), \tag{1.1}$$

where $\{a_n\}$ is a sequence of positive numbers converging to zero as n tends to infinity. The function $f_n(x)$ is called a kernel estimate of f(x). Convergence properties of this estimate have been extensively studied. We only mention here the works of Rosenblatt (1956, 1971), Parzen (1962), Nadaraya (1970, 1974), Bhattacharya (1967), Schuster (1969), Singh (1977a), Karunamuni (1991) and Karunamuni and Mehra (1991). Multivariate analogues of the estimate in (1.1) have been considered among others by Cacoullos (1966), Van Ryzin (1969), and Epanechnikov (1969).

For the one-dimensional case, let $r\geq 0$ be an integer and denote by $f^{(r)}$ the r-th order derivative of f where $f^{(0)}\equiv f$. Suppose that the kernel ϕ and its first r derivatives are functions of bounded variation and the Lebesgue integral $\int_{-\infty}^{\infty}|u|\phi(u)du$ is finite. Bhattacharya (1967) proposed and studied the asymptotic properties of the estimate of $f^{(i)}(x)$ given by

$$f_n^{(i)}(x)=\frac{1}{na_n^{i+1}}\sum_{i=1}^{n}\phi^{(i)}\left(\frac{x-X_i}{a_n}\right), \tag{1.2}$$

i=0, 1, 2, ..., r.. Schuster (1969) and Karunamuni and Mehra (1990) also studied the asymptotic properties of the estimates in (1.2) and obtained their rates of convergence.

We now mention briefly some instances where kernel estimates of density functions and their functionals have been applied to statistical problems.

Parzen (1962) and Nadaraya (1965) considered the problem of estimating the mode of a univariate pdf. Van Ryzin (1969) and Samanta (1973) considered the estimation of the mode of a multivariate density. Murthy (1965) applied the kernel method to the estimation of jumps, reliability and hazard rates. Using kernel estimate of the pdf, Bhattacharya (1967) studied a class of estimates of the Fisher information. Nadaraya (1965, 1970), Rosenblatt (1969), Schuster (1972), Schuster and Yakowitz (1979), and Gasser et al (1984) all considered the estimation of regression curves. Singh (1977b) considered a wide range of problems including those in econometrics which could be solved by using the kernel method. Rosenblatt (1975) devloped a test of independence using functionals of kernel estimates of density functions. Ahmad and Lin (1977) suggested estimates for a vector valued bivariate failure rate. Aitken and MacDonald (1979) applied kernel based density estimates to categorical data and Tetterington (1980) explored their use in this area. Copas and Fryer (1980) used kernel estimation techniques to determine the security risks among psychiatric patients. Ramlau-Hansen (1983) considered the use of kernel density estimation methods in estimating intensity functions in survival analysis. Thavaneswaran (1988) applied kernel functions to estimate signals for a semimartingale model.

In the next section we shall briefly discuss the problem of estimating the slope of a linear regression.

1.3 Efficient Nonparametric Estimation of the Slope of a Linear Regression

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be a sequence of independent and identically distributed two dimensional random variables such that $P[0 \le X_i \le 1] = 1$ for $i=1, 2, 3, \ldots$. Let the conditional distribution function of Y given X=x be denoted by $G(y/x)$, where $G(y/x)$ is absolutely continuous having the conditional probability density function $g(y/x)$. We suppose further that there is a probability density function f and two real numbers α and β, where $-\infty< \alpha<\infty$, $|\beta|<\gamma$ and γ is a known constant such that

$$g(y/x)=f(y-\alpha-\beta x) \tag{1.3}$$

for all $0\le x\le 1$ and $-\infty<y<\infty$. In such a case, if E(Y/X) exists, then it is a linear function in X with slope β. Even if E(Y/X) does not exist, it can be easily shown that, the conditional median (or any given quantile) of Y given X is a linear function in X with slope β. In this model β may be called the slope of the linear regression of Y on X in a sense more general than usual. In this dissertation we consider the problem of estimating and testing hypotheses about β. If f were known, then β could be estimated from (X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n) by the method of maximum likelihood by solving the pair of likelihood equations

$$\sum_{i=1}^{n}\frac{f^{(1)}(Y_i-a-bX_i)}{f(Y_i-a-bX_i)}=0 \tag{1.4}$$

and

$$\sum_{i=1}^{n}X_i\frac{f^{(1)}(Y_i-a-bX_i)}{f(Y_i-a-bX_i)}=0 \tag{1.5}$$

It is well known that under certain regularity conditions (see Cramer, 1946, p. 500) the likelihood equations have a solution (a_n^*, b_n^*) converging in probability to (α, β). Furthermore, $\sqrt{n}\,(b_n^*-\beta)$ converges in distribution to a normal random variable with mean zero and variance $1/\{\sigma^2 J(f)\}$, where $\sigma^2=\mathrm{var}(X_1)$ and J(f) is the Fisher information defined by

$$J(f)=\int_{-\infty}^{\infty}\frac{\{f^{(1)}(x)\}^2}{f(x)}dx. \tag{1.6}$$

Since $1/\{n\sigma^2 J(f)\}$ is the Cramer-Rao lower bound for the variance of all unbiased estimates of β the maximum likelihood estimate b_n^* is said to be asymptotically efficient.

The nonparametric counterpart of this problem arises when f is unknown. In the nonparametric setup, Theil (1950b) proposed an estimate b_n^+ of β given by

$$b_n^+ = \text{median}\{S_{ij}\}, \tag{1.7}$$

where $S_{ij} = (Y_j - Y_i)/(X_j - X_i)$ for $1 \leq i < j \leq n$. For testing the hypothesis $H_0: \beta = \beta_0$ against a suitable alternative, Theil (1950a) proposed a test based on Kendall's tau statistic. Other nonparametric procedures for making inferences about the regression coefficient have been proposed by Adichie (1967), Sen (1968), Srivastava and Saleh (1970), and Jureckova (1971) among others.

Samanta (1971) defined an empirical likelihood equation based on kernel estimates of a probability density function and its derivatives. Under some regularity conditions he proved that with a high probability, the empirical likelihood equation has a solution b_n which converges in probability to β and $\sqrt{n}\{b_n - \beta\}$ converges in distribution to a normal random variable with mean zero and variance $1/\{\sigma^2 J_{a_1,a_2}(f)\}$, where $\sigma^2 = \text{var}(X_1)$, a_1 and a_2 $(a_1 < a_2)$ are two arbitrarily fixed real numbers chosen in advance,

$$J_{a_1,a_2}(f) = \frac{\{f(a_1)\}^2}{F(a_1)} + \int_{a_1}^{a_2} \frac{\{f^{(1)}(x)\}^2}{f(x)} dx + \frac{\{f(a_2)\}^2}{1 - F(a_2)} \tag{1.8}$$

and F(.) is the distribution function of f(.). The quantity $J_{a_1,a_2}(f)$ has the property that although it is less than J(f) it converges to J(f) as a_1 and a_2 approach $-\infty$ and $+\infty$ respectively (see Samanta, 1974b).

In Chapter II we use kernel estimates of a probability density function and its derivatives to propose an estimate $\hat{b}_n$ of β and proceed to study its asymptotic distribution using the theory of a one-sample U-statistic. We show that the estimate $\hat{b}_n$ is asymptotically normally distributed with mean β and variance $1/\{n\sigma^2 J(f)\}$. Since $J_{a_1,a_2}(f) < J(f)$ the method proposed in Chapter II is an improvement upon a method developed by Samanta (1971).

1.4 Estimation of the Hodges-Lehmann Functional

It is well known that the functional $\Delta(f)= \int_{-\infty}^{\infty} f^2(x)dx$ appears in the expression for the asymptotic variance of the Hodges-Lehmann estimate of the shift parameter in the two sample location problem. The same functional also appears in the expressions for the asymptotic relative efficiencies of the rank tests for many problems like location shift, regression, dependence and analysis of variance (see Hodges and Lehmann, 1956, Hodges and Lehmann, 1963, Lehmann, 1963).

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variable having a pdf f(x). Bhattacharyya and Roussas (1969) proposed an estimate $\Delta_n(f)$ of the functional $\Delta(f)$ given by

$$\Delta_n(f)= \int_{-\infty}^{\infty} f_n^2(x)dx, \tag{1.9}$$

where $f_n(x)$ is the usual kernel estimator of f(x) as defined in (1.1). They established consistency in the quadratic mean of the estimate $\Delta_n(f)$. Dmitriev and Tarasenko (1973, 1974) proved the mean square convergence of another class of estimates of a general functional of the probability density function that includes $\Delta(f)$ as a special case.

In Chapter III we use the result in Schuster (1969) to study the rate of strong convergence of the estimate $\Delta_n(f)$. We then use the theory of U-statistic to establish the asymptotic normality of the estimate and derive explicit expressions for the mean square error (MSE) and MSE optimal smoothing parameter. Finally, we consider two numerical applications of these results: (a) the construction of a large sample confidence interval for the shift parameter in the two sample location problem and (b) the construction of a large sample confidence interval for the Hodges-Lehmann functional itself.

1.5 Nonparametric Estimation of a Probability Density Function and Its Functionals

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables having a probability density function f(x) and the distribution function F(x). Nadaraya (1964b) proposed and studied asymptotic properties of an estimator of the cumulative distribution function F(x) given by

$$F_n(x)=\frac{1}{n}\sum_{i=1}^{n}\Phi\left(\frac{x-X_i}{a_n}\right), \tag{1.10}$$

where $\Phi(t)=\int_{-\infty}^{t}\phi(u)du$ and ϕ is a kernel function defined on $(-\infty, \infty)$. Azzalini (1981) studied the second order properties of the estimator $F_n(x)$ in (1.10) together with the estimator, x_p, of the population quantile, ξ_p, that emerges as the solution of the equation $F_n(x)=p$.

In sections 4.2 through 4.4 of Chapter IV we derive, under certain regularity conditions, the asymptotic expressions for the mean square errors of the estimates of F(x), ξ_p and $f^{(i)}(x)$, i=0, 1, 2, ..., r. We then obtain expressions for the optimal smoothing parameter that minimize these mean square errors. Our results include the results in Azzalini (1981), Rosenblatt (1956) and Silverman (1986, p.70) as special cases. We also state the conditions under which we obtain the expression for the smoothing parameter that minimizes the mean square error of an estimate of the population mode.

1.6 Noparametric Estimation of the Conditional Probability Density Function and Its Functionals

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be independent and identically distributed two dimensional random variables with a joint probability density function f(x, y) and a joint distribution function $F(x, y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u, v)dudv$. The marginal density function of X and the conditional density function of Y given X=x are $g(x)=\int_{-\infty}^{\infty} f(x, y)dy$ and $f(y/x)=\frac{f(x,y)}{g(x)}$ respectively, provided $g(x)>0$. If E[Y] is finite, then the regression function of Y_1 on X_1 is defined as

$$r(x)=E[Y_1/X_1=x]. \tag{1.11}$$

The conditional distribution function of Y_1 given $X_1=x$ is defined by

$$F(y/x)=\int_{-\infty}^{y} f(u/x)du = \frac{\int_{-\infty}^{y} f(x, u)du}{g(x)}. \tag{1.12}$$

Let $\zeta_{p,x}$ denote the quantile of order p $(0<p<1)$ of the conditional distribution function F(y/x), i.e., a root of the equation $F(\zeta/x)=p$. We call $\zeta_{p,x}$ the population conditional quantile of order p and assume that $\zeta_{p,x}$ is unique. If we also assume that f(y/x) is uniformly continuous in y for each x, then the function f(y/x) is bounded and $\lim_{|y|\to\infty} f(y/x)=0$. Hence it follows that f(y/x) possesses a mode M(x) defined by

$$f(M(x)/x) = \max_{-\infty<y<\infty} f(y/x). \tag{1.13}$$

We call M(x) the population conditional mode and assume that M(x) is unique.

Let $f_n(x, y)$, $g_n(x)$, $f_n(y/x)$ and $F_n(y/x)$ be the kernel estimators of f(x, y), g(x), f(y/x) and F(y/x) respectively defined by

$$f_n(x,y)=\frac{1}{na_n^2}\sum_{j=1}^{n} \phi\left(\frac{x-X_j}{a_n}\right)\phi\left(\frac{y-Y_j}{a_n}\right) \tag{1.14}$$

$$g_n(x)=\frac{1}{na_n}\sum_{j=1}^{n}\phi\left(\frac{x-X_j}{a_n}\right) \tag{1.15}$$

$$f_n(y/x)=\frac{f_n(x, y)}{g_n(x)} \tag{1.16}$$

and

$$F_n(y/x)= \int_{-\infty}^{y} f_n(u/x)du = \frac{B_n(x, y)}{g_n(x)}, \tag{1.17}$$

where ϕ is a kernel function and $\{a_n\}$ is a sequence of positive numbers converging to zero, and

$$B_n(x, y)=\frac{1}{na_n}\sum_{j=1}^{n} \Phi\left(\frac{y-Y_j}{a_n}\right)\phi\left(\frac{x-X_j}{a_n}\right) \tag{1.18}$$

with $\Phi(y)= \int_{-\infty}^{y}\phi(u)du.$

Nadaraya (1964a), Watson (1964), Rosenblatt (1969) and Schuster (1972) have studied asymptotic properties of estimators of r(x) of the form

$$r_n(x)=\frac{\sum_{j=1}^{n} Y_j\phi\left(\frac{x-X_j}{a_n}\right)}{\sum_{j=1}^{n}\phi\left(\frac{x-X_j}{a_n}\right)}. \tag{1.19}$$

Rosenblatt (1969) also studied the properties of an estimate $f_n(y/x)$ of the conditional probability density function f(y/x). Samanta (1989) proved the strong consistency and asymptotic normality of an estimate $\overline{\zeta}_{p,x}$ of the conditional quantile $\zeta_{p,x}$ given by the root of the equation $F_n(\zeta/x)=p$. Samanta and Thavaneswaran (1990) also proved that under some regularity conditions an estimate $M_n(x)$ of the conditional mode M(x) given by $f_n(M_n(x)/x)= \max_{-\infty<y<\infty} f_n(y/x)$ is strongly consistent and asymptotically normally distributed.

In Chapter V, under some general conditions, we derive asymptotic expressions for the mean square errors of these estimates. We also obtain expressions for the smoothing parameter that minimize these mean square errors.

1.7 A General Concept of 'In Probability'

Suppose $\{Z_n\}$ is a sequence of random variables on $(-\infty, \infty)$ and $\{r_n\}$ is a sequence of positive numbers. It is customary to define o_p and O_p by adding probability requirements to the usual definitions of o and O as follows (see Pratt, 1959):

Definition 1. $Z_n=o_p(r_n)$ if, for every positive ε and η, for some N, for every n>N, $P\{|Z_n/r_n|\leq\eta\}\geq 1-\varepsilon$.

Definition 2. $Z_n=O_p(r_n)$ if, for every positive ε, for some positive η and N, for every n>N, $P\{|Z_n/r_n|\leq\eta\}\geq 1-\varepsilon$.

For an algebra of $o_p(1)$ and $O_p(1)$ we refer to Mann and Wald (1943) and Pratt (1959).

1.8 Outline of the Dissertation

In Chapter II we discuss the problem of estimating the slope of a linear regression function. In Chapter III we consider the problem of estimating the integral of the square of a probability density function (the Hodges-Lehmann functional). Chapter IV examines the problem of estimating a probability density function and its functionals including the derivatives, the mode, the distribution function, and the quantiles. Finally, in Chapter V we discuss the problem of estimating a conditional probability density function and its functionals such as the regression function, the conditional quantiles and the conditional mode. Unless otherwise stated all integrals in this dissertation will be understood to be Lebesgue integrals.

CHAPTER II

EFFICIENT NONPARAMETRIC ESTIMATION OF THE SLOPE OF A LINEAR REGRESSION

2.1 Introduction and Summary

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be independent and identically distributed two dimensional random variables such that $P[0 \leq X_i \leq 1] = 1$ for $i=1, 2, 3, \ldots$. Let $G(y/x)$ denote the conditional distribution function of Y given X=x and let us assume that for each x, $G(\cdot/x)$ is absolutely continuous having the conditional probability density $g(\cdot/x)$. We suppose further that there is a probability density function f and two real numbers α and β, where $-\infty < \alpha < \infty$, $|\beta| < \gamma$ and γ is a known constant such that

$$g(y/x) = f(y - \alpha - \beta x) \tag{2.1}$$

for all $0 \leq x \leq 1$ and $-\infty < y < \infty$. In such a case, if $E(Y/X)$ exists, then it is a linear function in X with slope β. Even if $E(Y/X)$ does not exist the conditional median (or any given quantile) of Y given X is a linear function in X with slope β. In this model, β may be called the slope of the linear regression of Y on X in a sense more general than usual.

In this dissertation we consider the problem of estimating and testing hypotheses about β. If f were known, then β could be estimated from $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ by the method of maximum likelihood by solving the pair of likelihood equations

$$\sum_{i=1}^{n} \frac{f^{(1)}(Y_i - a - bX_i)}{f(Y_i - a - bX_i)} = 0 \tag{2.2}$$

and

$$\sum_{i=1}^{n} X_i \frac{f^{(1)}(Y_i - a - bX_i)}{f(Y_i - a - bX_i)} = 0. \tag{2.3}$$

It is well known that, under certain regularity conditions (see Cramer, 1946, p. 500), the likelihood equations have a solution (a_n^*, b_n^*) converging in probability to (α, β).

Furthermore, $\sqrt{n}\,(b_n^*-\beta)$ converges in distribution to a normal random variable with mean zero and variance $1/\{\sigma^2 J(f)\}$, where $\sigma^2=\text{var}(X_1)$ and $J(f)$ is the Fisher information defined by

$$J(f)=\int_{-\infty}^{\infty}\frac{\{f^{(1)}(x)\}^2}{f(x)}dx. \tag{2.4}$$

The nonparametric counterpart of this problem arises when f is unknown. In this case we assume $\alpha=0$ without any loss of generality by writing $f(y)$ for $f(y-\alpha)$.

In this dissertation we propose a method for estimating β in the nonparametric setup by using kernel estimates of f and its derivatives (see Rosenblatt, 1956, Parzen, 1962, and Bhattacharya, 1967).

Let ϕ be a Borel measurable function and $\{a_n\}$ be a sequence of positive numbers converging to zero. We define the functions $\psi_n(x, y, b)$, $\Psi_n(x, y, b)$, $\psi_n^{(1)}(x, y, b)$ and $\Psi_n^{(1)}(x, y, b)$ by

$$\psi_n(x, y, b)=\frac{1}{na_n}\sum_{j=1}^{n}\phi\left(\frac{y-bx-Y_j+bX_j}{a_n}\right),$$

$$\Psi_n(x, y, b)=\frac{1}{n}\sum_{j=1}^{n}\Phi\left(\frac{y-bx-Y_j+bX_j}{a_n}\right), \tag{2.5}$$

$$\psi_n^{(1)}(x, y, b)=\frac{\partial}{\partial b}\psi_n(x, y, b),$$

and

$$\Psi_n^{(1)}(x, y, b)=\frac{\partial}{\partial b}\Psi_n(x, y, b)$$

respectively, where $\Phi(y)=\int_{-\infty}^{y}\phi(u)du$.

Let $\{d_n\}$ be a sequence of positive numbers converging to infinity and I_n be the indicator function of the interval $(-d_n, d_n)$. For any real number b and for any positive integer n, we define

$$L_n(b)=\frac{1}{n}\sum_{i=1}^{n}[I_n(Y_i)\frac{\psi_n^{(1)}(X_i, Y_i, b)}{\psi_n(X_i, Y_i, b)} + \Psi_n^{(1)}(X_i, -d_n, b) - \Psi_n^{(1)}(X_i, d_n, b)] \qquad (2.6)$$

and call $L_n(b)=0$ the empirical likelihood equation for estimating β. We briefly explain the motivation behind the function $L_n(b)$.

In the case when $\alpha=0$, if the function f were known, the maximum likelihood estimate of β on the basis of $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ is obtained by solving the likelihood equation

$$\frac{\partial}{\partial b}\sum_{i=1}^{n} \ln f(Y_i-bX_i)=0. \qquad (2.7)$$

Now, for every b, the conditional density function of $T_i=Y_i-bX_i$ given X_i is $f(t+bX_i-\beta X_i)$. Thus, when b is close to β the function

$$\frac{1}{na_n}\sum_{j=1}^{n}\phi\left(\frac{t-Y_j+bX_j}{a_n}\right) \qquad (2.8)$$

will closely approximate f(t) for large n. Since our object is to approximate the solution of (2.7) which is close to β with high probability when n is large, replacing f in (2.7) by its approximation (2.8) seems reasonable. With this replacement, the left hand side of (2.7) becomes

$$\sum_{i=1}^{n}\left\{\frac{\frac{1}{na_n^2}\sum_{j=1}^{n}(X_j-X_i)\,\phi^{(1)}\left(\frac{Y_i-bX_i-Y_j+bX_j}{a_n}\right)}{\frac{1}{na_n}\sum_{j=1}^{n}\phi\left(\frac{Y_i-bX_i-Y_j+bX_j}{a_n}\right)}\right\}. \qquad (2.9)$$

However, we modify the above function in order to study its convergence . For those Y_i's lying in the interval $(-d_n, d_n)$ we use the above form of the function and then add $\Psi_n^{(1)}(X_i, -d_n, b)$ and $-\Psi_n^{(1)}(X_i, d_n, b)$ to give rise to the function $L_n(b)$.

Let $\tilde{b}_n$ be any estimate of β such that $\tilde{b}_n-\beta=O_p(\frac{1}{\sqrt{n}})$. An estimate satisfying this condition can be obtained in the following way: We select λ_n pairs of equations from the n equations $Y_j=\alpha+\beta X_j$, j=1, 2, ..., n, where λ_n is the greatest integer less than or equal to $\frac{n}{2}$. We then obtain λ_n different estimates β_{jn}, j=1, 2, ..., λ_n of β by solving each pair of equations and take for $\tilde{b}_n$ the estimate of β defined by

$$\tilde{b}_n=\text{median}\{\beta_{jn},\ j=1, 2, ..., \lambda_n\}.$$

With $\tilde{b}_n$ as a first approximation for a solution of the equation $L_n(b)=0$ we get the next approximation $\hat{b}_n$ by going through exactly one iteration by the Newton-Raphson method, that is,

$$\hat{b}_n=\tilde{b}_n+\frac{L_n(\tilde{b}_n)}{-L_n^{(1)}(\tilde{b}_n)}. \tag{2.10}$$

In section 2.3 it is shown that under some regularity conditions $\sqrt{n}\,(\hat{b}_n-\beta)$ converges in distribution to a normal random variable with mean zero and variance $1/\{\sigma^2 J(f)\}$ (Theorem 2.1).

We now list the conditions on f, ϕ, the sequences $\{a_n\}$ and $\{d_n\}$ for which the asymptotic properties of the estimates proposed in this chapter are proved.

Condition 2.1: f and its first five derivatives exist and are bounded.

Condition 2.2: $J(f)=\int_{-\infty}^{\infty}\{\frac{d}{dx}ln\ f(x)\}^2 f(x)dx<\infty$.

Condition 2.3: There exist functions $M_1(x)$ and $M_2(x)$ integrable over $(-\infty,\infty)$ such that

$|f^{(1)}(x)|<M_1(x)$ and $|f^{(2)}(x)|<M_2(x)$ for all $x\in(-\infty,\infty)$.

Condition 2.4: There exists a strictly monotone increasing function H such that

$$\sup_{|x|\leq y} \frac{1}{f(x)} \leq H(y) \text{ for all y and } H(d_n+\gamma+2)=n^{1/25}.$$

Condition 2.5: ϕ and its first two derivatives are functions of bounded variation.

Condition 2.6: $\int_{-\infty}^{\infty}\phi(u)du=1,\ \int_{-\infty}^{\infty}u^r\phi(u)du=0,\ r=1, 2, 3, 4$ and

$$\int_{-\infty}^{\infty}|u^5\phi(u)|du<\infty.$$

Condition 2.7: $a_n=n^{-\delta},\ \frac{1}{8}<\delta<\frac{19}{150}.$

In the sequel, Conditions 2.1 through 2.7 will be referred to as Conditions A.2. Probability density functions which satisfy Conditions 2.1, 2.2 and 2.3 include among others the normal, the Cauchy, the contaminated normal density functions and the mixtures of the Cauchy and the normal probability density functions.

Condition 2.4 is similar to that in Bhattacharya (1967) and Dmitriev and Tarasenko (1973). If we let $H(y) =\sqrt{(2\pi)}\exp\left(\frac{y^2}{2}\right)$, then Condition 2.4 is satisfied if f is a standard normal density function or any other density function having flatter tail than the standard normal density function. An example of a kernel function satisfying Conditions 2.5 and 2.6 is

$$\phi(x)=\frac{15}{8}\left(1-\frac{2}{3}x^2+\frac{1}{15}x^4\right)\exp\left(\frac{-x^2}{2}\right) \quad \text{(see Nadaraya, 1989, p. 176).}$$

From Conditions 2.2 and 2.4 it follows that

$$J(f)= \int_{-\infty}^{\infty}\frac{\{f^{(1)}(x)\}^2}{f(x)}dx \geq \int_{-\infty}^{a}\frac{\{f^{(1)}(x)\}^2}{f(x)}dx$$

$$=F(a) \int_{-\infty}^{a}\left\{\frac{f^{(1)}(x)}{f(x)}\right\}^2\frac{f(x)}{F(a)}dx$$

$$\geq F(a)\left\{\int_{-\infty}^{a}\frac{f^{(1)}(x)}{f(x)}\frac{f(x)}{F(a)}dx\right\}^2$$

$$=\frac{\{f(a)\}^2}{F(a)}>0.$$

2.2 Convergence Properties of Kernel Estimates of a Density Function and Its Derivatives

Let $\{x_j\}$ be a bounded sequence of real numbers and $\{Z_j\}$ be a sequence of independent random variables with an unknown common distribution function F which is supposed to be absolutely continuous having a density function f. Let $\{c_j(x, x_1, x_2, ...)\}$ be a bounded sequence of real numbers depending on a real number x and the sequence $\{x_j\}$. In order to be specific we shall always consider $0 \leq x_j \leq 1$ and $|c_j| \leq 1$ for all j.

