NONPARAMETRIC AND ROBUST PROCEDURES FOR BIOEQUIVALENCE MODELS WITH MEASUREMENT ERRORS

by

Ho Kim

Department of Biostatistics
University of North Carolina

Institute of Statistics
Mimeo Series No. 2170T

October 1996

Nonparametric and Robust Procedures for Bioequivalence Models with Measurement Errors

by

Ho Kim

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics.

Chapel Hill

1996

Approved by:

P. K. Sen Advisor

Lloyd J. Edwards Reader

M. J. Symons Reader

Suchindran Reader

Bahjat Qaqish Reader

Abstract

In a biological assay (bioassay, for short), a new preparation and an old one are compared by means of the reactions that follow their applications to biological organisms, and the relative potency of the test preparation with respect to the standard one is the main interest under investigation. One typically assumes that the test preparation behaves as if it is a dilution (or concentration) of the standard one which is called the fundamental assumption of a bioassay. Our main interest is to estimate the relative potency and to verify the fundamental assumption. The usual maximum likelihood estimator has many optimal properties such as consistency and asymptotic normality under the assumed distribution. But these parametric estimates are generally not very robust against departures from the assumed form of distribution function. Rank based estimates of relative potency have been considered by Sen (1963), Shorack (1966) and Rao and Little (1976), among others. These estimates are invariant under the choice of any monotone transformation on the dose, and, besides being robust, are generally quite efficient for normal, log-normal, logistic or other common forms of the tolerance distributions. Under the (generalized) linear models set-up, to verify the fundamental assumption one typically tests the null hypothesis of the same slope for the two preparations. And to estimate the relative potency, one typically estimates the intercepts of the two dose-response curves from the two preparations. One of the difficulties of those analyses of the bioequivalence model is the assumed form of the expected value of the response as a function of the dose.

The conditional quantile process(CQP) is a very useful tool to verify the fundamental assumption in this situation. The k-nearest neighborhood estimator of the conditional quantile has been studied by Bhattacharya and Gangopadhyay (1988) and Sen and Gangopadhyay (1990). In our paper, these two methods are compared in terms of asymptotic properties. Also several bias correction methods are proposed. The proposed CQP is asymptotically a Gaussian process which is fully character-

ized by pointwise asymptotic normality and covariance structures for only shrinking neighborhood of fixed points. The proposed CQP has useful asymptotic properties and generates a smooth nonparametric dose-response curve. In addition, the CQP does not require any functional form of the expected value. To estimate the relative potency, a Kolmogorov-Smirnov type estimator is proposed along with.

In a clinical trial, some biological variables like cholesterol level can be considered as a function of some other variables like blood pressure. We don't have clear-cut evidence of linear normality of the response variables with respect to covariates. An additional problem is that the measurements of the covariates are not accurate; they include biases such as measurement bias, laboratory bias, worker bias and bias due to the data processing. In a classical parametric analysis, if we have measurement errors, it is well known that the estimates are biased. In this paper it is proved that the asymptotic properties of the CQP hold in the presence of small measurement errors. The difference of the quantiles between the two cdf's (one without measurement error and one with measurement error) is the order of $O(||\epsilon||^2)$ under the assumption of the uniform measurement error distribution ($U(-\epsilon, \epsilon)$).

Univariate CQP is generalized to the multivariate CQP for $(q \times 1)$ vector of covariates. We show that a similar expression (Bahadur type) of conditional quantile holds in this case by using contiguity. We apply our method to ARIC (Atherosclerosis Risk in Communities) study data. In the first of the two examples, we compare the hypertension group and the normotension group in terms of the carotid artery far wall thickness (WT). The WT curves over Body Mass Index (BMI) are compared. In the second example, we investigate the effect of smoking on the WT curve. Our conditional quantile method has useful applications and interpretations for the examples.

Acknowledgment

First and foremost I would like to thank Dr. Pranab K. Sen. This dissertation would not have been possible without his guidance and patience. Many thanks go to my committee, Drs. Chirayath M. Suchindran, Michael J. Symons, Bahjat F. Qaqish and Lloyd J. Edwards. I wish to extend my sincere appreciation to Drs. Lloyd E. Chambless and C. Ed Davis for providing me with the opportunity to work at the coordinating center. I also thank my fellow students, particularly my office mates: PaiLien Chen, Hildete Pinheiro, Jinwhan Jung and Charmaine Marquis; they have been a constant source of encouragement. The next round of applause is for Dr. Louis Wijnberg without whom I would have lost my sense of humor and my way through the days of my research. I would like to thank my parents and parents-in-law for their encouragement and support over the years. Last, but not least, I sincerely appreciate the support and understanding of my wife, Kyujeong, who has remained by my side.

Contents

List of Figures

List of Tables

Chapter 1

Literature Review and Introduction

1.1 Introduction

In a biological assay (*bioassay*), a new (test) preparation and an old (standard) one are compared by means of the reactions that follow their applications to biological organisms. The main interest under investigation is to examine how biological organisms react to the two preparations and to evaluate the relative potency of the test preparation with respect to the standard. To achieve this goal, a stimulus is applied to a subject, and changes in the organism (*response*) are measured. The magnitude of this response may depend on the amount of the stimulus (*dose*). Furthermore. the relationship between the dose and the response may not be exact, and it is subject to random variation in the application of the dose, in the manifestation of the response. or in both. The response to a stimulus may be classified as quantitative or quantal. For example, quantitative responses can be measured. as would be the case if we were studying weight changes following the use of a certain drug. On the other hand. quantal responses have two outcomes such as death or survival. There are three main types of biological assays that are commonly used for numerical evaluation of potencies: (1) Direct assays, (2) indirect assays with quantitative responses, and (3) indirect assays with quantal responses. In a *direct assay*, the basic principle is that doses of the standard and test preparations, sufficient to produce a specific response, are directly measured so that the response is certain while the dose is the random variable. The ratio of these two doses estimates the potency of the test preparation relative to the standard. Assays (2) and (3) are similar with respect to their statistical analysis. Both make use of *dose-response regression relationships*. In an indirect assay, for given levels of doses, the responses are measured. Thus, the response is a random variable whose distribution may depend on dose.

Direct assays have shortcomings, especially the bias introduced by the time-lag in producing the response. Even if this difficulty is overcome, it is not easy to administer the precise dosage that produces the desired response. Let x denote the dose and Y denote the response it induces. If S and T are abbreviations for standard and test preparations, respectively, then the regression relationships for the two preparations can be denoted by

$$E(Y_S) = h_S(x) \quad \text{and} \quad E(Y_T) = h_T(x). \tag{1.1}$$

From (1.1) one can obtain the doses x_T and x_S that would induce the average response y. More precisely, the doses x_S and x_T are equally effective. In a typical *indirect model*, one assumes that the test preparation behaves as if it is a dilution (or concentration) of the standard one, i.e. the potency ratio x_S/x_T is independent of y. Therefore,

$$x_S/x_T = \rho, \quad \text{hence} \tag{1.2}$$

$$h_T(x) = h_S(\rho x) \quad \text{for all } x, \text{ and } \rho > 0. \tag{1.3}$$

The positive constant ρ is termed *the relative potency* of the test preparation with respect to the standard one. The model (1.3) is termed *the fundamental assumption* of an assay. The main interest is to estimate the relative potency ρ and to verify the fundamental assumption. Parametric procedures for these problems are usually based on the assumptions of the distributions of Y (normal, lognormal, logistic or loglogistic), these procedures are discussed in detail in Finney (1964). Using (generalized) linear models, the usual maximum likelihood estimator has many optimal properties, such as consistency and asymptotic normality under the assumed distribution. However, these parametric estimates are generally not very robust against departures from the assumed form of the distribution function. In univariate linear models, testing equation (1.3) examines the equality of the slopes and the equality of the intercepts. Rank based estimates of relative potency have been considered by Sen (1963), Shorack (1966) and Rao and Little (1976), among others. These estimates are invariant under the choice of any monotone transformation on the dose, and, besides being robust, are generally quite efficient for normal, log-normal, logistic or other common forms of the tolerance distributions. Recently, a method to obtain a confidence interval of a conditional quantile was proposed by Sen and Gangopadhyay (1990). If we use the conditional median or quantile of the response instead of the expectation in (1.1), we can test the equivalence of a test preparation with respect to a standard one. Because the conditional quantile has a nonparametric flavor, it will be robust than the distribution-dependent expected value. We do not need to perform separate tests for slopes and intercepts as we usually do in dose-response parallel-line bioassay. After taking logarithms of the doses, the dose-response model becomes

$$M^T(x) = M^S(x+m) \text{ for all } x \text{ and for some } m \qquad (1.4)$$

where M^T and M^S are the average responses for the test and standard preparations. The fundamental assumption (1.4) can be better understood by Figure 1.1. To test (1.4) is to test parallelism of the two response curves, and to estimate m is to estimate the drift m under the assumption of (1.4).

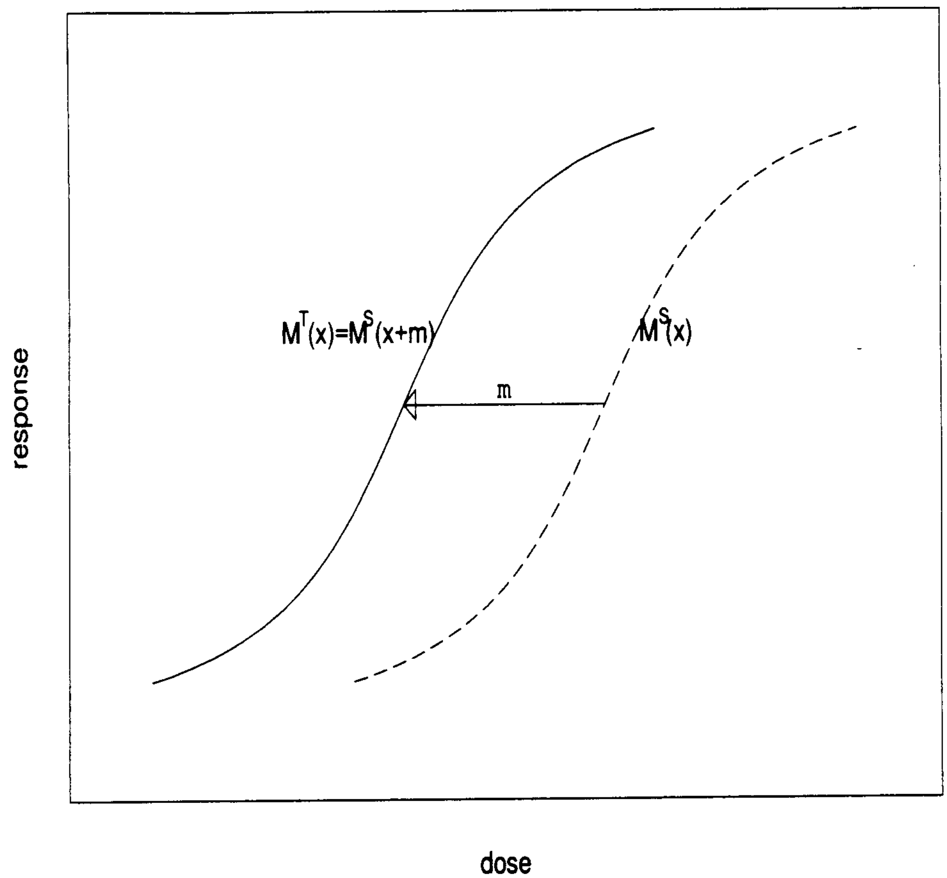

Figure 1.1: Graphical explanation of the bioequivalence model

1.2 Literature review

1.2.1 Parametric bioassay

Bioequivalence is an important area of pharmaceutical research. The key question is whether two or more formulations containing the same active ingredient are therapeutically equivalent. For example, a manufacturer of a new tabulate drug might be interested in knowing whether his or her tabulate formulation can be envisaged in comparing with the most readily available form, for example, a solution. Westlake (1979) and Metzler (1974) compared several statistical methods for bioequivalence study. The experimental design usually follows familiar paths employing either parallel groups. in which subjects in each group receive just one formulation. or the cross-over concept discussed by Grizzle (1965) and Wallenstein and Fisher (1977). The frequently used statistical method is repeated measurements experiments. Some of the more general methods are: the univariate split-plot ANOVA (Box (1950). Westlake (1974), Wallenstein and Fisher (1977)), MANOVA (Cole and Grizzle (1966)). analysis of growth curves (Grizzle and Allen (1969)) and analysis of principal components (Snee (1972)).

One of the main concerns of the bioequivalence study is to establish the equivalence between two formulations. Therefore, testing a null hypothesis of no difference is completely irrelevant to the demand of the situation. Thus, obtaining confidence intervals rather than testing a hypothesis would be more appropriate. Using confidence intervals leads to a different approach to the problems of choosing type I and type II error levels (Jones and Karson (1972), Westlake (1972)). The clinician or pharmacologist can specify that the bioequivalence of the new formulation relative to the standard one must be within a certain range. For example. the FDA guidelines for Bioavailability Studies state that "products whose rate and extent of absorption differ by 20 % or less generally considered bioequivalent" (Westlake (1979)).

To establish the bioequivalence, the most frequently used method is to estimate the confidence interval of the ratio of two responses, one from the test preparation and one from the standard preparation. Fieller's theorem referred by Govindarajulu (1988) is the classical method of obtaining a confidence interval of a ratio. Let μ. ν be two unknown parameters (here, they are doses of the test and standard preparations which produce a specific response) and let $\rho = \mu/\nu$. Let a and b be unbiased estimators for μ and ν, respectively. Assume that a and b are linear in observations that are normally distributed. Let $R = a/b$ be an estimate of ρ. Then. the upper and lower confidence limits for ρ are given by Fieller's Theorem (Fieller (1940) and Fieller (1944)), i.e.,

$$R_L, R_U = \frac{1}{1-g} \times$$
$$\left\{ R - \frac{g v_{12}}{v_{22}} \pm \frac{ts}{b} \left[v_{11} - 2Rv_{12} + R^2 v_{22} + R^2 v_{22} - g\left(v_{11} - \frac{v_{12}^2}{v_{22}} \right) \right]^{1/2} \right\}.$$

where $g = t^2 s^2 v_{22}/b^2$ and s^2 is an error mean square having m degrees of freedom and the estimates of the variances and covariance of a and b are $v_{11}s^2$, $v_{22}s^2$ and $s^2 v_{12}$, respectively; also $t = t_{m,1-\alpha/2}$. Therefore, we obtain a confidence interval of relative potency by Fieller's Theorem.

When we are interested in the difference rather than the ratio, we can use Behren's distribution. Let a and b be independent and unbiased estimates of μ and ν and be linear combinations of observations that are normally distributed. Also assume that estimates of Var(a) and Var(b) are given by

$$\hat{\text{Var}}(a) = s_1^2 v_{11}, \;\; \hat{\text{Var}}(b) = s_2^2 v_{22},$$

where s_1^2 and s_2^2 are independent mean squares with m_1, m_2 degrees of freedom, respectively. Then the estimated variance of $a - b$ is

$$s_{a-b}^2 = s_1^2 v_{11} + s_2^2 v_{22}.$$

Let

$$D = \frac{(a-b) - (\mu - \nu)}{(s_1^2 v_{11} + s_2^2 v_{22})^{1/2}}.$$

Then D has Behren's distribution (D is also called Sukhatme's D-statistic) which is tabulated by Finney (1964) . Its distribution is defined in terms of the degrees of freedom m_1 and m_2 and the angle θ where

$$\tan \theta = (s_1^2 v_{11}/s_2^2 v_{22})^{1/2}.$$

Next, let a and b denote independent estimates of $\mu = \nu$. Then we consider the weighted estimate given by

$$\begin{aligned} \bar{a} &= [a(s_1^2 v_{11})^{-1} + b(s_2^2 v_{22})^{-1}][(s_1^2 v_{11})^{-1} + (s_2^2 v_{22})^{-1}]^{-1} \text{ and} \\ s_{\bar{a}}^2 &= [(s_1^2 v_{11})^{-1} + (s_2^2 v_{22})^{-1}]^{-1}. \end{aligned}$$

Yates (1939) and Finney (1951) point out that

$$(\bar{a} - \mu)[(s_1^2 v_{11})^{-1} + (s_2^2 v_{22})^{-1}]^{-1/2}$$

follows the Behrens' distribution with d.f. m_1 and m_2 and an angle θ defined by

$$\tan\theta = (s_2^2 v_{22}/s_1^2 v_{11})^{1/2}.$$

Hence, one can test the significance of $\bar{a}$ from a theoretical value μ and also set up a confidence intervals of μ as

$$\bar{a} \pm d_{1-\alpha/2}[(s_1^2 v_{11})^{-1} + (s_2^2 v_{22})^{-1}]^{-1/2},$$

where $d_{1-\alpha/2}$ denotes the tabulated value for the probability $1 - \alpha/2$.

1.2.2 NN method

Let $\{(X_i, Y_i), i \geq 1\}$ be a sequence of i.i.d. random vectors with a distribution function $\pi(x, y), (x, y) \in R^2$ (here, X_i is the dose and Y_i is the response of ith subject). Let $F(x) = \pi(x, \infty)$, $x \in R$ and let $G(y|x)$ be the conditional d.f. of Y given $X = x$, for $y \in R$, $x \in R$. A conditional quantile function (of Y given $X = x$) is defined by

$$\xi_p(x) = \inf\{y : G(y|x) \geq p\},\ x \in R \ \ (0 < p < 1). \tag{1.5}$$

The following regularity conditions are assumed to be true:

(A1) The d.f. F admits an absolutely continuous density function f, such that

(a) $$f(x_0) = f_0 > 0$$

and

(b) $f''(x) = (d^2/dx^2)f(x)$ exists in a neighborhood of x_0, and there exist positive numbers ϵ and k_0, such that $|x - x_0| \le \epsilon$ implies that $|f''(x) - f''(x_0)| \le k_0|x - x_0|$.

(A2) The d.f. $G(y|x_0)$ has a continuous density function $g(y|x_0)$ for all y close to ξ_0. such that

(a) $$g(\xi_0|x_0) > 0 \text{ where } G(\xi_0|x_0) = p,$$

and

(b) the partial derivatives $g_y(y|x)$ and $g_{xx}(y|x)$ of $g(y|x)$ and $G_{xx}(y|x)$ of $G(y|x)$ exist in a neighborhood of (x_0, ξ_0), and there exist positive constants ϵ and k_0, such that $|x - x_0| \le \epsilon$ and $|y - \xi_0| \le \epsilon$ together imply that

$$\begin{aligned}
|g_y(y|x)| \le k_0,\ |g_x(y|x_0)| &\le k_0,\ |g_{xx}(y|x_0)| \le k_0, \\
|g_{xx}(y|x) - g_{xx}(y|x_0)| &\le k_0|x - x_0|, \\
|G_{xx}(y|x) - G_{xx}(y|x_0)| &\le k_0|x - x_0|.
\end{aligned}$$

In the sequel, we shall write $G(y|x_0) = G(y)$ and $g(y|x_0) = g(y)$.

For given p and x_0, let $\{(X_i, Y_i), i \ge 1\}$ be a sequence of i.i.d. random vectors. Consider the transformation $(X_i, Y_i) \to (Z_i, Y_i)$ where $Z_i = |X_i - x_0|$. For the collection $\{(Z_1, Y_1), \cdots, (Z_n, Y_n)\}$ of r.v.'s, let $Z_{n1} < \cdots < Z_{nn}$ be the order statistics corresponding to $Z_1, \cdots, Z_n$, and let $Y_{n1}, \cdots, Y_{nn}$ be the induced order statistics. (i.e.. $Y_{ni} = Y_j$ if $Z_{ni} = Z_j$, for $i, j = 1, \cdots, n$). For every positive integer $k (\le n)$, the $k - NN$ (nearest neighbor) empirical d.f. of Y (with respect to x_0) is given by

$$\hat{G}_{nk}(y) = k^{-1} \sum_{j=1}^{k} 1(Y_{nj} \le y),\ y \in R, \tag{1.6}$$

where $1(A)$ stands for the indicator function of the set A. The following estimators of ξ_0 are due to Bhattacharya and Gangopadhyay (1988).

(i) The $k-NN$ estimator of ξ_0 is

$$\begin{aligned}\hat{\xi}_{nk} &= \text{the } [kp]\text{-th order statistic of } Y_{n1},\cdots,Y_{nk} \\ &= \inf\{y : \hat{G}_{nk}(y) \geq k^{-1}[kp]\}. \qquad (1.7)\end{aligned}$$

(ii) The kernel estimator (with uniform kernel and bandwidth h) is

$$\tilde{\xi}_{nh} = \inf\{y : \hat{G}_{nK_n(h)}(y) \geq [K_n(h)]^{-1}[pK_n(h)]\} \qquad (1.8)$$

where

$$K_n(h) = \sum_{i=1}^{n} 1\left(Z_i \leq \frac{1}{2}h\right)$$

is a positive integer valued r.v. Note that by (1.7) and (1.8),

$$\tilde{\xi}_{nh} = \hat{\xi}_{nK_n(h)} \qquad (1.9)$$

where h is usually non-stochastic and $K_n(h)$ is stochastic in nature. In order that $\hat{\xi}_{nk}$ and $\tilde{\xi}_{nh}$ are consistent estimators of ξ_0, we need that as $n \to \infty$, $k = k_n \to \infty$ (and $h = h_n \downarrow 0$). For the asymptotic normality results, usually, one needs that $k_n = O(n^{\lambda}) = nh_n$, for some $\lambda \in (0,1)$. Bhattacharya and Gangopadhyay (1988) studied the case of $\lambda = 4/5$ which we call NN (nearest neighborhood) method. The following are some of their theorems stated without proofs.

Theorem 1.1 *For* $k \in I_n(a,b) = \{k : k_0 = [n^{4/5}a] \leq k \leq [n^{4/5}b] = k_1, \ 0 < a < b\}$

$$\hat{\xi}_{nk} - \xi = \beta(\xi)(k/n)^2 + \{kg(\xi)\}^{-1}\sum_{i=1}^{k}[1(Y_{ni}^* > \xi) - (1-p)] + R_{nk}$$

where

$$\begin{aligned}\beta(\xi) &= -[f(x_0)G_{xx}(\xi|x_0) + 2f'(x_0)G_x(\xi|x_0)]/\{23f^3(x_0)g(\xi)\}, \\ Y_{ni}^* &= G^{-1} \circ G^*(Y_{ni}|Z_{ni}), \ 1 \leq i \leq n, \ n \geq 1,\end{aligned}$$

$$\max_{k\in I_n(a,b)} |R_{nk}| = O(n^{-3/5}\log n), \quad a.s.$$

and

$$\begin{aligned}
G_x(y|x) &= (\partial/\partial x)G(y|x), \\
G_{xx}(y|x) &= (\partial^2/\partial x^2)G(y|x), \\
G^*(y|z) &= \frac{1}{f(x_0+z)+f(x_0-z)} \\
&\quad \times[f(x_0+z)G(y|x_0+z)+f(x_0-z)G(y|x_0-z)]. \\
G(y) &= G^*(y|0) = G(y|x_0).
\end{aligned}$$

Theorem 1.2 *For any* $0 < a < t < b < \infty$, *as* $n \to \infty$

$$\{n^{2/5}[\hat{\xi}_{n,[n^{4/5}t]} - \xi] - \beta t^2, a \le t \le b\} \xrightarrow{\mathcal{D}} \{\sigma t^{-1}B(t),\ a < t < b\},$$

where $\beta = \beta(\xi)$ *given in Theorem 1.1,* $\sigma^2 = p(1-p)/g^2(\xi)$ *and* $\{B(t), t > 0\}$ *denote a standard Brownian motion.*

Theorem 1.3 *For every* $t \in [a,b]$ *as* $n \to \infty$

$$n^{2/5}[\hat{\xi}_{n,[n^{4/5}t]} - \xi] \xrightarrow{\mathcal{D}} N(\beta t^2, p(1-p)t^{-1}/g(\xi)^{-2}).$$

1.2.3 δNN method

Sen and Gangopadhyay (1990) showed the bias term (βt^2) in theorems 1.2 and 1.3 can be eliminated when $k_n = O(n^{4/5-2\delta})$. We call their approach the δNN (δ nearest neighborhood) method. The followings are their main results.