We consider the function

$$k_n(x, z, s)=\frac{1}{n}\sum_{j=1}^{n}c_j(x, x_1, x_2, ...)f(z+sx_j). \tag{2.11}$$

In the sequel we shall write c_j for $c_j(x, x_1, x_2, ...)$ but its dependence on x and $\{x_j\}$ should always be kept in mind.

By $k_n^{(r)}$ we denote the r-th partial derivative of k_n with respect to z, that is,

$$k_n^{(r)}(x, z, s)=\frac{1}{n}\sum_{j=1}^{n}c_j f^{(r)}(z+sx_j) \tag{2.12}$$

which exist and are bounded for r up to 5 by virtue of Condition 2.1. We shall now examine certain estimates of these functions based on $Z_1, Z_2, ...$

Let ϕ be a Borel measurable function and $\{a_n\}$ be a sequence of positive numbers converging to zero. We define for each positive integer n

$$h_n(x, z, s)=\frac{1}{na_n}\sum_{j=1}^{n}c_j\phi\left(\frac{z-Z_j+sx_j}{a_n}\right) \tag{2.13}$$

and

$$h_n^{(r)}(x, z, s)=\frac{1}{na_n^{r+1}}\sum_{j=1}^{n} c_j\phi^{(r)}\left(\frac{z-Z_j+sx_j}{a_n}\right) \tag{2.14}$$

We study the convergence of $h_n^{(r)}(x, z, s)$ to $k_n^{(r)}(x, z, s)$ in the following lemmas:

Lemma 2.1. Under Conditions 2.1, 2.5 and 2.6

$$\sup_{\substack{-\infty<z<\infty \\ -\infty<s<\infty}} | E\{h_n^{(r)}(x, z, s)\}-k_n^{(r)}(x, z, s)| \leq C\, a_n^{5-r}, \quad r=0, 1, 2$$

where C is a positive constant.

Proof. For r=0, by making a change of variable, we have

$$E\{h_n(x, z, s)\}=\frac{1}{n}\sum_{j=1}^{n} c_j \int_{-\infty}^{\infty} f(z+sx_j-a_nu)\phi(u)du. \tag{2.15}$$

For r=1, using a change of variable followed by integration by parts, we obtain

$$E\{\frac{1}{a_n^2}\phi^{(1)}\left(\frac{z-Z_j+sx_j}{a_n}\right)\}= \int_{-\infty}^{\infty}\frac{1}{a_n^2}\phi^{(1)}\left(\frac{z-u+sx_j}{a_n}\right)f(u)du$$

$$= \int_{-\infty}^{\infty}\frac{1}{a_n}f(z+sx_j-a_nu)\phi^{(1)}(u)du$$

$$=\frac{1}{a_n}f(z+sx_j-a_nu)\phi(u)]_{-\infty}^{\infty}+ \int_{-\infty}^{\infty}f^{(1)}(z+sx_j-a_nu)\phi(u)du.$$

Using Conditions 2.1 and 2.5, we get

$$\frac{1}{a_n}f(z+sx_j-a_nu)\phi(u)]_{-\infty}^{\infty}=\lim_{u\to\infty}[\frac{1}{a_n}f(z+sx_j-a_nu)\phi(u)-\frac{1}{a_n}f(z+sx_j)\phi(0)]$$

$$+\lim_{u\to-\infty}[\frac{1}{a_n}f(z+sx_j)\phi(0)-\frac{1}{a_n}f(z+sx_j-a_nu)\phi(u)]=0.$$

Hence

$$E\{\frac{1}{a_n^2}\phi^{(1)}\left(\frac{z-Z_j+sx_j}{a_n}\right)\}=\int_{-\infty}^{\infty} f^{(1)}(z+sx_j-a_nu)\phi(u)du$$

and

$$E\{h_n^{(1)}(x, z, s)\}=\frac{1}{n}\sum_{j=1}^{n} c_j \int_{-\infty}^{\infty} f^{(1)}(z+sx_j-a_nu)\phi(u)du.$$

In a similar manner using a change of variable with repeated applications of integration by parts in conjunction with Conditions 2.1 and 2.5, we get

$$E\{h_n^{(2)}(x, z, s)\}=\frac{1}{n}\sum_{j=1}^{n} c_j \int_{-\infty}^{\infty} f^{(2)}(z+sx_j-a_nu)\phi(u)du.$$

Hence, for r=1, 2, we have

$$E\{h_n^{(r)}(x, z, s)\}=\frac{1}{n}\sum_{j=1}^{n} c_j \int_{-\infty}^{\infty} f^{(r)}(z+sx_j-a_nu)\phi(u)du. \tag{2.16}$$

Equations (2.15) and (2.16) are summarized to give

$$E\{h_n^{(r)}(x, z, s)\}=\frac{1}{n}\sum_{j=1}^{n} c_j \int_{-\infty}^{\infty} f^{(r)}(z+sx_j-a_nu)\phi(u)du \tag{2.17}$$

for r=0, 1, 2.

In the integral on the right hand side of (2.17) we expand $f^{(r)}(z+sx_j-a_nu)$ around $z+sx_j$ to the order of a_n^{5-r} for r=0, 1, 2. The lemma now follows by virtue of Conditions 2.1 and 2.6 and the boundedness of the sequence $\{c_j\}$.

Suppose $\{l_{in}\}$, i=1, 2 be two sequences of positive numbers such that $\lim_{n\to\infty} l_{in}=\infty$, i=1,2.

Lemma 2.2. Under Conditions 2.1, 2.5 and 2.6 for sufficiently large n and for r=0, 1, 2, if $a_n^{5-r}=o(\varepsilon_n)$, then

$$P\{\sup_{\substack{|z|\le l_{1n}\\ |s|<l_{2n}}} | h_n^{(r)}(x, z, s) - k_n^{(r)}(x, z, s)| \ge\varepsilon_n\}\le\{C_1 l_{1n} l_{2n}/(a_n^{2r+4}\varepsilon_n^2)\}\exp\{-na_n^{2r+2}\varepsilon_n^2/C_2\},$$

where C_1 and C_2 are constants not depending on $\{x_j\}$ and $\{c_j\}$.

Proof. Consider any r between 0, 1, 2. Let M be an upper bound for $|f^{(r+1)}(u)|$ and $|\phi^{(r+1)}(u)|$ and $\delta_n = a_n^{r+2}\varepsilon_n/8M$

Then

$$|k_n^{(r)}(x, z_1, s_1) - k_n^{(r)}(x, z_2, s_2)| \le \frac{\varepsilon_n}{4}$$

and

$$|h_n^{(r)}(x, z_1, s_1) - h_n^{(r)}(x, z_2, s_2)| \le \frac{\varepsilon_n}{4}$$

whenever $|z_1 - z_2| < \delta_n$ and $|s_1 - s_2| < \delta_n$. Let us divide the interval $[-l_{tn}, l_{tn}]$, t=1, 2 into $g_n^{(t)} = \frac{2l_{tn}}{\delta_n}$ consecutive intervals denoted by $J_{n1}^{(t)}, J_{n2}^{(t)}, \ldots, J_{ng_n^{(t)}}^{(t)}$, and z_{nl} and $s_{nr'}$ be arbitrary points in $J_{nl}^{(1)}$ and $J_{nr'}^{(2)}$ respectively. Since for r=0, 1, 2, $a_n^{5-r} = o(\varepsilon_n)$, it follows from Lemma 2.1 and from the properties of these intervals that for sufficiently large n,

$$\sup_{\substack{|z|<l_{1n}\\|s|<l_{2n}}} |h_n^{(r)}(x, z, s) - k_n^{(r)}(x, z, s)|$$

$$= \max_{\substack{l=1,2,\ldots,g_n^{(1)}\\ r'=1,2,\ldots,g_n^{(2)}}} \sup_{(z,s)\in J_{nl}^{(1)}\times J_{nr'}^{(2)}} |\{h_n^{(r)}(x, z, s) - h_n^{(r)}(x, z_{nl}, s_{nr'})\}$$

$$+\{h_n^{(r)}(x, z_{nl}, s_{nr'}) - Eh_n^{(r)}(x, z_{nl}, s_{nr'})\} + \{Eh_n^{(r)}(x, z_{nl}, s_{nr'}) - k_n^{(r)}(x, z_{nl}, s_{nr'})\}$$

$$-\{k_n^{(r)}(x, z, s) - k_n^{(r)}(x, z_{nl}, s_{nr'})\}|$$

$$\le \frac{\varepsilon_n}{2} + Ca_n^{5-r} + \max_{\substack{l=1,2,\ldots,g_n^{(1)}\\ r'=1,2,\ldots,g_n^{(2)}}} \sup_{(z,s)\in J_{nl}^{(1)}\times J_{nr'}^{(2)}} |h_n^{(r)}(x, z_{nl}, s_{nr'}) - Eh_n^{(r)}(x, z_{nl}, s_{nr'})|,$$

where C does not depend on $\{x_j\}$ and $\{c_j\}$. Hence for all large n

$$P\{\sup_{\substack{|z|\le l_{1n}\\ |s|\le l_{2n}}} | h_n^{(r)}(x, z, s) - k_n^{(r)}(x, z, s)| \ge \varepsilon_n\}$$

$$\le P\{\max_{\substack{l=1, 2, ..., g_n^{(1)}\\ r'=1, 2, ..., g_n^{(2)}}} |h_n^{(r)}(x, z_{nl}, s_{nr'}) - Eh_n^{(r)}(x, z_{nl}, s_{nr'})| \ge \frac{\varepsilon_n}{4}\}$$

$$\le \sum_{l=1}^{g_n^{(1)}} \sum_{r'=1}^{g_n^{(2)}} P\Big\{ \Big| \sum_{j=1}^{n} c_j \phi^{(r)}\Big(\frac{z_{nl}-Z_j+s_{nr'}x_j}{a_n}\Big) - E\Big[\sum_{j=1}^{n} c_j \phi^{(r)}\Big(\frac{z_{nl}-Z_j+s_{nr'}x_j}{a_n}\Big)\Big]\Big| \ge \frac{na_n^{r+1}\varepsilon_n}{4}\Big\}.$$

By a theorem of Hoeffding (1963) concerning the tail probabilities of independent and bounded random variables, each term in the above sum is bounded by $2\exp\{-na_n^{2r+2}\varepsilon_n^2/32M^2\}$. This completes the proof of the lemma with $C_1=512M^2$ and $C_2=32M^2$.

Lemma 2.3. Suppose (X_1, Y_1), (X_2, Y_2), ..., are independent and identically distributed two dimensional random variables such that the conditional distribution function of Y_i given $X_i=x$ is absolutely continuous almost everywhere having probability density function $g(y/x)$ given by $g(y/x)=f(y-\beta x)$. Then

(a) $Z_i=Y_i-\beta X_i$, $i=1, 2, ...$, are independent and identically distributed absolutely continuous random variables with probability density function f.

(b) The random variables $X_1, X_2, ..., Z_1, Z_2, ...$, are all independent.

Remark. In the above lemma the absolute continuity of X_i has not been assumed, but $Z_i=Y_i-\beta X_i$ is still absolutely continuous.

The proof of Lemma 2.3 follows from standard properties of conditional distributions and is omitted.

We now note that since the bounds given in Lemma 2.2 are independent of $\{x_j\}$ and $\{c_j\}$, the same bounds remain valid if $x, x_1, x_2, ...$, are replaced by $X, X_1, X_2, ...$, provided that the collection of random variables $\{Z_i, i=1, 2, ...\}$ is independent of the collection $\{X, X_i, i=1, 2, ...\}$. We state this fact formally as:

Lemma 2.4. If $\{X, X_1, X_2, ...\}$ is a sequence of random variables independent of the sequence $\{Z_1, Z_2, ...\}$, Lemma 2.2 remains valid when $\{x, x_1, x_2, ...\}$ is replaced by $\{X, X_1, X_2, ...\}$.

In the definition of the functions k_n, $k_n^{(r)}$, h_n and $h_n^{(r)}$ given by (2.11) to (2.14) we now replace x by X, x_j by X_j for all j. The functions ψ_n, $\psi_n^{(1)}$, $\Psi_n^{(1)}$ introduced in (2.6) then become special cases of the functions $h_n^{(r)}$, r=0, 1 when the sequences $\{c_j\}$ are suitably chosen. These relations are summarized in the following table.

Table 2.1: Relations among $h_n^{(r)}$ and the components of $L_n(b)$

		$c_j(X, X_1, X_2, ...)$	
r	$h_n^{(r)}(X, Y\text{-}bX, b\text{-}\beta)$	1	$X_j\text{-}X$
0	$h_n(X, Y\text{-}bX, b\text{-}\beta)$	$\psi_n(X, Y, b)$	$\Psi_n^{(1)}(X, Y, b)$
1	$h_n^{(1)}(X, Y\text{-}bX, b\text{-}\beta)$		$\psi_n^{(1)}(X, Y, b)$

We consider the functions $S_n(x, y, b)$ and $s_n(x, y, b)$ given by

$$S_n(x, y, b)=\frac{1}{n}\sum_{j=1}^{n}F(y\text{-}bx+(b\text{-}\beta)X_j) \tag{2.18}$$

$$s_n(x, y, b)=\frac{1}{n}\sum_{j=1}^{n}f(y\text{-}bx+(b\text{-}\beta)X_j). \tag{2.19}$$

We define

$$s_n^{(1)}(x, y, b)\dot{=}\frac{\partial}{\partial b}s_n(x, y, b)=\frac{1}{n}\sum_{j=1}^{n}(X_j\text{-}x)f^{(1)}(y\text{-}bx+(b\text{-}\beta)X_j) \tag{2.20}$$

$$S_n^{(1)}(x, y, b)=\frac{\partial}{\partial b}S_n(x, y, b)=\frac{1}{n}\sum_{j=1}^{n}(X_j-x)f(y-bx+(b-\beta)X_j). \tag{2.21}$$

For any real number b ($|b|<\gamma$) we now consider the functions $L_n^*(b)$ and $L_n^{*(1)}(b)$ defined by

$$L_n^*(b)=\frac{1}{n}\sum_{i=1}^{n}I_n(Y_i)\frac{s_n^{(1)}(X_i, Y_i, b)}{s_n(X_i, Y_i, b)} \tag{2.22}$$

and

$$L_n^{*(1)}(b)=\frac{d}{db}L_n^*(b). \tag{2.23}$$

The functions s_n, $s_n^{(1)}$ and $S_n^{(1)}$ given in (2.19), (2.20) and (2.21) become special cases of the functions $k_n^{(r)}$, r=0, 1 when the sequences $\{c_j\}$ are suitably chosen. These relations are summarized in the following table

Table 2.2: Relations among $k_n^{(r)}$ and the components of $L_n^*(b)$

r	$k_n^{(r)}(X, Y-bX, b-\beta)$	$c_j(X, X_1, X_2, \ldots)$: 1	$c_j(X, X_1, X_2, \ldots)$: X_j-X
0	$k_n(X, Y-bX, b-\beta)$	$s_n(X, Y, b)$	$S_n^{(1)}(X, Y, b)$
1	$k_n^{(1)}(X, Y-bX, b-\beta)$		$s_n^{(1)}(X, Y, b)$

From Lemma 2.4 we obtain the following lemma on the convergence of $L_n(b)$ to $L_n^*(b)$ and of $L_n^{(1)}(b)$ to $L_n^{*(1)}(b)$.

Lemma 2.5. Under Conditions 2.1, 2.4, 2.5, 2.6 and 2.7 for every $\varepsilon>0$

(a) $\lim_{n\to\infty} P\{\sup_{|b-\beta|\leq 1}|L_n(b)-L_n^*(b)|>\varepsilon\}=0$

(b) $\lim_{n\to\infty} P\{\sup_{|b-\beta|\leq 1} |L_n^{(1)}(b) - L_n^{*(1)}(b)|>\varepsilon\}=0.$

Proof. We outline the proof of part (a). The proof of part (b) is similar but slightly longer and we omit it.

Now,

$$\sup_{|b-\beta|\leq 1} L_n(b) - L_n^*(b)|$$

$$\leq \sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |\Psi_n^{(1)}(X_i, -d_n, b)-S_n^{(1)}(X_i, -d_n, b)|$$

$$+\sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |S_n^{(1)}(X_i, -d_n, b)|$$

$$+\sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |\Psi_n^{(1)}(X_i, d_n, b)-S_n^{(1)}(X_i, d_n, b)|$$

$$+\sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |S_n^{(1)}(X_i, d_n, b)|$$

$$+\sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} I_n(Y_i)\left|\frac{\psi_n^{(1)}(X_i, Y_i, b)}{\psi_n(X_i, Y_i, b)} - \frac{s_n^{(1)}(X_i, Y_i, b)}{s_n(X_i, Y_i, b)}\right|.$$

Suppose ε is an arbitrary small positive number. We have

$$P\{\sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |\Psi_n^{(1)}(X_i, d_n, b)-S_n^{(1)}(X_i, d_n, b)|>\varepsilon\}$$

$$\leq nP\{\sup_{|b-\beta|\leq 1} |\Psi_n^{(1)}(X_i, d_n, b)-S_n^{(1)}(X_i, d_n, b)|>\varepsilon\}.$$

Using the relations between the functions $\Psi_n^{(1)}$, $S_n^{(1)}$ and h_n, k_n from tables 2.1 and 2.2 we get

$$\sup_{|b-\beta|\leq 1} |\Psi_n^{(1)}(X_i, d_n, b)-S_n^{(1)}(X_i, d_n, b)|$$

$$=\sup_{|b-\beta|\leq 1} |h_n(X_i, d_n-bX_i, b-\beta)-k_n(X_i, d_n-bX_i, b-\beta)|$$

$$\leq \sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} |h_n(X_i, z, s)-k_n(X_i, z, s)|.$$

Hence,

$$\lim_{n\to\infty} P[\sup_{|b-\beta|\leq 1} \{\frac{1}{n}\sum_{i=1}^{n} |\Psi_n^{(1)}(X_i, d_n, b)-S_n^{(1)}(X_i, d_n, b)|\}>\varepsilon]$$

$$\leq \lim_{n\to\infty} nP\{\sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} |h_n(X_i, z, s)-k_n(X_i, z, s)|>\varepsilon\}=0, \text{ by Lemma 2.4.}$$

Next, we have with probability one

$$\sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |S_n^{(1)}(X_i, d_n, b)| \leq \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \sup_{|b-\beta|\leq 1} f(d_n-bX_i+(b-\beta)X_j). \tag{2.24}$$

Since for all i, j =1, 2, ..., n

$$P\{\lim_{n\to\infty} \sup_{|b-\beta|\leq 1} f(d_n-bX_i+(b-\beta)X_j)\}=0\}=1,$$

we conclude that

$$P\{\lim_{n\to\infty} \sup_{|b-\beta|\leq 1} \frac{1}{n}\sum_{i=1}^{n} |S_n^{(1)}(X_i, d_n, b)|=0\}=1.$$

It now suffices to show that for every $\varepsilon>0$

$$\lim_{n\to\infty} nP\{\sup_{|b-\beta|\leq 1} I_n(Y_i)|\frac{\psi_n^{(1)}(X_i, Y_i, b)}{\psi_n(X_i, Y_i, b)} - \frac{s_n^{(1)}(X_i, Y_i, b)}{s_n(X_i, Y_i, b)}|>\varepsilon\}=0.$$

To complete the proof we recall the relations between the functions s_n, $s_n^{(1)}$, ψ_n, $\psi_n^{(1)}$ with the functions k_n, $k_n^{(1)}$, h_n, $h_n^{(1)}$ as in Tables 2.1 and 2.2 above and we show that for every $\varepsilon>0$

$$\lim_{n\to\infty} nP\{\sup_{|b-\beta|\leq 1} I_n(Y_i)|\frac{h_n^{(1)}(X_i, Y_i-\beta X_i, b-\beta)}{h_n(X_i, Y_i-bX_i, b-\beta)} - \frac{k_n^{(1)}(X_i, Y_i-\beta X_i, b-\beta)}{k_n(X_i, Y_i-\beta X_i, b-\beta)}|>\varepsilon\}$$

$$\leq \lim_{n\to\infty} nP\{\sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} \left|\frac{h_n^{(1)}(X_i, z, s)}{h_n(X_i, z, s)} - \frac{k_n^{(1)}(X_i, z, s)}{k_n(X_i, z, s)}\right| > \varepsilon\} = 0.$$

If $\sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} |h_n(X_i, z, s)-k_n(X_i, z, s)|<\varepsilon_n$,

$$\sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} |h_n^{(1)}(X_i, z, s)-k_n^{(1)}(X_i, z, s)|<\varepsilon_n,$$

$$\sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} \{\frac{1}{n}\sum_{j=1}^{n} f(z+sX_j)\}^{-1} \leq \frac{1}{n}\sum_{j=1}^{n} \sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} \frac{1}{f(z+sX_j)}$$

$$\leq H(d_n+\gamma+1),$$

and $\lim_{n\to\infty} \varepsilon_n\{H(d_n+\gamma+1)\}=0$, then for sufficiently large n

$$\sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} \left|\frac{h_n^{(1)}(X_i, z, s)}{h_n(X_i, z, s)} - \frac{k_n^{(1)}(X_i, z, s)}{k_n(X_i, z, s)}\right|$$

$$\leq \sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} \left|\frac{h_n^{(1)}(X_i, z, s)-k_n^{(1)}(X_i, z, s)}{k_n(X_i, z, s)}\right|$$

$$+ \sup_{\substack{|s|\leq 1 \\ |z|\leq d_n+\gamma}} \frac{|h_n^{(1)}(X_i, z, s)||h_n(X_i, z, s)-k_n(X_i, z, s)|}{|h_n(X_i, z, s)||k_n(X_i, z, s)|}$$

$$\leq C\varepsilon_n\{H(d_n+\gamma+1)\}^2,$$

where C is a positive constant.

If we define $\varepsilon_n = \dfrac{\varepsilon}{C\{H(d_n+\gamma+1)\}^2}$, then

$$\lim_{n\to\infty} \varepsilon_n\{H(d_n+\gamma+1)\} = \lim_{n\to\infty} \frac{\varepsilon}{C\{H(d_n+\gamma+1)\}} = 0.$$

Hence, for all sufficiently large n,

$$P\{\sup_{\substack{|s|\le 1\\ |z|\le d_n+\gamma}} \left|\frac{h_n^{(1)}(X_i, z, s)}{h_n(X_i, z, s)} - \frac{k_n^{(1)}(X_i, z, s)}{k_n(X_i, z, s)}\right| > \varepsilon\}$$

$$\le P\{\sup_{\substack{|s|\le 1\\ |z|\le d_n+\gamma}} |h_n(X_i, z, s) - k_n(X_i, z, s)| \ge \frac{\varepsilon}{C\{H(d_n+\gamma+1)\}^2}\}$$

$$+ P\{\sup_{\substack{|s|\le 1\\ |z|\le d_n+\gamma}} |h_n^{(1)}(X_i, z, s) - k_n^{(1)}(X_i, z, s)| \ge \frac{\varepsilon}{C\{H(d_n+\gamma+1)\}^2}\}.$$

The proof is now completed by application of Lemma 2.4 and the hypothesis.

2.3 Asymptotic Efficiency of the Nonparametric Estimator of the Slope of a Linear Regression

We first prove the following lemmas:

Lemma 2.6. Under Conditions 2.2 and 2.3 we have

(a) $E\{X_i^r \dfrac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\}=0$, r=0, 1, 2, ...

(b) $E\{X_i^r \dfrac{f^{(2)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\}=0$, r=0, 1, 2, ...

(c) $E\{X_i^r [\dfrac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}]^2\}=J(f)E(X_i^r)$, r=0, 1, 2, ...

Proof. We note that by Lemma 2.3, the conditional density function of $Z_i=Y_i-\beta X_i$ given X_i is f. This implies that

$$E\{\frac{f^{(r)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}/ X_i\}=0 \text{ a.e, } r=1, 2.$$

and

$$E\{[\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}]^2/ X_i\}=J(f) \text{ a.e.}$$

Hence the lemma.

Lemma 2.7. Under Conditions 2.2 and 2.3

(a) $L_n^*(\beta)=o_p(1)$

(b) $L_n^{*(1)}(\beta)=-\sigma^2 J(f)+o_p(1)$.

Proof. (a) Define

$$\xi_n=\frac{1}{n}\sum_{i=1}^{n}\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}(\bar{X}_n-X_i)$$

$$=\bar{X}_n\frac{1}{n}\sum_{i=1}^{n}\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)} - \frac{1}{n}\sum_{i=1}^{n}X_i\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}.$$

By Khintchine's theorem (see e.g. Rao, 1973, p. 112) we have

$$\xi_n=\{E(X_1)+o_p(1)\}\{o_p(1)\}-o_p(1)=o_p(1).$$

Now, $E|L_n^*(\beta)-\xi_n|$

$$=E|\frac{1}{n}\sum_{i=1}^{n}I_n(Y_i)\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}(\bar{X}_n-X_i) - \frac{1}{n}\sum_{i=1}^{n}\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}(\bar{X}_n-X_i)|$$

$$\leq E\{\frac{1}{n}\sum_{i=1}^{n}|I_n(Y_i)-1|\frac{|f^{(1)}(Y_i-\beta X_i)|}{|f(Y_i-\beta X_i)|}|\bar{X}_n-X_i|\}$$

$$\leq E\{\frac{1}{n}\sum_{i=1}^{n}|I_n(Y_i)-1|\frac{|f^{(1)}(Y_i-\beta X_i)|}{|f(Y_i-\beta X_i)|}\}$$

$$\leq \int_{-\infty}^{-d_n+\gamma}|f^{(1)}(z)|dz + \int_{d_n-\gamma}^{\infty}|f^{(1)}(z)|dz.$$

By Condition 2.3 the right hand side of the above inequality tends to zero as n tends to infinity. Therefore, using the above computations, we get

$L_n^*(\beta)= \xi_n+o_p(1)=o_p(1)+o_p(1)=o_p(1).$

(b) Define

$$\eta_n=\frac{1}{n}\sum_{i=1}^{n}\{\frac{f^{(2)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\frac{1}{n}\sum_{j=1}^{n}(X_j-X_i)^2 - [\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}]^2(\bar{X}_n-X_i)^2\}$$

$$=\frac{1}{n}\sum_{i=1}^{n}X_i^2\{\frac{f^{(2)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\} + \frac{1}{n}\sum_{j=1}^{n}X_j^2\frac{1}{n}\sum_{i=1}^{n}\frac{f^{(2)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}$$

$$- 2\bar{X}_n\{\frac{1}{n}\sum_{i=1}^{n}X_i\frac{f^{(2)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\} - \bar{X}_n^2\frac{1}{n}\sum_{i=1}^{n}\{\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\}^2$$

$$-\frac{1}{n}\sum_{i=1}^{n}X_i^2\{\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\}^2 + 2\bar{X}_n[\frac{1}{n}\sum_{i=1}^{n}X_i\{\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)}\}^2].$$

By repeated applications of Khintchine's theorem in conjuction with results of Lemma 2.6, we get

$$\eta_n= o_p(1)+ \{E(X_1^2)+o_p(1)\}\{o_p(1)\} - 2\{E(X_1)+o_p(1)\}\{o_p(1)\}$$

$$- \{E(X_1)+o_p(1)\}^2\{J(f)+o_p(1)\} - \{E(X_1^2)J(f)+o_p(1)\}$$

$$+2\{E(X_1)+o_p(1)\}\{E(X_1)J(f)+o_p(1)\}$$

$$= -\sigma^2 J(f) + o_p(1).$$

Now,

$$E|L_n^{*(1)}(\beta) - \eta_n|$$

$$\leq E\{\frac{1}{n}\sum_{i=1}^{n} |I_n(Yi)-1| [\frac{|f^{(2)}(Y_i-\beta X_i)|}{|f(Y_i-\beta X_i)|} + (\frac{f^{(1)}(Y_i-\beta X_i)}{f(Y_i-\beta X_i)})^2]\}$$

$$\leq \int_{-\infty}^{-d_n+\gamma} |f^{(2)}(z)|dz + \int_{d_n-\gamma}^{\infty} |f^{(2)}(z)|dz + \int_{-\infty}^{-d_n+\gamma} \frac{\{f^{(1)}(z)\}^2}{f(z)} dz + \int_{d_n-\gamma}^{\infty} \frac{\{f^{(1)}(z)\}^2}{f(z)} dz.$$

By Conditions 2.2 and 2.3 the right hand side of this inequality tends to zero as n tends to infinity. Hence, using the above computations, we get

$$L_n^{*(1)}(\beta) = \eta_n + o_p(1) = -\sigma^2 J(f) + o_p(1).$$

Lemma 2.8. Let $\{A_i, i=1, 2, ...\}$ be a sequence of events such that $P(A_i)=1$ for each i, then $P(\bigcap_{i=1}^{\infty} A_i)=1$.

Proof. The proof is obvious.

Lemma 2.9. Under Conditions 2.1 and 2.4 there exists a positive constant C' not depending on n such that

$$P\{|L_n^{*(1)}(b) - L_n^{*(1)}(\beta)| \leq 2C'|b-\beta|[H(d_n+\gamma+2)]^4\}=1$$

for all $|b-\beta| \leq 1$.

Proof. We first note that

$$\sup_{\substack{|y| \leq d_n \\ 0 \leq x,\, x_1,\, x_2,\, ...,\, x_n \leq 1}} \{\frac{1}{n} \cdot \sum_{j=1}^{n} f(y-bx+(b-\beta)x_j)\}^{-1}$$

$$\le \frac{1}{n}\sum_{j=1}^{n} \sup_{\substack{|y|\le d_n \\ 0\le x,\, x_j\le 1 \\ |b-\beta|\le 1}} \frac{1}{f(y-bx+(b-\beta)x_j)}$$

$$\le H(d_n+\gamma+2).$$

Then, with probability one, we have

$$|L_n^{*(1)}(b) - L_n^{*(1)}(\beta)| \le \frac{1}{n}\sum_{i=1}^{n}(Q_{1in}+Q_{2in}),$$

where

$$Q_{1in}= I_n(Y_i)\left|\frac{s_n^{(2)}(X_i, Y_i, b)}{s_n(X_i, Y_i, b)} - \frac{s_n^{(2)}(X_i, Y_i, \beta)}{s_n(X_i, Y_i, \beta)}\right|$$

$$Q_{2in}= I_n(Y_i)\left|\left\{\frac{s_n^{(1)}(X_i, Y_i, b)}{s_n(X_i, Y_i, b)}\right\}^2 - \left\{\frac{s_n^{(1)}(X_i, Y_i, \beta)}{s_n(X_i, Y_i, \beta)}\right\}^2\right|.$$

With probability one, we have for $|b-\beta|\le 1$ and for any i,

$$Q_{1in}\le\{H(d_n+\gamma+2)\}^2 \sup_{|y|\le d_n} |s_n(X_i, y, \beta)s_n^{(2)}(X_i, y, b)-s_n^{(2)}(X_i, y, \beta)s_n(X_i, y, b)|$$

$$\le\{H(d_n+\gamma+2)\}^2 \sup_{|y|\le d_n} \{|s_n(X_i, y, \beta)||s_n^{(2)}(X_i, y, b)-s_n^{(2)}(X_i, y, \beta)|$$

$$+|s_n^{(2)}(X_i, y, \beta)||s_n(X_i, y, b)-s_n(X_i, y, \beta)|\}$$

$$\le M\{H(d_n+\gamma+2)\}^2 \sup_{|y|\le d_n} \{|s_n^{(2)}(X_i, y, b)-s_n^{(2)}(X_i, y, \beta)|+|s_n(X_i, y, b)-s_n(X_i, y, \beta)|\}$$

$$\le C_1\{H(d_n+\gamma+2)\}^2 |b-\beta|,$$

where C_1 is a positive constant. Similarly, with probability one, we have for $|b-\beta|\le 1$ and for any i,

$$Q_{2in}\le\{H(d_n+\gamma+2)\}^4 \sup_{|y|\le d_n} |\{s_n(X_i, y, \beta)s_n^{(1)}(X_i, y, b)\}^2$$

$$-\{s_n^{(1)}(X_i, y, \beta)s_n(X_i, y, b)\}^2|$$

$$\leq\{H(d_n+\gamma+2)\}^4M^2 \sup_{|y|\leq d_n} [|\{s_n^{(1)}(X_i, y, b)\}^2 - \{s_n^{(1)}(X_i, y, \beta)\}^2|$$

$$+|\{s_n(X_i, y, b)\}^2 - \{s_n(X_i, y, \beta)\}^2|]$$

$$\leq C_2\{H(d_n+\gamma+2)\}^4 |b-\beta|,$$

where C_2 is a positive constant.

Define $C'=\max(C_1, C_2)$. Hence for $|b-\beta|\leq 1$

$$P\{|L_n^{*(1)}(b) - L_n^{*(1)}(\beta)|\leq 2C'|b-\beta|[H(d_n+\gamma+2)]^4\}$$

$$\geq P\{\frac{1}{n}\sum_{i=1}^{n}(Q_{1in}+Q_{2in})\leq 2C'|b-\beta|[H(d_n+\gamma+2)]^4\}$$

$$\geq P\{Q_{1in}\leq C'|b-\beta|[H(d_n+\gamma+2)]^2,\ Q_{2in}\leq C'|b-\beta|[H(d_n+\gamma+2)]^4,\ i=1, 2, ..., n\}=1,$$

by Lemma 2.8.

Lemma 2.10. Under Conditions 2.1 and 2.4, for arbitrary $0<\delta<1$, there exist a positive constant C not depending on n such that

$$P\{\sup_{n^{4/25}|b-\beta|\leq\delta} |L_n^{*(1)}(b) - L_n^{*(1)}(\beta)|\leq C\delta\}=1.$$

Proof. Let C' be a constant for which Lemma 2.9 holds and define C=2C'. Let $\{b_v\}$ be a sequence of rational numbers such that

$$\{H(d_n+\gamma+2)\}^4|b_v-\beta|=n^{4/25}|b_v-\beta|\leq\delta.$$

Since $L_n^{*(1)}(b)$ is a continuous function in b for every sample sequence, it follows from Lemmas 2.8 and 2.9 that

$$P\{\sup_{n^{4/25}|b-\beta|\le\delta} |L_n^{*(1)}(b) - L_n^{*(1)}(\beta)|\le C\delta\}$$

$$= \quad P\{\sup_{n^{4/25}|b-\beta|\le\delta \text{ ,b rational}} |L_n^{*(1)}(b) - L_n^{*(1)}(\beta)|\le C\delta\}$$

$$\ge P\{|L_n^{*(1)}(b_v) - L_n^{*(1)}(\beta)|\le 2C'[H(d_n+\gamma+2)]^4 \,|b_v-\beta|,\ v=1, 2, \ldots\}$$

$$=1.$$

Lemma 2.11. Under Conditions 2.1 and 2.4, if $\{\beta_n\}$ is a sequence of random variables such that $n^{4/25}(\beta_n-\beta) = o_p(1)$, then

$$L_n^{*(1)}(\beta_n)=L_n^{*(1)}(\beta)+o_p(1).$$

Proof. The proof follows at once from Lemma 2.10 and the hypothesis.

Lemma 2.12. Under Conditions 2.1, 2.4, 2.5, 2.6 and 2.7, if $\{\beta_n\}$ is a sequence of random variables such that $\beta_n-\beta = o_p(1)$, then

$$L_n^{(1)}(\beta_n)=L_n^{*(1)}(\beta_n)+o_p(1).$$

Proof. We have for any $\varepsilon>0$

$$\lim_{n\to\infty} P\{|L_n^{(1)}(\beta_n)-L_n^{*(1)}(\beta_n)|<\varepsilon\}$$

$$\ge \lim_{n\to\infty} P\{[|L_n^{(1)}(\beta_n)-L_n^{*(1)}(\beta_n)|<\varepsilon]\cap[|\beta_n-\beta|\le 1]\}$$

$$\ge \lim_{n\to\infty} P\{[\sup_{|b-\beta|\le 1} |L_n^{(1)}(\beta_n)-L_n^{*(1)}(\beta_n)|<\varepsilon]\cap[|\beta_n-\beta|\le 1]\}$$

$=1$, by Lemma 2.5(b) and the hypothesis.

Lemma 2.13. Under Conditions A.2, if $\{\beta_n\}$ is a sequence of random variables such that $n^{4/25}(\beta_n-\beta) = o_p(1)$, then

$$L_n^{(1)}(\beta_n) = -\sigma^2 J(f)+o_p(1).$$

Proof. The proof follows at once since

$$\begin{aligned} L_n^{(1)}(\beta_n) &= L_n^{*(1)}(\beta_n)+o_p(1), \text{ by Lemma 2.12} \\ &= L_n^{*(1)}(\beta)+o_p(1)+o_p(1), \text{ by Lemma 2.11} \\ &= -\sigma^2 J(f)+o_p(1)+o_p(1), \text{ by Lemma 2.7} \\ &= -\sigma^2 J(f)+o_p(1). \end{aligned}$$

From (2.10) we get

$$\sqrt{n}\,(\hat{b}_n-\beta) = \sqrt{n}\,(\tilde{b}_n-\beta)+\frac{\sqrt{n}L_n(\tilde{b}_n)}{-L_n^{(1)}(\tilde{b}_n)}$$

$$=\sqrt{n}\,(\tilde{b}_n-\beta)+\frac{\sqrt{n}\,L_n(\beta)+\sqrt{n}\,(\tilde{b}_n-\beta)L_n^{(1)}(b_n')}{-L_n^{(1)}(\tilde{b}_n)}, \quad (2.25)$$

where $|b_n'-\beta| < |\tilde{b}_n-\beta|$. Since $\tilde{b}_n-\beta = O_p(\frac{1}{\sqrt{n}})$, we conclude that

$$n^{4/25}|b_n'-\beta| \leq n^{4/25}|\tilde{b}_n-\beta| = o_p(1).$$

Hence, by Lemma 2.13,

$$-L_n^{(1)}(b_n') = \sigma^2 J(f)+o_p(1) \quad (2.26)$$

and

$$-L_n^{(1)}(\tilde{b}_n) = \sigma^2 J(f)+o_p(1). \quad (2.27)$$

Using the results (2.26) and (2.27) in (2.25), we get

$$\sqrt{n}\,(\hat{b}_n-\beta) = \frac{\sqrt{n}L_n(\beta)}{\sigma^2 J(f)+o_p(1)} + \sqrt{n}\,(\tilde{b}_n-\beta)\{1 - \frac{\sigma^2 J(f)+o_p(1)}{\sigma^2 J(f)+o_p(1)}\}$$

$$= \frac{\sqrt{n}L_n(\beta)}{\sigma^2 J(f) + o_p(1)} + O_p(1)o_p(1)$$

$$= \frac{\sqrt{n}L_n(\beta)}{\sigma^2 J(f) + o_p(1)} + o_p(1). \qquad (2.28)$$

It is now clear from (2.28) that the large sample distribution of $\sqrt{n}\ (\hat{b}_n - \beta)$ depends on the large sample distribution of $\sqrt{n}\ L_n(\beta)$.

In the remaining part of this section we represent $\sqrt{n}\ L_n(\beta)$ as the sum of a function of a one sample U-statistic and a random variable which converges in probability to zero. We therefore first study the asymptotic properties of a one sample U-statistic.

For each n let $g_n((x_1, z_1), ..., (x_r, z_r))$ be a Borel measurable function of r real vector variables $(x_1, z_1), (x_2, z_2), ..., (x_r, z_r)$, where r is a given positive integer less than or equal to n.

We define

$$U_n((X_1, Z_1), ..., (X_n, Z_n)) = \frac{1}{n_{(r)}} \sum_P g_n((X_{i_1}, Z_{i_1}), ..., (X_{i_r}, Z_{i_r})), \qquad (2.29)$$

where

$n_{(r)} = \frac{n!}{(n-r)!}$ and the summation $\sum_P$ is over all permutations $(i_1, ..., i_r)$ of r distinct integers selected from (1, 2, ..., n). Let us define for each n another Borel measurable function $g_n^*((x_1, z_1), ..., (x_r, z_r))$ defined by

$g_n^*((x_1, z_1), ..., (x_r, z_r))$

$$= \frac{1}{r!} \sum_P g_n((x_{i_1}, z_{i_1}), ..., (x_{i_r}, z_{i_r})), \qquad (2.30)$$

where the summation $\sum_P$ is over all permutations $(i_1, ..., i_r)$ of the integers (1, 2, ..., r).

The random variable $U_n((X_1, Z_1), ..., (X_n, Z_n))$ defined by (2.29) can be written as

$U_n((X_1, Z_1), ..., (X_n, Z_n))$

$$=\frac{1}{\binom{n}{r}}\sum_{C} g_n^*((X_{i_1}, Z_{i_1}), \ldots, (X_{i_r}, Z_{i_r})), \qquad (2.31)$$

where $\binom{n}{r}=\frac{n!}{r!(n-r)!}$ and the summation $\sum_{C}$ extends over all combinations $(i_1, \ldots, i_r)$ of r distinct integers chosen from $(1, 2, \ldots, n)$.

In the following we shall write U_n for $U_n((X_1, Z_1), \ldots, (X_n, Z_n))$. The statistic U_n defined by (2.29) is called a one-sample U-statistic first studied by Hoeffding (1948), where the function g_n does not depend on n.

In the following we adopt the notation as used by Fraser (1957) but keep in mind the dependence of each expression on n.

Let us assume that $E[g_n^*((X_1, Z_1), \ldots, (X_r, Z_r))]^2$ exists for each n. This implies that $E[g_n^*((X_1, Z_1), \ldots, (X_r, Z_r))]$ also exists for each n and put

$$\theta_n=E[g_n^*((X_1, Z_1), \ldots, (X_r, Z_r))]. \qquad (2.32)$$

We then have

$$E[U_n]= \theta_n. \qquad (2.33)$$

We define

$$h_n((x_1, z_1), \ldots, (x_r, z_r))= g_n^*((x_1, z_1), \ldots, (x_r, z_r))-\theta_n \qquad (2.34)$$

and for each $c=0, 1, \ldots, r$ let $h_{n,c}((x_1, z_1), \ldots, (x_c, z_c))$ be the conditional expectation of $h_n((X_1, Z_1), \ldots, (X_r, Z_r))$ given $(X_1, Z_1)=(x_1, z_1), \ldots, (X_c, Z_c)=(x_c, z_c)$.

Since the random variables under consideration are all independent, we have

$$h_{n,c}((x_1, z_1), \ldots, (x_c, z_c))$$

$$=E[h_n((x_1, z_1), \ldots, (x_c, z_c), (X_{c+1}, Z_{c+1}), \ldots, (X_r, Z_r))]. \qquad (2.35)$$

We define

$$\zeta_{n,c}=E[h_{n,c}((X_1, Z_1), \ldots, (X_c, Z_c))]^2. \qquad (2.36)$$

Then

$$\zeta_{n,0} = 0 \tag{2.37}$$

and

$$\zeta_{n,r} = \mathrm{var}[g_n^*((X_1, Z_1), \ldots, (X_r, Z_r))]. \tag{2.38}$$

We assume the following:

Condition B.1: $\zeta_1 = \lim_{n\to\infty} \zeta_{n,1}$ exists.

Condition B.2: $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_{n,1}((X_i, Z_i))$ converges in distribution to a normal random variable with mean 0 and variance ζ_1.

The asymptotic distribution of U_n when g_n does not depend on n was obtained by Hoeffding (1948). The following result on the asymptotic distribution of U_n when the kernel g_n depends on n is a direct generalization of Hoeffding's result and its proof is omitted.

Lemma 2.14. (a) If Condition B.1 holds, and if for all c with the exception of c=1, $\zeta_{n,c} = o(n^{c-1})$, then

$$\mathrm{var}[\sqrt{n}U_n] = r^2\zeta_1 + o(1).$$

(b) If further, $\zeta_1 > 0$ and Condition B.2 holds, then $\sqrt{n}(U_n - \theta_n)$ converges in distribution to a normal random variable with mean 0 and variance $r^2\zeta_1$.

We now consider a one sample U-statistic $U_n((X_1, Z_1), \ldots, (X_n, Z_n))$ with r=3 and defined by

$$U_n((X_1, Z_1), \ldots, (X_n, Z_n))$$

$$= \frac{1}{n(n-1)(n-2)} \sum_{i\neq j\neq k=1}^{n} g_n((X_i, Z_i), (X_j, Z_j), (X_k, Z_k)), \tag{2.39}$$

where

$$g_n((x_1, z_1), (x_2, z_2), (x_3, z_3))$$

$$=I_n(z_1+\beta x_1)\frac{(x_2-x_1)}{f(z_1)}\frac{1}{a_n^2}\phi^{(1)}\left(\frac{z_1-z_2}{a_n}\right)\{2-\frac{\frac{1}{a_n}\phi\left(\frac{z_1-z_3}{a_n}\right)}{f(z_1)}\}$$

$$+(x_2-x_1)\frac{1}{a_n}\phi\left(\frac{-d_n-\beta x_1-z_2}{a_n}\right)-(x_2-x_1)\frac{1}{a_n}\phi\left(\frac{d_n-\beta x_1-z_2}{a_n}\right)$$

$$=I_{0n}((x_1, z_1), (x_2, z_2), (x_3, z_3))+I_{1n}((x_1, z_1), (x_2, z_2), (x_3, z_3))$$

$$-I_{2n}((x_1, z_1), (x_2, z_2), (x_3, z_3)), \qquad (2.40)$$

where

$I_{0n}((x_1, z_1), (x_2, z_2), (x_3, z_3))$

$$=I_n(z_1+\beta x_1)\frac{(x_2-x_1)}{f(z_1)}\frac{1}{a_n^2}\phi^{(1)}\left(\frac{z_1-z_2}{a_n}\right)\{2-\frac{\frac{1}{a_n}\phi\left(\frac{z_1-z_3}{a_n}\right)}{f(z_1)}\} \qquad (2.41)$$

$I_{1n}((x_1, z_1), (x_2, z_2), (x_3, z_3))$

$$=(x_2-x_1)\frac{1}{a_n}\phi\left(\frac{-d_n-\beta x_1-z_2}{a_n}\right) \qquad (2.42)$$

and

$I_{2n}((x_1, z_1), (x_2, z_2), (x_3, z_3))$

$$=(x_2-x_1)\frac{1}{a_n}\phi\left(\frac{d_n-\beta x_1-z_2}{a_n}\right) \qquad (2.43)$$

As explained before we can always write $U_n((X_1, Z_1), ..., (X_n, Z_n))$ in terms of the random variables $g_n^*((X_i, Z_i), (X_j, Z_j), (X_k, Z_k))$, i, j, k=1, 2, ..., n, $i \neq j \neq k$

where

$g_n^*((x_1, z_1), (x_2, z_2), (x_3, z_3))$

$$=\frac{1}{6}[g_n((x_1, z_1), (x_2, z_2), (x_3, z_3))+g_n((x_1, z_1), (x_3, z_3), (z_2, z_2))$$

$$+g_n((x_2, z_2), (x_1, z_1), (x_3, z_3))+g_n((x_2, z_2), (x_3, z_3), (x_1, z_1))$$

$$+g_n((x_3, z_3), (x_1, z_1), (x_2, z_2))+g_n((x_3, z_3), (x_2, z_2), (x_1, z_1))]. \qquad (2.44)$$

Hence

$$U_n=\frac{1}{\binom{n}{3}}\sum_{1\le i<j<k\le n} g_n^*((X_i, Z_i), (X_j, Z_j), (X_k, Z_k)). \qquad (2.45)$$

We now prove a lemma:

Lemma 2.15. Under Conditions 2.1, 2.4, 2.5 and 2.6, if $0\le x_1, x_2, x_3\le 1$, then

$$|Eg_n((x_1, Z_1), (x_2, Z_2), (x_3, Z_3))|=O[a_n^4+a_n^5\{H(d_n+\gamma)\}].$$

Proof. By Fubini's theorem (see e.g. Loeve, 1963, p. 136) and Lemma 2.1

$E[g_n((x_1, Z_1), (x_2, Z_2), (x_3, Z_3))]$

$$=(x_2-x_1)\{f(-d_n-\beta x_1)+O(a_n^5)\}-(x_2-x_1)\{f(d_n-\beta x_1)+O(a_n^5)\}$$

$$+\int_{-d_n-\beta x_1}^{d_n-\beta x_1}(x_2-x_1)\{\int_{-\infty}^{\infty} f^{(1)}(z_1-a_n u)\phi(u)du\}\{2-\frac{f(z_1)+O(a_n^5)}{f(z_1)}\}dz_1$$

$$=(x_2-x_1)\{f(-d_n-\beta x_1)+O(a_n^5)\}-(x_2-x_1)\{f(d_n-\beta x_1)+O(a_n^5)\}$$

$$+\int_{-d_n-\beta x_1}^{d_n-\beta x_1}(x_2-x_1)\{f^{(1)}(z_1)+a_n^4K_n(z_1)\}\{1+\frac{O(a_n^5)}{f(z_1)}\}dz_1, \qquad (2.46)$$

where

$$K_n(z_1)=\frac{1}{4!}\int_{-\infty}^{\infty} u^4 f^{(5)}(z_1-\alpha a_n u)\phi(u)du, \quad 0<\alpha<1.$$

We note that

$$\int_{-d_n-\beta x_1}^{d_n-\beta x_1}(x_2-x_1)\{f^{(1)}(z_1)\}\frac{O(a_n^5)}{f(z_1)}dz_1=O[a_n^5\{H(d_n+\gamma)\}].$$

Also, we have by interchanging the order of integration,

$$\left|\int_{-d_n-\beta x_1}^{d_n-\beta x_1}\int_{-\infty}^{\infty}u^4 f^{(5)}(z_1-\alpha a_n u)\phi(u)du\, dz_1\right|$$

$$=\left|\int_{-\infty}^{\infty}[f^{(4)}(d_n-\beta x_1-\alpha a_n u) - f^{(4)}(-d_n-\beta x_1-\alpha a_n u)]u^4\phi(u)du\right|<\infty.$$

The proof now follows from the above computations.

Corollary. Under the conditions of Lemma 2.15

$$|Eg_n((X_1, Z_1), (X_2, Z_2), (X_3, Z_3))|=O[a_n^4+a_n^5\{H(d_n+\gamma)\}].$$

Lemma 2.16. Under Conditions 2.1, 2.4, 2.5 and 2.6

$$\theta_n=O[a_n^4+a_n^5\{H(d_n+\gamma)\}].$$

Proof. The proof follows from the corollary of Lemma 2.15 and the definition of θ_n.