Theorem 1.4 *For* k_n *lying in*

$$I_n(a,b) = \{k|k_{n_0} = [n^{4/5-2\delta}a] \le k \le [n^{4/5-2\delta}b] = k_{n_1}\}$$

$$\hat{\xi}_{n,k_n}(x) - \xi(x) = \beta(\xi(x))(k_n/n)^2 + \{k_n g(\xi(x)|x)\}^{-1}\sum_{i=1}^{k}[1(Z^*_{ni} > \xi(x)) - (1-p)] + R_{nk_n}, \tag{1.10}$$

where

$$\beta(\xi(x)) = -[f(x)G_{xx}(\xi(x)|x) + 2f'(x)G_x(\xi(x)|x)]/\{24f^3(x)g(\xi(x)|x)\}, \tag{1.11}$$

$$Z^*_{ni} = G^{-1}\circ G^*(Z_{ni}|Y_{ni}); i \le n, \tag{1.12}$$

and

$$\max_{k\in I_n(a,b)} |R_{nk}| = O(n^{-3/5+3\delta/2}\log n) \ a.s.$$

Theorem 1.5 : *Under assumptions (A1) and (A2), for every k : $k = [tn^{4/5-2\delta}]$, $t \in [a.b]$,*

$$n^{2/5-\delta}[\hat{\xi}_{nk} - \xi_0] \xrightarrow{\mathcal{D}} N(0, t^{-1}p(1-p)/g^2(\xi_0)), \ \textit{as}\ n \to \infty.$$

Theorem 1.6 : *Under assumptions (A1) and (A2), for every $t \in [c,d]$ as $n \to \infty$*

$$n^{2/5-\delta}[\tilde{\xi}_{n[tn^{-1/5-2\delta}]} - \xi_0] \xrightarrow{\mathcal{D}} N(0, p(1-p)t^{-1}[f(x_0)g(\xi_0)]^{-2}).$$

They also considered the construction of (local) bootstrap confidence intervals for ξ_0 based on both the $k-NN$ and kernel methods, and their asymptotic properties. For this purpose, in the $k-NN$ procedure, a bootstrap sample $(Y^{**}_{n1}, \cdots, Y^{**}_{nk})$ is obtained from the empirical d.f. $\hat{G}_{nk}$, defined by (1.6). Thus given the induced order statistics $(Y^*_{n1}, \cdots, Y^*_{nk})$, the Y^{**}_{ni} are conditionally i.i.d. r.v. with the d.f. $\hat{G}_{nk}$. Similarly, in the kernel method, given the Y_{ni} and $h(>0)$, we define $K_n(h)$ by (1.9), and then a bootstrap sample of size $K_n(h)$ is obtained by resampling from all the Y_{ni}

for which $Z_i = |X_i - x_0| \leq h/2$. For the k-NN procedure, let $\hat{G}^*_{nk}$ be the bootstrap sample d.f. (based on $Y^*_{n1}, \cdots, Y^*_{nk}$). Then, the bootstrap estimator is

$$\begin{aligned}
\hat{\xi}^*_{nk} &= [kp]\text{-th order statistic of } Y^*_{n1}, \cdots, Y^*_{nk} \\
&= \inf\{y : \hat{G}^*_{nk}(y) \geq k^{-1}[kp]\}.
\end{aligned}$$

Similarly, in the kernel method, the bootstrap estimator is

$$\tilde{\xi}^*_{nk} = \inf\{y : \hat{G}^*_{nK_n(h)}(y) \geq \{K_n(h)\}^{-1}[pK_n(h)]\},$$

where $K_n(h) = \sum_{i=1}^n 1(Z_i \leq h/2)$ is conditionally held fixed. Moreover, paralleled to Theorem 1.5 and 1.6, we have the following asymptotic representation for the bootstrap estimates.

Theorem 1.7 : *Under (A1) and (A2), for each $t \in [a, b]$,*

$$n^{2/5-\delta}(\hat{\xi}^*_{n[tn^{4/5-2\delta}]} - \hat{\xi}_{n[tn^{4/5-2\delta}]}) \xrightarrow{\mathcal{D}} N(0, t^{-1}p(1-p)/g^2(\xi_0)),$$

for every $\delta : 0 < \delta < 1/10$, and every $0 < a < b < \infty$.

Theorem 1.8 : *Under (A1) and (A2), for each $t \in [c, d]$,*

$$n^{2/5-\delta}(\tilde{\xi}^*_{n[tn^{-1/5-2\delta}]} - \tilde{\xi}_{n[tn^{-1/5-2\delta}]}) \xrightarrow{\mathcal{D}} N(0, p(1-p)t^{-1}[f(x_0)g(\xi_0)]^{-2})$$

for every $\delta : 0 < \delta < 1/10$, and every $0 < c < d < \infty$.

One can generate a set of M bootstrap samples. From each sample, one can compute the bootstrap estimates, and use the empirical quantiles (for the $\hat{\xi}^*_{nk} - \hat{\xi}_{nk}$ or $\tilde{\xi}^*_{nh} - \tilde{\xi}_{nh}$) to set the desired bootstrap confidence intervals for ξ_0. The asymptotic properties of such bootstrap intervals would then follow Theorems 1.7 and 1.8.

1.3 Discussion

We note that the parametric and nonparametric bioassay of a dose-response model assume the functional forms of the expected values of response with respect to doses (i.e., the response is linear to the dose); however, there are several situations where such a model does not seem to be appropriate, especially from the point of robustness. In the present paper, we shall consider the methodology which uses the conditional quantile instead of conditional mean to provide a more robust approach.

In chapter 2, we will discuss our methodology: conditional quantile process and its asymptotic properties. And the test of the fundamental assumption and the estimation of the relative potency are of primary interests. Let $\xi_{pT}(Y|x)$ be the conditional $p \times 100\%$ quantile of the test response (Y) for given dose (x) as defined as in (1.5). Similarly, let $\xi_{pS}(Y|x)$ be the conditional quantile of the standard response. We will test the fundamental assumption that the test and standard preparations are parallel with a constant difference μ (relative potency) in terms of conditional quantiles as follows

$$Ho: \quad \xi_{pT}(Y|x) = \xi_{pS}(Y|x+\mu),$$

for all p, $(0 < p < 1)$ and x, $(x \in R)$ for some μ, $(\mu \in R)$. Or, it may be of interest to fix $p = 1/2$ to test medians. If Ho is not rejected, we want to make a statistical inference about μ which we introduced in earlier sections. One of the possible ways to get an estimate of μ is to take the distance between $\xi_{pT}(Y|x)$ and $\xi_{pS}(Y|x+\mu)$ by shifting x, i.e.,

$$\inf_{\mu} \left\{ \sup_{x} |m_{pT}(\cdot|x) - m_{pS}(\cdot|x+\mu)| \right\}. \tag{1.13}$$

Then $\hat{\mu}$ is the value which gives the infimum value of (1.13).

The δNN method $(k_n = O(n^{4/5-2\delta}))$ and the NN method $(k_n = O(n^{4/5}))$ are compared in terms of the asymptotic mean square errors. The δNN method gives unbiased estimate but larger variance. We propose bias correction methods for the

NN method which give smaller variance of the estimate.

In chapter 3, we discuss the measurement errors relating to the covariate. Measurement errors are important in environmental or animal experiments. For example, an investigator may conduct an experiment to measure the effects of fat and smoking on rats. He or she may observe rats in a chamber, controlling food and air condition to examine whether the given fat in the food and gas in the air are the exact same as the amounts in a rat's body. A rat may not eat some food and, it may not inhale the same amount of gas in the given air. In these situations, measurement errors may play a significant role in a statistical analysis. Within a small range of measurement errors, our procedure can be robust against the measurement errors by virtue of the nonparametric property. However when measurement errors are not small we may need a special procedure to control them. In other words, robustness or efficiency of our test depends on the measurement errors' characteristics; therefore, we need to explore the role of measurement errors in our robust bioassay procedures.

In general, we observe $X^* = X + u$ instead of X where u is the added error term. Consider $X^* = X + u$, where u is a continuous random variable. We have $Z_i^* = |X^* - x_0|$ instead of $Z_i = |X_i - x_0|$ and $Y_{n1}^*, \cdots, Y_{nn}^*$ corresponding to $Z_{n1}^*, \cdots, Z_{nn}^*$. And our estimates become

$$\begin{aligned}
\hat{G}_{nk}^*(y) &= k^{-1} \sum_{j=1}^{k} 1(Y_{nj} \leq y),\ y \in R, \\
\hat{\xi}_{nk}^* &= \inf\{y : \hat{G}_{nk}^*(y) \geq k^{-1}[kp]\}, \text{ and} \\
\tilde{\xi}_{nh}^* &= \inf\{y : \hat{G}_{nK_n(h)}^* \geq [K_{n(h)}]^{-1}[pK_{n(h)}]\}.
\end{aligned}$$

We also are interested in the effects of the measurement errors on our estimates under some assumptions of the distribution of u: normal or continuous and symmetric about zero. And we are also interested in the impact of the measurement errors on our asymptotic results.

In chapter 4, we will discuss the multivariate process. If we have a $(q \times 1)$ vector of covariates, we need to modify our method for this situation. We show that a similar expression (Bahadur type) of conditional quantile holds in this case by using contiguity.

In chapter 5, we will apply our method to ARIC (Atherosclerosis Risk in Communities) study data. In the first example, we compare the hypertension group and the control group in terms of the carotid artery far wall thickness (WT): the WT curves over Body Mass Index (BMI) are compared. In the second example, we investigate the effect of smoking on the WT curve. Our conditional quantile method has nice applications and interpretations as seen in the examples.

1.4 Significance of the study

In a classical parametric bioassay, several questions may arise. MLE has good asymptotic properties under the assumption of the distribution. However, it is well known that MLE's are not very robust against the departure of the assumed form of the distribution. Another difficulty of the parametric analysis is the form of the expected value. In some situations, such a model does not seem to be appropriate.

A nonparametric approach is needed to overcome such problems. The conditional quantile process is a very useful tool in this situation; it has good asymptotic properties and provides a clear dose-response curve without any assumption of parametric form. And it is proved that with some measurement errors, our asymptotic properties still hold.

In a clinical trial, some biological variables, such as cholesterol level, can be considered as a function of some other variables like blood pressure. We do not have clear evidence of linear normality of the response variables with covariates; other obstacle is that the measurements of the covariates are not accurate: they include biases such as measurement tool bias, laboratory bias, worker bias or bias due to the data processing. In a classical parametric analysis, if we have measurement errors, it is well known that the estimates are biased. But our method still works asymptotically with measurement errors. The proposed conditional quantile method works nicely in this situation, and have a clear interpretation without parametric assumption.

Chapter 2

Testing the Fundamental Assumption

2.1 Local asymptotic behavior of a conditional quantile process

Let $\xi(x)$ be the p-quantile of the d.f. $G(\cdot|x)$, $x \in [x_l, x_u] = \mathcal{X} \subset \mathbf{R}$, and let $\hat{\xi}_{n,k_n}(x)$ be the $k-NN$ estimator where Z_{ni} are defined with respect to the pivot x. Note that these Z_{ni} are the induced order statistics corresponding to the nearest neighbors of x (i.e. those X_i's which are closest to x). As such, if we consider two distinct points, say x_1 and x_2, then the two sets of X_i forming the corresponding nearest neighbors would have at most k_n members in common. As such, the proportion of the two subsets' common elements will be asymptotically $o(n^{-1}k_n)$, and we may immediately say that $\hat{\xi}_{n,k_n}(x_1)$ and $\hat{\xi}_{n,k_n}(x_2)$ are asymptotically independent. We set $W_n(x) = \{W_n(x), x \in [x_l, x_u]\} = \mathcal{X}$ by letting

$$W_n(x) = k_n^{1/2}\left[\hat{\xi}_{n,k_n}(x) - \xi(x)\right], \quad x \in \mathcal{X}.$$

Then the asymptotic behavior of W_n is depicted entirely by the pointwise asymptotic normality and the stochastic dependence pattern only in a shrinking neighborhood of fixed points.

We discussed the pointwise asymptotic normality in section 1.2.2 for $k_n = tn^{4/5-2\delta}$, $(0 < a < t < b < \infty)$. Then we have

$$W_n(x) \xrightarrow{\mathcal{D}} W(x), \text{ for every } x \in \mathcal{X},$$

where $W(x)$ is normal with mean 0 and variance $pq/(g^2(\xi(x)|x))$.

About the stochastic dependency, we refer to Gangopadhyay and Sen (1993). Shrinking neighborhoods are defined by the sets of x relative to a given x_0 as

$$I_n(x_0, K) = \{x : |x - x_0| \leq Kn^{-1}k_n\},$$

where K is a finite positive number. For any two points, say, x_n and x'_n, belonging to $I_n(x_0, K)$, let $k(x_n, x'_n)$ be the number of elements in common between the two neighbor sets. Consider the intersection set $I_n(x_n, x'_n)$:

$$[x_n - (2nf(x_0))^{-1}k_n, x_n + (2nf(x_0))^{-1}k_n] \cap [x'_n - (2nf(x_0))^{-1}k_n, x'_n + (2nf(x_0))^{-1}k_n].$$

Let $m_n(x_n, x'_n)$ be the length of this intersection set $I_n(x_n, x'_n)$. Then by Bahadur representation, we may again claim that

$$n^{-1}k(x_n, x'_n) - f(x_0)m_n(x_n, x'_n) = o(n^{-1}k_n) \ a.s. \text{ as } n \to \infty. \tag{2.1}$$

And we may conclude that $W_n(x_n)$ and $W_n(x'_n)$ have asymptotic correlation

$$nf(x_0)m_n(x_n, x'_n)/k_n, \tag{2.2}$$

and it is easy to show that for both x_n and x'_n belonging to $I_n(x_0, K)$, (2.2) has a limit (which vanishes when $m_n(x_n, x'_n) = 0$). In order to justify this asymptotic analysis, we need to extend (1.10) for a suitable convergence sequence $\{x_n\}$ (to x_0), where $x_n \in I_n(x_0, K)$. In this context, we need to replace in (1.10), x by x_n (everywhere) and to show that this holds uniformly in $I_n(x_0, K)$.

To facilitate this, we consider a stochastic process

$$W_n^0(t) = W_n(x_0 + n^{-1}k_n t),\ t \in [t_0, t_1].$$

Consider a compact interval $C = [-K, K]$ where $K(< \infty)$ is a positive number so chosen that $Kf(x_0) \leq 1$. In passing, we may remark that (2.2) posesses a limit, when we set $x_n = x_0 + n^{-1}k_n t_1$ and $x_n' = x_0 + n^{-1}k_n t_2$, with $t_1, t_2 \in C$. This is given by

$$[t_1 f(x_0) - 1/2, t_1 f(x_0) + 1/2] \cap [t_2 f(x_0) - 1/2, t_2 f(x_0) + 1/2]$$

which, for $t_2 > t_1$ is non-null when $(t_2 - t_1)f(x_0) \leq 1$, and that is the reason we set $Kf(x_0) \leq 1$.

Theorem 2.1 *If Z_{ni}^* in (1.12) has finite fourth moment then*

$$W_n^0(C) = \{W_n^0(t),\ t \in C\}$$

is asymptotically tight, where $W_n^0(t) = W_n(x_0 + n^{-1}k_n t)$ and $C = [-K, K]$ such that $Kf(x_0) \leq 1$, for $k_n = O(n^{4/5-2\delta})$.

Proof: Let $W_i = W_n^0(x_0 + n^{-1}k_n t_i)$ for $i = 1, 2, 3$ and $t_1 < t_2 < t_3$. From Theorem 2.1, we have

$$\begin{aligned} W_n(t_j) &= k_n^{1/2}\left(\hat{\xi}_{n,k_n}(x_0 + n^{-1}k_n t_j) - \xi(x_0 + n^{-1}k_n t_j)\right) \\ &= k_n^{1/2}\{\beta(\xi_{nj})(k_n/n)^2 + (k_n g(\xi_{nj}))^{-1}\sum_{i=1}^{k}[1(Z_{jni}^* > \xi_{nj}] \\ &\quad -(1-p)] + R_{jnk_n}\},\ \text{for } j = 1, 2, 3, \end{aligned}$$

where

$$\begin{aligned} \xi_{nj} &= \xi(x_0 + n^{-1}k_n t_j), \\ Z_{jni}^* &= G^{-1} \circ G^*(Z_{jni}|Y_{jni}); i \leq n, \end{aligned}$$

for

$$Y_{jni} = ||X_{ni} - (x_0 + n^{-1}k_n t_j)||$$

and

$$\max_{k\in I_n(a,b)} |R_{jnk}| = O(n^{-3/5-3\delta/2}\log n)\ a.s.$$

for $j = 1, 2$ and 3. And we have

$$\begin{aligned} W_n(t_2) - W_n(t_1) &= k_n^{1/2}\{(\beta(\xi_{n2}) - \beta(\xi_{n1}))(k_n/n)^2 \\ &\quad k_n^{-1}\sum_{i=1}^{k}[J_{2ni} - J_{1ni}] + (R_{2nk_n} - R_{1nk_n})\}, \end{aligned}$$

where

$$J_{jni} = \frac{1}{g(\xi_{nj}|x_0 + n^{-1}k_n t_j)}[1(Z^*_{jni} > \xi_{nj}) - (1-p)].$$

And

$$\begin{aligned} (W_n(t_2) - W_n(t_1))^2 &= k_n\{(\beta(\xi_{n2}) - \beta(\xi_{n1}))^2(k_n/n)^4 \\ &\quad + k_n^{-2}\left[\sum_{i=1}^{k}[J_{2ni} - J_{1ni}]\right]^2 + (R_{2nk_n} - R_{1nk_n})^2 \\ &\quad + 2(\beta(\xi_{n2}) - \beta(\xi_{n1}))(k_n/n)^2 k_n^{-2}\left[\sum_{i=1}^{k}[J_{2ni} - J_{1ni}]\right] \\ &\quad + 2(\beta(\xi_{n2}) - \beta(\xi_{n1}))(k_n/n)^2(R_{2nk_n} - R_{1nk_n}) \\ &\quad + 2k_n^{-2}\left[\sum_{i=1}^{k}[J_{2ni} - J_{1ni}]\right](R_{2nk_n} - R_{1nk_n})\}. \end{aligned}$$

In $E[(W_n(t_3) - W_n(t_2))^2(W_n(t_2) - W_n(t_1))^2]$ everything goes to zero as n goes to infinity other than

$$\frac{1}{k_n^2}E\left[\left(\sum_{i=1}^{k} J_{2ni} - J_{1ni}\right)^2\left(\sum_{i=1}^{k} J_{3ni} - J_{2ni}\right)^2\right]$$

$$= \frac{1}{k_n^2} E\left[\left(\sum_{i=1}^{k}(J_{2ni}-J_{1ni})^2+\sum_{i\neq j}(J_{2ni}-J_{1ni})(J_{2nj}-J_{1nj})\right)\right.$$

$$\left.\times\left(\sum_{i=1}^{k}(J_{3ni}-J_{2ni})^2+\sum_{i\neq j}(J_{3ni}-J_{2ni})(J_{3nj}-J_{2nj})\right)\right].$$

Let

$$E\left[\sum_{i=1}^{k}(J_{2ni}-J_{1ni})^2\sum_{i=1}^{k}(J_{3ni}-J_{2ni})^2\right]$$

$$= E\left[\sum_{i=1}^{k}(J_{2ni}-J_{1ni})^2(J_{3ni}-J_{2ni})^2+\sum_{i\neq j}(J_{2ni}-J_{1ni})^2(J_{3nj}-J_{2nj})^2\right]$$

$$= A + B,$$

$$E\left[\sum_{i=1}^{k}(J_{2ni}-J_{1ni})^2\sum_{i\neq j}(J_{3ni}-J_{2ni})(J_{3nj}-J_{2nj})\right]$$

$$= C,$$

and

$$E\left[\sum_{i\neq j}(J_{2ni}-J_{1ni})(J_{2nj}-J_{1nj})\sum_{i\neq j}(J_{3ni}-J_{2ni})(J_{3nj}-J_{2nj})\right]$$

$$= D.$$

And let κ_{nij} be the number of common elements around $x_0+n^{-1}k_nt_i$ and $x_0+n^{-1}k_nt_j$, then

$$\begin{aligned} A &= E\sum\left[J_{2i}^2J_{3i}^2+J_{2i}^4-2J_{2i}^2J_{3i}+J_{1i}^2J_{3i}^2+J_{1i}^2J_{3i}^2-2J_{1i}^2J_{2i}J_{3i}\right. \\ &\quad \left.-2J_{1i}J_{2i}J_{3i}^2-2J_{1i}J_{2i}^3+4J_{1i}J_{2i}^2J_{3i}\right] \\ &= O(k_n+\kappa_{n12}+\kappa_{n23}+\kappa_{n13}+\kappa_{n123}), \end{aligned}$$

$$
\begin{aligned}
B &= E\left[\sum_{i\neq j} J_{2i}^2 J_{3j}^2 + J_{2i}^2 J_{2j}^2 - 2J_{2i}^2 J_{2j} J_{3j} + J_{1i}^2 J_{3j}^2 + J_{1i}^2 J_{2j}^2 \right.\\
&\quad \left. -2J_{1i}^2 J_{2j} J_{3j} - 2J_{1i} J_{2i} J_{3j}^2 - 2J_{1i} J_{2i} J_{2j}^2 + 4J_{1i} J_{2i} J_{2j} J_{3j}\right]\\
&= O(k_n^2 + k_n \kappa_{n12} + k_n \kappa_{n23} + \kappa_{n12}\kappa_{n23})\\
&= O(\kappa_{n12}\kappa_{n23}),\\
C &\to 0\\
D &= E\left[\sum_{i\neq j}(J_{2i} - J_{1i})(J_{2j} - J_{1j}) \sum_{i\neq j}(J_{3i} - J_{2i})(J_{3j} - J_{2j})\right]\\
&= E\left[\sum_{i\neq j\neq k\neq l}(J_{2i} - J_{1i})(J_{2j} - J_{1j})(J_{3k} - J_{2k})(J_{3l} - J_{2l})\right]\\
&\quad + E\left[\sum_{i\neq j\neq k}(J_{2i} - J_{1i})(J_{2j} - J_{1j})(J_{3i} - J_{2i})(J_{3k} - J_{2k})\right]\\
&\quad + E\left[\sum_{i\neq j\neq k}(J_{2i} - J_{1i})(J_{2j} - J_{1j})(J_{3j} - J_{2j})(J_{3k} - J_{2k})\right]\\
&\quad + E\left[\sum_{i\neq j}(J_{2i} - J_{1i})(J_{2j} - J_{1j})(J_{3i} - J_{2i})(J_{3j} - J_{2j})\right]\\
&= 0 + E\left[\sum_{i\neq j}(J_{2i}J_{3i} - J_{1i}J_{3i} - J_{2i}^2 + J_{1i}J_{2i})(J_{2j}J_{3j} - J_{1j}J_{3j} - J_{2j}^2 + J_{1j}J_{2j})\right]\\
&= O(k_n^2 + k_n\kappa_{n23} + k_n\kappa_{n12} + k_n\kappa_{n23} + k_n\kappa_{n12} + \kappa_{n23}k_n\kappa_{n12}\\
&\quad + k_n\kappa_{n13} + \kappa_{n23}k_n\kappa_{n13})\\
&= O(\kappa_{n12}\kappa_{n23}).
\end{aligned}
$$

So

$$E\left[\left(\sum_{i=1}^{k} J_{2ni} - J_{1ni}\right)^2 \left(\sum_{i=1}^{k}(J_{3ni} - J_{2ni}\right)^2\right]$$

$$= O\left(\frac{\kappa_{n12}\kappa_{n23}}{k_n^2}\right)$$

$$= O((t_2 - t_1)(t_3 - t_2))$$

for $-K < t_1 < t_2 < t_3 < K$. By the theorem 2.3.3 by Sen (1981), we have shown the asymptotic tightness. $\diamond$

We have proved

1) $\{W_n(t_1), \cdots, W_n(t_m)\} \rightarrow \{W(t_1), \cdots, W(t_m)\}$ for all $t_j \in [-K, K]$, $j = 1, 2, \cdots, m$ as n goes to ∞.

2) Asymptotic tightness of $\{W_n(t), t \in [-K, K]\}$.

Those two conditions guarantee the weak convergence of $W_n(t)$ to $W(t)$ for all $t \in [-K, K]$ by Theorem 8.2.1 of Sen and Singer (1993).

2.2 Kolmogorov-Smirnov type estimator of the relative potency

The fundamental assumption of bioequivalence model is

$$\text{Ho} : \xi^T(x) = \xi^S(x + \mu) \text{ for all } x \in \mathcal{X} \subset \mathbf{R} \text{ and for some } \mu \in \mathcal{M} \subset \mathbf{R}. \qquad (2.3)$$

where $\mathcal{X}$ and $\mathcal{M}$ are closed intervals of $\mathbf{R}$. In this section, we discuss how to estimate μ and how to test the fundamental assumption. To estimate μ, we consider Kolmogorov-Smirnov type estimator μ which minimize

$$\sup_{x \in \mathcal{X}} |\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x + \mu)|.$$

We have

$$k_n^{1/2}\left(\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x+\mu)\right)$$

$$= k_n^{1/2}\left(\hat{\xi}^T(x) - \hat{\xi}^S(x+\mu)\right) + \left(W_n^T(x) - W_n^S(x+\mu)\right).$$

Under (2.3),

$$k_n^{1/2}\sup_{x\in\mathcal{X}}|\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x+\mu)|$$

$$= \sup_{x\in\mathcal{X}}|W_n^T(x) - W_n^S(x+\mu)|.$$

Let

$$a_n^*(\mu) = \sup_{x\in\mathcal{X}}|\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x+\mu)|,$$

then we need to find μ which minimize $a_n^*(\mu)$ with respect to μ which is equivalent to minimize

$$a_n(\mu) = \sup_{x\in\mathcal{X}}|W_n^T(x) - W_n^S(x+\mu)|$$

with respect to μ.

We can estimate μ by changing x and μ numerically.

Lemma 2.1 *Let*

$$a_n(\mu) = \sup_{x\in\mathcal{X}}|W_n^T(x) - W_n^S(x+\mu)| \text{ for } \mu \in \mathcal{M}$$

and

$$a(\mu) = \sup_{x\in\mathcal{X}}|W^T(x) - W^S(x+\mu)| \text{ for } \mu \in \mathcal{M}$$

then $a_n(\mu)$ converges to $a(\mu)$ in distribution for any $\mu \in \mathcal{M}$. So $\min_\mu a_n(\mu)$ converges to $\min_\mu a(\mu)$ in distribution.

Proof: For any $\mu \in \mathcal{M}$ and $a > 0$,

$$P(a_n(\mu) < a)$$

$$
\begin{aligned}
&= \mathrm{P}\left\{\sup_x |W_n^T(x) - W_n^S(x+\mu)| < a\right\} \\
&= \mathrm{P}\left\{\left|W_n^T(x) - W^T(x)\right| < a \text{ for all } x \in \mathcal{X}\right\} \\
&\to \mathrm{P}\left\{|W^T(x) - W^S(x+\mu)| < a \text{ for all } x \in \mathcal{X}\right\} \text{ as } n \to \infty \\
&= \mathrm{P}\left\{a(\mu) < a\right\}.
\end{aligned}
$$

◇

We may replace the conditional quantiles to the conditional d.f. because of the following lemma.

Lemma 2.2 $\xi^T(x) = \xi^S(x+\mu)$ *iff* $G_T(\cdot|x) = G_S(\cdot|x+\mu)$.

Proof: $(\Leftarrow)$

If $G_T(\cdot|x) = G_S(\cdot|x+\mu)$ then

$$
\begin{aligned}
& G_S(\xi^S(x)|x) = p \\
\Rightarrow\ & G_T(\xi^S(x)|x-\mu) = p \\
\Rightarrow\ & G_T(\xi^S(x+\mu)|x) = p \\
\Rightarrow\ & \xi^S(x+\mu) = \xi^T(x).
\end{aligned}
$$

$(\Rightarrow)$

If $\xi^T(x) = \xi^S(x+\mu)$ then

$$
\begin{aligned}
& G_S(\xi^S(x+\mu)|x+\mu) = p \\
\Rightarrow\ & G_S(\xi^T(x)|x+\mu) = p,
\end{aligned}
$$

and by the definition of the conditional d.f.

$$G_T(\xi^T(x-\mu)|x-\mu) = p.$$

So $G_T(\cdot|x-\mu) = G_S(\cdot|x)$. ◇

Let n, m be the sample sizes of the test and standard preparations. We may get Kolmogorov-Smirnov type estimator $\hat{\mu}$ which is defined by the equation

$$D_{n,m}(\hat{\mu}) = \min_{\mu} D_{n,m}(\mu),$$

where

$$D_{n,m}(\mu) = \sup_{x} |G_{n_T,k_{n_T}}(\cdot|x) - G_{m_S,k_{m_S}}(\cdot|x+\mu)|$$

where $G_{n_T,k_{n_T}}$ and $G_{m_S,k_{m_S}}$ are the $k-NN$ empirical df of Y^T and Y^S w.r.t. x and $x+\mu$, respectively. We may use the method proposed by Rao, Schuster and Little (1975) by replacing order statistics to induced order statistics, i.e., let

$$\begin{aligned} D^+_{n,m}(a) &= \sup_t [G_{m_S,k_{m_S}}(\cdot|t+a) - G_{n_T,k_{n_T}}(\cdot|t)] \\ D^-_{n,m}(a) &= \sup_t [G_{n_T,k_{n_T}}(\cdot|t) - G_{m_S,k_{m_S}}(\cdot|t+a)]. \end{aligned}$$

Then we may have the followings which are remained to future studies

(1) $\hat{\mu}_{n}$, is translation invariant; that is, for any real number c,

$$\hat{\mu}(Y_1^T,\cdots,Y_{k_n}^T;Y_1^S+c,\cdots,Y_{k_m}^S+c) = \hat{\mu}(Y_1^T,\cdots,Y_{k_n}^T;Y_1^S,\cdots,Y_{k_m}^S)+c.$$

(2) If $N=\min(n,m)$ and $\nu_N = O(N^{1/2-\delta})$ for some $\delta>0$, then $\nu_N(\hat{\mu}_{n,m}) \to 0$ as $N\to\infty$ with probability one.

2.3 Test of the fundamental assumption

Let $\hat{\xi}^T_{n,k_n}$ and $\hat{\xi}^S_{n,k_n}$ be the $k-NN$ estimators of the conditional medians of the test and the standard treatments. Then we have two independent stochastic processes

$$W_n^T(x) = k_n^{1/2}[\hat{\xi}^T_{n,k_n}(x) - \xi_T(x)] \xrightarrow{\mathcal{D}} W_T(x), \;\; x\in\mathcal{X} \text{ as } n\to\infty \qquad (2.4)$$

$$W_n^S(x) = k_n^{1/2}[\hat{\xi}^S_{n,k_n}(x) - \xi_S(x)] \xrightarrow{\mathcal{D}} W_S(x), \;\; x \in \mathcal{X} \text{ as } n \to \infty \tag{2.5}$$

where $W_T(x)$ and $W_S(x)$ are the Gaussian processes defined at previous section. Subtracting (2.5) from (2.4) gives

$$\begin{aligned} W_n^T(x) - W_n^S(x+\mu) &= k_n^{1/2}\left\{[\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x+\mu)] - [\xi_T(x) - \xi_S(x+\mu)]\right\} \\ &\xrightarrow{\mathcal{D}} W_T(x) - W_S(x+\mu) \;\; x \in \mathcal{X}. \end{aligned}$$

Under Ho : $\xi_T(x) = \xi_S(x+\mu)$,

$$\begin{aligned} W_n^T(x) - W_n^S(x+\mu) &= k_n^{1/2}\left\{\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x+\mu)\right\} \\ &\xrightarrow{\mathcal{D}} W_T(x) - W_S(x+\mu) \;\; x \in \mathcal{X}, \end{aligned}$$

where

$$W_T(x) - W_S(x+\mu) \sim N\left(0, pq\left(\frac{1}{g(\xi(x)|x)} + \frac{1}{g(\xi(x+\mu)|x+\mu)}\right)\right),$$

for all $x \in \mathcal{X}$. We may do the test by bootstrap sampling method. i.e. for m-dimentional $\boldsymbol{x} = (x_1, \cdots, x_m)$, we have m-dimentional multivariate normality after putting $\mu = \hat{\mu}$. We can estimate the variance martix by bootrapping method. And we may have asymptotic consistency of the bootstrap estimator of the variance estimator.

Let n, m be the sample sizes of the test and standard preparations. Let $m = c_1 n$ and $k_m = Ck_n$. Under Ho : $\xi^T(x) = \xi^S(x+\mu)$,

$$\begin{aligned} \hat{\xi}_T(x) - \hat{\xi}_S(x+\mu) &= k_n^{-1/2} W_n^T(x) - k_m^{-1/2} W_m^S(x+\mu) \\ &\quad + \xi^T(x) - \xi^S(x+\mu) \\ &= k_n^{-1/2}(W_n^T(x) - CW_m^S(x+\mu)). \end{aligned} \tag{2.6}$$

If the fundamental assumption holds, then we may use (2.6) to test it by similar resampling method.

We have

$$
\begin{aligned}
& P\left[\sup_x k_n^{1/2}|\hat{\xi}^T_{n,k_n}(x) - \hat{\xi}^S_{n,k_n}(x+\mu)| > C_{n,\alpha}\right] \\
= \; & P\left[\sup_{x\in\mathcal{X}} |W_n^T(x) - W_n^S(x+\mu)| > C_{n,\alpha}\right] \\
= \; & P\left[|W_n^T(x) - W_n^S(x+\mu)| > C_{n,\alpha} \text{ for some } x\right] \\
\to \; & P\left[|W^T(x) - W^S(x+\mu)| > C_\alpha \text{ for some } x\right] \text{ as } n \to \infty,
\end{aligned}
$$

provided

$$C_{n,\alpha} \to C_\alpha \text{ as } n \to \infty$$

by the week convergence.

2.4 Confidence intervals and powers of the tests: comparing the cardinality of the $k - NN$ estimator

Stute (1986) considered the case $k_n = O(n^{2/3})$, the cardinality of NN-estimator, while Bhattacharya and Gangopadhyay (1988) considered $k_n = O(n^{4/5})$. However, in the later case, there is generally a bias term of the order $n^{-2/5}$, and it may be necessary to eliminate the bias term in the drawing statistical conclusions. Sen and Gangopadhyay (1990) considered $k_n = O(n^{4/5-2\delta})$ for $\delta > 0$. And they showed that the bias term is eliminated. In this section, we compare the confidence intervals and the powers of our tests between $k_n = O(n^{4/5})$ and $k_n = O(n^{4/5-2\delta})$.

For $k_n = tn^{4/5-2\delta}$, $(0 < a < t < b < \infty,\ \delta > 0)$ and for fixed x and μ, we have

$$k_n^{1/2}\left[\hat{\xi}^T_n(x) - \hat{\xi}^S_n(x+\mu) - \left(\xi^T(x) - \xi^S(x+\mu)\right)\right] \to N(0, \sigma^2). \tag{2.7}$$

where

$$\sigma^2 = p(1-p)\left(\frac{1}{g^2(x)} + \frac{1}{g^2(x+\mu)}\right).$$

So $(1-\alpha) \times 100\%$ confidence interval is

$$\left[\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)\right] \pm Z_{\alpha/2}\sigma k_n^{-1/2}, \tag{2.8}$$

where Z_α is the α-percentile from the standard normal distribution. And the width of (2.8) is

$$2Z_{\alpha/2}\sigma k_n^{-1/2} = O(n^{-2/5+\delta}). \tag{2.9}$$

And the power under the alternative hypothesis Ha : $\xi^T(x) - \xi^S(x+\mu) = k_n^{-1/2}\sigma d(x)$ is

$$\begin{aligned}
&1 - \Pr\left\{-Z_{\alpha/2} < \frac{k_n^{1/2}}{\sigma}\left(\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)\right) < Z_{\alpha/2}\right\} \\
= \;& 1 - \Pr\Big\{-Z_{\alpha/2} < \frac{k_n^{1/2}}{\sigma}\left(\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu) - k_n^{-1/2}\sigma d(x) + k_n^{-1/2}\sigma d(x)\right) \\
& \qquad < Z_{\alpha/2}\Big\} \\
= \;& 1 - \Pr\Big\{-Z_{\alpha/2} - d(x) < \frac{k_n^{1/2}}{\sigma}\left((\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)) - (\xi^T(x) - \xi^S(x+\mu))\right) \\
& \qquad < Z_{\alpha/2} - d(x)\Big\} \\
= \;& 1 - \left(\Phi(Z_{\alpha/2} - d(x)) - \Phi(-Z_{\alpha/2} - d(x))\right).
\end{aligned} \tag{2.10}$$

To do the actual inferences for more than one point of $x \in \mathcal{X}$, we let $\boldsymbol{x} = (x_1, \cdots, x_m)'$ be the vector of the observed values. Then, as $n \to \infty$,

$$k_n^{1/2}\left[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x}) - \hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu) - \left(\boldsymbol{\xi}^T(\boldsymbol{x}) - \boldsymbol{\xi}^S(\boldsymbol{x}+\mu)\right)\right] \xrightarrow{\mathcal{D}} N_m(\mathbf{0}, \boldsymbol{\Sigma}), \tag{2.11}$$

where

$$\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x}) \;=\; (\hat{\xi}_n^T(x_1), \cdots, \hat{\xi}_n^T(x_m))',$$

$$\hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu) = (\hat{\xi}_n^S(x_1+\mu),\cdots,\hat{\xi}_n^S(x_m+\mu))',$$

$$\boldsymbol{\xi}^T(\boldsymbol{x}) = (\xi^T(x_1),\cdots,\xi^T(x_m))',$$

$$\boldsymbol{\xi}^S(\boldsymbol{x}+\mu) = (\xi^S(x_1+\mu),\cdots,\xi^S(x_m+\mu))',$$

and $\boldsymbol{\Sigma}$ is a diagonal matrix

$$\text{Diag}\left\{p(1-p)\left(\frac{1}{g^2(x_i)}+\frac{1}{g^2(x_i+\mu)}\right) \text{ for } i=1,\cdots,m\right\}.$$

And

$$\boldsymbol{\Sigma}^{-1/2}k_n^{1/2}\left[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x})-\hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu)-\left(\boldsymbol{\xi}^T(\boldsymbol{x})-\boldsymbol{\xi}^S(\boldsymbol{x}+\mu)\right)\right] \to N_m(\mathbf{0},\boldsymbol{I}), \tag{2.12}$$

for $\boldsymbol{\Sigma}^{-1/2}$ such that $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2}=\boldsymbol{I}$. To get $(1-\alpha)\times 100\%$ simultaneous confidence interval under the null hypothesis Ho : $\boldsymbol{\xi}^T(\boldsymbol{x})=\boldsymbol{\xi}^S(\boldsymbol{x}+\mu)$, we have

$$\Pr\left\{-Z_{\alpha/(m2)}<\boldsymbol{\Sigma}^{-1/2}k_n^{1/2}\left(\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x})-\hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu)\right)_i<Z_{\alpha/(m2)}\right\}=\frac{\alpha}{m}.$$

for $i=1,\cdots,m$. And

$$\Pr\left\{-Z_{\alpha/(m2)}<\boldsymbol{\Sigma}^{-1/2}k_n^{1/2}\left(\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x})-\hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu)\right)_i<Z_{\alpha/(m2)} \text{ for all } i=1,\cdots,m\right\}$$

is $\alpha/m\cdot m=\alpha$, where $\boldsymbol{A}_i$ is the i-th element of the vector $\boldsymbol{A}$. So the $(1-\alpha)\times 100\%$ simultaneous confidence interval for $\boldsymbol{\xi}^T(\boldsymbol{x})=\boldsymbol{\xi}^S(\boldsymbol{x}+\mu)$ is

$$\left[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x})-\hat{\boldsymbol{\xi}}^S(\boldsymbol{x}+\mu)\right]\pm k_n^{-1/2}Z_{\alpha/(m2)}\boldsymbol{\Sigma}^{1/2*}, \tag{2.13}$$

where $\boldsymbol{\Sigma}^{1/2*}=(\sigma_1,\cdots,\sigma_m)'$ where σ_i^2 is the i-th diagonal element of $\boldsymbol{\Sigma}$. And the power under the alternative hypothesis Ha : $\boldsymbol{\xi}^T(\boldsymbol{x})-\boldsymbol{\xi}^S(\boldsymbol{x}+\mu)=\boldsymbol{\Sigma}^{1/2}k_n^{-1/2}\boldsymbol{d}(\boldsymbol{x})$ where $\boldsymbol{d}(\boldsymbol{x})=(d(x_1),\cdots,d(x_m))'$ is

$$\Pr\left\{|\boldsymbol{\Sigma}^{-1/2}k_n^{1/2}\left(\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x})-\hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu)\right)_i|>Z_{\alpha/(m2)} \text{ for all } i=1,\cdots,m\right\}$$

$$=\Pr\left\{|\boldsymbol{\Sigma}^{-1/2}k_n^{1/2}\left(\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x})-\hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x}+\mu)-\boldsymbol{\Sigma}^{1/2}k_n^{-1/2}\boldsymbol{d}(\boldsymbol{x})\right)_i|\right.$$

$$> Z_{\alpha/(m2)} - \boldsymbol{d}(\boldsymbol{x}) \text{ for all } i = 1, \cdots, m\Big\}$$

$$= \prod_{i=1}^{m} \Big(1 - \Phi(Z_{\alpha/(m2)} - \boldsymbol{d}(\boldsymbol{x})_i) + \Phi(-Z_{\alpha/(m2)} - \boldsymbol{d}(\boldsymbol{x})_i)\Big). \tag{2.14}$$

For $k_n = tn^{4/5}$, $(0 < a < t < b < \infty)$ and for fixed x and μ, we have

$$k_n^{1/2}\left[\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu) - \left(\xi^T(x) - \xi^S(x+\mu)\right)\right]$$

$$\xrightarrow{\mathcal{D}} N\left(t^{5/2}(\beta(x) - \beta(x+\mu)), \sigma^2\right), \tag{2.15}$$

as $n \to \infty$, where

$$\sigma^2 = p(1-p)\left(\frac{1}{g^2(x)} + \frac{1}{g^2(x+\mu)}\right),$$

and

$$\beta(x) = \beta(\xi(x)) = \frac{-1}{24 f^3(x) g(\xi(x)|x)} \times [f(x)G_{xx}(\xi(x)|x) + 2f'(x)G_x(\xi(x)|x)].$$

as defined in (1.11). So $(1-\alpha)\times 100\%$ cofidence interval is

$$\left[\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)\right] - t^{2/5}(\beta(x) - \beta(x+\mu))k_n^{-1/2} \pm Z_{\alpha/2}\sigma k_n^{-1/2}. \tag{2.16}$$

And the width of (2.16) is

$$2Z_{\alpha/2}\sigma k_n^{-1/2} = O(n^{-2/5}) \tag{2.17}$$

And the power under the alternative hypothesis Ha : $\xi^T(x) - \xi^S(x+\mu) = k_n^{-1/2}\sigma d(x)$ is

$$1 - \Pr\Big\{-Z_{\alpha/2} < \frac{k_n^{1/2}}{\sigma}\left(\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)\right)$$

$$-\frac{t^{2/5}}{\sigma}(\beta(x) - \beta(x+\mu)) < Z_{\alpha/2} - d(x)\Big\}$$

$$= 1 - \Pr\Big\{-Z_{\alpha/2} < \frac{k_n^{1/2}}{\sigma}\left(\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu) - k_n^{-1/2}\sigma d(x) + k_n^{-1/2}\sigma d(x)\right)$$

$$-\frac{t^{2/5}}{\sigma}\left(\beta(x)-\beta(x+\mu)\right) < Z_{\alpha/2}\Big\}$$

$$= \quad 1-\Pr\left\{-Z_{\alpha/2}-d(x) < \frac{k_n^{1/2}}{\sigma}\left(\left(\hat{\xi}_n^T(x)-\hat{\xi}_n^S(x+\mu)\right)-\left(\xi^T(x)-\xi^S(x+\mu)\right)\right)\right.$$

$$\left.-\frac{t^{2/5}}{\sigma}\left(\beta(x)-\beta(x+\mu)\right) < Z_{\alpha/2}-d(x)\right\}$$

$$= 1-\left(\Phi(Z_{\alpha/2}-d(x))-\Phi(-Z_{\alpha/2}-d(x))\right). \tag{2.18}$$

Width of the confidence interval for $k_n = O(n^{4/5})$ (2.17) is shorter than that for $k_n = O(n^{4/5-2\delta})$ (2.9). But the confidence interval for $k_n = O(n^{4/5})$ has bias term $t^{2/5}(\beta(x)-\beta(x+\mu))$. So the question is the effect of the bias term in (2.16). If we have neglegible bias then $k_n = O(n^{4/5})$ gives the smaller confidence interval. but the bias term is not neglegible then we need to use $k_n = O(n^{4/5-2\delta})$ or consider how to adjust the bias term. We may say that the bias term is neglegible if the confidence interval for $k_n = O(n^{4/5})$ is a subset of that for $k_n = O(n^{4/5-2\delta})$. If we suppose $\hat{\xi}^T(x)-\hat{\xi}^S(x+\mu)$ are the same for the two k_n's then if

$$Z_{\alpha/2}\sigma_1 k_{n1}^{-1/2} > Z_{\alpha/2}\sigma_2 k_{n2}^{-1/2} - t^2\left(\beta(x)-\beta(x+\mu)\right)k_{n2}^{-1/2}, \tag{2.19}$$

and

$$-Z_{\alpha/2}\sigma_1 k_{n1}^{-1/2} < -Z_{\alpha/2}\sigma_2 k_{n2}^{-1/2} - t^{2/5}\left(\beta(x)-\beta(x+\mu)\right)k_{n2}^{-1/2}, \tag{2.20}$$

then $k_{n2} = tn^{4/5}$ is better where $k_{n1} = tn^{4/5-2\delta}$ and $k_{n2} = tn^{4/5}$. Condit (2.19) and (2.20) are equivalent to

$$(n^\delta - 1) > \left|\frac{t^{5/2}\left(\beta(x)-\beta(x+\mu)\right)}{Z_{\alpha/2}\sigma}\right|. \tag{2.21}$$

Left hand side of (2.21) goes to ∞ as $n \to \infty$, so we may say that $k_n = tn^{4/5}$ is asymptotically better than $k_n = tn^{4/5-2\delta}$ in terms of the width of the confidence interval.

If our test is to reject Ho : $\xi^T(x) - \xi^S(x+\mu) = 0$ if

$$\sup_{x\in\mathcal{X}} |\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)| > C_{n,\alpha},$$

we need a distribution of the supremeum of the difference of the processes. We may generate the critical values numerically. To calculate the power under the alternative hypothesis Ha : $\xi^T(x) - \xi^S(x+\mu) = d(x)$, we may need to generate the distribution based on the drift function $d(x)$. To do that, we may assume some parametric form of $d(x)$. e.g. quadratic form.

2.5 Bias correction for $k_n = O(n^{4/5})$

In the previous section, we suggested that $k_n = (n^{4/5})$ is asymptotically better than $k_n = O(k_n^{4/5-2\delta})$ in terms of the width of the confidence interval. But $k_n = O(n^{4/5})$ has bias term in the asymptotic mean, namely, βt^2 as in Theorem 1.2. In practice, we can adjust that bias by applying Jackknife method. In this section, we consider the bias correction for $k_n = O(n^{4/5})$ by applying Jackknife method for given data.

First method is Jackknife subsampling from the n original data (NBC). From Theorem 1.1 we have the following experssion for $\hat{\xi}_{k_n,n}$.

$$\hat{\xi}_{k_n,n} - \xi = \beta(\xi)(k_n/n)^2 + \{k_n g(\xi)\}^{-1} \sum_{i=1}^{k_n} [1(Y_{ni}^* > \xi) - (1-p)] + R_{nk_n}$$

where

$$\begin{aligned}
\beta(\xi) &= -[f(x_0)G_{xx}(\xi|x_0) + 2f'(x_0)G_x(\xi|x_0)]/\{23f^3(x_0)g(\xi)\}. \\
Y_{ni}^* &= G^{-1} \circ G^*(Y_{ni}|Z_{ni}), \;\; 1 \le i \le n, \; n \ge 1,
\end{aligned}$$

$$\max_{k_n \in I_n(a,b)} |R_{nk}| = O(n^{-3/5}\log n), \;\; a.s.$$

Similarly, for $\hat{\xi}_{k_{pn},pn}$, $(0 < p < 1)$

$$\hat{\xi}_{k_{pn},pn} - \xi = \beta(\xi)(k_{pn}/pn)^2 + \{k_{pn} g(\xi)/2\}^{-1} \sum_{i=1}^{[k_{pn}]} [1(Y_{pni}^* > \xi) - (1-p)] + R_{pnk_{pn}}.$$

Finally, we have

$$E\left[\hat{\xi}_{k_n,n} - \hat{\xi}_{k_{pn},pn}\right] = \beta(\xi)\left((k_n/n)^2 - (k_{pn}/pn)^2\right) + R$$
$$= \beta(\xi)n^{-2/5}t^2\left(1 - p^{-2/5}\right)$$

where

$$\max_{k_n \in I_n(a,b)} |R| = O(n^{-3/5}\log n), \quad a.s.,$$

for $k_n = tn^{4/5}$. We may use

$$\hat{\xi}_{k_n,n} - \frac{1}{t^2(1-p^{-2/5})} \cdot \text{ mean of } \left[\hat{\xi}_{k_n,n} - \hat{\xi}_{k_{pn},pn}\right] \tag{2.22}$$

as a bias reduced estimator. We have equivalent results for $-d$ subsampling method. Let $\hat{\xi}_{k_{n-d},n-d}$ be the conditional quantile estimator based on $n-d$ samples, then we have

$$\hat{\xi}_{k_{n-d},n-d} - \xi$$
$$= \beta(\xi)\left((k_{n-d})/(n-d)\right)^2 + \{(k_{n-d})g(\xi)\}^{-1}\sum_{i=1}^{k_{n-d}}[1(Y_{ni}^* > \xi) - (1-p)] + R_{n-d,k_{n-d}}.$$

Similarly, we have

$$E\left[\hat{\xi}_{k_n,n} - \hat{\xi}_{k_{n-d},n-d}\right] = \beta(\xi)\left[(k_n/n)^2 - (k_{n-d}/(n-d))^2\right] + R$$
$$= \beta(\xi)t^2n^{-2/5}\left[1 - \left(\frac{n-d}{n}\right)^{-2/5}\right],$$

for $k_n = tn^{4/5}$. We may use

$$\hat{\xi}_{k_n,n} - \frac{1}{t^2\left(1 - ((n-d)/n)^{-2/5}\right)} \cdot \text{ mean of } \left[\hat{\xi}_{k_n,n} - \hat{\xi}_{k_{n-d},n-d}\right]$$

as a bias reduced estimator.