Corollary. If further, Condition 2.7 is satisfied, then

$$\sqrt{n}\theta_n=o(1).$$

Lemma 2.17. Under Conditions 2.1, 2.4, 2.5 and 2.6, for arbitrary real number z_1 and real numbers x_i with $0 \le x_i \le 1$, i=1, 2, 3, there exists a positive constant C independent of x_1, x_2, x_3 and z_1 such that the following are true:

(a) $$\left|E[I_{0n}((x_1, z_1), (x_2, Z_2), (x_3, Z_3))] - I_n(z_1+\beta x_1)\frac{f^{(1)}(z_1)}{f(z_1)}(x_2-x_1)\right|$$

$$\le C[a_n^4\{H(d_n+\gamma)\}+a_n^5\{H(d_n+\gamma)\}^2]$$

(b) $$\left|E[I_{0n}((x_1, Z_1), (x_2, z_2), (x_3, Z_3))] - (x_2-x_1)\{\frac{1}{a_n}\phi\left(\frac{d_n-\beta x_1-z_2}{a_n}\right) - \frac{1}{a_n}\phi\left(\frac{-d_n-\beta x_1-z_2}{a_n}\right)\}\right|$$

$$\le C\,[a_n^4\{H(d_n+\gamma)\}]$$

(c) $$|E[I_{0n}((x_1, Z_1), (x_2, Z_2), (x_3, z_3))] - 2(x_2-x_1)\{f(d_n-\beta x_1)-f(-d_n-\beta x_1)\}$$

$$+(x_2-x_1)\{\frac{1}{a_n}\int_{-d_n-\beta x_1}^{d_n-\beta x_1}\frac{f^{(1)}(u)}{f(u)}\phi\left(\frac{u-z_3}{a_n}\right)du\}| \le C\,[a_n^4\{H(d_n+\gamma)\}]$$

(d) $|E[I_{1n}((x_1, z_1), (x_2, Z_2), (x_3, z_3))]-(x_2-x_1)f(-d_n-\beta x_1)| \le Ca_n^5$

(e) $|E[I_{2n}((x_1, z_1), (x_2, Z_2), (x_3, z_3))]-(x_2-x_1)f(d_n-\beta x_1)| \le Ca_n^5$.

Proof. We shall prove three of the above statements. The remaining statements can be similarly proved.

(a) By Fubini's theorem and Lemma 2.1, we have

$E[I_{0n}((x_1, z_1), (x_2, Z_2), (x_3, Z_3))]$

$$=E[I_n(z_1+\beta x_1)\frac{(x_2-x_1)}{f(z_1)}\frac{1}{a_n^2}\phi^{(1)}\left(\frac{z_1-Z_2}{a_n}\right)\{2 - \frac{\frac{1}{a_n}\phi\left(\frac{z_1-Z_3}{a_n}\right)}{f(z_1)}\}]$$

$$=I_n(z_1+\beta x_1)\frac{(x_2-x_1)}{f(z_1)}\{f^{(1)}(z_1)+O(a_n^4)\}\{1+\frac{O(a_n^5)}{f(z_1)}\}. \quad (2.47)$$

The proof of part (a) now follows from (2.47).

(b) By Fubini's theorem and Lemma 2.1 we have

$E[I_{0n}((x_1, Z_1), (x_2, z_2), (x_3, Z_3))]$

$$=E[I_n(Z_1+\beta x_1)\frac{(x_2-x_1)}{f(Z_1)}\frac{1}{a_n^2}\phi^{(1)}\left(\frac{Z_1-z_2}{a_n}\right)\{2-\frac{\frac{1}{a_n}\phi\left(\frac{Z_1-Z_3}{a_n}\right)}{f(Z_1)}\}]$$

$$=\int_{-d_n-\beta x_1}^{d_n-\beta x_1}(x_2-x_1)\frac{1}{a_n^2}\phi^{(1)}\left(\frac{u-z_2}{a_n}\right)\{1+\frac{O(a_n^5)}{f(u)}\}du$$

$$=(x_2-x_1)[\frac{1}{a_n}\phi\left(\frac{d_n-\beta x_1-z_2}{a_n}\right)-\frac{1}{a_n}\phi\left(\frac{-d_n-\beta x_1-z_2}{a_n}\right)]+O[a_n^4\{H(d_n+\gamma)\}]. \quad (2.48)$$

The proof of part (b) now follows from (2.48).

(e) By Lemma 2.1, we have

$E[I_{2n}((x_1, z_1), (x_2, Z_2), (x_3, z_3))]$

$$=E[(x_2-x_1)[\frac{1}{a_n}\phi\left(\frac{d_n-\beta x_1-Z_2}{a_n}\right)]$$

$$=(x_2-x_1)f(d_n-\beta x_1)+O(a_n^5). \quad (2.49)$$

The proof of part (e) now follows from (2.49).

We define $E(X_1)=m$.

For each n we define the functions $U_{in}(x, z)$, i=1, 2 and $r_n(z)$ by

$$U_{1n}(x, z)=I_n(z+\beta x)\frac{f^{(1)}(z)}{f(z)}$$

$$U_{2n}(x, z)= \int_{-d_n-\beta x_1}^{d_n-\beta x_1} \frac{f^{(1)}(u)}{f(u)} \frac{1}{a_n} \phi\left(\frac{u-z}{a_n}\right) du$$

and

$$r_n(z)=E\{(X_1-E(X_1))U_{2n}(X_1, z)\}.$$

Clearly, $\lim_{n\to\infty} U_{1n}(x, z)=\frac{f^{(1)}(z)}{f(z)}$ for all $0\le x\le 1$ and $-\infty<z<\infty$.

Lemma 2.18. Under Conditions 2.1, 2.4, 2.5 and 2.6, for arbitrary real number z and for real number x $0\le x\le 1$,

$$|h_{n,1}(x, z)+\frac{1}{3}(x-E(x))U_{1n}(x, z)+\frac{1}{3}r_n(z)|$$

$$\le C[a_n^4\{H(d_n+\gamma)\}+a_n^5\{H(d_n+\gamma)\}^2].$$

Proof. Using Lemmas 2.16 and 2.17 we obtain

$$|Eh_n((x, z), (x_2, Z_2), (x_3, Z_3))+\frac{1}{6}(2x-x_2-x_3)U_{1n}(x, z)+\frac{1}{6}(x_2-x_3)U_{2n}(x_2, z)$$

$$+\frac{1}{6}(x_3-x_2)U_{2n}(x_3, z)| \le C[a_n^4\{H(d_n+\gamma)\}+a_n^5\{H(d_n+\gamma)\}^2], \tag{2.50}$$

where C is a positive constant independent of x, x_2, x_3 and z. In (2.50) taking expectations with respect to X_2 and X_3 we have the desired conclusion.

Lemma 2.19. Under Conditions 2.1, 2.4 and 2.6, there exists a constant C independent of n such that

$$|U_{2n}(x, z) - \frac{f^{(1)}(z)}{f(z)} \int_{\frac{-d_n-\beta x-z}{a_n}}^{\frac{d_n-\beta x-z}{a_n}} \phi(t)dt| \le Ca_n\{H(d_n+\gamma)\}^2$$

for all $0 \le x \le 1$ and $-\infty < z < \infty$ such that $-d_n < z + \beta x < d_n$.

Proof. We have

$$U_{2n}(x, z) = \frac{1}{a_n} \int_{-d_n - \beta x_1}^{d_n - \beta x_1} \frac{f^{(1)}(u)}{f(u)} \phi\left(\frac{u-z}{a_n}\right) du.$$

Put $t = \frac{u-z}{a_n}$. Then

$$U_{2n}(x, z) = \int_{\frac{-d_n - \beta x - z}{a_n}}^{\frac{d_n - \beta x - z}{a_n}} \frac{f^{(1)}(z+a_n t)}{f(z+a_n t)} \phi(t) dt$$

$$= \int_0^{\frac{d_n - \beta x - z}{a_n}} \frac{f^{(1)}(z+a_n t)}{f(z+a_n t)} \phi(t) dt + \int_{\frac{-d_n - \beta x - z}{a_n}}^{0} \frac{f^{(1)}(z+a_n t)}{f(z+a_n t)} \phi(t) dt.$$

Now, expanding $\frac{f^{(1)}(z+a_n t)}{f(z+a_n t)}$ around z, we get

$$\int_0^{\frac{d_n - \beta x - z}{a_n}} \frac{f^{(1)}(z+a_n t)}{f(z+a_n t)} \phi(t) dt = \frac{f^{(1)}(z)}{f(z)} \int_0^{\frac{d_n - \beta x - z}{a_n}} \phi(t) dt$$

$$+ a_n \int_0^{\frac{d_n - \beta x - z}{a_n}} \left[\frac{f^{(2)}(z+\theta a_n t)}{f(z+\theta a_n t)} - \left\{\frac{f^{(1)}(z+\theta a_n t)}{f(z+\theta a_n t)}\right\}^2\right] t\, \phi(t) dt, \qquad (2.51)$$

where $0<\theta<1$. Since $0<t<\frac{d_n-\beta x-z}{a_n}$, $0<\theta<1$, and $-d_n<z+\beta x<d_n$ we have

$$-d_n-\beta x<z+\theta a_n t<d_n-\beta x.$$

Hence the second term in the right hand side of (2.51) is less than $Ca_n\{H(d_n+\gamma)\}^2$, where C is a positive constant. It can be shown in a similar manner that

$$\left| \int_{\frac{-d_n-\beta x-z}{a_n}}^{0} \frac{f^{(1)}(z+a_n t)}{f(z+a_n t)} \phi(t)dt - \frac{f^{(1)}(z)}{f(z)} \int_{\frac{-d_n-\beta x-z}{a_n}}^{0} \phi(t)dt \right| \leq Ca_n\{H(d_n+\gamma)\}^2,$$

where C is a positive constant. The desired conclusion now follows from the above computations.

Corollary. Under Conditions 2.1, 2.4, 2.6 and 2.7

$$\lim_{n\to\infty} U_{2n}(x, z)=\frac{f^{(1)}(z)}{f(z)} \text{ for all } 0\leq x\leq 1 \text{ and } -\infty<z<\infty.$$

The proof follows immediately from Lemma 2.19, since for any z and x, where $-\infty<z<\infty$ and $0\leq x\leq 1$, n can be chosen sufficiently large to satisfy the condition $-d_n<z+\beta x<d_n$ and the limits of integration in Lemma 2.19 $\frac{d_n-\beta x-z}{a_n}$ and $\frac{-d_n-\beta x-z}{a_n}$ approach $+\infty$ and $-\infty$ respectively as n tends to infinity.

Lemma 2.20. Under Conditions 2.1, 2.4, 2.6 and 2.7

$$\lim_{n\to\infty} r_n(z)=0 \text{ for all } -\infty<z<\infty.$$

Proof. Suppose $z=z_0$, where $-\infty<z_0<\infty$. By Lemma 2.19 and its Corollary

$$\lim_{n\to\infty} \{(x-m)U_{2n}(x, z_0)\}=(x-m)\frac{f^{(1)}(z_0)}{f(z_0)} \text{ for all } 0\leq x\leq 1$$

and

$|(x-m)U_{2n}(x, z_0)| \leq C|\frac{f^{(1)}(z_0)}{f(z_0)}|+1<\infty$ for all $0 \leq x \leq 1$ and for sufficiently large n. By an application of Lebesgue dominated convergence theorem (see e.g. Loeve, 1963, p. 125), we get

$$\lim_{n\to\infty} r_n(z_0)=\lim_{n\to\infty} E\{(X_1-m)U_{2n}(X_1, z_0)\}$$

$$=E[(X_1-m)\frac{f^{(1)}(z_0)}{f(z_0)}]=0. \quad (2.52)$$

Lemma 2.21. Under Conditions 2.1, 2.2, 2.3, 2.4, 2.6 and 2.7

(a) $\lim_{n\to\infty} \zeta_{n,1}=\frac{1}{9}\sigma^2 J(f)$

(b) $\lim_{n\to\infty} E[(X_1-m)\frac{f^{(1)}(Z_1)}{f(Z_1)}h_{n,1}(X_1, Z_1)]= - \frac{1}{3}\sigma^2 J(f)$

(c) $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{n,1}(X_i, Z_i)$ converges in distribution to a normal random variable with mean 0 and variance $\frac{1}{9}\sigma^2 J(f)$.

Proof. (a) Using Lemmas 2.18, 2.19 and 2.20, we first note that for all sufficiently large n, $|h_{n,1}(x, z)|^2$ is bounded above by $\{C|\frac{f^{(1)}(z)}{f(z)}|+1\}^2$ for all $0 \leq x \leq 1$ and for all $-\infty<z<\infty$, where C is a positive constant. We then note that

$$\lim_{n\to\infty} \{h_{n,1}(x, z)\}^2 = \frac{1}{9}(x-m)^2 \{\frac{f^{(1)}(z)}{f(z)}\}^2. \quad (2.53)$$

By an application of Lebesgue dominated convergence theorem, we get

$$\lim_{n\to\infty} \zeta_{n,1}=\lim_{n\to\infty} E\{h_{n,1}(X_1, Z_1)\}^2$$

$$=\frac{1}{9}E\{(X_1-m)^2 \frac{f^{(1)}(Z_1)}{f(Z_1)}\}^2$$

$$=\frac{1}{9}\sigma^2 J(f) \quad (2.54)$$

which completes the proof of part (a) of the lemma.

(b) As in part (a), using Lemmas 2.18, 2.19 and 2.20, we note that for all sufficiently large n,

$|(x-m)\frac{f^{(1)}(z)}{f(z)} h_{n,1}(x, z)|$ is bounded above by

$C|\frac{f^{(1)}(z)}{f(z)}|^2 + |\frac{f^{(1)}(z)}{f(z)}|$ for all $0 \le x \le 1$ and for all $-\infty < z < \infty$, where C is a positive constant. We then note that

$$\lim_{n\to\infty}\{(x-m)\frac{f^{(1)}(z)}{f(z)}h_{n,1}(x, z)\} = -\frac{1}{3}(x-m)^2\{\frac{f^{(1)}(z)}{f(z)}\}^2. \tag{2.55}$$

By an application of Lebesgue dominated convergence theorem, we get

$$\lim_{n\to\infty} E\{(X_1-m)\frac{f^{(1)}(Z_1)}{f(Z_1)}h_{n,1}(X_1, Z_1)\} = E[-\frac{1}{3}(X_1-m)^2\{\frac{f^{(1)}(Z_1)}{f(Z_1)}\}^2]$$

$$= -\frac{1}{3}\sigma^2 J(f). \tag{2.56}$$

(c) We have

$$E[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{n,1}(X_i, Z_i) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{3}(X_i-m)\frac{f^{(1)}(Z_i)}{f(Z_i)}]^2$$

$$= E[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{n,1}(X_i, Z_i)]^2 + \frac{1}{9}E[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i-m)\frac{f^{(1)}(Z_i)}{f(Z_i)}]^2$$

$$+ \frac{2}{3}[\frac{1}{n}\sum_{i=1}^{n} E(X_i-m)\frac{f^{(1)}(Z_i)}{f(Z_i)} h_{n,1}(X_i, Z_i)]$$

$$= \zeta_{n,1} + \frac{1}{9}\sigma^2 J(f) + \frac{2}{3} E\{(X_1-m)\frac{f^{(1)}(Z_1)}{f(Z_1)} h_{n,1}(X_1, Z_1)\}. \tag{2.57}$$

Using the results in part (a) and part (b), we get

$$\lim_{n\to\infty} E[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{n,1}(X_i, Z_i)+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{3}(X_i-m)\frac{f^{(1)}(Z_i)}{f(Z_i)}]^2$$

$$=\frac{1}{9}\sigma^2 J(f) +\frac{1}{9}\sigma^2 J(f) - \frac{2}{9}\sigma^2 J(f) =0. \quad (2.58)$$

We conclude that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{n,1}(X_i, Z_i) = -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{3}(X_i-m)\frac{f^{(1)}(Z_i)}{f(Z_i)} + o_p(1). \quad (2.59)$$

An application of central limit theorem completes the proof.

Lemma 2.22. Under Conditions 2.1, 2.4, 2.5, 2.6 and 2.7

$\zeta_{n,c}=O(n^{23/25})$ for c=2, 3.

Proof. By direct computation it can be shown that for all $0\le x_i\le 1$ and for all $-\infty<z_i<\infty$, i=1, 2, 3

$$g_n((x_1, z_1), (x_2, z_2), (x_3, z_3)) = O[\frac{1}{a_n^3}\{H(d_n+\gamma)\}^2]. \quad (2.60)$$

Hence

$$h_{n,c}((x_1, z_1), ..., (x_c, z_c))=O[\frac{1}{a_n^3}\{H(d_n+\gamma)\}^2] \quad (2.61)$$

for c=1, 2, 3. This implies that

$$\zeta_{n,c}=E\{h_{n,c}((X_1, Z_1), ..., (X_c, Z_c))\}^2$$

$$=O[\frac{1}{a_n^6}\{H(d_n+\gamma)\}^4]$$

$$=O(n^{23/25}). \quad (2.62)$$

Lemma 2.23. Under Conditions A.2, $\sqrt{n}U_n$ converges in distribution to a normal random variable with mean 0 and variance $\sigma^2 J(f)$.

Proof. By Lemmas 2.21 and 2.22, the hypotheses of Lemma 2.14 are satisfied. Hence, $\sqrt{n}(U_n-\theta_n)$ converges in distribution to a normal random variable with mean 0 and variance $9\zeta_1$. By Corollary of Lemma 2.16 and part (a) of Lemma 2.21 the proof now follows.

We note that

$$L_n(\beta)=\frac{1}{n}\sum_{i=1}^{n}[I_n(Y_i)\frac{\psi_n^{(1)}(X_i, Y_i, \beta)}{\psi_n(X_i, Y_i, \beta)} + \Psi_n^{(1)}(X_i, -d_n, \beta) - \Psi_n^{(1)}(X_i, d_n, \beta)]. \quad (2.63)$$

We define another random variable by

$$V_n=V_n((X_1, Y_1), ..., (X_n, Y_n))$$

$$=\frac{1}{n}\sum_{i=1}^{n}[I_n(Y_i)\frac{\psi_n^{(1)}(X_i, Y_i, \beta)}{f(Y_i-\beta X_i)}(2-\frac{\psi_n(X_i, Y_i, \beta)}{f(Y_i-\beta X_i)})$$

$$+ \Psi_n^{(1)}(X_i, -d_n, \beta) - \Psi_n^{(1)}(X_i, d_n, \beta)]. \quad (2.64)$$

Consider

$$L_n(\beta)-V_n$$

$$=\frac{1}{n}\sum_{i=1}^{n}I_n(Y_i)\frac{\psi_n^{(1)}(X_i, Y_i, \beta)}{\psi_n(X_i, Y_i, \beta)}\frac{\{\psi_n(X_i, Y_i, \beta)-f(Y_i-\beta X_i)\}^2}{\{f(Y_i-\beta X_i)\}^2}. \quad (2.65)$$

Lemma 2.24. Under Conditions A.2, for every $\varepsilon>0$,

$$\lim_{n\to\infty} P[\sqrt{n}|L_n(\beta)-V_n|>\varepsilon]=0.$$

Proof. We proceed along the line of proof of part (a) of Lemma 2.5. In view of relation (2.65) and relations between the functions s_n, $s_n^{(1)}$, ψ_n, $\psi_n^{(1)}$ with the functions k_n, $k_n^{(1)}$, h_n, $h_n^{(1)}$ as indicated in Tables 2.1 and 2.2 it suffices to show that

$$\lim_{n\to\infty} nP\left\{\sup_{|z|\le d_n+\gamma}\frac{|h_n^{(1)}(X_i, z, 0)|^{1/2}[n^{1/4}\{h_n(X_i, z, 0) - k_n(X_i, z, 0)\}]}{|h_n(X_i, z, 0)|^{1/2}\{k_n(X_i, z, 0)\}}>\sqrt{\varepsilon}\right\}=0.$$

Now, if $\sup_{|z|\le d_n+\gamma} |h_n^{(1)}(X_i, z, 0) - k_n^{(1)}(X_i, z, 0)|<1$,

$$\sup_{|z|\le d_n+\gamma} |h_n(X_i, z, 0) - k_n(X_i, z, 0)|<\varepsilon_n,$$

$$\sup_{|z|\le d_n+\gamma} \frac{1}{f(z)}\le H(d_n+\gamma)$$

and

$$\lim_{n\to\infty} \varepsilon_n H(d_n+\gamma)=0,$$

then

$$\sup_{|z|\le d_n+\gamma} \frac{|h_n^{(1)}(X_i, z, 0)|^{1/2}[n^{14}|h_n(X_i, z, 0) - k_n(X_i, z, 0)|]}{|h_n(X_i, z, 0)|^{1/2}|k_n(X_i, z, 0)|}$$

$$\le C\varepsilon_n\{H(d_n+\gamma)\}^{3/2}n^{1/4}, \qquad (2.66)$$

where C is a positive constant.

The proof can now be completed in a similar manner as in part (a) of Lemma 2.5 and using the results in Lemma 2.4.

We define another random variable $W_n=W_n((X_1, Y_1), ..., (X_n, Y_n))$ by

$$W_n=\frac{1}{n^3}[\sum_{i=1}^{n} I_{0n}((X_i, Z_i), (X_i, Z_i), (X_i, Z_i))+\sum_{i\neq j=1}^{n}\{I_{0n}((X_i, Z_i), (X_i, Z_i), (X_j, Z_j))$$

$$+I_{0n}((X_i, Z_i), (X_j, Z_j), (X_i, Z_i))+I_{0n}((X_i, Z_i), (X_j, Z_j), (X_j, Z_j))\}], \qquad (2.67)$$

where I_{0n} is as defined earlier. We have the following lemma on the random variable W_n:

Lemma 2.25. Under Conditions A.2

$$\sqrt{n}W_n=o_p(1).$$

Proof. We outline the proof. By definition we have

$$\frac{1}{n^{5/2}}\sum_{i=1}^{n} I_{0n}((X_i, Z_i), (X_i, Z_i), (X_i, Z_i)) \equiv 0 \tag{2.68}$$

and

$$\frac{1}{n^{5/2}}\sum_{i\neq j=1}^{n} I_{0n}((X_i, Z_i), (X_i, Z_i), (X_j, Z_j)) \equiv 0. \tag{2.69}$$

Using results in the proof of Lemma 2.22, it immediately follows that

$$\frac{1}{n^{5/2}}\sum_{i\neq j=1}^{n}\{I_{0n}((X_i, Z_i), (X_j, Z_j), (X_i, Z_i))+I_{0n}((X_i, Z_i), (X_j, Z_j), (X_j, Z_j))\}$$

$$=O_p[\frac{1}{n^{1/2}a_n^3}\{H(d_n+\gamma)\}^2]$$

$$=O_p(n^{-1/25})=o_p(1). \tag{2.70}$$

This completes the proof of the lemma.

Lemma 2.26. Under Conditions A.2, $\sqrt{n}L_n(\beta)$ converges in distribution to a normal random variable with mean 0 and variance $\sigma^2 J(f)$.

Proof. We define two random variables U_{1n}^* and U_{2n}^* by

$$U_{1n}^* = U_{1n}^*((X_1, Y_1), ..., (X_n, Y_n))$$

$$=\frac{1}{n(n-1)(n-2)}\sum_{i\neq j\neq k=1}^{n}\{I_{0n}((X_i, Z_i), (X_j, Z_j), (X_k, Z_k))\}$$

and

$$U_{2n}^* = U_{2n}^*((X_1, Y_1), ..., (X_n, Y_n))$$

$$=\frac{1}{n(n-1)}\sum_{i\neq j=1}^{n}\{I_{1n}((X_i, Z_i), (X_j, Z_j), (X_k, Z_k))-I_{2n}((X_i, Z_i), (X_j, Z_j), (X_k, Z_k))\},$$

where I_{0n}, I_{1n}, and I_{2n} are as defined in (2.41), (2.42) and (2.43) respectively.

Then, $U_{1n}^* + U_{2n}^* = U_n$ and it follows from a straight forward computation that

$$\frac{1}{\sqrt{n}}U_{2n}^{*}=O_p\left(\frac{1}{a_n\sqrt{n}}\right)=o_p(1). \tag{2.71}$$

Using the above results we have

$$\sqrt{n}L_n(\beta)=\sqrt{n}V_n+o_p(1), \text{ by Lemma 2.24}$$

$$=\sqrt{n}\left[\frac{(n-1)(n-2)}{n^2}U_{1n}^{*}+\frac{(n-1)}{n}U_{2n}^{*}+W_n\right]+o_p(1)$$

$$=\sqrt{n}\left[\frac{(n-1)(n-2)}{n^2}\{U_n-U_{2n}^{*}\}+\frac{(n-1)}{n}U_{2n}^{*}\right]+o_p(1)+o_p(1), \text{ by Lemma 2.25}$$

$$=\sqrt{n}\left[\{U_n(1+o(1))+\frac{2(n-1)}{n^2}U_{2n}^{*}\right]+o_p(1)$$

$$=\sqrt{n}U_n+\sqrt{n}U_n.o(1)+\frac{2(n-1)}{n}\frac{U_{2n}^{*}}{\sqrt{n}}+o_p(1)$$

$$=\sqrt{n}U_n+O_p(1).o(1)+2(1+o(1)).o_p(1)+o_p(1)$$

$$=\sqrt{n}U_n+o_p(1), \text{ by Lemma 2.23.}$$

Another application of Lemma 2.23 completes the proof.

We now prove the main theorem of this chapter.

Theorem 2.1. Under Conditions A.2, the estimate $\hat{b}_n$ is asymptotically normally distributed with mean β and variance $1/\{n\sigma^2 J(f)\}$.

Proof. The proof follows at once from relation (2.28) in conjunction with Lemma 2.26.

2.4 Miscellaneous Remarks

(i) Choice of the sequence $\{d_n\}$.

The choice of the sequence $\{d_n\}$ requires knowledge of the tails of the density function. If the density function has flatter tails than the standard normal probability density function, then d_n may be chosen to be equal to $\sqrt{\{\frac{2}{25}\ln n\}}$ for sufficiently large n. The density functions which posses flatter tails than the standard normal density function include among others the Cauchy density function, the mixture of the standard normal and the Cauchy density functions, and the density function of the random variable t with degrees of freedom $1<v<\infty$.

(ii) Estimation of J(f).

Following Bhattacharya (1967), we propose the following estimate of J(f) given by

$$\hat{J}(f)=\int_{-d_n}^{d_n}\frac{\{f_n^{(1)}(x)\}^2}{f_n(x)}dx.$$

The consistency of $\hat{J}(f)$ can be proved in a similar manner as in Dmitriev and Tarasenko (1973)

(iii) Large sample tests of hypotheses about β.

Suppose we want to test the null hypothesis H_o: $\beta=\beta_o$ against the alternative H_a: $\beta\neq\beta_o$. We have shown that under H_o, $\sqrt{n}\ (\hat{b}_n-\beta_o)$ is asymptotically normally distributed with mean 0 and variance $1/\{\sigma^2 J(f)\}$. Using a consistent estimate $\hat{\sigma}$ of σ and a consistent estimate $\hat{J}(f)$ of J(f), we now define the statistic $Z_n=\{n\hat{J}(f)\}^{1/2}\,\hat{\sigma}\ \{\hat{b}_n-\beta_o\}$ which is asymptotically distributed as a standard normal random variable when H_o is true. Z_n can be used as a test statistic for testing H_o against the alternative H_a and the critical region $|Z_n|\geq R^{-1}(1-\frac{\alpha}{2})$ will have an approximate level of significance α for large n, where R is the standard normal distribution function. In a similar manner we can construct a large sample confidence interval for the unknown slope β for a prescribed confidence coefficient. The statistic Z_n can also be used for testing the hypothesis H_o: $\beta\leq\beta_o$ (or $\beta\geq\beta_o$) against H_a: $\beta>\beta_o$ (or $\beta<\beta_o$). It can be shown that the asymptotic efficiency (ARE) of these tests relative to the test based on the maximum likelihood estimate is 1.

CHAPTER III

NONPARAMETRIC ESTIMATION OF THE HODGES-LEHMANN FUNCTIONAL

3.1 Introduction

Suppose $X_1, X_2, ..., X_n$ are independent and identically distributed random variables with a pdf f(x). The usual kernel estimator of f(x) is $f_n(x)$ defined by $f_n(x)= \frac{1}{na_n}\sum_{j=1}^{n}\phi\left(\frac{x-X_j}{a_n}\right)$, where ϕ is a probability density function and $\{a_n\}$ is a sequence of positive numbers converging to zero as n tends to infinity. Bhattacharyya and Roussas (1969) first considered the estimation of the functional $\Delta(f)= \int_{-\infty}^{\infty} f^2(x)dx$ and have shown that under certain conditions the estimate $\Delta_n(f)= \int_{-\infty}^{\infty} f_n^2(x)dx$ is consistent for $\Delta(f)$ in the quadratic mean. In this dissertation we show that, under some regularity conditions, the estimate $\Delta_n(f)$ is strongly consistent and asymptotically normally distributed. We obtain an expression for the smoothing parameter a_n for which the mean square error $E\{\Delta_n(f)-\Delta(f)\}^2$ is minimized and consider some applications of the asymptotic results.

3.2 Strong Consistency and Asymptotic Normality of $\Delta_n(f)$

We first prove the following theorem:

Theorem 3.1. Suppose ϕ is a function of bounded variation with

$\int_{-\infty}^{\infty} t\,\phi(t)dt=0$ and $\int_{-\infty}^{\infty} t^2\phi(t)dt<\infty$. Further, suppose that the first two derivatives of f exist and $f^{(2)}$ is a bounded function. If the sequences $\{a_n\}$ and $\{b_n\}$ are such that $a_n^2 b_n=o(1)$ and the infinite series $\sum_{n=1}^{\infty} \exp(-\gamma n a_n^2/b_n^2)$ converges for every positive value of γ, then $b_n\{\Delta_n(f)-\Delta(f)\}$ converges to zero with probability one as n tends to infinity.

Proof. The proof follows at once from an obvious modification of theorem 2.5 in Schuster (1969) and the following result:

$$b_n|\Delta_n(f)-\Delta(f)| = b_n\left| \int_{-\infty}^{\infty} f_n^2(x)dx - \int_{-\infty}^{\infty} f^2(x)dx\right|$$

$$\leq 2b_n \sup_x |f_n(x)-f(x)|.$$

Corollary 3.1. If $a_n=n^{-1/6}$ and $b_n=n^{\delta}$, $0<\delta<1/3$, then $b_n\,|\Delta_n(f)-\Delta(f)|$ converges to zero with probability one as n tends to infinity.

Remark. If ϕ is the standard normal pdf, then the hypothesis of the above theorem is satisfied.

We now list the conditions on f and ϕ for which the asymptotic normality of the estimator proposed in this dissertation will be proved:

Condition (3.2a): f and its first three derivatives exist and are bounded.

Condition (3.2b): $\int_{-\infty}^{\infty} |f^{(i)}(x)|dx<\infty$, i=2,3.

Condition (3.2c): $0<\xi_1=\{ \int_{-\infty}^{\infty} f^3(x)dx-[\int_{-\infty}^{\infty} f^2(x)dx]^2\}<\infty$.

Condition (3.2d): $\phi(u)$ is bounded.

Condition (3.2e): $\int_{-\infty}^{\infty} u\phi(u)du=0.$

Condition (3.2f): $\int_{-\infty}^{\infty} |u|^3\phi(u)du<\infty.$

In Condition (3.2c) the positiveness of ξ_1 can be verified using an application of Cauchy-Schwarz inequality.

We first note that the estimator $\Delta_n(f)$ can be expressed as

$$\Delta_n(f)=\sum_{i=1}^{n}\sum_{j=1}^{n} g_n(X_i,X_j)$$

$$=(1-\frac{1}{n})U_n+\frac{1}{na_n}\int_{-\infty}^{\infty}\phi^2(t)dt, \tag{3.1}$$

where

$$g_n(x_i,x_j)=\frac{1}{a_n^2}\int_{-\infty}^{\infty}\phi(\frac{x-x_i}{a_n})\phi(\frac{x-x_j}{a_n})dx \tag{3.2}$$

and

$$U_n=\frac{1}{\binom{n}{2}}\sum_{1\le i<j\le n}\sum g_n(X_i,X_j) \tag{3.3}$$

is a one sample U-statistic first studied by Hoeffding (1948).

Thus the asymptotic distribution of $\Delta_n(f)$ depends on that of U_n. We define

$g_{1,n}(x_1)=E[g_n(X_1,X_2)|X_1=x_1]$

$h_n(x_1,x_2)=g_n(x_1,x_2)-\eta_n,$

where $\eta_n=E[g_n(X_1,X_2)]$

$h_{1,n}(x_1)=E[h_n(X_1,X_2)|X_1=x_1]$

$\xi_{1,n}=E[h_{1,n}^2(X_1)]$

$\xi_{2,n}=E[h_n^2(X_1,X_2)]$

$\rho_n^3=E|h_{1,n}(X_1)|^3.$

We prove the following lemmas:

Lemma 3.1. Under Conditions (3.2a), (3.2b), (3.2e) and (3.2f)

$$\eta_n=\Delta(f)+a_n^2\{\int_{-\infty}^{\infty}t^2\phi(t)dt\}\{\int_{-\infty}^{\infty}f(x)f^{(2)}(x)dx\}+o(a_n^2), \tag{3.4}$$

where $\Delta(f)$ is as defined before.

Proof. We have, using Fubini's theorem

$$\eta_n=E\{\frac{1}{a_n^2}\int_{-\infty}^{\infty}\phi(\frac{x-x_1}{a_n})\phi(\frac{x-x_2}{a_n})dx\}$$

$$=\int_{-\infty}^{\infty}[\int_{-\infty}^{\infty}\phi(t)f(x-ta_n)dt]^2dx.$$

Expanding $f(x-ta_n)$ by Taylor's formula around x to the order of $(ta_n)^3$ and using the integral form of the remainder term (see Singh, 1977) and Condition (3.2e) we have

$$\eta_n=\int_{-\infty}^{\infty}\{f(x)+\frac{a_n^2}{2}f^{(2)}(x)\int_{-\infty}^{\infty}t^2\phi(t)dt+\frac{1}{2}\int_{-\infty}^{\infty}\phi(t)\int_{x}^{x-ta_n}(x-ta_n-z)^2f^{(3)}(z)dz\,dt\}^2dx$$

$$=\int_{-\infty}^{\infty}\{f(x)+\frac{a_n^2}{2}f^{(2)}(x)\int_{-\infty}^{\infty}t^2\phi(t)dt+\frac{a_n^3}{2}\int_{-\infty}^{\infty}t^3\phi(t)\int_{0}^{1}(1-u)^2f^{(3)}(x-ta_nu)du\,dt\}^2dx. \tag{3.5}$$

Expanding the integrand in (3.5) and using Conditions (3.2a), (3.2b) and (3.2f) we complete the proof of the lemma.

Corollary 3.1. Under the conditions of Lemma 3.1

$$n^{1/2}\{E(U_n)-\Delta(f)\}=O(n^{1/2}a_n^2). \tag{3.6}$$

Lemma 3.2. Under Conditions (3.2a), (3.2b), (3.2d), (3.2e) and (3.2f)

(i) $\xi_{1,n}=\xi_1+2a_n^2\{\int_{-\infty}^{\infty}t^2\phi(t)dt\}[\int_{-\infty}^{\infty}f^2(x)f^{(2)}(x)dx-\{\int_{-\infty}^{\infty}f^2(x)dx\}$

$$.\{\int_{-\infty}^{\infty}f(x)f^{(2)}(x)dx\}]+o(a_n^2) \qquad (3.7)$$

(ii) $\xi_{2,n}=\frac{K}{a_n}-\eta_n^2,$ (3.8)

where K is some constant and η_n is as defined earlier.

Proof. (i) We have

$$g_{1,n}(x_1)=E[g_n(x_1,X_2)]$$

$$=\frac{1}{a_n^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\frac{x-x_1}{a_n})\phi(\frac{x-u}{a_n})f(u)dxdu$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(w)\phi(t)f(x_1+a_nw-a_nt)dtdw$$

Hence

$$E[g_{1,n}^2(X_1)]=\int_{-\infty}^{\infty}\{[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(w_1)\phi(t_1)f(x+a_nw_1-a_nt_1)dt_1dw_1]$$

$$.[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(w_2)\phi(t_2)f(x+a_nw_2-a_nt_2)dt_2dw_2]\}f(x)dx.$$

Expanding $f(x+a_nw_1-a_nt_1)$ and $f(x+a_nw_2-a_nt_2)$ by Taylor's formula around x to the order of $[a_n(w_1-t_1)]^3$ and $[a_n(w_2-t_2)]^3$ respectively, with the integral form for the remainder term and simplifying, we get

$$E[g_{1,n}^2(X_1)]=\int_{-\infty}^{\infty}\{[f(x)+a_n^2f^{(2)}(x)\int_{-\infty}^{\infty}t^2\phi(t)dt$$

$$+\frac{a_n^3}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(w_1-t_1)^3\phi(w_1)\phi(t_1)\int_0^1(1-u_1)^2 f^{(3)}(x+a_n(w_1-t_1)u_1)du_1dw_1dt_1]$$

$$.[f(x)+a_n^2 f^{(2)}(x)\int_{-\infty}^{\infty}t^2\phi(t)dt$$

$$+\frac{a_n^3}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(w_2-t_2)^3\phi(w_2)\phi(t_2)\int_0^1(1-u_2)^2 f^{(3)}(x+a_n(w_2-t_2)u_2)du_2dw_2dt_2]\}$$

$$.f(x)dx$$

$$=\int_{-\infty}^{\infty}f^3(x)dx+2a_n^2\{\int_{-\infty}^{\infty}t^2\phi(t)dt\}\{\int_{-\infty}^{\infty}f^2(x)f^{(2)}(x)dx\}+o(a_n^2). \qquad (3.9)$$

From (3.5) and (3.9), we get

$$\xi_{1,n}=E[h_{1,n}^2(x_1)]$$

$$=E[g_{1,n}^2(x_1)]-\eta_n^2$$

$$=\xi_1+2a_n^2\{\int_{-\infty}^{\infty}t^2\phi(t)dt\}[\int_{-\infty}^{\infty}f^2(x)f^{(2)}(x)dx$$

$$-\{\int_{-\infty}^{\infty}f^2(x)dx\}\{\int_{-\infty}^{\infty}f(x)f^{(2)}(x)dx\}]+o(a_n^2).$$

(ii) We have

$$\xi_{2,n}=E[g_n^2(X_1,X_2)]-\eta_n^2$$

$$=\frac{1}{a_n^4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\frac{u_1-x_1}{a_n})\phi(\frac{u_1-x_2}{a_n})\phi(\frac{u_2-x_1}{a_n})\phi(\frac{u_2-x_2}{a_n})f(x_1)f(x_2)dx_1dx_2du_1du_2-\eta_n^2$$

$$=\frac{1}{a_n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(t_1)\phi(t_2)\phi(t_3)\phi(t_2+t_3-t_1)f(t_4+t_2a_n-t_1a_n)f(t_4)dt_1dt_2dt_3dt_4-\eta_n^2 \qquad (3.10)$$

$$=\frac{K}{a_n}-\eta_n^2,$$

where K is the multiple integral in the right hand side of (3.10).

Lemma 3.3. Under the conditions of Lemma 3.2

(i) $\lim_{n\to\infty}\xi_{1,n}= \xi_1$

(ii) $\xi_{2,n}=O(1/a_n)$

(iii) $\lim_{n\to\infty}\rho_n^3$ exists.

Proof. The proofs of part (i) and part (ii) follow directly from Lemma 3.2

(iii) We have

$$\rho_n^3=E|g_{1,n}(X_1)-\eta_n|^3$$

$$\leq E|g_{1,n}(X_1)|^3+3E|g_{1,n}(X_1)|^2|\eta_n|+3E|g_{1,n}(X_1)||\eta_n|^2+|\eta_n|^3.$$

Using Condition (3.2a) we conclude that

$|g_{1,n}(x_1)|=|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(w)\phi(t)f(x_1+a_nw-a_nt)dtdw|$ is bounded, and hence $E|g_{1,n}(X_1)|^r$

is finite for r=1, 2, 3.

The results obtained in Lemma 3.3 enable us to use Hoeffding's (1948) theorem on the asymptotic distribution of U_n and we have the following lemma:

Lemma 3.4. Under the conditions of Lemma 3.3, if $\lim_{n\to\infty}na_n=\infty$, then $var[n^{1/2}U_n]=4\xi_1+o(1)$. If further $\xi_1>0$, then $n^{1/2}\{U_n-E(U_n)\}$ converges in distribution to a normal random variable with mean 0 and variance $4\xi_1$.

Lemma 3.5. Under Conditions (3.2a) through (3.2f), if $\lim_{n\to\infty} na_n=\infty$ and $\lim_{n\to\infty} na_n^4=0$, then $n^{1/2}\{U_n-\Delta(f)\}$ converges in distribution to a normal random variable with mean 0 and variance $4\xi_1$.

Proof. Using relation (3.4) we have

$$
\begin{aligned}
n^{1/2}\{U_n-\Delta(f)\}&=n^{1/2}\{U_n-E(U_n)\}+\{E(U_n)-\Delta(f))\}\\
&=n^{1/2}\{U_n-E(U_n)\}+O(n^{1/2}a_n^2)\\
&=n^{1/2}\{U_n-E(U_n)\}+o(1). \qquad (3.11)
\end{aligned}
$$

The proof now follows from Lemma 3.4.

We now prove the main theorem in this section.

Theorem 3.2. Under Conditions (3.2a) through (3.2f), if $\lim_{n\to\infty} na_n^2=\infty$ and $\lim_{n\to\infty} na_n^4=0$, then the estimate $\Delta_n(f)=\int_{-\infty}^{\infty} f_n^2(x)dx$ is asymptotically normally distributed with mean $\Delta(f)=\int_{-\infty}^{\infty} f^2(x)dx$ and variance

$$\frac{4\xi_1}{n}=\frac{4}{n}[\int_{-\infty}^{\infty} f^3(x)dx-\{\int_{-\infty}^{\infty} f^2(x)dx\}^2].$$

Proof. Using relation (3.1) and the hypothesis, we have

$$
\begin{aligned}
n^{1/2}\{\Delta_n(f)-\Delta(f)\}&=n^{1/2}\{(1-\frac{1}{n})U_n+\frac{1}{na_n}\int_{-\infty}^{\infty}\phi^2(t)dt-\Delta(f)\}\\
&=(1-\frac{1}{n})n^{1/2}(U_n-\Delta(f))-\frac{\Delta(f)}{n^{1/2}}+\frac{1}{n^{1/2}a_n}\int_{-\infty}^{\infty}\phi^2(t)dt\\
&=n^{1/2}(U_n-\Delta(f))+o(1)O_p(1)+o(1)+o(1), \text{ by Lemma 3.5}.
\end{aligned}
$$

Another application of Lemma 3.5 completes the proof of the theorem.

Note: Schweder (1975) and Ahmad (1976) have proved an identical theorem for the estimate $\Delta_{1n}(f)=\int_{-\infty}^{\infty} f_n(x)dF_n(x)$, where $F_n(x)$ is the empirical distribution function.

3.3 Optimal choice of the smoothing parameter

In this section we prove the following theorem:

Theorem 3.3. Under Conditions (3.2a), (3.2b), (3.2d), (3.2e) and (3.2f) in section 3.2, if further $\lim_{n\to\infty} na_n^{5/2}=\infty$ and $\lim_{n\to\infty} na_n^3=c$, where c is a constant not equal to zero, then an expression for a_n that minimizes $E\{\Delta_n(f)-\Delta(f)\}^2$ is given by

$$a_n=\left(\frac{-\lambda_2}{n\lambda_1}\right)^{1/3}, \tag{3.12}$$

where

$$\lambda_1=\{\int_{-\infty}^{\infty} t^2\phi(t)dt\}\{\int_{-\infty}^{\infty} f(x)f^{(2)}(x)dx\} \tag{3.13}$$

and

$$\lambda_2=\int_{-\infty}^{\infty}\phi^2(t)dt. \tag{3.14}$$

The smallest asymptotic mean square error is

$$\min_{a_n} E\{\Delta_n(f)-\Delta(f)\}^2 \approx \frac{4}{n}\xi_1. \tag{3.15}$$

Proof. From (3.1) we have using Lemma 3.1

$$E\{\Delta_n(f)\}=(1-\frac{1}{n})\{\Delta(f)+\lambda_1a_n^2+o(a_n^2)\}+\frac{\lambda_2}{na_n}$$
$$=\Delta(f)+\lambda_1a_n^2+\frac{\lambda_2}{na_n}+o(a_n^2), \tag{3.16}$$

where λ_1 and λ_2 are as defined in (3.13) and (3.14) respectively.

Using formula (5.6) in Fraser (1957, p. 225) and Lemmas 3.1 and 3.2, we get

$$\mathrm{var}\{\Delta_n(f)\}=(1-\frac{1}{n})^2\mathrm{var}(U_n)$$

$$=(1-\frac{1}{n})^2\{\frac{4}{n}(1-\frac{1}{n-1})\xi_{1,n}+\frac{2}{n(n-1)}\ \xi_{2,n}\}$$

$$=(1-\frac{1}{n})^2\{\frac{4}{n}(1-\frac{1}{n-1})[\xi_1+O(a_n^2)]+\frac{2}{n(n-1)}[O(\frac{1}{a_n})+O(1)]\}$$

$$\approx \frac{4}{n}\xi_1+O(\frac{a_n^2}{n})+O(\frac{1}{n^2a_n}). \tag{3.17}$$

Hence, using the choice of a_n as given in the hypothesis, we have

$$E\{\Delta_n-\Delta\}^2=\mathrm{var}\{\Delta_n(f)\}+\{E[\Delta_n(f)]-\Delta(f)\}^2$$

$$\approx \frac{4}{n}\xi_1+(\lambda_1a_n^2+\frac{\lambda_2}{na_n})^2+O(\frac{a_n^2}{n})+O(\frac{1}{n^2a_n})+o(a_n^4)+o(\frac{a_n}{n}))$$

$$=\frac{4}{n}\xi_1+(\lambda_1a_n^2+\frac{\lambda_2}{na_n})^2+o((\lambda_1a_n^2+\frac{\lambda_2}{na_n})^2). \tag{3.18}$$

The expression in (3.12) is obtained by letting $\lambda_1a_n^2+\frac{\lambda_2}{na_n}=0$. Substituting this value of a_n in (3.18) and neglecting terms of "higher order of smallness" gives the expression for the minimum mean square error in (3.15).

3.4 Some Applications

In this section we present two examples.

Example 3. Let $X_1, X_2, ..., X_{m_N}$ and $Y_1, Y_2, ..., Y_{n_N}$ be independent random samples from the probability density functions f and g respectively, where $m_N+n_N=N$ for each N. Suppose there exists a probability density function h and two real numbers θ_1 and θ_2 such that

$$f(x)=h(x-\theta_1) \text{ and } g(x)=h(x-\theta_2) \tag{3.19}$$

for all real numbers x. In such a case, the number $\theta=\theta_2-\theta_1$ is said to be the shift between the densities f and g. If h is unknown, then the point estimator of θ is given by

$$\hat{\theta}=\text{median}\{Y_j-X_i,\ i=1, 2, ..., m_N,\ j=1, 2, ..., n_N\}. \tag{3.20}$$

Suppose m_N and n_N tend to infinity in such a way that $\lim_{n\to\infty} \frac{m_N}{m_N+n_N}=\lambda$, $0<\lambda<1$. Then, under some regularity conditions (see Hodges and Lehmann, 1963), $\hat{\theta}$ is asymptotically normally distributed with mean θ and variance $1/\{12\lambda(1-\lambda)[\int_{-\infty}^{\infty} f^2(x)dx]^2\}$. Hence, using a consistent estimate of $\int_{-\infty}^{\infty} f^2(x)dx$, we can construct a large sample confidence interval for the shift parameter θ.

We now assume that $h(x)=\frac{1}{\pi}\frac{1}{1+x^2}$, $-\infty<x<\infty$. Under this assumption we propose an estimator of θ_1 and θ_2 as

$$\hat{\theta}_1=\text{median}\{X_1, X_2, ..., X_{m_N}\} \tag{3.21}$$

and

$$\hat{\theta}_2=\text{median}\{Y_1, Y_2, ..., Y_{n_N}\} \tag{3.22}$$

respectively. We also construct a best asymptotic normal (BAN) estimators of θ_1 and θ_2 by

$$\theta_1^*=\hat{\theta}_1+\frac{L_n(\hat{\theta}_1)}{-L_n^{(1)}(\hat{\theta}_1)} \tag{3.23}$$

and

$$\theta_2^*=\hat{\theta}_2+\frac{L_n(\hat{\theta}_2)}{-L_n^{(1)}(\hat{\theta}_2)} \tag{3.24}$$

respectively, where $L_n(\theta)=-\frac{1}{n}\sum_{i=1}^{n}\frac{h^{(1)}(Z_i-\theta)}{h(Z_i-\theta)}$, $L_n^{(1)}(\theta)=\frac{\partial}{\partial\theta}L_n(\theta)$ and $h^{(1)}(Z-\theta)=\frac{\partial}{\partial\theta}h(Z-\theta)$. We call the three estimators of θ given by $\hat{\theta}$, $\hat{\hat{\theta}}=\hat{\theta}_2-\hat{\theta}_1$ and $\theta^*=\theta_2^*-\theta_1^*$ the Hodges-Lehmann estimator, the 'median' estimator and the BAN estimator respectively. It follows that the estimator $\hat{\hat{\theta}}=\hat{\theta}_2-\hat{\theta}_1$ is asymptotically normally distributed with mean θ and variance $\frac{\pi^2}{4\lambda(1-\lambda)N}$. It is also known that the BAN estimator $\theta^*=\theta_2^*-\theta_1^*$ has asymptotic variance $\frac{2}{\lambda(1-\lambda)N}$.

We now draw a random sample of size $m_N=100$ from the Cauchy pdf $f(x)=\frac{1}{\pi}\frac{1}{1+x^2}$ and another independent random sample of size $n_N=100$ from the Cauchy pdf $g(x)=\frac{1}{\pi}\frac{1}{1+(x-5)^2}$. Using a consistent estimate of $\int_{-\infty}^{\infty} f^2(x)dx$, we construct a 100(1-$\alpha$) percent nonparametric large sample confidence interval for θ (where θ is the shift between the two densities and is equal to 5 in this example) as

$$\hat{\theta} \pm z_{\alpha/2}\{\sqrt{600}\int_{-\infty}^{\infty} f_n^2(x)dx\}^{-1},$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of a standard normal distribution.

We also consider a large sample confidence interval for the shift parameter based on the Mann-Whitney statistic as given in Noether (1972).

We can also obtain parametric 100(1-α) percent large sample confidence intervals for θ as

$$\hat{\hat{\theta}} \pm z_{\alpha/2}\pi\{\sqrt{200}\}^{-1}$$

and

$$\theta^* \pm z_{\alpha/2}\{\sqrt{25}\}^{-1}$$

from the 'median' and the BAN estimators respectively.

Random samples following Cauchy distribution were simulated by generating uniform random variates (SAS Institute, 1982) and applying the appropriate inverse transformation. The medians were obtained using PROC UNIVARIATE in SAS.

The following table gives 95 percent confidence intervals for the four methods outlined above:

Table 3.1: Large sample Confidence Intervals for the shift between two Cauchy distributions

Estimator	95% C. I.
[1]Nonparametric	(4.66, 5.60)
[2]Nonparametric*	(4.69, 5.60)
Median	(4.75, 5.62)
BAN	(4.69, 5.47)

[1]The interval is based on the Mann-Whitney statistic. [2]The interval is based on the Hodges-Lehmann estimator.