Second method is Jackknife subsampling method from the k_n nearest neighborhood sample (KBC). For $\hat{\xi}_{pk_n,n}$, $(0 < p < 1)$

$$\hat{\xi}_{pk_n,n} - \xi = \beta(\xi)(k_n/n)^2p^2 + \{pk_ng(\xi)/2\}^{-1}\sum_{i=1}^{[pk_n]}[1(Y_{ni}^* > \xi) - (1-p)] + R_{npk_n}.$$

Finally, we have

$$E\left[\hat{\xi}_{k_n,n} - \hat{\xi}_{pk_n,n}\right] = \beta(\xi)(k_n/n)^2(1-p^2) + R,$$

where

$$\max_{k_n \in I_n(a,b)} |R| = O(n^{-3/5}\log n), \quad a.s.$$

We may use

$$\hat{\xi}_{k_n,n} - \frac{1}{1-p^2}\left[\hat{\xi}_{k_n,n} - \hat{\xi}_{pk_n,n}\right] \tag{2.23}$$

as a bias reduced estimator. Similarly, let $\hat{\xi}_{k_n-d,n}$ be the conditional quantile estimator based on $k_n - d$ samples, then we have

$$\hat{\xi}_{k_n-d,n} - \xi = \beta(\xi)\left((k_n-d)/n\right)^2 + \{(k_n-d)g(\xi)\}^{-1}\sum_{i=1}^{k_n-d}[1(Y_{ni}^* > \xi) - (1-p)] + R_{n,k_n-d}.$$

Similarly, we have

$$E\left[\hat{\xi}_{k_n,n} - \hat{\xi}_{k_n-d,n}\right] = \beta(\xi)(k_n/n)^2\left[1 - \left(\frac{k_n-d}{k_n}\right)^2\right] + R.$$

We may use

$$\hat{\xi}_{k_n,n} - \frac{1}{1-\left(\frac{k_n-d}{k_n}\right)^2}\cdot\left[\hat{\xi}_{k_n,n} - \hat{\xi}_{k_n-d,n}\right]$$

as a bias reduced estimator.

Chapter 3

NN-Estimator of a Conditional Quantile with Measurement Errors

3.1 Asymptotic properties of k-NN estimator with measurement errors on the covariate variable

3.1.1 Approximation of the conditional quantile

With uniform measurement error on X, we observe $(T_1, Y_1), (T_2, Y_2), \cdots, (T_n, Y_n)$ of iidrv. where $T = X + U$, $U \sim U(-\epsilon, \epsilon)$ and X and U are assumed to be independent. Then

$$P(T \leq t) \quad = \quad P(X + U \leq t)$$

$$
\begin{aligned}
&= \int_{-\epsilon}^{\epsilon} P(X+U \le t | U = u) f_U(u) \mathrm{d}u \\
&= \int_{-\epsilon}^{\epsilon} P(X \le t-u | U = u) f_U(u) \mathrm{d}u \\
&= \int_{-\epsilon}^{\epsilon} G_X(t-u) \frac{1}{2\epsilon} \mathrm{d}u, \\
&= \int_{t-\epsilon}^{t+\epsilon} G_X(u) \frac{1}{2\epsilon} \mathrm{d}u \\
&\to G_X(t) \text{ as } \epsilon \to 0+,
\end{aligned}
$$

$$
\begin{aligned}
f_T(t) &= \int_{-\epsilon}^{\epsilon} f_X(t-u) \frac{1}{2\epsilon} \mathrm{d}u. \\
&= \int_{t-\epsilon}^{t+\epsilon} f_X(u) \frac{1}{2\epsilon} \mathrm{d}u \\
&\to f_X(t) \text{ as } \epsilon \to 0+.
\end{aligned}
$$

And

$$
\begin{aligned}
f_{X,T}(x,t) &= f_{X,X+U}(x,t) \\
&= f_{X,U}(x, t-x) \\
&= f_X(x) f_U(t-x) \\
&= f_X(x) I_{(x-\epsilon<t<x+\epsilon)} 1/(2\epsilon),
\end{aligned}
$$

and we assume non-differential measurement error, i.e.,

$$P(Y|T,X) = P(Y|X).$$

So.

$$f_{Y|T}(y|t) = \int f_{Y,X|T}(y,x|t) \mathrm{d}x$$

$$
\begin{aligned}
&= \int \left(f_{Y,X,T}(y,x,t)/f_T(t)\right) \mathrm{d}x \\
&= \int \Big(f_{Y|X,T}(y|x,t) f_{X,T}(x,t)/f_T(t)\Big) \mathrm{d}x \\
&= \int \Big(f_{Y|X}(y|x) f_{X,T}(x,t)/f_T(t)\Big) \mathrm{d}x \\
&= \frac{1}{f_T(t)} \frac{1}{2\epsilon} \int_{t-\epsilon}^{t+\epsilon} f_{Y|X}(y|x) f_X(x) \mathrm{d}x, \\
G^{\epsilon}_{Y|T}(y|t) &= \frac{1}{f_T(t)} \frac{1}{2\epsilon} \int_{-\infty}^{y} \int_{t-\epsilon}^{t+\epsilon} f_{Y|X}(z|x) f_X(x) \mathrm{d}x \mathrm{d}z. \\
&= \frac{1}{f_T(t)} \frac{1}{2\epsilon} \int_{t-\epsilon}^{t+\epsilon} G_{Y|X}(y|x) f_X(x) \mathrm{d}x \\
&= \frac{\int_{t-\epsilon}^{t+\epsilon} G_{Y|X}(y|x) f_X(x) \mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon} f_X(x) \mathrm{d}x}, \qquad (3.1)
\end{aligned}
$$

and

$$
g^{\epsilon}_{Y|T}(y|t) = \frac{\int_{t-\epsilon}^{t+\epsilon} g_{Y|X}(y|x) f_X(x) \mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon} f_X(x) \mathrm{d}x}.
$$

It is easy to show that

$$
\left.\frac{\partial}{\partial \epsilon} G^{\epsilon}_{Y|T}(y|t)\right|_{\epsilon=0+} = \frac{2G_{Y|X}(y|t) f_X^2(t) - 2G_{Y|X}(y|t) f_X^2(t)}{f_X^2(t)} = 0.
$$

and let

$$
\left.\frac{\partial^2}{\partial \epsilon^2} G^{\epsilon}_{Y|T}(y|t)\right|_{\epsilon=0+} = \frac{1}{3} \frac{G_{xx}(y|t) f_X(t) + 2G_x(y|t) f_X'(t)}{f_X^2(t)} = A.
$$

where

$$
G_x(y|x) = \frac{\partial}{\partial x} G_{Y|X}(y|x) \text{ and } G_{xx}(y|x) = \frac{\partial^2}{\partial x^2} G_{Y|X}(y|x).
$$

So for small ϵ,

$$
G^{\epsilon}_{Y|T}(y|t) = G_{Y|X}(y|t) - \frac{1}{3} \frac{G_{xx}(y|t) f_X(t) + 2G_x(y|t) f_X'(t)}{f_X^2(t)} \epsilon^2/2 + o(||\epsilon||^2). \qquad (3.2)
$$

and

$$G_{Y|T}(\tau|t) = G_{Y|X}(\theta|t) + f_{Y|X}(\theta|t)(\tau - \theta) + O(||\tau - \theta||^2). \tag{3.3}$$

If we put (3.3) into (3.2) then we have

$$\begin{aligned} G^{\epsilon}_{Y|T}(\tau|t) &= G_{Y|X}(\theta|t) + f_{Y|X}(\theta|t)(\tau - \theta) + O(||\tau - \theta||^2) \\ &\quad -\frac{1}{3}\frac{G_{xx}(\tau|t)f_X(t) + 2G_x(\tau|t)f'_X(t)}{f_X^2(t)}\epsilon^2/2 + o(||\epsilon||^2) \text{ as } \epsilon \to 0+ \end{aligned} \tag{3.4}$$

Due to the Lemma 3.1 (3.4) becomes

$$\begin{aligned} G^{\epsilon}_{Y|T}(\tau|t) &= G_{Y|X}(\theta|t) + f_{Y|X}(\theta|t)(\tau - \theta) \\ &\quad -\frac{1}{3}\frac{G_{xx}(\tau|t)f_X(t) + 2G_x(\tau|t)f'_X(t)}{f_X^2(t)}\epsilon^2/2 + o(||\epsilon||^2) \text{ as } \epsilon \to 0+ \end{aligned} \tag{3.5}$$

If we let

$$G^{\epsilon}_{Y|T}(\tau|t) = G_{Y|X}(\theta|t) = p,$$

then

$$\begin{aligned} & G^{\epsilon}_{Y|T}(\tau|t) - G_{Y|X}(\theta|t) \\ = \; & f_{Y|X}(\theta|t)(\tau - \theta) - \frac{1}{3}\frac{G_{xx}(\tau|t)f_X(t) + 2G_x(\tau|t)f'_X(t)}{f_X^2(t)}\epsilon^2/2 + o(||\epsilon||^2) = 0 \end{aligned} \tag{3.6}$$

as $\epsilon \to 0+$. So we have the following theorem.

Theorem 3.1 *Let θ and τ be the p-quantile from $G_{Y|X}(\cdot|x)$ and $G_{Y|T}(\cdot|t)$, respectively, where $T = X + U$, $U \sim U(-\epsilon, \epsilon)$ and T and X are assumed to be independent, then*

$$\tau = \theta + \frac{1}{f_{Y|X}(\theta|t)} \cdot \frac{1}{3}\frac{G_{xx}(\tau|t)f_X(t) + 2G_x(\tau|t)f'_X(t)}{f_X^2(t)}\epsilon^2/2 + o(||\epsilon||^2) \tag{3.7}$$

as $\epsilon \to 0+$.

Lemma 3.1 *Let* $p = G^{\epsilon}_{Y|T}(\tau|t) = G_{Y|X}(\theta|t)$ *and if there exists* $G^{-1}_{\epsilon,t}(p)$ *such that* $G^{\epsilon}_{Y|T}(y|t) = p \Leftrightarrow G^{-1}_{\epsilon,t}(p) = y$ *for* $\epsilon < \epsilon_1$, $y \in R$, $t \in R$ *and* $p \in (0,1)$ *. Then*

$$||\tau - \theta|| = O(||\epsilon||^2)$$

as $\epsilon_1 \to 0+$ *.*

Proof: Let

$$\begin{aligned} G^{\epsilon}_{Y|T}(\tau|t) &= G_{Y|X}(\tau|t) + \frac{\partial^2}{\partial\epsilon^2} G^{\epsilon}_{Y|T}(\tau|t)\bigg|_{\epsilon=0} \epsilon^2/2 + O(||\epsilon||^3) \\ &= G_{Y|X}(\tau|t) + A\epsilon^2/2 + O(||\epsilon||^3) \\ &= p. \end{aligned}$$

Then

$$\begin{aligned} G_{Y|X}(\tau|t) &= p - A\epsilon^2/2 - O(||\epsilon||^3) \\ \tau = G^{-1}_{\epsilon,t}(p) &= G^{-1}_t(p - A\epsilon^2/2 - O(||\epsilon||^3)), \end{aligned}$$

for $G^{-1}_t(\cdot)$ such that $G_{Y|T}(\theta|t) = p \Leftrightarrow \theta = G^{-1}_t(p)$. So,

$$\begin{aligned} \lim_{\epsilon\to 0+} G^{-1}_{\epsilon,t}(p) &= G^{-1}_t(p) = \theta \\ \lim_{\epsilon\to 0+} \frac{\partial}{\partial\epsilon} G^{-1}_{\epsilon,t}(p) &= 0 \end{aligned}$$

Therefore,

$$\tau = G^{-1}_{\epsilon,t}(p) = \theta + O(||\epsilon||^2).$$

And the lemma is proved. ◇

3.1.2 Asymptotic normality of the conditional quantile estimator with measurement error

We take $W_i = |T_i - x_0|$ for $i = 1, \cdots, n$ then $W_{n1} < \cdots < W_{nn}$ are the order statistics of $W_1, \cdots, W_n$ and $Y^{\epsilon}_{n1}, \cdots, Y^{\epsilon}_{nn}$ are the induced order statistics. The conditional quantile, $\xi^\epsilon = \xi_p^\epsilon(x_0)$ is satisfying

$$G^\epsilon_{Y|T}(\xi^\epsilon|x_0) = p,$$

for given p, $(0 < p < 1)$ and $x_0 \in R^1$. The contaminated k-NN empirical cdf of Y (with respect to x_0) is defined as

$$\hat{G}^\epsilon_{nk}(y) = k^{-1}\sum_{i=1}^{k} 1(Y^\epsilon_{ni} \le y).$$

The k-NN estimator of ξ^ϵ can now be defined as

$$\begin{aligned} \hat{\xi}^\epsilon_{nk} &= \text{the } [kp]\text{th order statistics of } Y^\epsilon_{n1}, \cdots, Y^\epsilon_{nk} \\ &= \inf\{y : \hat{G}^\epsilon_{nk}(y) \ge [kp]/k\}. \end{aligned}$$

Then we have the following theorems similar to Bhattacharya and Gangopadhyay (1988)'s

Theorem 3.2 *For $\epsilon > 0$ and $k \in I_n(a,b)$*

$$\hat{\xi}^\epsilon_{nk} - \xi^\epsilon = \beta(\xi^\epsilon)(k/n)^2 + \{kg^\epsilon(\xi^\epsilon)\}^{-1}\sum_{i=1}^{k}[1(Y^{*\epsilon}_{ni} > \xi^\epsilon) - (1-p)] + R^\epsilon_{nk}$$

where

$$\begin{aligned} \beta(\xi^\epsilon) &= -[f^\epsilon(x_0)G^\epsilon_{tt}(\xi^\epsilon|x_0) + 2f^{\epsilon'}(x_0)G^\epsilon_t(\xi^\epsilon|x_0)]/\{23f^{\epsilon 3}(x_0)g^\epsilon(\xi^\epsilon)\}. \\ Y^{*\epsilon}_{ni} &= (G^\epsilon)^{-1} \circ G^{*\epsilon}(Y^\epsilon_{ni}|W_{ni}), \quad 1 \le i \le n,\ n \ge 1, \end{aligned}$$

$$\max_{k \in I_n(a,b)} |R^\epsilon_{nk}| = O(n^{-3/5}\log n), \quad a.s.$$

and

$$\begin{aligned} G^\epsilon_t(y|t) &= (\partial/\partial t)G^\epsilon(y|t) \\ G^\epsilon_{tt}(y|t) &= (\partial^2/\partial t^2)G^\epsilon(y|t) \\ G^{*\epsilon}(y|w) &= \frac{\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} G_{Y|X}(y|x)f_x(x)\,dx + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} G_{Y|X}(y|x)f_x(x)\,dx}{\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} f_x(x)\,dx + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} f_x(x)\,dx} \\ G^\epsilon(y) &= G^\epsilon_{Y|T}(y|x_0). \end{aligned}$$

Theorem 3.3 *Let* $T_n^\epsilon(t) = \hat{\xi}^\epsilon_{n,[n^{4/5}t]}$. *Then for any* $0 < a < b$.

$$\{n^{2/5}[T_n^\epsilon(t) - \xi^\epsilon] - \beta t^2, a \le t \le b\} \xrightarrow{\mathcal{D}} \{\sigma^\epsilon t^{-1}B(t),\ a < t < b\},$$

where $\beta = \beta(\xi^\epsilon)$ *given in Theorem 3.4*, $(\sigma^\epsilon)^2 = p(1-p)/(g^\epsilon)^2(\xi^\epsilon)$ *and* $\{B(t), t > 0\}$ *denote a standard Brownian motion.*

Theorem 3.4 *For* $\epsilon = \epsilon_n = o(n^{-1/5})$,

$$\begin{aligned}
\hat{\xi}^\epsilon_{nk} - \xi &= (\hat{\xi}^\epsilon_{nk} - \xi^\epsilon) + (\xi^\epsilon - \xi) \\
&= \beta(\xi^\epsilon)(k/n)^2 + \{kg^\epsilon(x_0)\}^{-1}\sum_{i=1}^{k}[1(Y^{*\epsilon}_{ni} > \xi^\epsilon) - (1-p)] + R^{\epsilon *}_{nk}
\end{aligned}$$

where

$$\begin{aligned}
\beta(\xi^\epsilon) &= -[f^\epsilon(x_0)G^\epsilon_{tt}(\xi^\epsilon|x_0) + 2f^{\epsilon'}(x_0)G^\epsilon_t(\xi^\epsilon|x_0)]/\{23f^{\epsilon 3}(x_0)g^\epsilon(\xi^\epsilon)\} \\
&\to -[f(x_0)G_{xx}(\xi|x_0) + 2f'(x_0)G_x(\xi|x_0)]/\{23f^3(x_0)g(\xi)\} \\
&= \beta(\xi) \ \ as \ \ n \to \infty,
\end{aligned}$$

$$Y^{*\epsilon}_{ni} = (G^\epsilon)^{-1} \circ G^{*\epsilon}(Y^\epsilon_{ni}|W_{ni}),\ \ 1 \le i \le n,\ n \ge 1,$$

and

$$\max_{k \in I_n(a,b)} |R^{\epsilon *}_{nk}| = o(n^{-2/5}),\ \ a.s.$$

Theorem 3.5 *Let* $T_n^\epsilon(t) = \hat{\xi}^\epsilon_{n,[n^{4/5}t]}$. *Then for any* $0 < a < b$ *and* $\epsilon_n = o(n^{-1/5})$.

$$\{n^{2/5}[T_n^\epsilon(t) - \xi] - \beta t^2, a \le t \le b\}$$

$$= \{n^{2/5}[T_n^\epsilon(t) - \xi^\epsilon] - \beta t^2 + n^{2/5}[\xi^\epsilon - \xi], a \le t \le b\} \xrightarrow{\mathcal{D}} \{\sigma t^{-1}B(t),\ a < t < b\},$$

where $\beta = \beta(\xi^\epsilon)$ *given in Theorem 3.4*, $\sigma^2 = p(1-p)/g^2(\xi)$ *and* $\{B(t), t > 0\}$ *denote a standard Brownian motion.*

Proof of Theorem 3.4

$$\begin{aligned}\max_{k\in I_n(a,b)}|R^{\epsilon*}_{nk}| &= \max_{k\in I_n(a,b)}|R^{\epsilon}_{nk}+R^{\epsilon}| \\ &= O(n^{-3/5}\log n)+o(n^{-2/5}) \\ &= o(n^{-2/5})\ \ a.s.\end{aligned}$$

And

$$G^{\epsilon}_{Y|T}(y|t)=\frac{\int_{t-\epsilon}^{t+\epsilon}G_{Y|X}(y|x)f_X(x)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon}f_X(x)\mathrm{d}x},$$

so

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}G^{\epsilon}_{Y|T}(y|t) &= \Big\{\Big(G_{Y|X}(y|t+\epsilon)f_X(t+\epsilon)-G_{Y|X}(y|t-\epsilon)f_X(t-\epsilon)\Big)\int_{t-\epsilon}^{t+\epsilon}f_X(x)\mathrm{d}x \\ &\quad -\int_{t-\epsilon}^{t+\epsilon}G_{Y|X}(y|x)f_X(x)\mathrm{d}x(f_X(t+\epsilon)-f_X(t-\epsilon))\Big\} \\ &\quad /\ \Big(\int_{t-\epsilon}^{t+\epsilon}f_X(x)\mathrm{d}x\Big)^2,\end{aligned}$$

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}G^{\epsilon}_{Y|T}(y|t) &= \Big\{\Big\{\Big[G_x(y|t+\epsilon)f_X(t+\epsilon)+G_{Y|X}(y|t+\epsilon)f'_X(t+\epsilon) \\ &\quad -G_x(y|t-\epsilon)f_X(t-\epsilon)-G_{Y|X}(y|t-\epsilon)f'_X(t-\epsilon)\Big]\int_{t-\epsilon}^{t+\epsilon}f_X(x)\mathrm{d}x \\ &\quad +\Big\{\Big[G_{Y|X}(y|t+\epsilon)f_X(t+\epsilon)-G_{Y|X}(y|t-\epsilon)f_X(t-\epsilon)\Big] \\ &\quad -\Big[G_{Y|X}(y|t+\epsilon)f_X(t+\epsilon)-G_{Y|X}(y|t-\epsilon)f_X(t-\epsilon)\Big]\Big\} \\ &\quad \times(f_X(t+\epsilon)-f_X(t-\epsilon)) \\ &\quad -\int_{t-\epsilon}^{t+\epsilon}G_{Y|X}(y|x)f_X(x)\mathrm{d}x(f'_X(t+\epsilon)-f'_X(t-\epsilon))\Big\}\end{aligned}$$

$$
\begin{aligned}
& \times \left(\int_{\epsilon-t}^{\epsilon+t} f_X(x)\mathrm{d}x\right)^2 \\
- \quad & \Big\{ \Big(G_{Y|X}(y|t+\epsilon) f_X(t+\epsilon) - G_{Y|X}(y|t-\epsilon) f_X(t-\epsilon) \Big) \int_{t-\epsilon}^{t+\epsilon} f_X(x)\mathrm{d}x \\
& - \int_{t-\epsilon}^{t+\epsilon} G_{Y|X}(y|x) f_X(x)\mathrm{d}x (f_X(t+\epsilon) - f_X(t-\epsilon) \Big\} \\
& \times 2 \int_{t-\epsilon}^{t+\epsilon} f_X(x)\mathrm{d}x (f_X(t+\epsilon) - f_X(t-\epsilon)) \Big\} \\
/ \quad & \left(\int_{t-\epsilon}^{t+\epsilon} f_X(x)\mathrm{d}x\right)^4 .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& f^\epsilon(x_0) G_{tt}^\epsilon(\xi^\epsilon|x_0) + 2 f^{\epsilon'}(x_0) G_t^\epsilon(\xi^\epsilon|x_0) \\
= \quad & \frac{1/(2\epsilon) \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x}{(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)^4} \Big\{ \Big[(G_x(\xi^\epsilon|x_0+\epsilon) f_X(x_0+\epsilon) + G_{Y|X}(\xi^\epsilon|x_0+\epsilon) f'_X(x_0+\epsilon) \\
& - G_x(\xi^\epsilon|x_0-\epsilon) f_X(x_0-\epsilon) - G_{Y|X}(\xi^\epsilon|x_0-\epsilon) f'_X(x_0-\epsilon)) \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x \\
& - \int_{x_0-\epsilon}^{x_0+\epsilon} G_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x (f'_X(x_0+\epsilon) - f'_X(x_0-\epsilon) \Big] (\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)^2 \\
& - \Big[(G_{Y|X}(\xi^\epsilon|x_0+\epsilon) f_X(x_0+\epsilon) - G_{Y|X}(\xi^\epsilon|x_0-\epsilon) f_X(x_0-\epsilon)) \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x \\
& - \int_{x_0-\epsilon}^{x_0+\epsilon} G_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x (f_X(x_0+\epsilon) - f_X(x_0-\epsilon)) \Big] \\
& \times 2 \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x \times (f_X(x_0+\epsilon) - f_X(x_0-\epsilon)) \Big\} \\
+ \quad & 2\frac{1}{2\epsilon}(f_X(x_0+\epsilon) - f_X(x_0-\epsilon)) \\
& \Big[(G_{Y|X}(\xi^\epsilon|x_0+\epsilon) f_X(x_0+\epsilon) - G_{Y|X}(\xi^\epsilon|x_0-\epsilon) f_X(x_0-\epsilon)) \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x
\end{aligned}
$$

$$
- \int_{x_0-\epsilon}^{x_0+\epsilon} G_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x (f_X(x_0+\epsilon) - f_X(x_0-\epsilon))\Big]
$$

$$
\Big/ \left(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x\right)^2
$$

$$
= \frac{1}{2\epsilon} \times \frac{1}{(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)^2} \Big\{ (G_x(\xi^\epsilon|x_0+\epsilon) f_X(x_0+\epsilon) + G_{Y|X}(\xi^\epsilon|x_0+\epsilon) f'_X(x_0+\epsilon)
$$

$$
-G_x(\xi^\epsilon|x_0-\epsilon) f_X(x_0-\epsilon) - G_{Y|X}(\xi^\epsilon|x_0-\epsilon) f'_X(x_0-\epsilon))(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)^2
$$

$$
\Big[- \int_{x_0-\epsilon}^{x_0+\epsilon} G_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x (f'_X(x_0+\epsilon) - f'_X(x_0-\epsilon)\Big] (\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)\Big\}.
$$

$$
\frac{f^\epsilon(x_0) G^\epsilon_{tt}(\xi^\epsilon|x_0) + 2 f^{\epsilon'}(x_0) G^\epsilon_t(\xi^\epsilon|x_0)}{f^{\epsilon 3}(x_0) g^\epsilon(\xi^\epsilon)}
$$

$$
= \frac{1}{2\epsilon} \frac{1}{(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)^2} \Big\{ (\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x)^2 \big[(G_x(\xi^\epsilon|x_0+\epsilon) f_X(x_0+\epsilon)
$$

$$
+G_{Y|X}(\xi^\epsilon|x_0+\epsilon) f'_X(x_0+\epsilon) - G_x(\xi^\epsilon|x_0-\epsilon) f_X(x_0-\epsilon)
$$

$$
- G_{Y|X}(\xi^\epsilon|x_0-\epsilon) f'_X(x_0-\epsilon)\big]
$$

$$
- \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x \int_{x_0-\epsilon}^{x_0+\epsilon} G_{Z|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x (f'_X(x_0+\epsilon) - f'_X(x_0-\epsilon))\Big\}
$$

$$
\Big/ \quad \left(\frac{1}{2\epsilon}\right)^3 \left(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x\right)^3 \frac{\int_{x_0-\epsilon}^{x_0+\epsilon} g_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x}
$$

$$
= \frac{1}{2\epsilon} \Big[(G_x(\xi^\epsilon|x_0+\epsilon) f_X(x_0+\epsilon) + G_{Y|X}(\xi^\epsilon|x_0+\epsilon) f'_X(x_0+\epsilon)
$$

$$
- G_x(\xi^\epsilon|x_0-\epsilon) f_X(x_0-\epsilon) - G_{Y|X}(\xi^\epsilon|x_0-\epsilon) f'_X(x_0-\epsilon)\Big]
$$

$$
- \frac{\int_{x_0-\epsilon}^{x_0+\epsilon} G_{Z|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x} (f'_X(x_0+\epsilon) - f'_X(x_0-\epsilon))
$$

$$
\begin{aligned}
/ & \left(\frac{1}{2\epsilon}\right)^3 \left(\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x\right)^2 \left(\int_{x_0-\epsilon}^{x_0+\epsilon} g_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x\right) \\
= & \frac{1}{2\epsilon}\left[G_x(\xi^\epsilon|x_0+\epsilon)f_X(x_0+\epsilon) - G_x(\xi^\epsilon|x_0-\epsilon)f_X(x_0-\epsilon)\right] \\
& +\frac{1}{2\epsilon}\left[G_{Y|X}(\xi^\epsilon|x_0+\epsilon)f'_X(x_0+\epsilon) - G_{Y|X}(\xi^\epsilon|x_0-\epsilon)f'_X(x_0-\epsilon)\right] \\
& -\frac{1}{2\epsilon}\frac{\int_{x_0-\epsilon}^{x_0+\epsilon} G_{Y|X}(\xi^\epsilon|x)f_X(x)\mathrm{d}x}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x}\left(f'_X(x_0+\epsilon) - f'_X(x_0-\epsilon)\right) \\
/ & \left(\frac{1}{2\epsilon}\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x\right)^2 \left(\frac{1}{2\epsilon}\int_{x_0-\epsilon}^{x_0+\epsilon} g_{Y|X}(\xi^\epsilon|x) f_X(x)\mathrm{d}x\right) \qquad (3.8) \\
\to & \Big[G_{xx}(\xi|x_0)f_X(x_0) + 2f'_X(x_0)G_x(\xi|x_0) + G_{Y|X}(\xi|x_0)f''_X(x_0) \\
& -(G_{Y|X}(\xi|x_0)f_X(x_0)/f_X(x_0))f''_X(x_0)\Big] \\
/ & f_X^3(x_0)g_{Y|X}(\xi|x_0) \\
= & \frac{G_{xx}f + 2f'G_x}{f^3 g} \text{ as } \epsilon \to 0.
\end{aligned}
$$

◇

3.1.3 Asymptotic variance and bias of the process with measurement error

In this section, we investigate the asymptotic variance and bias of our process with measurement error. We call the variance and the bias from the uniform measurement error model as contaminated bias and contaminated variance. We will show the convergences of contaminated variance and bias to the original (uncontaminated) variance and bias. And sufficient conditions for the condition that the contaminated variance and bias are greater than the unocontaminated ones are derived.

From the Theorems 1.4 and Theorem 1.5, we let $\beta = \beta(\xi)$ be the asymptotic bias ($ABIAS$) and let $p(1-p)/g^2(\xi)$ be the asymptotic vaiance ($AVAR$), where

$$
\begin{aligned}
\beta(\xi) &= -[f(x_0)G_{xx}(\xi|x_0) + 2f'(x_0)G_x(\xi|x_0)]/\{23f^3(x_0)g(\xi)\}, \\
g(\xi) &= g_{Y|X}(\xi|x_0).
\end{aligned}
$$

Similarly, from the Theorem 3.2 and Theorem 3.3, we let $\beta^\epsilon = \beta(\xi^\epsilon)$ be the asymptotic bias ($AVAR_\epsilon$) and let $p(1-p)/(g^\epsilon)^2(\xi^\epsilon)$ be the asymptotic vaiance ($ABIAS_\epsilon$), where

$$
\begin{aligned}
\beta(\xi^\epsilon) &= -[f^\epsilon(x_0)G^\epsilon_{tt}(\xi^\epsilon|x_0) + 2f^{\epsilon'}(x_0)G^\epsilon_t(\xi^\epsilon|x_0)]/\{23f^{\epsilon 3}(x_0)g^\epsilon(\xi^\epsilon)\}, \\
g^\epsilon(\xi^\epsilon) &= g^\epsilon_{Y|T}(\xi^\epsilon|t) = \frac{\int_{x_0-\epsilon}^{x_0+\epsilon} g_{Y|X}(\xi^\epsilon|x)f_X(x)\mathrm{d}x}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x}.
\end{aligned}
$$

To prove Theorem 3.6, we need Lemma 3.2.

Lemma 3.2

$$
\frac{\int_{x_0-\epsilon}^{x_0+\epsilon}(x-x_0)f_X(x)dx}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)dx} = \frac{1}{3}\frac{f'_X(x_0)}{f_X(x_0)}\epsilon^2 + O(||\epsilon||^4)
$$

Proof: By Taylor expansion,

$$
\begin{aligned}
\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x &= \int_{x_0-\epsilon}^{x_0+\epsilon}\Big(f(x_0) + f'(x_0)(x-x_0) + f''(x_1)(x-x_1)^2/2\Big)\,\mathrm{d}x \\
&= 2\epsilon f(x_0) + O(||\epsilon||^3) \text{ for } |x-x_1| < |x-x_0|,
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_{x_0-\epsilon}^{x_0+\epsilon}(x-x_0)f_X(x)\mathrm{d}x \\
&= \int_{x_0-\epsilon}^{x_0+\epsilon}\Big(f(x_0)(x-x_0) + f'(x_0)(x-x_0)^2 + f''(x_1)(x-x_1)^3/2\Big)\,\mathrm{d}x \\
&= 2/3f'(x_0)\epsilon^3 + O(||\epsilon||^5) \text{ for } |x-x_1| < |x-x_0|,
\end{aligned}
$$

◇

Let $g(y|x) = g_{Y|X}(y|x)$ be the conditional pdf of Y given $X = x$, then

$$\begin{aligned} g(\xi^\epsilon|x) - g(\xi|x_0) &= g(\xi^\epsilon|x) - g(\xi^\epsilon|x_0) + g(\xi^\epsilon|x_0) - g(\xi|x_0) \\ &= g_x(\xi^\epsilon|x_0)(x - x_0) + O(||x - x_0||^2) \\ &\quad + g'(\xi|x_0)(\xi^\epsilon - \xi) + O(||\xi^\epsilon - \xi||^2) \end{aligned} \tag{3.9}$$

where

$$\begin{aligned} g_x(y|x) &= \frac{\partial}{\partial x} g_{Y|X}(y|x) \\ g'(y|x) &= \frac{\partial}{\partial y} g_{Y|X}(y|x). \end{aligned}$$

Theorem 3.6 *If $||\xi^\epsilon - \xi||^2 = O(||\epsilon||^2)$ then,*

$$g^\epsilon(\xi^\epsilon) = g(\xi) + g'(\xi|x_0)(\xi^\epsilon - \xi) + O(||\epsilon||^2). \tag{3.10}$$

So.

$$AVAR_\epsilon \to AVAR \text{ as } \epsilon \to 0.$$

And if

$$g'(\xi|x_0)(\xi^\epsilon - \xi) < 0 \text{ for all } \epsilon > 0, \tag{3.11}$$

then

$$AVAR_\epsilon > AVAR \text{ for small } \epsilon.$$

Proof: From (3.9),

$$\begin{aligned} & g^\epsilon(\xi^\epsilon) \times \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x \\ = & \int_{x_0-\epsilon}^{x_0+\epsilon} g(\xi^\epsilon|x) f(x)\mathrm{d}x \\ = & \ g(\xi|x_0) \int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x + g_x(\xi^\epsilon|x_0) \int_{x_0-\epsilon}^{x_0+\epsilon} (x - x_0) f_X(x)\mathrm{d}x \end{aligned}$$

$$+g'(\xi|x_0)(\xi^\epsilon - \xi)\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x$$

$$+\int_{x_0-\epsilon}^{x_0+\epsilon}\left(O(||x-x_0||^2)+O(||\xi^\epsilon-\xi||^2)\right)f(x)\mathrm{d}x,$$

and by Lemma 3.2,

$$\begin{aligned}
g^\epsilon(\xi^\epsilon) &= g(\xi|x_0)+g_x(\xi^\epsilon|x_0)\left(\frac{1}{3}\frac{f'(x_0)}{f(x_0)}\epsilon^2+O(||\epsilon||^4)\right) \\
&\quad +g'(\xi|x_0)(\xi^\epsilon-\xi)+O(||xi^\epsilon-\xi||^2) \\
&= g(\xi|x_0)+g'(\xi|x_0)(\xi^\epsilon-\xi)+(||\epsilon||^2).
\end{aligned}$$

◇

If $g_{Y|X}(y|x)$ (conditional distribution of Y) given $X=x$ is unimodal and has maximum at median then the difference between $g^\epsilon(\xi^\epsilon)$ and $g(\xi)$ at the median is the order of ϵ^2, so the difference is neglectable up to order ϵ^2 . So our procedure is robust for the measurement error; our procedure is not sensitive to the measurement error.

From (3.8), let

$$\begin{aligned}
-2\beta\beta(\xi^\epsilon) &= \frac{1}{\left(\frac{1}{2\epsilon}\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x\right)^2\int_{x_0-\epsilon}^{x_0+\epsilon} g(\xi^\epsilon|x)f(x)\mathrm{d}x}\times \\
&\quad \{[G_x(\xi^\epsilon|x_0+\epsilon)f(x+\epsilon)-G_x(\xi^\epsilon|x_0-\epsilon)f(x-\epsilon)] \\
&\quad +[G(\xi^\epsilon|x_0+\epsilon)f'(x+\epsilon)-G(\xi^\epsilon|x_0-\epsilon)f'(x-\epsilon)] \\
&\quad -\frac{\int_{x_0-\epsilon}^{x_0+\epsilon} G(\xi^\epsilon|x)f(x)\mathrm{d}x}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x}[f'(x_0+\epsilon)-f'(x_0-\epsilon)]\} \\
&= \frac{A+B+C}{D}.
\end{aligned}$$

$$D = \left(\frac{1}{2\epsilon}\right)^2\left(2\epsilon f(x_0)+O(||\epsilon||^3)\right)^2$$

$$\cdot \Big[g(\xi|x_0) + g'(\xi|x_0)(\xi^\epsilon - \xi)(2\epsilon f(x_0) + O(||\epsilon||^3) + O(||\epsilon||^3)\Big]$$

$$= \quad f^2(x_0)\,(g(\xi|x_0) + g'(\xi|x_0)(\xi - \xi^\epsilon)) \times 2f(x_0)\epsilon + O(||\epsilon||^2)$$

$$= \quad 2f^3(x_0)g(\xi|x_0)\epsilon + 2f^3(x_0)g'(\xi|x_0)(\xi^\epsilon - \xi)\epsilon + O(||\epsilon||^2).$$

$$\begin{aligned} G(\xi^\epsilon|x_0 \pm \epsilon) \quad - \quad & G(\xi|x_0) \\ = \quad & G(\xi^\epsilon|x_0 \pm \epsilon) - G(\xi|x_0 \pm \epsilon) + G(\xi|x_0 \pm \epsilon) - G(\xi|x_0) \\ = \quad & g(\xi|x_0 \pm \epsilon)(\xi^\epsilon - \xi) + O(||\xi^\epsilon - \xi||^2) \\ & \pm G_x(\xi|x_0)\epsilon + O(||\epsilon||^2), \end{aligned}$$

and

$$f'(x_0 \pm \epsilon) = f'(x_0) \pm f''(x_0)\epsilon + O(||\epsilon||^2).$$

Therefore,

$$\begin{aligned} & G(\xi^\epsilon|x_0 \pm \epsilon)f'(x_0 \pm \epsilon) \\ & = G(\xi|x_0)f'(x_0) \pm [G_x(\xi|x_0)f'(x_0) + G(\xi|x_0)f''(x_0)]\,\epsilon \\ & + g(\xi|x_0)f'(x_0)(\xi^\epsilon - \xi) \\ & (\pm g_x(\xi|x_0)f'(x_0) + g(\xi|x_0)f''(x_0))\,\epsilon(\xi^\epsilon - \xi) + O(||\epsilon||^2). \end{aligned}$$

$$\begin{aligned} B \quad = \quad & 2\,[G_x(\xi|x_0)f'(x_0) + G(\xi|x_0)f''(x_0)]\,\epsilon \\ & + 2\,[g_x(\xi|x_0)f'(x_0) + g(\xi|x_0)f''(x_0)]\,\epsilon(\xi^\epsilon - \xi) + O(||\epsilon||^2). \end{aligned}$$

Similarly,

$$\begin{aligned} A \quad = \quad & 2\,[G_{xx}(\xi|x_0)f(x_0) + G_x(\xi|x_0)f'(x_0)]\,\epsilon \\ & + 2\,[G''_x(\xi|x_0)f(x_0) + G'_x(\xi|x_0)f'(x_0)]\,\epsilon(\xi^\epsilon - \xi) + O(||\epsilon||^2), \end{aligned}$$

where

$$G'_x = \frac{\partial}{\partial y} G_x(y|x) \text{ and } G''_x = \frac{\partial}{\partial y} G_x^{(1)}(y|x).$$

And

$$\begin{aligned}
C &= \frac{1}{\int_{x_0-\epsilon}^{x_0+\epsilon} f_X(x)\mathrm{d}x[2f''(x_0)\epsilon + O(||\epsilon||^3)]} \\
&\quad \times \int_{x_0-\epsilon}^{x_0+\epsilon} \Big(G(\xi|x_0) + G_x(\xi^\epsilon|x)(x-x_0) + O(||x-x_0||^2) \\
&\quad g(\xi|x_0)(\xi^\epsilon - \xi) + O(||\xi^\epsilon - \xi||^2) \Big) f(x)\mathrm{d}x \\
&= 2f''(x_0)G(\xi|x_0)\epsilon + 2f''(x_0)g(\xi|x_0)(\xi^\epsilon - \xi)\epsilon + O(||\epsilon||^2).
\end{aligned}$$

$$\begin{aligned}
A+B+C &= 2\epsilon\left[G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0)\right] \\
&\quad +2\epsilon(\xi^\epsilon - \xi)\left[g_x(\xi|x_0)f'(x_0) + G'_x(\xi|x_0)f'(x_0) + G''_x(\xi|x_0)f(x_0)\right] \\
&\quad +O(||\epsilon||^2).
\end{aligned}$$

Therefore,

$$\begin{aligned}
-23\beta(\xi^\epsilon) &= \frac{G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0)}{f^3(x_0)g(\xi|x_0)} \\
&\quad + \frac{1}{f^3(x_0)g(\xi|x_0)}\Big\{g_x(\xi|x_0)f'(x_0) + G'_x(\xi|x_0)f'(x_0) + f(x_0)G''_x(\xi|x_0) \\
&\quad - \frac{(G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0))\, f^3(x_0)g'(\xi|x_0)}{f^3(x_0)g(\xi|x_0)}\Big\}(\xi^\epsilon - \xi) \\
&\quad +O(||\epsilon||^2) \\
&= -23\beta(\xi) \\
&\quad + \frac{1}{f^3(x_0)g(\xi|x_0)}\{g_x(\xi|x_0)f'(x_0) + G'_x(\xi|x_0)f'(x_0) + G''_x(\xi|x_0)f(x_0)
\end{aligned}$$

$$- \left(G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0)\right)\frac{g'(\xi|x_0)}{g(\xi|x_0)}\Bigg\} (\xi^\epsilon - \xi)$$

$$+O(||\epsilon||^2).$$

Theorem 3.7 *If* $||\xi^\epsilon - \xi||^2 = O(||\epsilon||^2)$ *then,*

$$ABIAS_\epsilon^2 - ABIAS^2$$

$$= 23\left(-\frac{1}{23f^3(x_0)g(\xi|x_0)}\{g_x(\xi|x_0)f'(x_0) + G_x'(\xi|x_0)f'(x_0) + G_x''(\xi|x_0)f(x_0)\right)$$

$$- \left(G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0)\right)\frac{g'(\xi|x_0)}{g(\xi|x_0)}\Bigg\} (\xi^\epsilon - \xi) + O(||\epsilon||^2)$$

$$= -\frac{2}{23}\cdot\frac{G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0)}{f^3(x_0)g(\xi|x_0)}$$

$$\times\frac{1}{23f^3(x_0)g(\xi|x_0)}\{g_x(\xi|x_0)f'(x_0) + G_x'(\xi|x_0)f'(x_0) + G_x''(\xi|x_0)f(x_0)$$

$$- \left(G_{xx}(\xi|x_0)f(x_0) + 2G_x(\xi|x_0)f'(x_0)\right)\frac{g'(\xi|x_0)}{g(\xi|x_0)}\Bigg\} (\xi^\epsilon - \xi) + O(||\epsilon||^2).$$

3.2 Approximation of Conditional Quantiles with Measurement Erros

3.2.1 First order approximation

Let $G(\cdot|x)$ be the conditional c.d.f. of Y given x then

$$G(\cdot|x+\epsilon) = G(\cdot|x) + G'(\cdot|x)\epsilon + O(||\epsilon||^2) \qquad (3.12)$$

as $\epsilon \to 0$. where $G'(\cdot|x)$ is the first derivetive of $G(\cdot|x)$ with respect to x. And as $||\tau - \theta|| \to 0$,

$$G(\tau|x) = G(\theta|x) + g(\theta|x)(\tau - \theta) + O(||\tau - \theta||^2), \qquad (3.13)$$

where $g(\theta|x) = (\partial/\partial\tau)(G(\tau|x))|_{\tau=\theta}$. Substitute (3.13) into (3.12), we get

$$G(\tau|x+\epsilon) = G(\theta|x) + g(\theta|x)(\tau-\theta) + O(||\tau-\theta||^2) + g(\tau|x)\epsilon + O(||\epsilon||^2),$$

as $\epsilon \to 0$. If we let $G(\tau|x+u) = G(\theta|x) = p$,

$$g(\theta|x)(\tau-\theta) + O(||\tau-\theta||^2) + G'(\tau|x)\epsilon + O(||\epsilon||^2) = 0,$$

as $\epsilon \to 0$. So,

$$\tau = \theta - \frac{G'(\tau|x)}{g(\theta|x)}\epsilon + O(||\epsilon||^2) \tag{3.14}$$

as $\epsilon \to 0$ if

$$O(||\tau-\theta||^2) = O(||\epsilon||^2). \tag{3.15}$$

3.2.2 2nd order approximation

Let $G(\cdot|x)$ be the conditional c.d.f. of Y given x then

$$G(\cdot|x+\epsilon) = G(\cdot|x) + G'(\cdot|x)\epsilon + G''(\cdot|x)\epsilon^2/2 + o(||\epsilon||^2) \tag{3.16}$$

as $\epsilon \to 0$, where $G'(\cdot|x)$ and $G''(\cdot|x)$ are the first and second derivetives of $G(\cdot|x)$ with respect to x. And as $||\tau-\theta|| \to 0$,

$$G(\tau|x) = G(\theta|x) + g(\theta|x)(\tau-\theta) + g'(\theta|x)(\tau-\theta)^2/2 - o(||\tau-\theta||^2). \tag{3.17}$$

where $g(\theta|x) = (\partial/\partial\tau)(G(\tau|x))|_{\tau=\theta}$ and $g'(\theta|x) = (\partial^2/\partial\tau^2)(G(\tau|x))|_{\tau=\theta}$. Substitute (3.17) into (3.16), we get

$$\begin{aligned} G(\tau|x+\epsilon) &= G(\theta|x) + g(\theta|x)(\tau-\theta) + g'(\theta|x)(\tau-\theta)^2/2 + o(||\tau-\theta||^2) \\ &\quad + G'(\tau|x)\epsilon + G''(\tau|x)\epsilon^2/2 + o(||\epsilon||^2), \end{aligned}$$

as $\epsilon \to 0$. Since $G(\tau|x+\epsilon) = G(\theta|x) = p$,

$$g(\theta|x)(\tau-\theta) + g'(\theta|x)(\tau-\theta)^2/2 + o(||\tau-\theta||^2) + G'(\tau|x)\epsilon + G''(\tau|x)\epsilon^2/2 + o(||\epsilon||^2) = 0.$$

as $\epsilon \to 0$. If

$$||\tau - \theta||^2 = o(||\epsilon||^2) \text{ as } ||\tau - \theta|| \to 0, \tag{3.18}$$

then we have

$$G(\tau|x+\epsilon) = G(\theta|x) + g(\theta|x)(\tau - \theta) + G'(\tau|x)\epsilon + G''(\tau|x)\epsilon^2/2 + o(||\epsilon||^2).$$

as $\epsilon \to 0$. So,

$$\tau = \theta - \frac{G'(\tau|x)}{g(\theta|x)}\epsilon - \frac{G''(\tau|x)}{g(\theta|x)}\epsilon^2/2 + o(||\epsilon||^2), \tag{3.19}$$

as $\epsilon \to 0$.

3.2.3 2nd order approximation with symmetric argument

Let $G(\cdot|x)$ be the conditional c.d.f. of Y given x then

$$\begin{aligned} G(\cdot|x+\epsilon) &= G(\cdot|x) + G'(\cdot|x)\epsilon + G''(\cdot|x)\epsilon^2/2 + o(||\epsilon||^2) \\ G(\cdot|x-\epsilon) &= G(\cdot|x) - G'(\cdot|x)\epsilon + G''(\cdot|x)\epsilon^2/2 + o(||\epsilon||^2), \end{aligned}$$

as $\epsilon \to 0$, where $G'(\cdot|x)$ and $G''(\cdot|x)$ are the first and second derivetives of $G(\cdot|x)$ with respect to x. And

$$(G(\cdot|x+\epsilon) + G(\cdot|x-\epsilon))/2 = G^*(\cdot|x,\epsilon) = G(\cdot|x) + G''(\cdot|x)\epsilon^2/2 + o(||\epsilon||^2), \tag{3.20}$$

as $\epsilon \to 0$. $G^*(\cdot|x,\epsilon)$ can be considered the conditional c.d.f. of Y given x and ϵ. Remember that $T \overset{\mathcal{D}}{=} X + U \overset{\mathcal{D}}{=} X - U$ under the symmetric assumption of U. And as $||\tau - \theta|| \to 0$,

$$G(\tau|x) = G(\theta|x) + g(\theta|x)(\tau - \theta) + O(||\tau - \theta||^2), \tag{3.21}$$

where $g(\theta|x) = (\partial/\partial\tau)(G(\tau|x))|_{\tau=\theta}$. Substitute (3.21) into (3.20), we get

$$G^*(\tau|x,\epsilon) = G(\theta|x) + g(\theta|x)(\tau - \theta) + O(||\tau - \theta||^2) + G''(\tau|x)\epsilon^2/2 + o(||\epsilon||^2).$$

Since $G^*(\tau|x,\epsilon) = G(\theta|x) = p$,

$$\begin{aligned}\tau &= \theta - \frac{G''(\tau|x)}{g(\theta|x)}\epsilon^2/2 + o(||\epsilon||^2) + O(||\tau-\theta||^2) \\ &= \theta - \frac{G''(\tau|x)}{g(\theta|x)}\epsilon^2/2 + o(||\epsilon||^2),\end{aligned}$$

as $\epsilon \to 0$ if

$$||\tau-\theta||^2 = o(||\epsilon||^2) \text{ as } ||\tau-\theta|| \to 0. \tag{3.22}$$

So the question is the relationship between the magnitude of measurement error (ϵ) and the difference between the two quantiles ($\tau - \theta$).