*The functional $\int_{-\infty}^{\infty} f^2(x)dx$ was estimated by a consistent estimate $\int_{-\infty}^{\infty} f_n^2(x)dx$ based on the standard normal density as the kernel. In this case

$$\int_{-\infty}^{\infty} f_n^2(x)dx=\{\sqrt{2}\, a_n\, n^2\}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \phi\left(\frac{X_i-X_j}{\sqrt{2}\, a_n}\right), \tag{3.25}$$

where $\phi(x)=(2\pi)^{-1/2} \exp(-x^2/2)$ (see Bhattacharyya and Roussas, 1969). Following Theorem 3.3 we used $a_n=n^{-1/3}$ in the right hand side of (3.25). Thus, using simulated data from Cauchy distributions, we find that although the confidence interval based on our nonparametric procedure is wider than the one based on the BAN estimator, the difference in lengths of these intervals being .13 is small. We also see that in this particular situation the present nonparametric procedure yields a shorter interval than the one based on the Mann-Whitney statistic.

Example 3.2. In this example we construct a large sample nonparametric confidence interval for $\Delta(f)=\int_{-\infty}^{\infty} f^2(x)dx$ using a consistent estimate of the variance of an estimate of $\Delta(f)$ as given in Theorem 3.2.

We first consider the following lemmas:

Lemma 3.6. If f and $f^{(1)}$ are bounded, ϕ is a function of bounded variation, and the sequences $\{a_n\}$ and $\{l_n\}$ are such that $a_n=n^{-\delta}$, $\frac{1}{8}<\delta<\frac{3}{8}$, and $l_n=n^{\gamma}$, $0<\gamma<\frac{1}{8}$, then $\sup_x n^{\gamma}|f_n(x)-f(x)|$ converges to zero with probability one as n tends to infinity.

Proof. The proof follows at once from Theorem 2.5 in Schuster (1969).

Lemma 3.7. Under the conditions of Lemma 3.6

(i) $\sup_x |f_n(x)|=O_p(1)$

(ii) $\sup_x |f_n^2(x)|=O_p(1)$.

Proof. (i) We have

$$\sup_x |f_n(x)|\leq\sup_x |f_n(x)-f(x)|+\sup_x |f(x)|.$$

The proof follows from the above inequality in conjunction with Lemma 3.6.

The proof of part (ii) follows from Lemma 3.6 and part (i) by noting that

$$\sup_x |f_n^2(x)|\leq\sup_x |f_n^2(x)-f^2(x)|+\sup_x |f^2(x)|\leq\sup_x |f_n(x)-f(x)| \sup_x |f_n(x)+f(x)|+\sup_x |f^2(x)|.$$

Lemma 3.8. Under the conditions of Lemma 3.6

$\left|\int_{-l_n}^{l_n} f_n^3(x)dx-\int_{-\infty}^{\infty} f^3(x)dx\right|$ converges in probability to zero as n tends to infinity.

Proof. For arbitrarily small $\varepsilon>0$ and sufficiently large n, we have

$$\left|\int_{-l_n}^{l_n} f_n^3(x)dx-\int_{-\infty}^{\infty} f^3(x)dx\right|\leq\left|\int_{-l_n}^{l_n} f_n^3(x)dx-\int_{-l_n}^{l_n} f^3(x)dx\right|+\frac{\varepsilon}{2}$$

$$\leq 2 \sup_{|x|<l_n} l_n |f_n^3(x)-f^3(x)|+\frac{\varepsilon}{2}$$

$$\leq 2 \sup_{|x|<l_n} l_n |f_n(x)-f(x)| \sup_{|x|<l_n} |f_n^2(x)+f^2(x)+f_n(x)f(x)|+\frac{\varepsilon}{2}.$$

An application of Lemmas 3.6 and 3.7 completes the proof.

We now propose to construct a large sample 100(1-α) percent confidence interval for $\Delta(f)$ as

$$\Delta_n(f) \pm z_{\alpha/2}\sqrt{\{\frac{4}{n}(\int_{-l_n}^{l_n} f_n^3(x)dx-[\int_{-\infty}^{\infty} f_n^2(x)dx]^2)\}}.$$

Random samples from Cauchy and logistic distributions were simulated by generating uniform random variates (SAS Institute, 1982) and applying appropriate inverse transformations. A random sample following normal distribution was simulated by the normal random variate generator (SAS Institute, 1982). Table 3.2 gives nonparametric confidence intervals for the Hodges-Lehmann functional corresponding to a confidence coefficient 0.95 and a sample size 100.

Table 3.2: Large sample Confidence Intervals for $\Delta(f)=\int_{-\infty}^{\infty} f^2(x)dx$ for some distributions

pdf	$\Delta(f)$	95% C. I.
$(2\pi)^{-1/2}\exp(-x^2/2)$	0.28	(0.23, 0.31)
$\frac{1}{\pi}\frac{1}{1+x^2}$	0.16	(0.13, 0.21)
$\frac{\exp(-x)}{\{1+\exp(-x)\}^2}$	0.17	(0.15, 0.21)

Remark: The functionals $\int_{-\infty}^{\infty} f^2(x)dx$ and $\int_{-\infty}^{\infty} f^3(x)dx$, appearing in the variance for the estimate of $\Delta(f)$, are estmated by $\int_{-\infty}^{\infty} f_n^2(x)dx$ and $\int_{-l_n}^{l_n} f_n^3(x)dx$ respectively, where $f_n(x)$ is a kernel estimate of f(x) obtained by using the standard normal probability function as a kernel function. In this case the estimate $\int_{-l_n}^{l_n} f_n^3(x)dx$ is given by

$$\int_{-l_n}^{l_n} f_n^3(x)dx=\{n^3\}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\int_{-l_n}^{l_n}\phi\left(\frac{x-X_i}{a_n}\right)\phi\left(\frac{x-X_j}{a_n}\right)\phi\left(\frac{x-X_k}{a_n}\right)dx, \quad (3.26)$$

where $\phi(x)=(2\pi)^{-1/2}\exp(-x^2/2)$. By straight forward integration, we get

$$\int_{-l_n}^{l_n} f_n^3(x)dx$$

$$=\{2\pi\sqrt{3}a_n^2 n^3\}^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\exp[-\frac{1}{3a_n^2}(X_i^2+X_j^2+X_k^2-X_iX_j-X_iX_k-X_jX_k)]$$

$$.[\Phi(\frac{\sqrt{3}}{a_n}(l_n-\frac{X_i+X_j+X_k}{3})-\Phi(\frac{\sqrt{3}}{a_n}(-l_n-\frac{X_i+X_j+X_k}{3})], \quad (3.27)$$

where

$\Phi(t)=\int_{-\infty}^{t}\frac{1}{\sqrt{(2\pi)}}\exp\{-y^2/2\}dy$. Following Theorem 3.3 and Lemmas 3.6 through 3.8 we used $a_n=n^{-1/3}$ and $l_n=n^{1/9}$ in the right hand side of (3.27).

CHAPTER IV

NONPARAMETRIC ESTIMATION OF A PROBABILITY DENSITY FUNCTION AND ITS FUNTIONALS

4.1 Introduction

Suppose $X_1, X_2, ..., X_n$ are independent and identically distributed random variables having the pdf f(x) where f is unknown. An estimator $f_n(x)$ of f(x) proposed by Rosenblatt (1956) and later studied by Parzen (1962) and others is given by:

$$f_n(x)=\frac{1}{na_n}\sum_{i=1}^{n}\phi\left(\frac{x-X_i}{a_n}\right), \tag{4.1}$$

where ϕ is a known probability density function and $\{a_n\}$ is a sequence of positive numbers converging to zero as n tends to infinity. ϕ is called a kernel function (or window function) and the sequence $\{a_n\}$ is called a smoothing parameter (or bandwidth). The estimator $f_n(x)$ is called a kernel estimator (window estimator) of f(x). The choice of $\{a_n\}$ in defining the estimator $f_n(x)$ has been studied among others by Rosenblatt (1956), Woodroofe (1970), Nadaraya (1974) and Silverman (1986). "The bandwidth determines, among other things, how far away the observations are allowed to be far from x and still contributes to the estimation of f(x). The bandwidth also governs the peakedness of the kernel and, hence, the degree of dependence of the estimator on information near x. Small values of a_n will result in rougher (wigglier) estimators that rely heavily on the data near x. In contrast, larger a_n's allow more averaging to occur and thereby give smoother estimators." (see Eubank, 1988, p. 112). It is well known, (see Rosenblatt, 1956, Parzen, 1962, Nadaraya, 1965 and Schuster, 1969), that the asymptotic behavior of $f_n(x)$ depends on the smoothness of f and on the sequence $\{a_n\}$. In Rosenblatt (1956) it is shown that the optimal choice of a_n, that minimizes the asymptotic mean square error

(MSE) of $f_n(x)$, also depends on the smoothness of f near x. That is, an optimal value of a_n, denoted by a_{1n}, is given by:

$a_{1n}=q(n, f, f^{(1)}, f^{(2)}, ..., x, \phi)$, where q is a known function of its arguments and is still unknown to the statistician since f, $f^{(1)}, f^{(2)}$, ... are unknown. Woodroofe (1970) has proposed to estimate a_{1n} by using consistent estimates of f, $f^{(1)}, f^{(2)}$, ... and has shown that

$$\lim_{n\to\infty}\left[\frac{E\{f_n(x;a_{2n}) - f(x)\}^2}{E\{f_n(x;a_{1n}) - f(x)\}^2}\right]=1, \tag{4.2}$$

where a_{2n} is an estimate of a_{1n}. Rosenblatt (1956) and Silverman (1986) have investigated the second order and fourth order properties of a class of estimators of the density function, and Azzalini (1981) has studied the second order properties of a class of estimators of the distribution function and population quantiles. The optimum value of a_n that minimizes the asymptotic MSE has been obtained in each case.

In this chapter we derive expressions for the mean square errors of kernel estimates of the derivative of the probability density function f, the population mode θ, the distribution function F and the population quantile of order p ($0<p<1$), ζ_p. In each case we shall obtain expressions for the optimum value of a_n which minimize the mean square error. It will be shown that the results obtained in Rosenblatt (1956), Silverman (1986) and Azzalini (1981) are special cases of our results.

4.2 Estimation of the Derivatives of a Probability Density Function.

As in Nadaraya (1974) we shall say that ϕ is a kernel function of class C_s (s is an even integer greater than or equal to 2) if it satisfies the following conditions:

(A) $\phi(x)=\phi(-x)$, $\sup_{x\varepsilon R}|\phi(x)|\leq K<\infty$.

(B) $\alpha_i= \int_{-\infty}^{\infty} x^i\phi(x)dx= 1, \quad \text{if } i=0$

$$= 0, \quad \text{if } i=1, 2, ..., s\text{-}1$$

$$\alpha_s \neq 0, \qquad \int_{-\infty}^{\infty} |x^{s+1}\phi(x)|dx < \infty.$$

For example, a kernel function of class C_s is given by

$$\phi(x)=\{H_0(x)+d_2H_2(x)+d_4H_4(x)+ \ldots +d_{s-2}H_{s-2}(x)\}z(x), \tag{4.3}$$

where z(x) is the standard normal pdf, $d_s=\frac{(-1)^{s/2}}{2^{s/2}(s/2)!}$ and $H_r(x)$ is the Chebychev-Hermite polynomial of degree r defined by

$$H_r(x)=x^r - \frac{r^{[2]}}{2.1!}x^{r-2} + \frac{r^{[4]}}{2^2.2!}x^{r-4} - \frac{r^{[6]}}{2^3.3!}x^{r-6} + - \ldots$$

and $H_0(x)\equiv 1$ (see Kendall and Stuart 1969, p. 155).

Following Bhattacharya (1967) and Schuster (1969), an estimator of the r-th derivative $f^{(r)}(x)$ of a pdf f(x) is given by

$$f_n^{(r)}(x)=\frac{1}{na_n^{r+1}}\sum_{i=1}^{n}\phi^{(r)}\left(\frac{x-X_i}{a_n}\right) \tag{4.4}$$

We shall assume the following regularity conditions:

Condition (4.2a): f and its first (r+s+1) derivatives exist and are bounded and $f^{(r+s)}(x)\neq 0$.

Condition (4.2b): $\phi(x)$ is a kernel function belonging to class C_s.

Condition (4.2c): $\phi^{(i)}(x)$ is absolutely integrable for i=0, 1, 2, ..., r.

Condition (4.2d): $\int_{-\infty}^{\infty} t^i[\phi^{(r)}(t)]^2dt < \infty$ for i=0, 1.

The kernel function given in (4.3) also satisfies Conditions (4.2c) and (4.2d).

By Conditions (4.2a) and (4.2b), we have

$$E\{f_n^{(r)}(x)\} = \int_{-\infty}^{\infty} \frac{1}{a_n^{r+1}} \phi^{(r)}\left(\frac{x-u}{a_n}\right) f(u)du$$

$$= \int_{-\infty}^{\infty} \frac{1}{a_n^{r}} \phi^{(r)}(t) f(x-ta_n)dt.$$

Integrating by parts successively, we get after using Conditions (4.2a) and (4.2c)

$$E\{f_n^{(r)}(x)\} = \frac{1}{a_n^{r}} f(x-ta_n)\phi^{(r-1)}(t)]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{a_n^{r-1}} \phi^{(r-1)}(t) f^{(1)}(x-ta_n)dt$$

$$= \frac{1}{a_n^{r-1}} f^{(1)}(x-ta_n)\phi^{(r-2)}(t)]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{a_n^{r-2}} \phi^{(r-2)}(t) f^{(2)}(x-ta_n)dt$$

....

....

$$= \int_{-\infty}^{\infty} \phi(t) f^{(r)}(x-ta_n)dt \qquad (4.5)$$

Now, expanding $f^{(r)}(x-ta_n)$ around x to the order of $(ta_n)^{s+1}$ (where s is an even integer) and using Conditions (4.2a) and (4.2b), we have

$$E\{f_n^{(r)}(x)\} = f^{(r)}(x) + a_n^s \left[\frac{1}{s!} f^{(r+s)}(x) \int_{-\infty}^{\infty} t^s \phi(t)dt + o(1) \right]. \qquad (4.6)$$

Using computation in (4.6) we have

$$na_n^{2r+1} \mathrm{var}\{f_n^{(r)}(x)\} = \int_{-\infty}^{\infty} \frac{1}{a_n} \left[\phi^{(r)}\left(\frac{x-u}{a_n}\right) \right]^2 f(u)du$$

$$- a_n^{2r+1}\left[\int_{-\infty}^{\infty}\frac{1}{a_n^{r+1}}\phi^{(r)}\left(\frac{x-u}{a_n}\right)f(u)du\right]^2$$

$$= \int_{-\infty}^{\infty}[\phi^{(r)}(t)]^2 f(x-ta_n)dt - a_n^{2r+1}\{f^{(r)}(x)+o(1)\}. \tag{4.7}$$

In the right hand side of (4.7), expanding $f(x-ta_n)$ around x to the order of ta_n and using Condition (4.2d), we have

$$\operatorname{var}\{f_n^{(r)}(x)\} = \frac{f(x)}{na_n^{2r+1}}\int_{-\infty}^{\infty}[\phi^{(r)}(t)]^2 dt + o\left(\frac{1}{na_n^{2r+1}}\right) \tag{4.8}$$

Hence

$$E\{f_n^{(r)}(x)-f^{(r)}(x)\}^2 \approx \frac{A_1}{na_n^{2r+1}} + a_n^{2s}B_1 + o\left(\frac{1}{na_n^{2r+1}} + a_n^{2s}\right), \tag{4.9}$$

where $A_1 = f(x)\int_{-\infty}^{\infty}[\phi^{(r)}(t)]^2 dt$ and $B_1 = \left(\frac{1}{s!}f^{(r+s)}(x)\int_{-\infty}^{\infty}t^s\phi(t)dt\right)^2$.

Differentiating the dominating terms in the right hand side of (4.9) with respect to a_n and equating to zero gives

$$a_n = \left(\frac{(2r+1)A_1}{2snB_1}\right)^{1/[2(s+r)+1]}. \tag{4.10}$$

The corresponding minimum MSE is given by

$$E_m\{f_n^{(r)}(x)-f^{(r)}(x)\}^2 \approx \left\{\frac{2(s+r)+1}{2r+1}\right\}\left\{\frac{(2r+1)A_1}{2sn}\right\}^{2s/[2(r+s)+1]} B_1^{(2r+1)/[2(r+s)+1]}. \tag{4.11}$$

The results in Rosenblatt (1956) and Silverman (1986, p. 67) are special cases of the above results when r=0 and s=2, 4.

4.3 Estimation of the Population Mode

In this section we shall consider the estimaton of the population mode.

If the pdf f(x) is uniformly continuous in x, then f(x) possesses a mode θ defined by

$$f(\theta) = \max_{-\infty<x<\infty} f(x) \tag{4.12}$$

We assume that θ given by (4.12) is unique. If the kernel function ϕ, appearing in the expression for the estimate $f_n(x)$ of f(x) in (4.1), is so chosen that $\phi(u)$ tends to zero as u tends to $\pm\infty$, then for every sample sequence $f_n(x)$ is continuous and tends to zero as x tends to $\pm\infty$. Consequently, there is a random variable θ_n such that

$$f_n(\theta_n) = \max_{-\infty<x<\infty} f_n(x) \tag{4.13}$$

We call θ_n the sample mode and consider it as an estimate of θ.

Parzen (1962) and Nadaraya (1965) have shown that, under some regularity conditions, θ_n is strongly consistent for θ and is asymptotically normally distributed. In this section, we use a method developed by Azzallini (1981) to obtain an expression for the mean square error $E\{\theta_n-\theta\}^2$ and find the optimum value of the smoothing parameter a_n for which the mean square error is a minimum.

We assume the following conditions:

Condition (4.3a): f and its first (s+2) derivatives are bounded.

Condition (4.3b): $f^{(1)}(\theta)=0$, $f^{(2)}(\theta)<0$, $f^{(1+s)}(\theta)\neq 0$.

Condition (4.3c): $\phi(x)$ is a kernel function belonging to class C_s.

Condition (4.3d): $\phi^{(i)}(x)$ is a function of bounded variation for i=0, 1, 2.

Condition (4.3e): $\int_{-\infty}^{\infty} t^i[\phi^{(1)}(t)]^2 dt < \infty$ for i=0, 1.

Condition (4.3f): $a_n=n^{-\delta}$, $\frac{1}{2s+3} \leq \delta < \frac{1}{6}$.

The kernel function given in (4.3) also satisfies Conditions (4.3d) and (4.3e).

Let $f_n(x)$ be an estimate of $f(x)$ as defined in (4.1). If θ_n is the mode of $f_n(x)$, then

$$f_n^{(1)}(\theta_n) = 0, \quad f_n^{(2)}(\theta_n) < 0. \tag{4.14}$$

Expanding $f_n^{(1)}(\theta_n)$ around θ, we get

$$0 = f_n^{(1)}(\theta_n) = f_n^{(1)}(\theta) + (\theta_n - \theta) f_n^{(2)}(\theta_n^*), \tag{4.15}$$

where $|\theta_n^* - \theta| < |\theta_n - \theta|\cdot$ From relation (4.15) we may write

$$(\theta_n - \theta) = -\frac{E\{f_n^{(1)}(\theta)\}}{f_n^{(2)}(\theta_n^*)} - \frac{f_n^{(1)}(\theta) - E\{f_n^{(1)}(\theta)\}}{f_n^{(2)}(\theta_n^*)}. \tag{4.16}$$

Lemma 4.1. Under Conditions (4.3a), (4.3b), (4.3c), (4.3d) and (4.3e)

$$\text{(i) } E\{f_n^{(1)}(\theta)\} = \frac{a_n^s}{s!} f^{(1+s)}(\theta) \int_{-\infty}^{\infty} t^s \phi(t)dt + O(a_n^{1+s}) \tag{4.17}$$

$$\text{(ii) } \operatorname{var}\{f_n^{(1)}(\theta)\} = \frac{f(\theta)}{na_n^3} \int_{-\infty}^{\infty} [\phi^{(1)}(t)]^2 dt + o\left(\frac{1}{na_n^3}\right) \tag{4.18}$$

Proof. (i) The proof follows at once from (4.6) by letting $r=1$ in conjunction with relation $f^{(1)}(\theta)=0$.

(ii) The proof is a special case of the computation in (4.8) with $r=1$.

Lemma 4.2. Under Conditions (4.3a), (4.3d) and (4.3f)

(i) $f_n^{(2)}(\theta_n^*) = f^{(2)}(\theta) + o_p(1)$

(ii) $(na_n^3)^{1/2}\{f_n^{(1)}(\theta) - E[f_n^{(1)}(\theta)]\}$ converges in distribution to a normal random variable with mean 0 and variance $f(\theta) \int_{-\infty}^{\infty} [\phi^{(1)}(t)]^2 dt$.

Proof. For a proof of these results see Samanta (1973).

In the following we shall assume Conditions (4.3a) through (4.3f).

Using part (i) of Lemma 4.1 and part (i) of Lemma 4.2 in equation (4.16), we get

$$f^{(2)}(\theta)\{\theta_n-\theta\}=\{-\frac{a_n^s}{s!}f^{(1+s)}(\theta)\int_{-\infty}^{\infty}t^s\,\phi(t)dt+O(a_n^{1+s})\}\{1+o_p(1)\}$$

$$-\{f_n^{(1)}(\theta)\;-\;E[f_n^{(1)}(\theta)]\}\{1+o_p(1)\}. \qquad (4.19)$$

Dividing both sides of (4.19) by the asymptotic standard deviation of $f_n^{(1)}(\theta)$ obtained in part (ii) of Lemma 4.1, we get

$$\frac{(na_n^3)^{1/2}\{f^{(2)}(\theta)(\theta_n-\theta)+\frac{a_n^s}{s!}f^{(s+1)}(\theta)\int_{-\infty}^{\infty}t^s\phi(t)dt\}}{\{f(\theta)\int_{-\infty}^{\infty}[\phi^{(1)}(t)]^2dt\}^{1/2}}$$

$$=-\frac{(na_n^3)^{1/2}\{f_n^{(1)}(\theta)\;-\;E[f_n^{(1)}(\theta)]\}\{1+o_p(1)\}}{\{f(\theta)\int_{-\infty}^{\infty}[\phi^{(1)}(t)]^2dt\}^{1/2}}+o_p(1). \qquad (4.20)$$

Using part (ii) of Lemma 4.2 we conclude that

$$\frac{(na_n^3)^{1/2}\{f^{(2)}(\theta)(\theta_n-\theta)+\frac{a_n^s}{s!}f^{(s+1)}(\theta)\int_{-\infty}^{\infty}t^s\,\phi(t)dt\}}{\{f(\theta)\int_{-\infty}^{\infty}[\phi^{(1)}(t)]^2dt\}^{1/2}}$$

converges in distribution to a normal random variable with mean 0 and variance 1.

From this result, we get an asymptotic expression for the mean square error as

$$E\{\theta_n-\theta\}^2=\mathrm{var}\{\theta_n\}+\{E[\theta_n]-\theta\}^2$$

$$\approx \frac{A_2}{na_n^3} + a_n^{2s}B_2 + o(\frac{1}{na_n^3} + a_n^{2s}), \tag{4.21}$$

where $A_2 = \frac{f(\theta)}{[f^{(2)}(\theta)]^2} \int_{-\infty}^{\infty} [\phi^{(1)}(t)]^2 dt$ and $B_2 = \left\{ \frac{f^{(1+s)}(\theta)}{s!f^{(2)}(\theta)} \int_{-\infty}^{\infty} t^s \phi(t)dt \right\}^2$.

Differentiating the dominating terms in the right hand side of (4.21) with respect to a_n and equating to zero gives the optimal choice of a_n as

$$a_{n(op)} = \left\{ \frac{3A_2}{2snB_2} \right\}^{1/[2s+3]} . \tag{4.22}$$

With this choice of a_n the smallest mean square error is

$$E\{\theta_n - \theta\}^2 \approx \left(\frac{2s+3}{3}\right)\left(\frac{3A_2}{2sn}\right)^{2s/[2s+3]} B_2^{3/[2s+3]} . \tag{4.23}$$

4.4 Estimation of the Distribution Function and the Quantiles

In this section we shall assume the following regularity conditions:

Condition (4.4a): f and its first s derivatives are bounded and $f^{(s-1)}(x) \neq 0$.

Condition (4.4b): ϕ is a kernel function belonging to class C_s and having finite support $[-h, h]$.

Condition (4.4c): $\lim_{n\to\infty} na_n^{2(s+1)} = 0$.

Kernel functions having finite support [-1, 1] and belonging to class C_s are derived in Müller (1984).