Note that

$$\begin{aligned}\theta(x+\epsilon) &= \theta(x) + \theta'(x)\epsilon + \theta''(x)\epsilon^2/2 + o(||\epsilon||^2) \\ \theta(x-\epsilon) &= \theta(x) - \theta'(x)\epsilon + \theta''(x)\epsilon^2/2 + o(||\epsilon||^2),\end{aligned}$$

as $\epsilon \to 0$. By similar argument, we let

$$\tau \cong \frac{\theta(x+\epsilon)+\theta(x-\epsilon)}{2} = \theta(x) + \theta''(x)\epsilon^2/2 + o(||\epsilon||^2),$$

so that (3.18) and (3.22) hold.

3.2.4 Bivariate symmetric measurement error

Let $Y, X \sim f(Y,X)$ and

$$T = \begin{cases} X+\epsilon \text{ wp } 1/2 \\ X-\epsilon \text{ wp } 1/2 \end{cases}.$$

Then

$$\begin{aligned}f(Y|T=t) &= \frac{f_{Y,T}(Y,T=t)}{f_T(T=t)} \\ &= \frac{(f_{Y,X}(Y,t+\epsilon)+f_{Y,X}(Y,t-\epsilon))/2}{(f_X(t+\epsilon)+f_X(t-\epsilon))/2}\end{aligned}$$

$$= \frac{f_{Y,X}(Y,t+\epsilon) + f_{Y,X}(Y,t-\epsilon)}{f_X(t+\epsilon) + f_X(t-\epsilon)}.$$

And

$$\begin{aligned}
G^{\epsilon}_{Y|T}(y|t)\Big|_{\epsilon=0+} &= \left.\frac{\int_{-\infty}^{y} f_{Y,X}(z,t+\epsilon) + f_{Y,X}(z,t-\epsilon)\mathrm{d}z}{f_X(t+\epsilon) + f_X(t-\epsilon)}\right|_{\epsilon=0+} \\
&= \frac{\int_{-\infty}^{y} f_{Y,X}(z,t)\mathrm{d}z}{f_X(t)} \\
&= G_{Y|X}(y|t), \\
(\partial/\partial\epsilon)G^{\epsilon}_{Y|T}(y|t)\Big|_{\epsilon=0+} &= \left.\int_{-\infty}^{y} \frac{\partial}{\partial\epsilon}\left[\frac{f_{Y,X}(z,t+\epsilon) + f_{Y,X}(z,t-\epsilon)}{f_X(t+\epsilon) + f_X(t-\epsilon)}\right]\mathrm{d}z\right|_{\epsilon=0+} \\
&= 0,
\end{aligned}$$

and let

$$\begin{aligned}
(\partial^2/\partial\epsilon^2)G^{\epsilon}_{Y|T}(y|t) &= \left.\int_{-\infty}^{y} \frac{\partial^2}{\partial\epsilon^2}\left[\frac{f_{Y,X}(z,t+\epsilon) + f_{Y,X}(z,t-\epsilon)}{f_X(t+\epsilon) + f_X(t-\epsilon)}\right]\mathrm{d}z\right|_{\epsilon=0} \\
&= \int_{-\infty}^{y} \frac{4}{f_X^2}\left(f''(z,t)f_X(t) - f(z,t)f''(t)\right)\mathrm{d}z \\
&= A(y|t),
\end{aligned}$$

where

$$\begin{aligned}
f''(z,t) &= \left.\frac{\partial^2}{\partial\epsilon^2} f_{Y|X}(z,t+\epsilon)\right|_{\epsilon=0+} \\
f''(t) &= \left.\frac{\partial^2}{\partial\epsilon^2} f_X(t+\epsilon)\right|_{\epsilon=0+},
\end{aligned}$$

then we have

$$G^{\epsilon}_{Y|T}(y|t) = G_{Y|X}(y|t) + G''_{Y|X}(y|t)\Big|_{\epsilon=0}\epsilon^2/2 + o(||\epsilon||^2).$$

And

$$\frac{\partial}{\partial y}G(y|x) = g(y|x),$$

$$\frac{\partial^2}{\partial y^2}G(y|x) = g'(y|x).$$

And

$$G_{Y|X}(\tau|t) = G_{Y|X}(\theta|t) + g(\theta|t)(\tau-\theta) + g'(\theta|t)(\tau-\theta)^2/2 + o(||\tau-\theta||^2).$$

Therefore, if

$$||\tau-\theta||^2 = O(||\epsilon||^2), \tag{3.23}$$

the relationship between θ and τ is

$$\begin{aligned} G^{\epsilon}_{Y|T}(\tau|t) &= G_{Y|X}(\theta|t) + g(\theta|t)(\tau-\theta) + g'(\theta|t)(\tau-\theta)^2/2 + A(\tau|t)\epsilon^2/2 + O(||\epsilon||^3). \\ 0 &= g(\theta|t)(\tau-\theta) + g'(\theta|t)(\tau-\theta)^2/2 + A(\tau|t)\epsilon^2/2 + O(||\epsilon||^3). \end{aligned}$$

Example 3.1 *In the normal case, (3.23) becomes,*

$$\begin{aligned} G(\tau|t,\epsilon) &= G(\theta|t) + g(\theta|t)(\tau-\theta) \\ &\quad -g(\theta|t)\frac{\theta-\mu}{\sigma}(\tau-\theta)^2/2 + A(\tau,t)\epsilon^2/2 + O(||\epsilon||^3), \\ 0 &= g(\theta|t)(\tau-\theta) - g(\theta|t)\frac{\theta-\mu}{\sigma}(\tau-\theta)^2/2 + A(\tau|t)\epsilon^2/2 + O(||\epsilon||^3) \end{aligned} \tag{3.24}$$

where μ and σ are the contitional mean and standard deviation of Y given $X+x$.

3.2.5 Relationship between two quantiles: under normal assumptions

Let

$$Y|x \sim N(a+bx, \sigma^2),$$

then for small ϵ (>0)

$$Y|x+\epsilon \sim N(a+b(x+\epsilon), \sigma^2). \tag{3.25}$$

Let $\theta(x)$ be the $p \times 100$ % quantile of $Y|x$, and $\tau(x+\epsilon)$ be the $p \times 100$ % quantile of $Y|x+\epsilon$, i.e.,

$$\Phi\left(\frac{\theta(x)-a-bx}{\sigma}\right) = \Phi\left(\frac{\tau(x+\epsilon)-a-bx-b\epsilon}{\sigma}\right) = p.$$

So in the normal case, we have

$$||\theta(x)-\tau(x+\epsilon)|| = O(||\epsilon||).$$

Therefore, (3.14) and (3.23) hold in the normal case.

Example 3.2 *Under the assumption of (3.25), (3.14) becomes*

$$\tau = \theta + \frac{a}{\sigma}\frac{\phi((\tau-ax-b)/\sigma)}{\phi((\theta-ax-b)/\sigma)}\epsilon + O(||\epsilon||^2).$$

Let $Y, X \sim BN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho_{YX})$ and $T = X + U$ where $U \sim N(0, \sigma_U^2)$ and X and U are assumed to be independent. Then $T \sim N(\mu_X, \sigma_X^2 + \sigma_U^2)$. We are interested in the difference between the conditional quantiles of Y given X and given T. Under these assumptions,

$$\begin{aligned} Y|X = x &\sim N\left(\mu_Y + \rho_{XY}\frac{\sigma_Y}{\sigma_X}(x-\mu_X), \sigma_Y^2(1-\rho_{XY}^2)\right) \\ Y|T = t &\sim N\left(\mu_Y + \rho_{TY}\frac{\sigma_Y}{\sqrt{\sigma_X^2+\sigma_U^2}}(t-\mu_X), \sigma_Y^2(1-\rho_{TY}^2)\right) \end{aligned}$$

Our parameters, conditional quantiles θ, τ are expressed as

$$\Phi(\theta|x) = p = \Phi(\tau|t).$$

If we restrict to $p = 1/2$, i.e., median, then we have

$$\begin{aligned} \theta &= \mu_Y + \rho_{XY}\frac{\sigma_Y}{\sigma_X}(x-\mu_X) \\ \tau &= \mu_Y + \rho_{TY}\frac{\sigma_Y}{\sqrt{\sigma_X^2+\sigma_U^2}}(t-\mu_X). \end{aligned}$$

So,

$$\theta - \tau = \sigma_Y \left(\rho_{XY} \frac{x-\mu_X}{\sigma_X} - \rho_{TY} \frac{t-\mu_X}{\sqrt{\sigma_X^2 + \sigma_U^2}} \right),$$

$$\theta = \tau + \frac{\sigma_Y}{\sigma_X}(\rho_{XY} x - \rho_{TY} t) - \frac{\sigma_Y}{\sigma_X}\mu_X \left(\rho_{XY} - \frac{\rho_{TY}}{\sqrt{1+\sigma_U^2/\sigma_X^2}} \right).$$

So.

$$\theta(x) - \tau(t) = \frac{\sigma_Y}{\sigma_X}(\rho_{XY} x - \rho_{TY} t) - \frac{\sigma_Y}{\sigma_X}\mu_X \left(\rho_{XY} - \frac{\rho_{TY}}{\sqrt{1+\sigma_U^2/\sigma_X^2}} \right).$$

In general. we may assume $\rho_{XY} < \rho_{TY}$, so we may let $\rho_{TY} = c\rho_{XY}$, $(0 < c < 1)$ then

$$\theta(x) - \tau(t) = \frac{\sigma_Y}{\sigma_X}\rho_{XY}(x - ct) - \frac{\sigma_Y}{\sigma_X}\mu_X \rho_{XY} \left(1 - \frac{c}{\sqrt{1+\sigma_U^2/\sigma_X^2}} \right).$$

For general p, $(0 < p < 1)$, we have

$$\int_{-\infty}^{\theta(x)} \phi_{Y|x}(y)\mathrm{d}y = \int_{-\infty}^{\tau(t)} \phi_{Y|t}(y)\mathrm{d}y = p$$

$$\Phi_{Y|x}(\theta(x)) = \Phi_{Y|t}(\tau(t)) = p$$

Again. let $\rho_{TY} = c\rho_{XY}$, $(0 < c < 1)$ then

$$\Phi\left(\frac{\theta - \mu_Y - \rho_{XY}\sigma_Y/\sigma_X(x-\mu_X)}{\sigma_Y\sqrt{1-\rho_{XY}^2}} \right) = \Phi\left(\frac{\tau - \mu_Y - c\rho_{XY}\sigma_Y/\sqrt{\sigma_X^2+\sigma_U^2}(t-\mu_X)}{\sigma_Y\sqrt{1-c^2\rho_{XY}^2}} \right)$$

$$\begin{aligned} \sqrt{1-c^2\rho_{XY}^2}\theta - \sqrt{1-\rho_{XY}^2}\tau &= \rho_{XY}\frac{\sigma_Y}{\sigma_X}\left(\sqrt{1-c^2\rho_{XY}^2}x - \frac{\sqrt{1-\rho_{XY}^2}}{\sqrt{1+\sigma_U^2/\sigma_X^2}}\tau \right) \\ &\quad + \mu_X\rho_{XY}\frac{\sigma_Y}{\sigma_X}\left(\sqrt{\frac{1-\rho_{XY}^2}{1+\sigma_U^2/\sigma_X^2}} - \sqrt{1-c^2\rho_{XY}^2} \right). \end{aligned}$$

Let $Y, X \sim BN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$ and

$$T = \begin{cases} X + \epsilon \text{ wp } 1/2 \\ X - \epsilon \text{ wp } 1/2 \end{cases}$$

Then

$$
\begin{aligned}
f(Y|T) &= \frac{f(Y,T)}{f(T)} \\
&= \frac{(f(Y,T+\epsilon)+f(Y,T-\epsilon))/2}{(f(T+\epsilon)+f(T-\epsilon))/2} \\
&= \frac{\text{Mixture of two bivariate normals}}{\text{Mixture of two univariate normals}}.
\end{aligned}
$$

$$
\begin{aligned}
G(y|T=t) &= \int_{-\infty}^{y} f(z|T=t)\mathrm{d}z \\
&= \int_{-\infty}^{y} \frac{\phi_{Y,X}(z,t+\epsilon)+\phi_{Y,X}(z,t-\epsilon)}{\phi_X(t+\epsilon)+\phi_X(t-\epsilon)}\mathrm{d}z \\
&= \frac{1}{\phi_X(t+\epsilon)+\phi_X(t-\epsilon)} \\
&\quad \times \int_{-\infty}^{y} \left(\frac{\phi_{Y,X}(z,t+\epsilon)}{\phi_X(t+\epsilon)}\phi_X(t+\epsilon)+\frac{\phi_{Y,X}(z,t-\epsilon)}{\phi_X(t-\epsilon)}\phi_X(t-\epsilon)\mathrm{d}z\right) \\
&= \frac{1}{\phi_X(t+\epsilon)+\phi_X(t-\epsilon)} \\
&\quad \times \int_{-\infty}^{y} \phi_{Y|X}(z|t+\epsilon)\phi_X(t+\epsilon)+\phi_{Y|X}(z|t-\epsilon)\phi_X(t-\epsilon)\mathrm{d}z \\
&= \frac{1}{\phi_X(t+\epsilon)+\phi_X(t-\epsilon)} \\
&\quad \times \left(\Phi_{Y|X}(z|t+\epsilon)\phi_X(t+\epsilon)+\Phi_{Y|X}(z|t-\epsilon)\phi_X(t-\epsilon)\right).
\end{aligned}
$$

So the relationship is

$$
G(\tau|t) = \frac{1}{\phi_X(t+\epsilon)+\phi_X(t-\epsilon)}
$$

$$\times \left(\Phi_{Y|X}(z|t+\epsilon)\phi_X(t+\epsilon) + \Phi_{Y|X}(z|t-\epsilon)\phi_X(t-\epsilon)\right)$$

$$= \Phi_{Y|X}(\theta|x) = p,$$

$$\Phi\left(\frac{\theta - a - bx}{\sigma}\right) = \frac{1}{\phi_X(t+\epsilon) + \phi_X(t-\epsilon)}$$
$$\times \left\{\Phi\left(\frac{\tau - a - bt - b\epsilon}{\sigma}\right) + \Phi\left(\frac{\tau - a - bt + b\epsilon}{\sigma}\right)\right\}.$$

where $Y|X = x \sim N(a + bx, \sigma^2)$ for $a = \mu_Y - \rho\sigma_Y/\sigma_X\mu_X, b = \rho\sigma_Y/\sigma_X$ and $\sigma^2 = \sigma_Y^2(1-\rho^2)$.

3.2.6 Approximation of conditional cdf with bivariate normality and uniform measurement error

Condition 3.1 *Let (Y, X) has bivariate normal distribution and $T = X + U$. $U \sim U(-\epsilon, \epsilon)$. And X and U are assumed to be independent. Let $G_\epsilon(\cdot|t)$ be the conditional cdf of Y given $T = t$.*

Theorem 3.8 *Under condition 3.1, as $\epsilon \to 0+$,*

$$G_\epsilon(y|t) = \Phi\left(\frac{y - at - b}{\sigma_1}\right) - \phi\left(\frac{y - at - b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)$$
$$\times \left[\frac{\phi((t-\epsilon-\mu_X)/\sigma_X) - \phi((t+\epsilon-\mu_X)/\sigma_X)}{\Phi((t+\epsilon-\mu_X)/\sigma_X) - \Phi((t-\epsilon-\mu_X)/\sigma_X)}\sigma_X + \mu_X - t\right]$$
$$+o(||\epsilon||^2)$$
$$= \Phi_{Y|X}(y|t) - \phi_{Y|X}(y|t)\left(\frac{1}{\sigma_1}\right)$$
$$\times \left[\frac{\phi_X(t-\epsilon) - \phi_X(t+\epsilon)}{\Phi_X(t+\epsilon) - \Phi_X(t-\epsilon)}\sigma_X^2 + \mu_X - t\right] + o(||\epsilon||^2),$$

and

$$\lim_{\epsilon \to 0+} G_\epsilon(y|t) = \Phi_{Y|X}(y|t)$$

as $\epsilon \to 0+$. *where* $\Phi(\cdot)$ *and* $\phi(\cdot)$ *are the cdf and pdf of standard normal distribution.* $ax + b$ *and* σ_1 *are the conditional mean and the conditional standard deviation of* Y *given* $X = x$, μ_X *and* σ_X *are the marginal mean and the standard deviation of* X.

Proof: Let

$$\Phi(y|x) = \Phi\left(\frac{y-(ax+b)}{\sigma_1}\right),$$

where $a = \rho_{XY}\sigma_Y/\sigma_X$, $b = \mu_Y - \rho_{XY}\sigma_Y/\sigma_X\mu_X$ and $\sigma_1 = \sqrt{\sigma_Y^2(1-\rho_{XY}^2)}$. Then

$$\begin{aligned}
\Phi\left(\frac{y-(ax+b)}{\sigma_1}\right)\bigg|_{x=t} &= \Phi\left(\frac{y-at-b}{\sigma_1}\right)\\
\frac{\partial}{\partial x}\Phi\left(\frac{y-(ax+b)}{\sigma_1}\right)\bigg|_{x=t} &= -\phi\left(\frac{y-at-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\\
\frac{\partial^2}{\partial x^2}\Phi\left(\frac{y-(ax+b)}{\sigma_1}\right)\bigg|_{x=t} &= -\phi\left(\frac{y-at-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)^2\left(\frac{y-at-b}{\sigma_1}\right)
\end{aligned}$$

So

$$\begin{aligned}
\Phi(y|x) &= \Phi\left(\frac{y-at-b}{\sigma_1}\right) - \phi\left(\frac{y-ax-b}{\sigma_1}\right)\left(\frac{a}{\sigma}\right)(x-t)\\
&\quad -\phi\left(\frac{y-at-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)^2\left(\frac{y-at-b}{\sigma_1}\right)(x^*-t)^2/2, \qquad (3.26)
\end{aligned}$$

for $|x^* - t| < |x - t|$. Note that

$$\frac{\partial}{\partial x}\phi(ax+b) = -\phi(ax+b)(ax+b)a, \qquad (3.27)$$

$$\frac{\partial}{\partial x}\phi_X(x) = -\frac{(x-\mu_X)}{\sigma_X^2}\phi_X(x), \qquad (3.28)$$

and

$$\frac{\int_{t-\epsilon}^{t+\epsilon} x\phi_X(x)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon}\phi_X(x)\mathrm{d}x} = \frac{\int_{t-\epsilon}^{t+\epsilon}\left(-\phi_X'(x)\sigma_X^2 + \mu_X\phi_X(x)\right)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon}\phi_X(x)\mathrm{d}x}$$

$$= \frac{-\left[\phi_X(t+\epsilon) - \phi_X(t-\epsilon)\right]\sigma_X^2}{\Phi_X(t+\epsilon) - \Phi_X(t-\epsilon)} + \mu_X$$

$$= \frac{-\left[\phi((t+\epsilon-\mu_X)/\sigma_X) - \phi((t-\epsilon-\mu_X)/\sigma_X)\right]}{\Phi((t+\epsilon-\mu_X)/\sigma_X) - \Phi((t-\epsilon-\mu_X)/\sigma_X)}\sigma_X + \mu_X \quad (3.29)$$

And by (3.1) and (3.26),

$$\begin{aligned} G_\epsilon(y|t) &= \frac{\int_{t-\epsilon}^{t+\epsilon}\Phi(y|x)\phi_X(x)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon}\phi_X(x)\mathrm{d}x} \\ &= \Phi\left(\frac{y-at-b}{\sigma_1}\right) \\ &\quad -\phi\left(\frac{y-at-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\left[\frac{\int_{t-\epsilon}^{t+\epsilon}x\phi_(x)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon}\phi_X(x)\mathrm{d}x} - t\right] \\ &\quad +R, \end{aligned}$$

where

$$\begin{aligned} |R| &< \text{Constant}\left|\frac{\int_{t-\epsilon}^{t+\epsilon}(x^*-t)^2\phi_X(x)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon}\phi_X(x)\mathrm{d}x}\right| \\ &< \text{Constant}\int_{t-\epsilon}^{t+\epsilon}|x-t|^2\mathrm{d}x \\ &= \text{Constant}\int_{t-\epsilon}^{t+\epsilon}(x^2-2xt+t^2)\mathrm{d}x \\ &= \text{Constant}\left[\frac{2}{3}\epsilon^3\right]. \end{aligned}$$

So. we have

$$|R| = o(||\epsilon||^2).$$

And

$$\begin{aligned} &\lim_{\epsilon\to 0+}\left[\frac{\phi_X(t-\epsilon) - \phi_X(t+\epsilon)}{\Phi_X(t+\epsilon) - \Phi_X(t-\epsilon)}\sigma_X^2 + \mu_x - t\right] \\ = &\lim_{\epsilon\to 0+}\left[\frac{\phi_X(t-\epsilon)((t-\epsilon-\mu_X)/\sigma_X^2) + \phi_X(t+\epsilon)((t+\epsilon-\mu_X)/\sigma_X^2)}{\phi_X(t+\epsilon) + \phi_X(t-\epsilon)}\sigma_X^2 + \mu_x - t\right] \end{aligned}$$

$$= \quad t - \mu_x + \mu_x - t$$

$$= \quad 0$$

The theorem is proved. $\diamond$

Lemma 3.3 *Under condition 3.1, theorem 3.8 becomes*

$$\begin{aligned} G_\epsilon(y|t) &= \Phi_{Y|X}(y|t) - \phi_{Y|X}(y|t)\left(\frac{1}{\sigma_1}\right) \\ &\quad \times \left[\frac{\phi_X(t-\epsilon) - \phi_X(t+\epsilon)}{\Phi_X(t+\epsilon) - \Phi_X(t-\epsilon)}\sigma_X^2 + \mu_X - t\right] + o(||\epsilon||^2) \\ &= \Phi_{Y|X}(y|t) - \phi_{Y|X}(y|t)\left(\frac{1}{\sigma_1}\right) \\ &\quad \times \left[\frac{\phi_X(t)}{\Phi_X(t))}\sigma_X^2 + \mu_X - t\right] + O(||\epsilon||^2), \end{aligned}$$

as $\epsilon \to 0+$.

Extend these to general $G_X(\cdot)$ and $g(x) = g((x-\mu)/\sigma)$.

What about

$$\int_a^b \phi'(x)\mathrm{d}x = \phi(a) - \phi(b) \ \ ?$$

3.2.7 Conditional quantile from general cdf with uniform measure error

In this section, we consider general cdf of (Y, X) and uniform measurement error. Let $Y, X \sim G_{YX}(y, x)$. $X \sim G_X(\cdot)$ and $T = X + U$ where $U \sim \mathrm{U}(-\epsilon, \epsilon)$ and X and U are assumed to be independent. Then, under non-differential ME assumption, we have

$$G_T(t) \quad = \quad \int_{t-\epsilon}^{t+\epsilon} G_X(u)\frac{1}{2\epsilon}\mathrm{d}u$$

$$
\begin{aligned}
g_T(t) &= \int_{t-\epsilon}^{t+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u \\
G^{\epsilon}_{Y|T}(y|t) &= \frac{\int_{t-\epsilon}^{t+\epsilon} G_{Y|X}(y|x)g_X(x)\mathrm{d}x}{\int_{t-\epsilon}^{t+\epsilon} g_X(x)\mathrm{d}x}.
\end{aligned}
$$

Without ME, let $Z = |X - x_0|$, then we have

$$
\begin{aligned}
f_Z(z) &= g_X(x_0+z) + g_X(x_0-z), \\
g(y|z) &= \left[g_X(x_0+z)g(y|x_0+z) + g_X(x_0-z)g(y|x_0-z)\right]/f_Z(z), \\
G_{Y|Z}(y|z) &= \left[g_X(x_0+z)G_{Y|X}(y|x_0+z) + g_X(x_0-z)G_{Y|X}(y|x_0-z)\right]/f_Z(z).
\end{aligned}
$$

If we let $W = |T - x_0|$, then we have the following results,

$$
\begin{aligned}
f_W(w) &= g_T(x_0+w) + g_T(x_0-w) \\
&= \int_{x_0+w-\epsilon}^{x_0+w+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u, \\
g(y|w) &= \left[g_T(x_0+w)g(y|x_0+w) + g_T(x_0-w)g(y|x_0-w)\right]/f_W(w), \\
G^{\epsilon}_{Y|W}(y|w) &= \left[g_T(x_0+w)G^{\epsilon}_{Y|T}(y|x_0+w) + g_T(x_0-w)G^{\epsilon}_{Y|T}(y|x_0-w)\right]/f_W(w) \\
&= \left[\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u \times \frac{\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} G_{Y|X}(y|x)g_X(x)\mathrm{d}x}{\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} g_X(x)\mathrm{d}x}\right. \\
&\quad \left. + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u \times \frac{\int_{x_0-w-\epsilon}^{x_0-w+\epsilon} G_{Y|X}(y|x)g_X(x)\mathrm{d}x}{\int_{x_0-w-\epsilon}^{x_0-w+\epsilon} g_X(x)\mathrm{d}x}\right] \\
&\quad / \left[\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} g_X(u)\frac{1}{2\epsilon}\mathrm{d}u\right] \\
&= \left[\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} G_{Y|X}(y|x)g_X(x)\mathrm{d}x + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} G_{Y|X}(y|x)g_X(x)\mathrm{d}x\right] \\
&\quad / \left[\int_{x_0+w-\epsilon}^{x_0+w+\epsilon} g_X(u)\mathrm{d}u + \int_{x_0-w-\epsilon}^{x_0-w+\epsilon} g_X(u)\mathrm{d}u\right].
\end{aligned}
$$

Let

$$G^{\epsilon}_{Y|W}(y|w) = \frac{1}{\int_A g_X(u)\mathrm{d}u + \int_B g_X(u)\mathrm{d}u} \times \left[\int_A G_{Y|X}(y|x)g_X(x)\mathrm{d}x + \int_B G_{Y|X}(y|x)g_X(x)\mathrm{d}x\right]$$

.

If we do Taylor Expansion, we have

$$\begin{aligned}\lim_{\epsilon\to 0+} G^{\epsilon}_{Y|W}(y|w) &= \left[G_{Y|X}(y|x_0+\epsilon)g_X(x_0+\epsilon)G_{Y|X}(y|x_0-\epsilon)g_X(x_0-\epsilon)\right] \\ &\quad \Big/ \left[g_X(x_0+\epsilon)+g_X(x_0-\epsilon)\right] \\ &= G_{Y|Z}(y|w),\end{aligned}$$

and

$$\lim_{\epsilon\to 0+} \frac{\partial}{\partial\epsilon} G^{\epsilon}_{Y|W}(y|w) = 0.$$

So, we get

$$G^{\epsilon}_{Y|W}(y|w) = G_{Y|Z}(y|w) + \frac{\partial^2}{\partial\epsilon^2} G^{\epsilon}_{Y|W}(y|w)\bigg|_{\epsilon\to 0+} \epsilon^2/2 + o(||\epsilon||^2).$$

Theorem 3.9 *Let $(Y,X) \sim BN$, $T = X+U$, $U \sim U(-\epsilon,\epsilon)$ and $W = |T-x_0|$. And X and U are assumed to be independent. Let $G^{\epsilon}_{Y|W}(y|w)$ be the conditional cdf of Y given $W = w$, then*

$$G^{\epsilon}_{Y|W}(y|w) = G^{*\epsilon}_{Y|T}(y|t) + (A^* + B^*) + o(||\epsilon||^2),$$

where

$$G^{*\epsilon}_{Y|T}(y|t) = \frac{1}{(\int_A \phi_X(x)dx + \int_B \phi_X(x)dx)} \times \left(\Phi_{Y|X}(y|x_0+w)\int_A \phi_X(x)dx + \Phi_{Y|X}(y|x_0-w)\int_B \phi_X(x)dx\right)$$

$$= \frac{\phi_X(x_0+w)+\phi_X(x_0-w)}{\Phi_{Y|X}(y|x_0+w)\phi_X(x_0+w)+\Phi_{Y|X}(y|x_0-w)\phi_X(x_0-w)} + O(||\epsilon||^2)$$

$$= G_{Y|T}(y|t) + O(||\epsilon||^2)$$

$$\begin{aligned}
A^* &= \frac{1}{(\int_A \phi_X(x)dx + \int_B \phi_X(x)dx)} \\
&\times - \left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\int_A (-\phi'(x)+\mu_X\phi_X(x))\,dx\right. \\
&\left. + \phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\int_B (-\phi'(x)+\mu_X\phi_X(x))\,dx\right] \\
&= \frac{1}{(\phi_X(x_0+w)+\phi_X(x_0-w))} \\
&\times \left\{\left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi'_X(x_0+w)\right.\right. \\
&\left. +\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi'_X(x_0-w)\right] \\
&-\mu_X\left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi_X(x_0+w)\right. \\
&\left.\left. + \phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi'_X(x_0-w)\right]\right\} + O(||\epsilon||^2). \\
&= \frac{1}{(\phi_X(x_0+w)+\phi_X(x_0-w))} \times \left\{\left[\phi_{Y|X}(y|x_0+w)a\phi'_X(x_0+w)\right.\right. \\
&\left. +\phi_{Y|X}(y|x_0-w)a\phi'_X(x_0-w)\right] - \mu_X\left[\phi_{Y|X}(y|x_0+w)a\phi_X(x_0+w)\right. \\
&\left.\left. + \phi_{Y|X}(y|x_0-w)a\phi'_X(x_0-w)\right]\right\} + O(||\epsilon||^2).
\end{aligned}$$

$$
\begin{aligned}
B^* &= \frac{1}{(\int_A \phi_X(x)dx + \int_B \phi_X(x)dx)} \\
&\quad \times \left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0+w)\int_A \phi_X(x)dx \right. \\
&\quad \left. + \phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0-w)\int_B \phi_X(x)dx\right] \\
&= \frac{1}{(\phi_X(x_0+w)+\phi_X(x_0-w))} \\
&\quad \times \left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0+w)\phi_X(x_0+w) \right. \\
&\quad \left. +\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0-w)\phi_X(x_0-w)\right] + O(||\epsilon||^2). \\
&= \frac{1}{(\int_A \phi_X(x)dx + \int_B \phi_X(x)dx)} \\
&\quad \times \left[\phi_{Y|X}(y|x_0+w)a\phi_X(x_0+w)\phi_{Y|X}(y|x_0)a\phi_X(x_0-w)\right] + O(||\epsilon||^2).
\end{aligned}
$$

$A = (x_0+w-\epsilon, x_0+w+\epsilon)$, $B = (x_0-w-\epsilon, x_0-w+\epsilon)$ and $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and pdf of standard normal distribution, $ax+b$ and σ_1 are the conditional mean and the conditional standard deviation of Y given $X = x$, μ_X and σ_X are the marginal mean and the standard deviation of X.

Proof: Let

$$\Phi_{Y|X}(y|x) = \Phi\left(\frac{y-(ax+b)}{\sigma_1}\right),$$

where $a = \rho_{XY}\sigma_Y/\sigma_X$, $b = \mu_Y - \rho_{XY}\sigma_Y/\sigma_X\mu_X$ and $\sigma_1 = \sqrt{\sigma_Y^2(1-\rho_{XY}^2)}$. Then

$$\Phi_{Y|X}(y|x)\Big|_{x=x_0+w} = \Phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)$$

$$\frac{\partial}{\partial x}\Phi\left(\frac{y-(ax+b)}{\sigma_1}\right)\Bigg|_{x=x_0+w} = -\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)$$

So

$$\begin{aligned}
G^{\epsilon}_{Y|W}(y|w) & \\
&= \frac{1}{(\int_A \phi_X(x)\mathrm{d}x + \int_B \phi(x)\mathrm{d}x)} \\
&\times\Bigg\{\int_A\Bigg[\Phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right) \\
&-\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma}\right)(x-x_0-w)+R\Bigg]\phi_X(x)\mathrm{d}x \\
&+\int_B\Bigg[\Phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right) \\
&-\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma}\right)(x-x_0-w)+R\Bigg]\phi_X(x)\mathrm{d}x\Bigg\}
\end{aligned}$$

And

$$\begin{aligned}
G^{*\epsilon}_{Y|Z}(y|w) &= \left(\Phi_{Y|X}(y|x_0+w)\int_A \phi_X(x)\mathrm{d}x + \Phi_{Y|X}(y|x_0-w)\int_B \phi_X(x)\mathrm{d}x\right) \\
&\quad \Big/\left(\int_A \phi_X(x)\mathrm{d}x + \int_B \phi_X(x)\mathrm{d}x\right) \\
&\to G_{Y|Z}(y|w) \text{ as } \epsilon \to 0+.
\end{aligned}$$

$$\begin{aligned}
A &= -\Bigg[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\int_A x\phi_X(x)\,\mathrm{d}x \\
&+\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\int_B x\phi_X(x)\mathrm{d}x\Bigg] \\
&\Big/\left(\int_A \phi_X(x)\mathrm{d}x + \int_B \phi_X(x)\mathrm{d}x\right)
\end{aligned}$$

$$
\begin{aligned}
&= -\left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\int_A(-\phi'(x)+\mu_X\phi_X(x))\,\mathrm{d}x\right.\\
&\quad\left.+\,\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\int_B(-\phi'(x)+\mu_X\phi_X(x))\,\mathrm{d}x\right]\\
&\quad\Big/\left(\int_A\phi_X(x)\mathrm{d}x+\int_B\phi_X(x)\mathrm{d}x\right)\\
&\to -\left[\phi\left(\frac{y-(ax_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(-\phi'(x_0+w)+\mu_X\phi_X(x_0+w))\right.\\
&\quad\left.+\,\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(-\phi'(x_0-w)+\mu_X\phi_X(x_0-w))\,\mathrm{d}x\right]\\
&\quad\Big/(\phi_X(x_0+w)+\phi_X(x_0-w))\\
&= \left\{\left[\phi\left(\frac{y-(ax_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi'_X(x_0+w)\right.\right.\\
&\quad\left.+\phi\left(\frac{y-(ax_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi'_X(x_0-w)\right]\\
&\quad-\mu_X\left[\phi\left(\frac{y-(ax_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi_X(x_0+w)\right.\\
&\quad\left.\left.+\,\phi\left(\frac{y-(ax_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)\phi'_X(x_0-w)\right]\right\}\\
&\quad\Big/(\phi_X(x_0+w)+\phi_X(x_0-w)) \text{ as } \epsilon\to 0+
\end{aligned}
$$

$$
\begin{aligned}
B &= \left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0+w)\int_A\phi_X(x)\mathrm{d}x\right.\\
&\quad\left.+\,\phi\left(\frac{y-a(x_0-w)-b}{\sigma 1}\right)\left(\frac{a}{\sigma_1}\right)(x_0-w)\int_B\phi_X(x)\mathrm{d}x\right]\\
&\quad\Big/\left(\int_A\phi_X(x)\mathrm{d}x+\int_B\phi_X(x)\mathrm{d}x\right)
\end{aligned}
$$

$$
\begin{aligned}
\to \quad & \left[\phi\left(\frac{y-a(x_0+w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0+w)\phi_X(x_0+w)\right. \\
& \left.+\phi\left(\frac{y-a(x_0-w)-b}{\sigma_1}\right)\left(\frac{a}{\sigma_1}\right)(x_0-w)\phi_X(x_0-w)\right] \\
& \Big/ \left(\phi_X(x_0+w)+\phi_X(x_0-w)\right)
\end{aligned}
$$

$$
\begin{aligned}
|R| & < \int_A (x-(x_0+w))^2 \mathrm{d}x + \int_B (x-(x_0-w))^2 \mathrm{d}x \\
& = o(||\epsilon||^2)
\end{aligned}
$$

Chapter 4

Multivariate Conditional Quantile Process

4.1 Multivariate conditional quantile

In a multivariate set-up, suppose that $(Y_i, \boldsymbol{X}_i)$, $i = 1, \cdots, n$ are independent and identically distributed (i.i.d) random vectors with a $(q+1)$-dimensional distribution function (d.f.) $\Pi(\cdot)$.

For every $\boldsymbol{x}_0 \in \mathrm{R}^q$ and $\boldsymbol{x} \in \mathrm{R}^q$,

$$\begin{aligned} \rho(\boldsymbol{x}, \boldsymbol{x}_0) &: \mathrm{R}^q \times \mathrm{R}^q \to \mathrm{R}^+ \\ \rho(\boldsymbol{x}_i, \boldsymbol{x}_0) &= D_i \text{ say for } i = 1, \cdots, n. \end{aligned}$$

We choose a sequence $\{k_n\}$, such that k_n is increasing in n and $n^{-1}k_n$ converges to 0 as $n \to \infty$. We denote the order statistics of D_i by $D_{n:1}, \cdots, D_{n:n}$. If the distribution of D is continuous, ties among the D_i may be neglected with probability one.

We define the anti-ranks S_i, $i = 1, \cdots, n$ by letting $D_{n:i} = D_{S_i}$, for $i = 1, \cdots, n$. Thus, $(S_1, \cdots, S_n)$ is a random permutation of the numbers $(1, \cdots, n)$. Consider the set of observations:

$$(Y_{S_i}, \boldsymbol{X}_{S_i}),\ i = 1, \cdots, k_n.$$

The k_n-nearest neighbor (k_n-NN) empirical d.f. of Y with respect to $\boldsymbol{x}_0$ and $\rho(\cdot)$ is defined as

$$G_{nk_n}(y) \quad = \quad k_n^{-1} \sum_{i=1}^{k} I(Y_{S_i} \leq y),\ y \in \mathrm{R}. \tag{4.1}$$

$$\hat{\xi}_{nk_n}(\boldsymbol{x}_0) \quad = \quad \inf\{y : G_{nk_n}(y) \geq [pk_n]/k_n\}\ ;\ p \in (0,1)$$

The Bhattachatya (1974) lemma yields that the Y_{S_i} are conditionally (given the $\boldsymbol{X}_{S_i}$) independent r.v.'s, although they are not necessarily identically distributed. However. the very construction of the neighborhood makes the marginal (conditional) distributions of the Y_{S_i}in this set nearly identical, a fact that has been established through contiguity of (conditional) probability measures by Gangopadhyay and Sen (1992). We consider the function in a general mold as

$$\Xi(C) = \{\xi(\boldsymbol{x}) : \boldsymbol{x} \in C\},$$

where C is (usually a compact) subset of R^q which includes the essential support of the distribution of the concomitant variable $\boldsymbol{X}$.

The following regularity conditions are assumed as Sen (1994):

[A1] Let $F(\boldsymbol{x})$, $\boldsymbol{x} \in R^q$ be the joint d.f. of the covariate $\boldsymbol{X}$. We assume that F admits an absolutely continuous (q-variate) density function $f(\boldsymbol{x})$, $\boldsymbol{x} \in R^q$. such that

$$f(\boldsymbol{x}) > 0, \text{ for every } \boldsymbol{x} \in C, \tag{4.2}$$

and the Hessian $(\partial^2/\partial\boldsymbol{x}\partial\boldsymbol{x}')(f(\boldsymbol{x}))$ exists for all $\boldsymbol{x} \in C$. Moreover, there exist positive and finite numbers ϵ, K_0 and α, such that $||\boldsymbol{x} - \boldsymbol{x}^*|| < \epsilon$ implies that

$$||(\partial^2/\partial\boldsymbol{x}\partial\boldsymbol{x}')f(\boldsymbol{x})|_{\boldsymbol{x}} - (\partial^2/\partial\boldsymbol{x}\partial\boldsymbol{x}')f(\boldsymbol{x})|_{\boldsymbol{x}^*}|| \leq K_0||\boldsymbol{x} - \boldsymbol{x}^*||^\alpha, \text{ uniformly in } C.$$

Note that the D_i are defined with respect to a pivot $\boldsymbol{x}_0$, and hence, both $F_D(d)$ and $f_D(d)$ depend on this pivot $\boldsymbol{x}_0$. We therefore write $f_D(d)$ as $f_D(d; \boldsymbol{x}_0)$. Note that by definition,

$$f_D(d; \boldsymbol{x}_0) = \int_{\{\boldsymbol{x}; \rho(\boldsymbol{x}, \boldsymbol{x}_0)=d\}} f(\boldsymbol{x}) d\nu(\boldsymbol{x}; \boldsymbol{x}_0), \; d \in \mathrm{R}^+, \tag{4.3}$$

where $\nu(\boldsymbol{x}; \boldsymbol{x}_0)$ stands for the measurement on the segment $\{\boldsymbol{x}; \rho(\boldsymbol{x}, \boldsymbol{x}_0) = d\}$.

[A2] We denote the conditional distribution of Y_i given $\boldsymbol{X}_i = \boldsymbol{x}_0$, by $g(y|\boldsymbol{x}_0)$ and assume that $G(\cdot)$ has a continuous density $g(y|\boldsymbol{x}_0)$ for almost all $(y, \boldsymbol{x}_0)$. And $(\partial/\partial \boldsymbol{x}) g(y|\boldsymbol{x})$ and $(\partial^2/\partial \boldsymbol{x} \partial \boldsymbol{x}') g(y|\boldsymbol{x})$ exist in a neighborhood of $\boldsymbol{x}_0$. The conditional d.f. and density of Y_i given $D_i = d$(with respect to the pivot $\boldsymbol{x}_0$), are respectively given by

$$G_D(y|\boldsymbol{x}_0, d) = \left\{ \int_{\{\boldsymbol{x}; \rho(\boldsymbol{x}, \boldsymbol{x}_0)=d\}} G(y|\boldsymbol{x}) f(\boldsymbol{x}) d\nu(\boldsymbol{x}; \boldsymbol{x}_0) \right\} \Big/ f_D(d; \boldsymbol{x}_0). \tag{4.4}$$

and

$$g_D(y|\boldsymbol{x}_0, d) = \left\{ \int_{\{\boldsymbol{x}; \rho(\boldsymbol{x}, \boldsymbol{x}_0)=d\}} g(y|\boldsymbol{x}) f(\boldsymbol{x}) d\nu(\boldsymbol{x}; \boldsymbol{x}_0) \right\} \Big/ f_D(d; \boldsymbol{x}_0),$$

where $\boldsymbol{x}_0 \in C$ and $y \in R$.

[A3] We assume that there exists a positive d_0, such that the first and second order partial derivatives of $g_D(y|\boldsymbol{x}_0, d)$ with respect to d, denoted by $g'_D(\cdot)$ and $g''_D(\cdot)$ respectively, exist for almost all $y, \boldsymbol{x}_0$ and $d : 0 < d \leq d_0$, and these are dominated by some Lebesgue measurable functions. Let

$$|g'_D(y|\boldsymbol{x}_0, d)| \leq u_1(y) \tag{4.5}$$

$$|g''_D(y|\boldsymbol{x}_0, d)| \leq u_2(y). \tag{4.6}$$

Moreover,

$$|g''_D(y|\boldsymbol{x}_0, d) - g''_D(y|\boldsymbol{x}_0, 0)| \leq u_3(y) d^{\alpha}, \tag{4.7}$$

where $u_3(y)$ is a Lebesgue measurable, α is a positive number and (4.7) holds for all $d \leq d_0$ and $\boldsymbol{x}_0 \in C$.

[**A4**] We assume that

$$I(y|\boldsymbol{x}_0, d) = E[\{(\partial/\partial D)\log g_D(Y|\boldsymbol{x}_0, d)\}^2 \,|D = d] < \infty,$$

uniformly in $d : 0 < d \leq d_0$ and $\boldsymbol{x}_0 \in C$. We may write (for $\boldsymbol{x}$ in a neighborhood of $\boldsymbol{x}_0$)

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + (\boldsymbol{x} - \boldsymbol{x}_0)'[(\partial/\partial\boldsymbol{x})f(\boldsymbol{x})]_{\boldsymbol{x}_0} \\ &\quad +(1/2)(\boldsymbol{x} - \boldsymbol{x}_0)'\mathbf{A}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + o(||\boldsymbol{x} - \boldsymbol{x}_0||^2), \end{aligned}$$

where

$$\mathbf{A}(\boldsymbol{x}_0) = (\partial^2/\partial\boldsymbol{x}\partial\boldsymbol{x}')f(\boldsymbol{x})|_{\boldsymbol{x}_0}. \tag{4.8}$$

Thus, whenever $\rho(\boldsymbol{x}, \boldsymbol{x}_0) = d$ yield a closed contour which is symmetric around $\boldsymbol{x}_0$. we have from (4.3) and (4.8) that for $d \to 0$,

$$f_D^0(d) = f(\boldsymbol{x}_0)a(d) + (1/2)Q_1(d) + o(d^2 a(d)), \quad \text{as } d \to 0, \tag{4.9}$$

where $a(d)$ is the surface area of the contour $\{\boldsymbol{x}; \rho(\boldsymbol{x}, \boldsymbol{x}_0) = d\}$ and

$$Q_1(d) = \int_{\{\boldsymbol{x};\rho(\boldsymbol{x},\boldsymbol{x}_0)=d\}} (\boldsymbol{x} - \boldsymbol{x}_0)'\mathbf{A}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)d\nu(\boldsymbol{x}; \boldsymbol{x}_0). \tag{4.10}$$

Note that by the Courant theorem on quadratic forms, $(\boldsymbol{x} - \boldsymbol{x}_0)'\mathbf{A}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)/||\boldsymbol{x} - \boldsymbol{x}_0||^2 \leq a^*(\boldsymbol{X}_0)$ = largest characteristic root of $\mathbf{A}(\boldsymbol{x}_0)$, so that we obtain that $Q_1(d)/a(d) = O(d^1)$, as $d \to 0$.

For the conditional d.f. $G(y|\boldsymbol{x})$, we make similar expansion around $\boldsymbol{x}_0$ and obtain that for small $||\boldsymbol{x} - \boldsymbol{x}_0||$,

$$\begin{aligned} G(y|\boldsymbol{x}) &= G(y|\boldsymbol{x}_0) + (\boldsymbol{x} - \boldsymbol{x}_0)'[(\partial/\partial\boldsymbol{x})G(y|\boldsymbol{x})]_{\boldsymbol{x}_0} \\ &\quad +(1/2)(\boldsymbol{x} - \boldsymbol{x}_0)'\mathbf{B}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + o(||\boldsymbol{x} - \boldsymbol{x}_0||^2), \end{aligned} \tag{4.11}$$

where $\mathbf{B}(\boldsymbol{x}_0)$ is the Hessian of $G(y|\boldsymbol{x})$ with respect to $\boldsymbol{x}$ at $\boldsymbol{x}_0$. And let us also denote by

$$\mathbf{C}(\boldsymbol{x}_0) = [(\partial/\partial\boldsymbol{x})G(y|\boldsymbol{x})]_{\boldsymbol{x}_0}[(\partial/\partial\boldsymbol{x}')f(\boldsymbol{x})]_{\boldsymbol{x}_0}.$$

From (4.4), (4.9) and (4.11), we have for small d,

$$f_D(d;\boldsymbol{x}_0)G_D(y|\boldsymbol{x}_0,d) = f(\boldsymbol{x}_0)G(y|\boldsymbol{x}_0)a(d)+(1/2)G(y|\boldsymbol{x}_0)Q_1(d)$$
$$+(1/2)f(\boldsymbol{x}_0)Q_2(d)+Q_3(d)+o(Q_1(d)+Q_2(d)+Q_3(d)).$$

where $Q_2(d)$ and $Q_3(d)$ are defined as in (4.10) with $\mathbf{A}(\boldsymbol{x}_0)$ being replaced by $\mathbf{B}(\boldsymbol{x}_0)$ and $\mathbf{C}(\boldsymbol{x}_0)$ respectively. Integrating over $0 \le d \le r$, for small r, we have the average conditional d.f.

$$\begin{aligned} G_r(y|\boldsymbol{x}_0) &= \left\{\int_0^r f_D(d;\boldsymbol{x}_0)G_D(y|\boldsymbol{x}_0,d)\mathrm{d}d\right\} \Big/ \left\{\int_0^r f_D(d;\boldsymbol{x}_0)\mathrm{d}d\right\} \\ &= \frac{1}{f(\boldsymbol{x}_0)v(r)+(1/2)Q_1^*(r)+o(Q_1^*(r))} \\ &\quad \times [f(\boldsymbol{x}_0)G(y|\boldsymbol{x}_0)v(r)+(1/2)G(y|\boldsymbol{x}_0)Q_1^*(r) \\ &\quad +(1/2)f(\boldsymbol{x}_0)Q_2^*(r)+Q_3^*(r)+o(Q_1^*(r))] \\ &= G(y|\boldsymbol{x}_0)+\{f(\boldsymbol{x}_0)Q_2^*(r)+2Q_3^*(r)\}\,/\,\{2f(\boldsymbol{x}_0)v(r)+Q_1^*(r)\} \\ &\quad +o((Q_1^*(r))/v(r)), \end{aligned}$$

where

$$v(r)=\int_0^r a(d)\mathrm{d}d$$

and

$$Q_j^*(r)=\int_0^r Q_j(d)\mathrm{d}d, \quad j=1,2,3, \quad \text{for } r>0.$$

Note that $Q_j^*(r)$ and $v(r)$ all converges to 0 as $r \to 0$, we may apply L'Hopital's rule and obtain that as $r \to 0$,

$$\{f(\boldsymbol{x}_0)Q_2^*(r)+2Q_3^*(r)\}\,/\,\{2f(\boldsymbol{x}_0)v(r)+Q_1^*(r)\}=(r^2/2)h(y,\boldsymbol{x}_0)+o(r^2). \quad (4.12)$$

Thus, for $r \to 0$,

$$G_r(y|\boldsymbol{x}_0)=G(y|\boldsymbol{x}_0)+(r^2/2)h(y,\boldsymbol{x}_0)+o(r^2). \quad (4.13)$$

In a similar manner, we have

$$g_r(y|\boldsymbol{x}_0) = g(y|\boldsymbol{x}_0) + (r^2/2)h^*(y,\boldsymbol{x}_0) + o(r^2), \quad \text{as } r \to 0. \tag{4.14}$$

In the case of vector-valued concomitant variates, the rate of r_n is depend on q, the number of components. In this context, we may note that for the Euclidean norm, corresponding to the value of $d(>0)$, the volume of the sphere $\{\boldsymbol{x}; \rho(\boldsymbol{x}, \boldsymbol{x}_0) = d\}$ is $O(d^q)$, for $q \geq 1$, so that we have $\nu(r) = O(r^q)$ as $r \to 0$. A similar order holds for general norms. On the other hand, if we choose a nearest neighborhood of k_n members close to the point $\boldsymbol{x}_0$, then under (4.2), we have

$$n^{-1}k_n = O(r_n^q) \text{ when } k_n/n \to 0 \text{ as } n \to \infty.$$

As such, if we let $r_n \sim n^{-\lambda}$, for some $\lambda > 0$, we may set

$$k_n \sim n^{1-q\lambda}, \quad \text{as } n \to \infty,$$

and this naturally put an upper bound for λ, namely, $1/q$. Moreover, the bias term arising in the conditional quantile estimation, as can be dictated by (4.13), is $O(r_n^2)$, so that the contribution of the bias term in the asymptotic mean square error of the estimator is $O(r_n^4) = O(n^{-4\lambda})$, while the variance term is $O(k_n^{-1}) = O(n^{q\lambda - 1})$. Hence we should select λ in such a way that

$$1 - q\lambda \leq 4\lambda \text{ i.e. } \lambda \geq 1/(q\lambda),$$

so that

$$k_n = O(n^{4/(q+4)}), \quad \text{as } n \to \infty.$$

If we choose $k_n = o(n^{4/(4+q)})$, the contribution of the bias term in the asymptotic mean square error will be negligible, but the order of the mean square error will be $o(n^{-4/(4+q)})$ so that we have a somewhat slower rate of convergence.