Define an estimator, $F_n(x)$, of a distribution function $F(x) = \int_{-\infty}^{x} f(u)du$ by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} W\left(\frac{x-X_i}{a_n}\right),$$

where

$$W(t)=\int_{-\infty}^{t}\phi(u)du$$

Now,

$$E\{F_n(x)\}=\int_{-\infty}^{\infty}W\left(\frac{x-u}{a_n}\right)f(u)du$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{x}\frac{1}{a_n}\phi\left(\frac{t-u}{a_n}\right)f(u)dt\,du$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{x}\phi(z)f(t-a_nz)dt\,dz$$

$$=\int_{-\infty}^{\infty}\phi(z)F(x-a_nz)dz.$$

Expanding $F(x-a_nz)$ around x to the order of $(a_nz)^{s+1}$ and using Conditions (4.4a) and (4.4b), we get

$$E\{F_n(x)\}=F(x)+a_n{}^s\frac{f^{(s-1)}(x)}{s!}\int_{-h}^{h}t^s\phi(t)\,dt+O(a_n{}^{s+1}). \tag{4.24}$$

Next, using the fact that ϕ vanishes outside the interval [-h, h], we have

$$E\left\{W\left(\frac{x-X_1}{a_n}\right)\right\}^2=\int_{-\infty}^{\infty}\left\{W\left(\frac{x-u}{a_n}\right)\right\}^2f(u)du$$

$$=\int_{-\infty}^{x-ha_n}f(u)du+\int_{x-ha_n}^{x+ha_n}\left\{W\left(\frac{x-u}{a_n}\right)\right\}^2f(u)du$$

$$=F(x-ha_n)+a_n \int_{-h}^{h}\{W(t)\}^2 f(x-ta_n)dt. \quad (4.25)$$

Expanding $F(x-ha_n)$ around x to the order of $(ha_n)^2$ and expanding $f(x-ta_n)$ around x to the order of (ta_n) then using (4.24), we get

$$\text{var}\{F_n(x)\}=\frac{F(x)[1-F(x)]}{n} - \frac{ua_n}{n} + O\left(\frac{a_n^2}{n}\right), \quad (4.26)$$

where

$$u = f(x)\left\{h - \int_{-h}^{h} W^2(t)dt\right\}.$$

Hence,

$$E\{F_n(x)-F(x)\}^2 \approx \frac{F(x)[1-F(x)]}{n} - \frac{ua_n}{n} + \nu_s a_n^{2s}, \quad (4.27)$$

where

$$\nu_s = \left\{\frac{f^{(s-1)}(x)}{s!} \int_{-h}^{h} t^s \phi(t)dt\right\}^2.$$

Differentiating the right hand side of (4.27) with respect to a_n and equating to zero gives

$$a_n = \left\{\frac{u}{2sn\nu_s}\right\}^{1/(2s-1)}. \quad (4.28)$$

The corresponding minimum MSE is

$$MSE_{a_n}\{F_n(x)\} \approx \frac{F(x)[1-F(x)]}{n} - \frac{(2s-1)}{\nu_s^{1/(2s-1)}}\left(\frac{u}{2sn}\right)^{2s/(2s-1)}. \quad (4.29)$$

Suppose ζ_p is the p-th quantile of the distribution under consideration, and is defined as the unique root of the equation F(x)= p. We consider an estimate x_p of ζ_p given by the root of the equation

$$F_n(x)= p. \tag{4.30}$$

That is, $F_n(x)=p$ if and only if $x=x_p$. Expanding $F_n(x_p)$ around ζ_p, we get

$$F_n(x_p)=F_n(\zeta_p)+(x_p-\zeta_p)f_n(\xi), \tag{4.31}$$

where $|\xi-\zeta_p|<|x_p-\zeta_p|$. Replacing x by ζ_p in (4.24) and (4.26), we get

$$E\{F_n(\zeta_p)\}=p+a_n{}^s v_s{}^{*1/2}+O(a_n{}^{s+1}) \tag{4.32}$$

and

$$\mathrm{var}\{F_n(\zeta_p)\}=\frac{p[1-p]}{n}-\frac{u^* a_n}{n}+O(\frac{a_n{}^2}{n}), \tag{4.33}$$

where the values of u^* and $v_s{}^*$ are the respective values of u and v_s evaluated at $x=\zeta_p$. From (4.30) and (4.31) we may write

$$(x_p-\zeta_p)=\frac{p-F_n(\zeta_p)}{f_n(\xi)}$$

$$=\frac{p-E\{F_n(\zeta_p)\}}{f_n(\xi)}-\frac{F_n(\zeta_p)-E\{F_n(\zeta_p)\}}{f_n(\xi)}. \tag{4.34}$$

Using the computation in (4.32) and the fact that $f_n(\xi)$ converges to $f(\zeta_p)$ in probability as n tends to infinity (see Nadaraya, 1964), we get

$$(x_p-\zeta_p)=-\frac{a_n{}^s v_s{}^{*1/2}+O(a_n{}^{s+1})}{f(\zeta_p)+o_p(1)}-\frac{F_n(\zeta_p)-E\{F_n(\zeta_p)\}}{f(\zeta_p)+o_p(1)}. \tag{4.35}$$

Hence

$$\frac{n^{1/2}\{f(\zeta_p)(x_p-\zeta_p)+a_n{}^s v_s{}^{*1/2}\}}{\{p(1-p)-u^* a_n\}^{1/2}}$$

$$= -\frac{n^{1/2}\{F_n(\zeta_p)-E[F_n(\zeta_p)]\}}{\{p(1-p) - u^* a_n\}^{1/2}}\{1+ o_p(1)\}+O(n^{1/2}a_n^{s+1})\{1+op(1)\}. \quad (4.36)$$

It can be easily shown, using Liapounov's version of the central limit theorem (see Loeve 1963, p. 275), that $\frac{F_n(\zeta_p)-E\{F_n(\zeta_p)\}}{[var\{F_n(\zeta_p)\}]^{1/2}}$ converges in distribution to a normal random variable with mean 0 and variance 1 as n tends to infinity. Using this result in conjunction with the computations in (4.33) and (4.36), we conclude that $\frac{n^{1/2}\{f(\zeta_p)(x_p-\zeta_p)+a_n^s v_s^{*1/2}\}}{\{p(1-p) - u^* a_n\}^{1/2}}$ converges in distribution to a normal random variable with mean 0 and variance 1 as n tends to infinity, since $\lim_{n\to\infty} na_n^{2(s+1)}=0$ by Condition (4.4c).

From this result we now get an asymptotic expression for the mean square error as:

$$E\{x_p-\zeta_p\}^2=var\{x_p\}+\{E[x_p] - \zeta_p\}^2$$

$$\approx \frac{p(1-p)}{nf^{*2}} - \frac{u^* a_n}{nf^{*2}} + \frac{a_n^{2s} v_s^*}{f^{*2}}, \quad (4.37)$$

where $f^*=f(\zeta_p)$. Differentiating the right hand side of (4.37) with respect to a_n and equating to zero gives

$$a_n^* = \left\{\frac{u^*}{2snv_s^*}\right\}^{1/(2s-1)}. \quad (4.38)$$

Under this optimal choice of a_n, we have

$$E\{x_p\} \approx \zeta_p - \frac{sgn(f^{*(s-1)})}{f^* v_s^{*1/2(2s-1)}}\left\{\frac{u^*}{2sn}\right\}^{s/(2s-1)} \quad (4.39)$$

and

$$E\left\{x_p-\zeta_p\right\}^2 \approx \frac{p(1-p)}{nf^{*2}} - \frac{(2s-1)}{f^{*2} v_s^{*1/(2s-1)}}\left(\frac{u^*}{2sn}\right)^{2s/(2s-1)}. \quad (4.40)$$

The results in Azzalini (1981) are special cases of the results obtained in (4.29) and (4.40) when s=2.

CHAPTER V

NONPARAMETRIC ESTIMATION OF THE CONDITIONAL DENSITY FUNCTION AND ITS FUNCTIONALS

5.1 Introduction

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be independent and identically distributed two dimensional random variables with a joint probability density function f(x, y) and a joint distribution function $F(x, y)= \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)dudv$. The marginal density function of X and the conditional density function of Y given X=x are $g(x)= \int_{-\infty}^{\infty} f(x, y)dy$ and $f(y/x)= \frac{f(x,y)}{g(x)}$ respectively, provided g(x)>0. If E[Y] is finite, then the regression function of Y_1 on X_1 is defined as

$$r(x)=E[Y_1/X_1=x]. \tag{5.1}$$

The conditional distribution function of Y_1 given X_1 =x is defined by

$$F(y/x)= \int_{-\infty}^{y} f(u/x)du = \frac{\int_{-\infty}^{y} f(x, u)du}{g(x)}. \tag{5.2}$$

Let $\zeta_{p,x}$ denote the quantile of order p (0<p<1) of the conditional distribution function F(y/x), i.e., a root of the equation $F(\zeta/x)=p$. We call $\zeta_{p,x}$ the population conditional quantile of order p and assume that $\zeta_{p,x}$ is unique. If we also assume that f(y/x) is uniformly continuous in y for each x, then it follows that f(y/x) possesses a mode M(x) defined by

$$f(M(x)/x) = \max_{-\infty<y<\infty} f(y/x). \tag{5.3}$$

We call M(x) the population conditional mode, and assume that M(x) is unique. We also define

$$w(x)= \int_{-\infty}^{\infty} y f(x,y)dy$$

$$v(x)= \int_{-\infty}^{\infty} y^2 f(x,y)dy,$$

and note that the regression function can now be written as $r(x)=\frac{w(x)}{g(x)}$ if $g(x)>0$.

Let $f_n(x, y)$, $g_n(x)$, $f_n(y/x)$ and $F_n(y/x)$ be the estimators of $f(x, y)$, $g(x)$, $f(y/x)$ and $F(y/x)$ respectively, defined by

$$f_n(x,y)=\frac{1}{na_n^2}\sum_{j=1}^{n}\phi\left(\frac{x-X_j}{a_n}\right)\phi\left(\frac{y-Y_j}{a_n}\right) \tag{5.4}$$

$$g_n(x)=\frac{1}{na_n}\sum_{j=1}^{n}\phi\left(\frac{x-X_j}{a_n}\right), \tag{5.5}$$

$$f_n(y/x)=\frac{f_n(x, y)}{g_n(x)} \tag{5.6}$$

and

$$F_n(y/x)= \int_{-\infty}^{y} f_n(u/x)du = \frac{B_n(x, y)}{g_n(x)}, \tag{5.7}$$

where ϕ is a kernel function and $\{a_n\}$ is a sequence of positive numbers converging to zero and

$$B_n(x, y)=\frac{1}{na_n}\sum_{j=1}^{n}\Phi\left(\frac{y-Y_j}{a_n}\right)\phi\left(\frac{x-X_j}{a_n}\right) \tag{5.8}$$

with $\Phi(y)= \int_{-\infty}^{y}\phi(u)du$.

Nadaraya (1964a), Watson (1964), Rosenblatt (1969) and Schuster (1972) have studied asymptotic properties of estimators of r(x) of the form

$$r_n(x)=\frac{\sum_{j=1}^{n} Y_j\phi\left(\frac{x-X_j}{a_n}\right)}{\sum_{j=1}^{n}\phi\left(\frac{x-X_j}{a_n}\right)}. \tag{5.9}$$

Rosenblatt (1969) also studied the properties of an estimate $f_n(y/x)$ of the conditional probability density function $f(y/x)$. Samanta (1989) proved the strong consistency and asymptotic normality of an estimate $\overline{\zeta}_{p,x}$ of the conditional quantile $\zeta_{p,x}$ given by the root of the equation $F_n(\zeta/x)=p$. Samanta and Thavaneswaran (1990) also proved that, under some regularity conditions, an estimate $M_n(x)$ of the conditional mode $M(x)$ given by $f_n(M_n(x)/x)=\max_{-\infty<y<\infty} f_n(y/x)$ is strongly consistent and asymptotically normally distributed.

In the next three sections we compute the mean square errors of the estimates described above. We know, in Chapter IV, that an example of a kernel function belonging to class Cs (s an even number≥2) is given by

$$\phi(x)=\{H_0(x)+d_2H_2(x)+d_4H_4(x)+ \dots +d_{s-2}H_{s-2}(x)\}z(x), \tag{5.10}$$

where $z(x)$ is the standard normal pdf and $H_r(x)$ is the Chebychev-Hermite polynomial of degree r defined by

$$H_r(x)=x^r - \frac{r^{[2]}}{2.1!}x^{r-2} + \frac{r^{[4]}}{2^2.2!}x^{r-4} - \frac{r^{[6]}}{2^3.3!}x^{r-6} + - \dots$$

with $H_0(x)=1$, $r^{[a]}=r(r-1)(r-2)\dots(r-a+1)$ and $d_s=\frac{(-1)^{s/2}}{2^{s/2}(s/2)!}$

(see Cramer 1946, p. 133).

Similarly an example of a kernel function belonging to class C_s (s an odd number ≥ 3) is given by

$$\phi(x)=\{H_0(x)+d_2H_2(x)+d_4H_4(x)+ \dots +d_{s-1}H_{s-1}(x)+\lambda_sH_s(x)\}z(x), \tag{5.11}$$

where $\lambda_s\neq0$ and coefficients d's are as defined before.

5.2 Estimation of the Conditional Probability Density Function and the Regression Function.

We state the regularity conditions under which all asymptotic results in this section will be proved.

Condition (5.2a): $f^{(i,j)}(x,y)$ exist and are bounded for $0 \leq i+j \leq s+1$.

Condition (5.2b): $g^{(i)}(x)$ exist and are bounded for $0 \leq i \leq s+1$.

Condition (5.2c): $w^{(i)}(x)$ exist and are bounded for $0 \leq i \leq s+1$.

Condition (5.2d): $\nu(x)$ and $\nu^{(1)}(x)$ exist and are bounded.

Condition (5.2e): ϕ is a kernel function belonging to class C_s (s an even number≥ 2).

We first prove two simple lemmas:

Lemma 5.1. Under Conditions (5.2a), (5.2b) and (5.2e)

$$\text{(i) } E\{f_n(x,y)\}=f(x,y)+\frac{a_n^s}{s!}[f^{(s,0)}(x,y)+f^{(0,s)}(x,y)] \int_{-\infty}^{\infty} t^s\, \phi(t)dt + O(a_n^{s+1}) \qquad (5.12)$$

$$\text{(ii) } E\{g_n(x)\}=g(x)+\frac{a_n^s}{s!}g^{(s)}(x) \int_{-\infty}^{\infty} t^s\, \phi(t)dt + O(a_n^{s+1}). \qquad (5.13)$$

Poof. We outline the proof of part (i) of the lemma. The proof of part (ii) is similar and is omitted.

We have

$$E\{f_n(x,y)\}=\frac{1}{a_n^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi\left(\frac{x-u}{a_n}\right)\phi\left(\frac{y-v}{a_n}\right)f(u,v)dudv$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(t_1)\phi(t_2)f(x-a_nt_1,y-a_nt_2)dt_1dt_2.$$

The proof is completed by expanding $f(x-a_nt_1,y-a_nt_2)$ around (x,y) to the order of a_n^{s+1} and using Conditions (5.2a) and (5.2e).

Lemma 5.2. Under the conditions of Lemma 5.1, we have

(i) $\mathrm{var}\{f_n(x,y)\}=\frac{f(x,y)}{na_n^2}\{\int_{-\infty}^{\infty}\phi^2(t)dt\}^2+o(\frac{1}{na_n^2})$ (5.14)

(ii) $\mathrm{var}\{g_n(x)\}=\frac{g(x)}{na_n}\int_{-\infty}^{\infty}\phi^2(t)dt+o(\frac{1}{na_n})$ (5.15)

(iii) $\mathrm{cov}\{f_n(x,y),g_n(x)\}=\frac{f(x,y)}{na_n}\int_{-\infty}^{\infty}\phi^2(t)dt+o(\frac{1}{na_n})$. (5.16)

We indicate the proof of part (iii) of the lemma. The proofs of parts (i) and (ii) of the lemma are similar and are omitted.

We have

$$na_n\mathrm{cov}\{f_n(x,y),g_n(x)\}=\frac{1}{a_n^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi^2(\frac{x-u}{a_n})\phi(\frac{y-v}{a_n})f(u,v)dudv$$

$$-a_n[\frac{1}{a_n^2}E\{\phi(\frac{x-X_1}{a_n})\phi(\frac{y-Y_1}{a_n})\}][\frac{1}{a_n}E\{\phi(\frac{x-X_1}{a_n})\}]$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi^2(t_1)\phi(t_2)f(x-a_nt_1,y-a_nt_2)dt_1dt_2+O(a_n),\text{ by Lemma 5.1.}$$

Expanding $f(x-a_nt_1,y-a_nt_2)$ around (x,y) to the order of a_n and using Conditions (5.2a) and (5.2e), we get

$$na_n\,\mathrm{cov}\{f_n(x,y),g_n(x)\}=f(x,y)\int_{-\infty}^{\infty}\phi^2(t)dt+O(a_n).$$

Using the delta method (see Rao, 1973, p. 388) in conjunction with Lemmas 5.1 and 5.2, we get

$$E\left\{\frac{f_n(x,y)}{g_n(x)}-\frac{f(x,y)}{g(x)}\right\}^2\approx\frac{1}{\{g(x)\}^2}E\{(f_n(x,y)-f(x,y)\}^2+\frac{\{f(x,y)\}^2}{\{g(x)\}^4}E\{g_n(x)-g(x)\}^2$$

$$-2\frac{f(x,y)}{\{g(x)\}^3}E\{[(f_n(x,y)-f(x,y)][(g_n(x)-g(x)]\}$$

$$= \frac{1}{\{g(x)\}^2}\Big[\mathrm{var}\{f_n(x,y)\}+\{Ef_n(x,y)-f(x,y)\}^2\Big]$$

$$+\frac{\{f(x,y)\}^2}{\{g(x)\}^4}\Big[\mathrm{var}\{g_n(x)\}+\{Eg_n(x)-g(x)\}^2\Big]$$

$$-\ 2\frac{f(x,y)}{\{g(x)\}^3}\Big[\mathrm{cov}\{f_n(x,y),g_n(x)\}+\{(Ef_n(x,y)\ -\ f(x,y)\}\{Eg_n(x)\ -\ g(x)\}\Big]$$

$$= \frac{A_1}{na_n^2} + a_n^{2s}B_1 + o\Big(\frac{1}{na_n^2} + a_n^{2s}\Big), \tag{5.17}$$

where $A_1 = \frac{f(x,y)}{\{g(x)\}^2}\Big\{\int_{-\infty}^{\infty}\phi^2(t)dt\Big\}^2$ and

$$B_1 = \frac{1}{\{g(x)\}^2(s!)^2}\Big\{\int_{-\infty}^{\infty}t^s\phi(t)dt\Big\}^2\Big[f^{(s,0)}(x,y)+f^{(0,s)}(x,y)\ -\ \frac{f(x,y)g^{(s)}(x)}{g(x)}\Big]^2.$$

Differentiating the dominating terms in the right hand side of (5.17) with respect to a_n and equating to zero gives the optimal value of a_n as

$$a_{n(op)} = \left\{\frac{A_1}{snB_1}\right\}^{1/(2s+2)}. \tag{5.18}$$

With this choice of a_n the minimum mean square error becomes

$$\min_{a_n} E\Big\{\frac{f_n(x,y)}{g_n(x)} - \frac{f(x,y)}{g(x)}\Big\}^2 \approx (s+1)\left(\frac{A_1}{sn}\right)^{s/(s+1)} B_1^{\ 1/(s+1)}. \tag{5.19}$$

In the rest of this section we discuss the estimates of the regression function. We first note that the estimate of the regression function $r_n(x)$ can be written as

$$r_n(x) = \frac{w_n(x)}{g_n(x)},$$

where $\quad w_n(x)=\frac{1}{na_n}\sum_{j=1}^{n}Y_j\phi\left(\frac{x-X_j}{a_n}\right)$

We now prove the following lemma:

Lemma 5.3. Under Conditions (5.2b), (5.2c), (5.2d) and (5.2e), we have

$$\text{(i) } E\{w_n(x)\}=w(x)+\frac{a_n^s}{s!}w^{(s)}(x)\int_{-\infty}^{\infty}t^s\,\phi(t)dt+O(a_n^{s+1}) \qquad (5.20)$$

$$\text{(ii) } \mathrm{var}\{w_n(x)\}=\frac{v(x)}{na_n}\int_{-\infty}^{\infty}\phi^2(t)dt+o(\frac{1}{na_n}) \qquad (5.21)$$

$$\text{(iii) } \mathrm{cov}\{w_n(x),g_n(x)\}=\frac{w(x)}{na_n}\int_{-\infty}^{\infty}\phi^2(t)dt+o(\frac{1}{na_n}). \qquad (5.22)$$

We shall prove part (i) and part (iii) of the lemma. The proof of part (ii) can be completed in a similar manner.

(i) Using Fubini's theorem, we have

$$E\{w_n(x)\}=\frac{1}{a_n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}u\phi(\frac{x-z}{a_n})f(z,u)dz\,du$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}u\phi(t)f(x-a_nt,u)dt\,du$$

$$=\int_{-\infty}^{\infty}w(x-a_nt)\phi(t)dt.$$

The proof of part (i) of the lemma is completed by expanding $w(x-a_nt)$ around x to the order of a_n^{s+1} and using Conditions (5.2c) and (5.2e).

(iii) Using part (ii) of Lemma 5.1 and part (i) of Lemma 5.3 and Fubini's theorem we get

$$na_n\ \mathrm{cov}\{w_n(x),g_n(x)\}=a_n\,\mathrm{cov}\left\{\frac{1}{a_n}Y_1\phi(\frac{x-X_1}{a_n}),\frac{1}{a_n}\phi(\frac{x-X_1}{a_n})\right\}$$

$$=\frac{1}{a_n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\phi^2(\frac{x-z}{a_n})f(z,y)dz\,dy$$

$$-a_n\Big[E\{\frac{Y_1}{a_n}\phi(\frac{x-X_1}{a_n})\}E\{\frac{1}{a_n}\phi(\frac{x-X_1}{a_n})\}\Big]$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\phi^2(t)f(x-a_nt,y)dt\,dy+O(a_n)$$

$$=\int_{-\infty}^{\infty}\phi^2(t)w(x-a_nt)dt+O(a_n).$$

Expanding $w(x-a_nt)$ around x to the order of a_n and using Conditions (5.2c) and (5.2e) we now complete the proof of part (iii) of the lemma. Using the delta method in conjunction with part (ii) of Lemma 5.1, part (ii) of Lemma 5.2 and Lemma 5.3, we get

$$E\{r_n(x)-r(x)\}^2\approx\frac{1}{\{g(x)\}^2}E\{w_n(x) - w(x)\}^2+\frac{\{w(x)\}^2}{\{g(x)\}^4}E\{g_n(x)-g(x)\}^2$$

$$-2\frac{w(x)}{\{g(x)\}^3}E\{[w_n(x)-w(x)][(g_n(x)-g(x)]\}$$

$$=\frac{A_2}{na_n}+a_n^{2s}B_2+o(\frac{1}{na_n}+a_n^{2s}), \qquad (5.23)$$

where $A_2=\frac{1}{\{g(x)\}^2}\int_{-\infty}^{\infty}\phi^2(t)dt\Big[v(x)-\frac{\{w(x)\}^2}{g(x)}\Big]$

and

$$B_2=\frac{1}{\{g(x)\}^2(s!)^2}\Big\{\int_{-\infty}^{\infty}t^s\,\phi(t)dt\Big\}^2\Big[w^{(s)}(x)-\frac{w(x)g^{(s)}(x)}{g(x)}\Big]^2.$$

Differentiating the dominating terms in the right hand side of (5.23) with respect to a_n and equating to zero gives the optimal value of a_n as

$$a_{n(op)}=\Big\{\frac{A_2}{2snB_2}\Big\}^{1/(2s+1)}, \qquad (5.24)$$

With this choice of a_n, the expression for the asymptotic minimum value of $E\{r_n(x)-r(x)\}^2$ is

$$\min_{a_n} E\{r_n(x)-r(x)\}^2 \approx (2s+1)\left(\frac{A_2}{2sn}\right)^{2s/(2s+1)} B_2^{1/(2s+1)}. \tag{5.25}$$

We now consider the computation of the mean square errors of the estimates of the conditional mode and the conditional quantile.

5.3 Estimation of the Conditional Mode

In this section we assume that s is an odd number greater than or equal to 3, and compute an approximate expression for the mean square error of the proposed nonparametric estimator of the conditional mode under the following conditions:

Condition (5.3a): $g(x)$ is uniformly continuous.

Condition (5.3b): $f^{(i,j)}(x,y)=\dfrac{\partial^{i+j}f(x,y)}{\partial x^i \partial y^j}$ exist and are bounded for $0 \le i+j \le s+2$.

Condition (5.3c): ϕ and its first two derivatives are functions of bounded variation.

Condition (5.3d): $\lim_{|u|\to\infty} |u^2\phi^{(i)}(u)| = 0$, $i=0,1$.

Condition (5.3e): ϕ is a kernel function belonging to class C_s.

Condition (5.3f): $a_n=n^{-\delta}$, $\dfrac{1}{2(s+2)} \le \delta < \dfrac{1}{8}$.

We note that the kernel function defined in (5.11) also satisfies Conditions (5.3c) and (5.3d).

The following lemmas will be needed in carrying out the necessary computation:

Lemma 5.4. Under Conditions (5.3b), (5.3c) and (5.3e)

$$E\{f_n^{(0,1)}(x,y)\}=f^{(0,1)}(x,y) - \frac{a_n^s}{s!}[f^{(0,s+1)}(x,y)+f^{(s,1)}(x,y)] \int_{-\infty}^{\infty} t^s\phi(t)dt + O(a_n^{s+1}). \tag{5.26}$$

Proof. Using Fubini's theorem and Conditions (5.3b) and (5.3c), we have by integration by parts

$$E\{f_n^{(0,1)}(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(r)\phi(w)f^{(0,1)}(x-ra_n,y-wa_n)drdw.$$

The proof of the lemma is completed by expanding $f^{(0,1)}(x-ra_n,y-wa_n)$ around (x,y) to the order of a_n^{s+1} and using Conditions (5.3b) and (5.3e).

In Lemma 5.5 below we denote by C(f) the set of continuity points of f.

Lemma 5.5. Under Conditions (5.3c), (5.3d) and (5.3f), if $(x,y)\in C(f)$ and $\lim_{n\to\infty} na_n^2=\infty$, then

$(na_n^4)^{1/2}\left[f_n^{(0,1)}(x,y)-E\{f_n^{(0,1)}(x,y)\}\right]$ converges in distribution to a normal random variable with mean zero and variance $f(x,y)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\phi(u)\phi^{(1)}(v)\}^2dudv.$

Proof. See Samanta and Thavaneswaran (1990).

In the remaining part of this section we shall assume all Conditions from (5.3a) through (5.3f).

We first note that $f^{(0,1)}(x,M(x))=0$ and $f^{(0,2)}(x,M(x))<0$. Expanding $f_n^{(0,1)}(x,M_n(x))$ around (x,M(x)), we get

$$0=f_n^{(0,1)}(x,M_n(x))=f_n^{(0,1)}(x,M(x))+\{M_n(x)-M(x)\}f_n^{(0,2)}(x,M_n^*(x))$$

where $|M_n^*(x) - M(x)| < |M_n(x)-M(x)|$

Hence,

$$\{M_n(x)-M(x)\} = -\frac{f_n^{(0,1)}(x,M(x))}{f_n^{(0,2)}(x,M_n^*(x))}$$

$$=\frac{\left[-E\{f_n^{(0,1)}(x,M(x))\} - \{f_n^{(0,1)}(x,M(x)) -E\{f_n^{(0,1)}(x,M(x))\}\}\right]}{\{f^{(0,2)}(x,M(x))+o_p(1)\}}, \qquad (5.27)$$

by Lemma 4 in Samanta and Thavaneswaran (1990). Using Lemma 5.4 and multiplying both sides of (5.27) by $f^{(0,2)}(x,M(x))$, we get

$$f^{(0,2)}(x,M(x))\{M_n(x)-M(x)\}=$$

$$\Big[\frac{a_n^s}{s!}[f^{(0,s+1)}(x,M(x))+f^{(s,1)}(x,M(x))]\int_{-\infty}^{\infty} t^s\phi(t)dt+O(a_n^{s+1})\Big]\{1+o_p(1)\}$$

$$-\Big[f_n^{(0,1)}(x,M(x))-E\{f_n^{(0,1)}(x,M(x))\}\Big]\{1+o_p(1)\}.$$

Hence

$$\frac{(na_n^4)^{1/2}\Big[f^{(0,2)}(x,M(x))\{M_n(x)-M(x)\}-\frac{a_n^s}{s!}[f^{(0,s+1)}(x,M(x))+f^{(s,1)}(x,M(x))]\int_{-\infty}^{\infty} t^s\phi(t)dt\Big]}{\Big[f(x,M(x))\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\phi(u)\phi^{(1)}(v)\}^2dudv\Big]^{1/2}}$$

$$=\frac{-(na_n^4)^{1/2}\Big[f_n^{(0,1)}(x,M(x))-E\{f_n^{(0,1)}(x,M(x))\}\Big]\{1+o_p(1)\}}{\Big[f(x,M(x))\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\phi(u)\phi^{(1)}(v)\}^2dudv\Big]^{1/2}}+o_p(1).$$

Using Lemma 5.5 we now conclude that $M_n(x)$ is asymptotically normally distributed with mean

$$E\{M_n(x)\}\approx M(x)+\frac{a_n^s}{s!f^{(0,2)}(x,M(x))}[f^{(0,s+1)}(x,M(x))+f^{(s,1)}(x,M(x))]\int_{-\infty}^{\infty} t^s\,\phi(t)dt \tag{5.28}$$

and variance

$$\mathrm{Var}\{M_n(x)\}\approx\frac{f(x,M(x))\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\phi(u)\phi^{(1)}(v)\}^2dudv}{na_n^4\{f^{(0,2)}(x,M(x))\}^2} \tag{5.29}$$

Hence

$$E\{M_n(x)-M(x)\}^2=\mathrm{Var}\{M_n(x)\}+\{E[M_n(x)]-M(x)\}^2$$

$$\doteq\frac{A_3}{na_n^4}+a_n^{2s}B_3+o\Big(\frac{1}{na_n^4}+a_n^{2s}\Big), \tag{5.30}$$

where

$$A_3=\frac{f(x,M(x))}{\{f^{(0,2)}(x,M(x))\}^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\phi(u)\phi^{(1)}(v)\}^2dudv$$

and

$$B_3=\left[\frac{\{f^{(0,s+1)}(x,M(x))+f^{(s,1)}(x,M(x))\}}{s!f^{(0,2)}(x,M(x))}\int_{-\infty}^{\infty}t^s\ \phi(t)dt\right]^2.$$

Differentiating the dominating terms in the right hand side of (5.30) with respect to a_n and equating to zero gives the optimal value of a_n as

$$a_{n(op)}=\left\{\frac{2A_3}{snB_3}\right\}^{1/(2s+4)}, \qquad (5.31)$$

With this choice of a_n, an expression for the asymptotic minimum value of $E\{M_n(x)-M(x)\}^2$ is

$$\min_{a_n} E\left\{M_n(x)-M(x)\right\}^2\approx\frac{(s+2)}{2}\left(\frac{2A_3}{sn}\right)^{s/(s+2)}\ B_3^{\ 2/(s+2)}. \qquad (5.32)$$

5.4 Estimation of the Conditional Quantile

In this section we assume that s is an even number greater than or equal to 2. We list the regularity conditions under which the mean square error of the proposed nonparametric estimator of a conditional quantile is derived.

Condition (5.4a): $F^{(1+i,j)}(x,y)=\frac{\partial^{1+i+j}F(x,y)}{\partial x^{1+i}\partial y^j}$ exist and are bounded for $0\leq i+j\leq s+1$.

Condition (5.4b): $g^{(i)}(x)=\int_{-\infty}^{\infty}\{\partial^i f(x,y)/\partial x^i\}dy$ exist and are bounded for $i=1, 2, ..., s+1$.

Condition (5.4c): ϕ is a function of bounded variation.

Condition (5.4d): ϕ is a kernel function belonging to class C_s.

Condition (5.4e): $a_n = n^{-\delta}$, $\frac{1}{2s+1} \leq \delta < \frac{1}{4}$.

We note that the kernel function defined in (5.10) also satisfies Conditions (5.4c).

We first prove two lemmas:

Lemma 5.6 Under Conditions (5.4a), (5.4b) and (5.4d)

$$\text{(i) } E\{g_n(x)\} = g(x) + \frac{a_n^s}{s!} g^{(s)}(x) \int_{-\infty}^{\infty} t^s \phi(t)dt + O(a_n^{s+1}) \tag{5.33}$$

$$\text{(ii) } E\{B_n(x,y)\} = F^{(1,0)}(x,y) + \frac{a_n^s}{s!}[F^{(1+s,0)}(x,y) + F^{(1,s)}(x,y)] \int_{-\infty}^{\infty} t^s \phi(t)dt + O(a_n^{s+1}). \tag{5.34}$$

Proof The proof of the first part of the lemma follows from direct computation and is omitted. To prove the second part we have by using Fubini's theorem

$$E\{B_n(x,y)\} = \frac{1}{a_n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi\left(\frac{y-v}{a_n}\right)\phi\left(\frac{x-u}{a_n}\right) f(u,v)dudv$$

$$= \frac{1}{a_n^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{y} \phi\left(\frac{t-v}{a_n}\right)\phi\left(\frac{x-u}{a_n}\right) f(u,v)dtdudv$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{y} \phi(z_1)\phi(z_2) f(x-a_n z_1, t-a_n z_2)dtdz_1dz_2$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(z_1)\phi(z_2) F^{(1,0)}(x-a_n z_1, y-a_n z_2)dz_1dz_2.$$

Expanding $F^{(1,0)}(x-a_n z_1, y-a_n z_2)$ around (x,y) to the order of a_n^{s+1} and using Conditions (5.4a) and (5.4d) completes the proof.

Corollary. Under the conditions of Lemma 5.6

$$\frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}} = p + \frac{a_n^s}{g(x)s!}\int_{-\infty}^{\infty} t^s\phi(t)dt\Big[F^{(1+s,0)}(x,\zeta_{p,x}) + F^{(1,s)}(x,\zeta_{p,x}) - pg^{(s)}(x)\Big] + O(a_n^{s+1}).$$

(5.35)

Lemma 5.7. Under Conditions (5.4a), (5.4b), (5.4d) and (5.4e)

$F_n(\zeta_{p,x}/x) = \frac{\{B_n(x,\zeta_{p,x})\}}{\{g_n(x)\}}$ is asymptotically normally distributed with mean $\frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}}$ and variance $\frac{p(1-p)}{na_n g(x)}\int_{-\infty}^{\infty}\{\phi(t)\}^2 dt.$

Proof. We omit the proof which can be completed along the same line as in Lemma 8 and Theorem 2 in Samanta (1989).

In the remaining part of this section we shall assume all Conditions from (5.4a) through (5.4e).

Expanding $F_n(\overline{\zeta}_{p,x}/x)$ around $\zeta_{p,x}$, we get

$$p = F_n(\zeta_{p,x}/x) + (\overline{\zeta}_{p,x} - \zeta_{p,x})f_n(\zeta/x), \tag{5.36}$$

where ζ is some random point between $\overline{\zeta}_{p,x}$ and $\zeta_{p,x}$. Hence,

$$(\overline{\zeta}_{p,x} - \zeta_{p,x}) = \frac{1}{f_n(\zeta/x)}\Big[p - \frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}} - \{F_n(\zeta_{p,x}/x) - \frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}}\}\Big]$$

$$= \frac{1}{\{f(\zeta_{p,x}/x) + o_p(1)\}}\Big[p - \frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}} - \{F_n(\zeta_{p,x}/x) - \frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}}\}\Big], \tag{5.37}$$

by Lemma 6 in Samanta (1989). Using the corollary of Lemma 5.6 and multiplying both sides of (5.37) by $f(\zeta_{p,x}/x)$, we get

$f(\zeta_{p,x}/x)\{\overline{\zeta}_{p,x} - \zeta_{p,x}\} =$

$$\left\{-\frac{a_n^s}{g(x)s!}\int_{-\infty}^{\infty} t^s\,\phi(t)dt\left[F^{(1+s,0)}(x,\zeta_{p,x})+F^{(1,s)}(x,\zeta_{p,x}) - pg^{(s)}(x)\right]+O(a_n^{s+1})\right\}\{1+o_p(1)\}$$

$$-\left\{F_n(\zeta_{p,x}/x) - \frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}}\right\}\{1+o_p(1)\}.$$

Hence

$$\frac{\{na_n g(x)\}^{1/2}}{\{p(1-p)\int_{-\infty}^{\infty}[\phi(t)]^2dt\}^{1/2}}\left[f(\zeta_{p,x}/x)\{\bar{\zeta}_{p,x} - \zeta_{p,x}\}+\right.$$

$$\left.\frac{a_n^s}{g(x)s!}\int_{-\infty}^{\infty} t^s\phi(t)dt\{F^{(1+s,0)}(x,\zeta_{p,x})+F^{(1,s)}(x,\zeta_{p,x}) - pg^{(s)}(x)\}\right]$$

$$=\frac{-\{na_n g(x)\}^{1/2}}{\{p(1-p)\int_{-\infty}^{\infty}[\phi(t)]^2dt\}^{1/2}}\left[F_n(\zeta_{p,x}/x) - \frac{E\{B_n(x,\zeta_{p,x})\}}{E\{g_n(x)\}}\right]\{1+o_p(1)\} + o_p(1).$$

Using Lemma 5.7, we now conclude that $\bar{\zeta}_{p,x}$ is asymptotically normally distributed with mean

$$E\{\bar{\zeta}_{p,x}\}\approx\zeta_{p,x} - \frac{a_n^s}{s!f(x,\zeta_{p,x})}\int_{-\infty}^{\infty} t^s\phi(t)dt\left[F^{(1+s,0)}(x,\zeta_{p,x})+F^{(1,s)}(x,\zeta_{p,x}) - pg^{(s)}(x)\right] \tag{5.38}$$

and variance

$$\mathrm{Var}\{\bar{\zeta}_{p,x}\}\approx\frac{p(1-p)g(x)}{na_n\{f(x,\zeta_{p,x})\}^2}\int_{-\infty}^{\infty}\{\phi(t)\}^2dt. \tag{5.39}$$

Hence

$$E\{\bar{\zeta}_{p,x} - \zeta_{p,x}\}^2=\mathrm{Var}\{\bar{\zeta}_{p,x}\}+\{E[\bar{\zeta}_{p,x}] - \zeta_{p,x}\}^2$$

$$\approx\frac{A_4}{na_n} + a_n^{2s}B_4+o\left(\frac{1}{na_n} + a_n^{2s}\right), \tag{5.40}$$

where

$$A_4 = \frac{p(1-p)g(x)}{\{f(x,\zeta_{p,x})\}^2} \int_{-\infty}^{\infty} \{\phi(t)\}^2 dt$$

$$B_4 = \frac{1}{\{s!f(x,\zeta_{p,x})\}^2} \left\{ \int_{-\infty}^{\infty} t^s \phi(t) dt [F^{(1+s,0)}(x,\zeta_{p,x}) + F^{(1,s)}(x,\zeta_{p,x}) - pg^{(s)}(x)] \right\}^2.$$

Differentiating the dominating terms in the right hand side of (5.40) with respect to a_n and equating to zero gives the optimal choice of a_n as

$$a_{n(op)} = \left\{ \frac{A_4}{2snB_4} \right\}^{1/(2s+1)}. \tag{5.41}$$

With this choice of a_n, an expression for the asymptotic minimum value of $E\{\bar{\zeta}_{p,x} - \zeta_{p,x}\}^2$ is

$$\min_{a_n} E\left\{ \bar{\zeta}_{p,x} - \zeta_{p,x} \right\}^2 \approx (2s+1)\left(\frac{A_4}{2sn}\right)^{2s/(2s+1)} B_4^{1/(2s+1)}. \tag{5.42}$$

REFERENCES

1. Adichie, J.N. (1967). Asymptotic Efficiency of a Class of Nonparametric Tests for Regression Parameters, *Annals of Mathematical Statistics*, 38, 884-893.
2. Ahmad, I.A. (1976). On Asymptotic Properties of an Estimate of a Functional of a Probability Density, *Scandinavian Actuarial Journal*, 176-181.
3. Ahmad, I.A. and Lin, Pi-erh (1977). Nonparametric Estimation of Vector-valued Bivariate Failure Rate, *Annals of Statistics,* 5, 1027-1038.
4. Aitken, C. and MacDonald, D.G. (1979). An Application of Discrete Kernel Methods to Forensic Odontology, *Applied Statistics*, 28, 55-67.
5. Azzalini, A. (1981). A Note on the Estimation of a Distribution Function and Quantiles by a Kernel Method, *Biometrika*, 68, 326-328.
6. Bhattacharyya, G.K. and Roussas, G.G. (1969). Estimation of a Certain Functional of a Probability Density Function, *Skandinavisk Aktuarietidskrift*, 3-4, 201-206.
7. Bhattacharya, P.K. (1967). Estimation of a Probability Density Function and Its Derivatives, *Sankhyā Series A* , 29, 373-382.
8. Cacoullos, T. (1966). Estimation of a Multivariate Density, *Annals of the Institute of Statistical Mathematics*, 18, 178-189.
9. Copas, J.B. and Fryer, M.J. (1980). Density Estimation and Suicide Risks in Psychiatric Treatment, *Journal of the Royal Statistical Society Series A*, 143, 167-176.
10. Cramer, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press.
11. Devroye, Luc (1987). *A Course in Density Estimation*, Birkhäuser, Boston.
12. Devroye, Luc and Györfi, L. (1985). *Nonparametric Density Estimation: The L_1 View*, John Wiley and Sons, Toronto.
13. Dmitriev, Yu. G. and Tarasenko, F.P. (1973). On the Estimation of Functionals of the Probability Density and Its Derivative, *Theory of Probabilty and Its*

Applications, 18, 628-633.

14. Dmitriev, Yu. G. and Tarasenko, F.P. (1974). On a Class of Non-parametric Estimates of Non-linear Functionals of Density, *Theory of Probabilty and Its Applications*, 19, 390-394.

15. Eddy, W.F. (1980). Optimum Kernel Estimators of the Mode, *Annals of Statistics*, 8 870-882.

16. Epanechnikov, V.A. (1969). Nonparametric Estimates of Multivariate Probability Density, *Theory of Probabilty and Its Applications*, 14, 153-158.

17. Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York.

18. Fraser, D.A.S. (1957). *Nonparametric Methods in Statistics*, J. Wiley and Sons, New York.

19. Gasser, T., Müller, H.G., Köhler, W., Molinari, L., and Prader, A. (1984). Nonparametric Regression Analysis of Growth Curves, *Annals of Statistics*, 12, 210-229.

20. Hodges, J.L., Jr and Lehmann, E.L. (1956). Efficiency of Some Nonparametric Competitors of the t-test, *Annals of Mathematical Statistics*, 27, 324-335.

21. Hodges, J.L. and Lehmann, E.L. (1963). Estimates of Location Based on Rank Tests, *Annals of Mathematical Statistics,* 34, 598-611.

22. Hoeffding, W. (1948). A Class of Statistics with Asymptotically Normal Distribution, *Annals of Mathematical Statistics*, 19, 293-325.

23. Hoeffding, W. (1963). Probability Inequalities for Sums of Bounded Random Variables, *Journal of the American Statistical Association*, 58, 13-30.

24. Jureckova, J. (1971). Nonparametric Estimate of Regression Coefficients, *Annals of Mathematical Statistics*, 42, 1328-1338.

25. Karunamuni, R.J. (1991). A New Smoothness Quantification in Kernel Density Estimation, *Journal of Statistical Planning and Inference*, 27, 361-373.

26. Karunamuni, R.J. and Mehra, K.L. (1990). Improvements on Strong Uniform Consistency of some known Kernel Estimates of a Density and Its Derivatives, *Statistics & Probabilty Letters*, 9, 133-140.

27. Karunamuni, R.J. and Mehra, K.L. (1991). Optimal Convergence Properties of Kernel Density Estimators without Differentiability Conditions, *Annals of the Institute of Statistical Mathematics*, 43, 327-346.

28. Kendall, M.G. and Stuart, A. (1969). *The Advanced Theory of Statistics* , Vol.1. Hafner Pub. Co., New York.

29. Lehmann, E.L. (1963). Nonparametric Confidence Interval for a Shift Parameter, *Annals of Mathematical Statistics*, 34, 1507-1512.

30. Loeve, M. (1963). *Probability Theory*, Third Edition, Von Nostrand, New York.

31. Mann, H.B. and Wald A. (1943). On Stochastic Limit and Order Relationships, *Annals of Statistics*, 14, 217-226.

32. Müller, H.G. (1984). Smooth Optimum Kernel Estimators of Densities, Regression Curves and Modes, *Annals of Statistics*, 12, 766-774.

33. Müller, H.G. (1988). *Nonparametric Regression Analysis of Longitudinal Data.: Lecture Notes in Statistics* , No. 40, Springer-Verlag, Heidelberg.

34. Murthy, V.K. (1965). Estimation of Jumps, Reliability and Hazard Rate, *Annals of Mathematical Statistics*, 36, 1032-1040.

35. Nadaraya, È.A. (1964a). On Estimating Regression, *Theory of Probility and Its Applications,* 9, 141-142.

36. Nadaraya, È.A. (1964b). Some New Estimates for Distribution Functions, *Theory of Probability and Its Applications*, 15, 497-500.

37. Nadaraya, È.A. (1965). Non-parametric Estimates of Density Functions and Regression Curves, *Theory of Probabilty and Its Applications*, 10, 186-190.

38. Nadaraya, È.A. (1970). Remarks on Nonparametric Estimates for Density Functions and Regression Curves, *Theory of Probability and Its Applications*, 15,

134-136.

39. Nadaraya, È.A. (1974). On the integral Mean Square Error of Some Nonparametric Estimates For Density Function, *Theory of Probability and Its Applications*, 19, 133-141.

40. Nadaraya, È.A. (1989). *Nonparametric Estimation of Probability Densities and Regression Curves*, Kluwer Academic Publishers, Boston.

41. Noether, G.E. (1972). Distribution- Free Confidence Intervals, *The American Statistician*, 26, 39-41.

42. Parzen, E. (1962). On Estimation of a Probability Density Function and Mode, *Annals of Mathematical Statistics*, 33, 1065-1076.

43. Prakasa Rao, B.L.S. (1983) *Nonparametric Functional Estimation*, Academic Press, New York.

44. Pratt, J.W. (1959). On A General Concept of 'In Probability', *Annals of Mathematical Statistics*, 30, 549-558.

45. Ramlau-Hansen, H. (1983). Smoothing Counting Process Intensities by means of Kernel Functions, *Annals of Statistics,* 11, 453-466.

46. Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*, Wiley, New York.

47. Rosenblatt, M. (1956). Remarks on Some Non-parametric Estimates of a Density Function, *Annals of Mathematical Statistics*, 27, 832-837.

48. Rosenblatt, M. (1969). Conditional Probability Density and Regression Estimates, *Multivariate Analysis* II (Krishnaiah, P.R., Editor), 25-31.

49. Rosenblatt, M. (1975). A Quadratic Measure of Deviation of Two Dimensional Density Estimates and Tests of Independence, *Annals of Statistics*, 3, 1-14.

50. SAS Institute, Inc. (1982). *SAS User's Guide: Statistics*, Cary, NC: Author.

51. Samanta, M. (1971). Optimal Inference of the Partial Regression Coefficients, *Calcutta Bulletin of Statistical Association*, 20, 109-133.

52. Samanta, M. (1973). Nonparametric Estimation of the Mode of a Multivariate Density, *South African Statistical Journal*, 7, 109-117.

53. Samanta, M. (1974a). A Note on Uniform Strong Convergence of Bivariate Density Estimates, *Z. Wearscheinlichkeitsttheorie Verw. Gebiete*, 28, 85-88.

54. Samanta, M. (1974b). Efficient Nonparametric Estimation of a Shift Parameter, *Sankhyā Series A*, 36, 273-292.

55. Samanta, M. (1989). Non-parametric Estimation of Conditional Quantiles, *Statistics & Probability Letters*, 7, 407-412.

56. Samanta,.M. and Thavaneswaran, A. (1990). Non-parametric Estimation of the Conditional Mode, *Communications in Statistics Series A- Theory and Methods*, 19, 4515-4524.

57. Schuster, E.F. (1969). Estimation of a Probability Density Function and its Derivatives, *Annals of Mathematical Statistics*, 40, 1187-1195.

58. Schuster, E.F. (1972). Joint Asymptotic Distribution of the Estimated Regression Function at a Finite Number of Distinct Points, *Annals of Mathematical Statistics*, 43, 84-88.

59. Schuster, E.F. and Yakowitz, S. (1979). Contribution to Theory of Nonparametric Regression with Applications to System Identification, *Annals of Statistics,* 7, 139-145.

60. Schweder, T. (1975). Window Estimation of the Asymptotic Variance of Rank Estimators of Location, *Scandinavian Journal of Statistics*, 2, 113-126.

61. Sen, P.K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau, *Journal of the American Statistical Association,* 63, 1379 -1389.

62. Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis,* Chapman and Hall, London.

63. Singh, R.S. (1977a). Improvement on some known Nonparametric

Uniformly Consistent Estimators of Derivatives of a Density, *Annals of Statistics,* 5, 394-400.

64. Singh, R.S. (1977b). Application of Estimators of Density Function and Its Derivatives to certain Statistical Problems, *Journal of the Royal Statistical Society Series B*, 39, 359-363.

65. Singh, R.S. (1979). On Necessary and Sufficient Conditions for Uniform Strong Consistency of Estimators of a Density and its Derivatives, *Journal Multivariate Analysis*, 9, 157-164.

66. Srivastava M.S. and Saleh, A.K.Md.E. (1970). On a Class of Nonparametric Estimates for Regression parameters, *Journal of Statistical Research*, 4, 133-139.

67. Tetterington, D.M. (1980). A Comparative study of Kernel-Based Density Estimates for Categorical Data, *Technometrics*, 22, 259-269.

68. Thavaneswaran, A. (1988). Smoothing Signals for Semimartingales, *Stochastic Processes and Applications,* 28, 81-89.

69. Theil, H. (1950a). Rank-Invariant Method of Linear and Polynomial Regression Analysis I, *Nederlandse Akademie van Wetenschappen Proceedings, Series A*, 53, 386-392.

70. Theil, H. (1950b). Rank-Invariant Method of Linear and Polynomial Regression Analysis III, *Nederlandse Akademie van Wetenschappen Proceedings, Series A*, 53, 1397-1412.

71. Van Ryzin, J. (1969). On Strong Consistency of Density Estimates, *Annals of Mathematical Statistics*, 40, 1765-1772.

72. Watson, G.S. (1964). Smooth Regression Analysis, *Sankhyā Series A*, 26, 359-372.

73. Woodroofe, M.B. (1970). On Choosing a Delta-Sequence, *Annals of Mathematical Statistics*, 41, 1665-1671.

BIOGRAPHICAL DATA

Leo Odiwuor Odongo was born in Uholo, Siaya District, Kenya, on February 21, 1955. He spent his first eleven years of education at Rangala, and Ramba both in Siaya District, Kenya, and a further two years at Amukura High School in Busia District, Kenya. In September, 1977, he joined Kenyatta University College, then a constituent college of the University of Nairobi, from where he graduated in May, 1980, with a B.Ed (First Class Honors), majoring in mathematics and statistics. In September, 1980, he joined the University of Nairobi for an M.Sc. degree in Mathematical Statistics and graduated in 1983. From September, 1982 to July, 1984, he taught at Kenya Polytechnic in Nairobi. In August, 1984, he joined Kenyatta University as an Assistant Lecturer until August, 1988, when he came to Canada. He joined the University of Manitoba in September, 1988, and graduated from there with an M.Sc. in statistics in 1990. Since then he has been working for his Ph.D. degree in statistics at the University of Manitoba.