4.2 Asymptotic properties of multivariate conditional quantile through contiguity

In this section, we will extend the Bahadur type representation for conditional quantile given by Gangopadhyay and Sen (1993) to multivariate case through contiguity.

Theorem 4.1 *For k_n lying in (k_{n0}, k_{n1}) where $k_{n0} = [an^{4/(q+4)}]$, $k_{n1} = [bn^{4/(q+4)}]$ and $0 < a < b < \infty$, as n incerases,*

$$\hat{\xi}_{n,k_n}(\boldsymbol{x}_0) - \xi(\boldsymbol{x}_0) = \beta(\boldsymbol{x}_0)(n^{-1}k_n)^2 + (k_n g(\xi(\boldsymbol{x}_0)|\boldsymbol{x}_0))^{-1} \times \sum_{i=1}^{k_n}[I(Y^*_{ni} > \xi(\boldsymbol{x}_0)) - (1-p)] + R_{n,k_n}. \qquad (4.15)$$

where

$$\max\{|R_{n,k_n} : k_{n0} \le k_n \le k_{n1}\} = O(n^{-3/(q+4)}\log n)$$

and letting $G(\cdot)$ for $G(y|\boldsymbol{x}_0)$ and $G_D(\cdot)$ for $G_D(y|\boldsymbol{x}_0, d)$, the Y^*_{ni} are conditionally independent r.v.'s defined by

$$Y^*_{ni} = G^{-1}G_D(Y_{S_i}|\boldsymbol{X}_{S_i}), \quad i = 1, \cdots, n;$$

and $\beta(\boldsymbol{x}_0) = -h(\xi, \boldsymbol{x}_0)\{24f^2(\boldsymbol{x}_0)g(\xi)\}^{-1}$ where $h(\cdot,\cdot)$ is in (4.12).

Note that the induced order statistics Y_{S_i}'s are conditionally independent given $\mathcal{A} = \sigma(S_1, S_2, \cdots)$ with Y_{S_i} having conditional c.d.f. $G_{D_{n_i}}(\cdot|\boldsymbol{x}_0) = G_D(\cdot|\boldsymbol{x}_0, D_{S_i})$. Suppose ξ_{nk} be the p-th quantile of the random c.d.f. $\overline{G}_{nk}(y) = (1/k)\sum_{i=1}^{k} G_{D_{n_i}}(\cdot|\boldsymbol{x}_0) = (1/k)\sum_{i=1}^{k} G_{ni}(\cdot)$; i.e.,

$$\overline{G}_{nk}(\xi_{nk}) = p,$$

then one can think of $\hat{\xi}$ as an estimator of ξ_{nk}. The difference between $\overline{G}_{nk}(\cdot)$ and $G(\cdot)$ of which ξ is the p-quantile gives rise to the bias term, as well as error of small magnitude. The difference between $\hat{G}(\cdot)$ and $\overline{G}(\cdot)$ provides the empirical process, which is the main feature of the Bahadur type representation.

For every $k(\leq n)$; let q_{nk} and Q_{nk} denote the joint conditional density and distribution of $Y_{S_1}, \cdots, Y_{S_k}$, given $\boldsymbol{X}_{S_1}, \cdots, \boldsymbol{X}_{S_k}$. Using the conditional independence of $(Y_{S_1}, \cdots, Y_{S_k})$, we have

$$q_{nk}(\cdot) = \prod_{i=1}^{k} g_{D_{n_i}}(\cdot|\boldsymbol{x}_0) = \prod_{i=1}^{k} g_{ni}(\cdot)$$

and

$$Q_{nk}(\cdot) = \prod_{i=1}^{k} G_{D_{n_i}}(\cdot|\boldsymbol{x}_0) = \prod_{i=1}^{k} G_{ni}(\cdot).$$

Now let $Y^0_{S_1}, \cdot, Y^0_{S_k}$ be i.i.d. r.v.'s with a density and d.f. $g(y)$ and $G(y)$, and let

$$p_{nk}(\cdot) = \prod_{i=1}^{k} g(\cdot) \text{ and } P_{nk} = \prod_{i=1}^{k} G(\cdot).$$

Thus, q_{nk} is the true conditional density of $(Y_{S_1}, \cdots, Y_{S_k})$ given $(\boldsymbol{X}_{S_1}, \cdots, \boldsymbol{X}_{S_k})$, while p_{nk} and P_{nk} stands for the joint density and distribution of an i.i.d. model. In this set-up, we may have

Theorem 4.2 *Let $k = [tn^{\lambda}]$ for some $\lambda(0 < \lambda \leq 4/(q+4))$ and for some $t(0 < a < t < b < \infty)$. Then under [A1]-[A4], the densities q_{nk} are contiguous to the densities p_{nk} a.e. (F)*

Lemma 4.1 *Let $W_1, \cdots, W_k$ i.i.d. with distribution H, where $k \in I_n(a, b)$, and let η be the p-quantile of H, i.e. $H(\eta) = p$. Assume that the distribution H is twice differentiable with $H'(\eta) = h(\eta) > 0$. Let a_n be a sequence of constants and that $a_n \sim n^{-2/(q+4)} \log n$. Then*

$$\begin{aligned} V_n^* &= \max_{k \in I_n(a,b)} \sup_{|\omega - \eta| \leq a_n} |(\hat{H}_{nk}(\omega) - \hat{H}_{nk}(\eta)) - (H(\omega) - H(\eta))| \\ &= O(n^{-3/(q+4)} \log n) \ a.s. \end{aligned}$$

where $\hat{H}_{nk}(\cdot)$ is the empirical distribution function of $(W_1, \cdots, W_k)$.

Proof: As in (Bahadur (1966)), choose $b_n = n^{1/(q+4)}$ and divide the interval $[\eta - a_n, \eta + a_n]$ into $2b_n$ intervals $J_{nk,r} = [\eta + r(a_n/b_n), \eta + (r-1)(a_n/b_n)] = \tau_{nk,r}, \tau_{nk,r+1}]$, for $r = -b_n, \cdots, b_n + 1$.

Thus for $\omega \in J_{nk,r}$,

$$
\begin{aligned}
V_n^* = \max_{k \in I_n(a,b)} V_{nk}^* &= \max_{k \in I_n(a,b)} \sup_{|\omega - \eta| \leq a_n} |V_{nk}(\omega)| \\
&\leq \max_{k \in I_n(a,b)} \max_{-b_n \leq r \leq b_n} |V_{nk}(\tau_{nk,r})| + \max_{k \in I_n(a,b)} \max_{-b_n \leq r \leq b_n} |\alpha_{nk,r}|,
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_{nk,r} &= |H(\tau_{nk,r+1}) - H(\tau_{nk,r})| \\
&\leq \sup_{|\omega - \eta| \leq a_n} |H'(\omega)| n^{-3/(q+4)} \log n,
\end{aligned}
$$

for $-b_n \leq r \leq b_n - 1$. Thus

$$
\max_k \max_r \alpha_{nk,r} = O(n^{3/(q+4)} \log n).
$$

Finally, using Bernstein's inequality,

$$
\begin{aligned}
& P(\max_k \max_r |V_{nk}(\tau_{nk,r})| > Mn^{-3/(q+4)} \log n) \\
\leq & \sum_k \sum_k P(|V_{nk}(\tau_{nk,r})| > Mn^{-3/(q+4)} \log n) \\
\leq & \sum_k \sum_r \exp[-CM^2 \log n] \\
= & \ 2(b-a) n \exp[-CM^2 \log n],
\end{aligned}
$$

for large constant C. Thus

$$
\sum_n P(\max_k \max_r |V_{nk}(\tau_{nk,r})| > Mn^{-3/(q+1)} \log n) < \infty,
$$

for sufficiently large M, which completes the proof. $\diamond$

Lemma 4.2 *Let For $B > b/f(\boldsymbol{x}_0)$ and for large n*

$$P[D_{n,[n^{4/(q+4)}b]} > n^{-1/5}B] \leq \exp[-2n^{3/5}(Bf(\boldsymbol{x}_0) - b)^2].$$

Proof: Apply modified version of Lemma 2 of Bhattacharya and Mack (1987).

For $B > b/f(\boldsymbol{x})$ and sifficiently large n

$$P[Y_{\phi b} > n^{-1}\phi^{(q+4)/5}(n)B] \leq \exp[-2n^{-1}\phi^{2(q+4)/5}(n)\{Bf(\boldsymbol{x}) - b\}^2],$$

put $\phi = 4/(q+4)$. ◇

Lemma 4.3 $k^{-1}\sum_{i=1}^{k} D_{n:i}^2 = \{Kf(\boldsymbol{x}_0)\}^2 (k/n)^2 + R_{nk}$ *where*

$$\max_{k\leq[n^{4/(q+4)}b]} |R_{nk}| = O(n^{-3/(q+4)}) \quad a.s.$$

and

$$K = 3 \times \frac{\partial}{\partial u} g(u)\bigg|_{u=0},$$

where $u = F_D(d)$ and $d = F_D^{-1}(u) = g(u)$.

Proof: Apply Lemma 2 of Bhattacharya and Gangopadhyay (1988). In scalar case.

$$\frac{\partial}{\partial u} g(u)\bigg|_{u=0} = \frac{1}{f(x_0+d) + f(x_0-d)}\bigg|_{d=0} = \frac{1}{2f(x_0)}.$$

$$\max_{k\leq n^{4/(q+4)}b} |R_{nk}(3)| = O(n^{4/(q+4)-2}) = O(n^{-(2-4/(q+4))}).$$

And $3/(q+4) < 4/(q+4)$, $3/(q+4) < 2 - 4/(q+4)$ and $(n^{-7/10}\sqrt{\log\log n}) \sim O(n^{-3/(q+4)})$. So we get Lemma 4.4. ◇

Lemma 4.4 *For $r \to 0$,*

$$G_r(y|\boldsymbol{x}_0) = G(y|\boldsymbol{x}_0) + (r^2/2)h(y, \boldsymbol{x}_0) + r^3 R(y, \boldsymbol{x}_0).$$

Proof: (4.13). ◇

Lemma 4.5 *For every B, there exist N and C such that in the sample space of infinite sequences $\{((d_1, y_1), (d_2, y_2), \cdots) : d_i > 0,\ y_i \text{ real}\}$, $D_{n,[n^{4/(q+4)}b]} \leq Bn^{-1/5}$ implies $\max_{k \in I_n(a,b)} |\xi_{nk} - \xi| \leq Cn^{-2/(q+4)}$ for all $n \geq N$.*

Corollary 4.1 *For $0 < a < b$*

$$\max_{k \in I_n(a,b)} |\xi_{nk} - \xi| = O(n^{-2/(q+4)}) \quad a.s.$$

Lemma 4.6 *For $0 < a, b$,*

$$\max_{k \in I_n(a,b)} |\xi_{nk} - \xi - \beta(\xi)(k/n)^2| = O(n^{-3/(q+4)}),\ a.s.,$$

where $\beta(\xi) = -h(\xi)\left\{24 f^2(\boldsymbol{x}_0) g(\xi)\right\}^{-1}$.

Lemma 4.7 *Under [A1]-[A4],*

$$\begin{aligned} D_n^* &= \max_{k n I_n(a,b)} \sup_{|z - \xi_{nk}| \leq a_n} |\left\{\hat{G}_{nk}(z) - \hat{G}_{nk}(\xi_{nk})\right\} - \left\{\overline{G}_{nk}(z) - \overline{G}_{nk}(\xi_{nk})\right\}| \\ &= O_p(n^{-3/(q+4)} \log n). \end{aligned}$$

Proof: We have shown in lemma 1 that under p_{nk} measure, D_n^* is $O(n^{-3/(q+4)} \log n)$ a.s. Now, since q_{nk} is contiguous to p_{nk} (Theorem 4.2), under q_{nk} measure, D_n^* is $O_p(n^{-3/(q+4)} \log n)$. ◇

Lemma 4.8 *For $a_n = n^{-2/(q+4)} \log n$,*

$$P[\max_{k \in I_n(a,b)} |\xi_{nk} - \xi| > a_n \ \ i.o.] = 0.$$

Now, lemma 4.7 and 4.8 implies

$$p - \hat{G}_{nk}(\xi_{nk}) = \overline{G}_{nk}(\hat{\xi}_{nk}) - \overline{G}_{nk}(\xi_{nk}) + R_{nk}^*$$

where

$$\max_{k \in I_n(a,b)} |R_{nk}^*| = O_p(n^{-3/(q+4)} \log n).$$

But

$$\overline{G}_{nk}(\hat{\xi}_{nk}) - \overline{G}_{nk}(\xi_{nk}) = (\hat{\xi}_{nk} - \xi_{nk})\overline{g}_{nk}(\xi^*_{nk}),$$

where ξ^*_{nk} lies between $\hat{\xi_{nk}}$ and ξ_{nk} and $\overline{g}(\cdot) = (1/k)\sum_{i=1}^k g(\cdot|\boldsymbol{X}_{S_i})$. Note that

$$\max_{k\in I_n(a,b)} |\xi_{nk}{}^* - \xi| \leq \max_{k\in I_n(a,b)} |\hat{\xi}_{nk} - \xi_{nk}| + \max_{k\in I_n(a,b)} |\xi_{nk} - \xi|.$$

and hence by lemmas 4.6 and 4.8

$$\max_{k\in I_n(a,b)} |\xi_{nk}{}^* - \xi| = O(n^{-2/(q+4)} \log n), \text{ a.s.}$$

Consequently, lemmas 4.3 and 4.5 imply

$$\max_{k\in I_n(a,b)} |\overline{g}_{nk}(\xi_{nk}{}^*) - g(\xi)| + O(n^{-2/(q+4)} \log n), \text{ a.s.}$$

Also, by Hoeffding's inequality we have

$$p - \hat{G}_{nk}(\xi_{nk}) = O(n^{-2/(q+4)} \log n), \quad \text{a.s.}$$

Thus, we have

$$\hat{\xi}_{nk} = \xi_{nk} + \{kg(\xi)\}^{-1} \sum_{i=1}^{k} [1(Y_{ni} > \xi) - \{1 - G_{ni}(\xi)\}] + R_{nk} \tag{4.16}$$

where

$$\max_{k\in I_n(a,b)} |R_{nk}| + O_p(n^{-3/(q+4)} \log n).$$

(4.16) can be expressed as

$$\hat{\xi}_{nk} = \xi_{nk} + \{kg(\xi)\}^{-1} \sum_{i=1}^{k} [1(Y^*_{ni} > \xi) - (1 - p)] + R_{nk}, \tag{4.17}$$

where $\max_{k\in I_n(a,b)} |R_{nk}| = O_p(n^{-3/(q+4)} \log n)$ and $Y^*_{ni} = G^{-1}G_{ni}(Y_{ni})$ and $G(\xi) = p$. Now (4.17) along with lemma 4.7 completes the proof of Theorem 4.1.

We also have following asymptotic normalities for multivariate case. The following theorem is the extension of Theorem N2 by (Bhattacharya and Gangopadhyay (1988)).

Theorem 4.3 *Let $T_n(t) = \hat{\xi}_{n,[n^{4/(q+4)}t]}$. Then for any $0 < a < t < b < \infty$,*

$$\left\{n^{2/(q+4)}[T_n(t) - \xi(\boldsymbol{x}_0)] - \beta t^2,\ a \le t \le b\right\} \xrightarrow{\mathcal{D}} \left\{\sigma t^{-1} B(t),\ a \le t \le b\right\},$$

where $\sigma^2 = p(1-p)/g^2(\xi), \{B(t), t > 0\}$ denote a standard Brownian motion for $q \ge 2$.

Proof: From theorem 1.1, we have

$$n^{2/(q+4)}(\hat{\xi} - \xi) - \beta t^2 n^{-2(q-1)/(q+4)} = \frac{\sqrt{p(1-p)}}{g(\xi)} t^{-1} n^{-2/(q+4)} \sum_{i=1}^{[n^{4/(q+4)}t]} W_{ni} + R(t),$$

where

$$W_{ni} = \frac{1(Y_{ni}^* > \xi(\boldsymbol{x}_0)) - (1-p)}{\sqrt{p(1-p)}}, 1 \le i \le n,$$

are iid with mean 0 and variance 1 for each n and $R(t) = 0_p(1)$. Thus we have, with $\sigma = \sqrt{p(1-p)}/g(\xi)$,

$$n^{2/(q+4)}(\hat{\xi} - \xi) - \beta t^2 n^{-2(q-1)/(q+4)} = \sigma t^{-1} n^{-2/(q+4)} \sum_{i=1}^{[n^{4/(q+4)}t]} W_{ni} + o_p(1),$$

uniformly in $a \le t \le b$. Now theorem 1, page 452 of Gikhman and Skorokhod (1969) to see that $\{n^{-2/(q+4)} \sum_1^{[4/(q+4)t]} W_{ni},\ \ a \le t \le b\} \xrightarrow{\mathcal{D}} \{B(t),\ \ a \le t \le b\}$. This proves the theorem. ◇

It is shown that if for some $\delta > 0$, $k_n = O(n^{4/5-\delta})$, then the bootstrap method proposed by Sen and Gangopadhyay (1990) works out neatly and the leading bias term may be eliminated in the univariate case. We extend their bootstrap method with $k_n = O(n^{4/(q+4)-\delta})$ for some $\delta > 0$ in the multivariate case. A bootstrap sample $(Y_{n1}^*, \cdots, Y_{nk}^*)$ is obtained from the empirical d.f. $\hat{G}_{nk}$ defined by (4.1). Thus, given the induced order statistics $(Y_{n1}, \cdots, Y_{nk})$, the Y_{ni}^* are conditionally i.i.d.r.v. with the d.f. $\hat{G}_{nk}$. Let $\hat{G}_{nk}^*$ be the bootrap sample d.f. (based on $Y_{n1}^* \cdots, Y_{nk}^*$). Then the bootrap estimator is

$$\begin{aligned} \hat{\xi}_{nk}^* &= [kp]\text{-th order statistic of } Y_{n1}^*, \cdots, Y_{nk}^* \\ &= \inf\{y : \hat{G}_{nk}^*(y) \ge k^{-1}[kp]\}. \end{aligned}$$

The following two theorems can be proven without major changes to their original ones.

Theorem 4.4 *Under [A1] and [A2], for any* $k \in I_n(a,b) = \{k : k_0 = [an^{4/(q+4)-2\delta} \leq k \leq [bn^{4/(q+4)-2\delta}] = k_1\}$, $0 < a < b < \infty$, *as* $n \to \infty$,

$$\hat{\xi}^*_{nk} - \hat{\xi}_{nk} = \{kg(\xi_0)\}^{-1} \sum_{i=1}^{k} \{1(Y^{0*}_{ni} > \xi_0) - (1-p)\} + R^{*0}_{nk}$$

where

$$Y^{0*}_{ni} = G^{-1}\hat{G}_{nk}(Y^*_{ni}),\ i = 1, \cdots, k$$

are i.i.d.r.v. with d.f. $G(\cdot)$ *and as* $n \to \infty$

$$\max_{k \in I_n(a,b)} \left| R^{*0}_{nk} \right| = O(n^{-3/(q+4)+3\delta/2} \log n) \quad a.s..$$

Theorem 4.5 *Under the hypothesis of Theorem 4.4, as* $n \to n$

$$\left\{n^{2/(q+4)-\delta}(\hat{\xi}^*_{n,[tn^{4/(q+4)-2\delta}]} - \hat{\xi}_{n,[tn^{4/(q+4)-2\delta}]}),\ \ t \in [a,b]\right\}$$

$$\xrightarrow{\mathcal{D}} \{p(1-p)\}^{1/2}\{g(\xi)\}^{-1}\{t^{-1}B(t),\ t \in [a,b]\}$$

Hence, for each t

$$n^{2/(q+4)-\delta}(\hat{\xi}^*_{n,[tn^{4/(q+4)-2\delta}]} - \hat{\xi}_{n,[tn^{4/(q+4)-2\delta}]}) \xrightarrow{\mathcal{D}} N(0, t^{-1}p(1-p)/g^2(\xi)).$$

One can generate a set of M bootstrap samples and from each one, compute the bootstrap estimates, and then use the empirical quantile (for the $\hat{\xi}^*_{nk} - \hat{\xi}_{nk}$) to set the desired confidence interval for ξ_0. The asymptotic properties of such bootstrap intervals would then follow from Theorems 4.4 and 4.5.

4.3 Conditional Independence and Contiguity

In this section, we discuss the conditional independence of Y_{S_i}'s and contiguity for proving theorem 2.

Lemma 4.9 *Under the continuity assumption of $F_D(d;\boldsymbol{x}_0)$, the marginal distribution of D, for every n and almost all $(D_1,\cdots,D_n)$, $Y_{S_1},\cdots,Y_{S_n}$ are conditionally independent given $\boldsymbol{X}_{S_1},\cdots,\boldsymbol{X}_{S_n}$ with conditional cdf's $G_{D_{S_1}}(\cdot|\boldsymbol{x}_0),\cdots,G_{D_{S_n}}(\cdot|\boldsymbol{x}_0)$ respectively.*

Proof: (Lemma 1 of Bhattachatya (1974)) For any $\mathbf{d}_n = (d_1,\cdots,d_n)$ no two coordinates of which are equal, let $S_k = S_{k,\mathbf{d}_n}$ be the antirank as defined before. By the continuity assumption, $S_{k,\mathbf{D}_n}$, $k = 1,\cdots,n$ are defined a.s. and $D_{nk} = D_{S_{k,\mathbf{D}_n}}$, $Y_{nk} = Y_{S_{k,\mathbf{D}_n}}$, $k = 1,\cdots,n$. Hence the conditional joint cdf of $Y_{n1},\cdots,Y_{nn}$ given $\boldsymbol{X}_{S_1},\cdots,\boldsymbol{X}_{S_n}$ is the same as the conditional joint cdf of $Y_{S_{k,\mathbf{D}_n}}$, $k = 1,\cdots,n$, given $\boldsymbol{X}_{S_k}$, $k = 1,\cdots,n$, which is easily seen to be the product $\prod_{k=1}^{n} G_{D_{S_i}}(\cdot|\boldsymbol{x}_0) = \prod_{i=1}^{k} G_{D_{nk}}$ due to the independence of Y_j and D_i for every $i \neq j = 1,\cdots,n$. $\diamond$

Consider a sequence $\{p_n, q_n\}$ of simple hypothesis p_n and simple alternatives q_n defined on measure spaces $(\mathcal{X}_n, \mathcal{A}_n, \mu_n)$, $n \geq 1$, respectively.

Definition 4.1 *If for any sequence of events $\{A_n\}$, $A_n \in \mathcal{A}_n$,*

$$[P_n(A_n) \to 0] \Rightarrow [Q_n(A_n) \to 0]$$

holds, we says that the densities q_n are contiguous to the densities p_n, where $dP_n = p_n d\mu_n$, $dQ_n = q_n d\mu_n$.

Contiguity implies that any sequence of random variables converging to zero in P_n-probability converges to zero in Q_n-probability, $n \to \infty$.

Let us denote by $g_{D_{ni}}(y|\boldsymbol{x}_0) = g_{ni}(y)$ and $G_{D_{ni}}(y|\boldsymbol{x}_0)$. Also for every $k(\leq n)$; let q_{nk} and Q_{nk} denote the joint conditional density and distribution of $Y_{S_1},\cdots,Y_{S_k}$, given $\boldsymbol{X}_{S_1},\cdots,\boldsymbol{X}_{S_k}$. Using the conditional independence of $(Y_{S_1},\cdots,Y_{S_k})$, we have

$$q_{nk}(\cdot) = \prod_{i=1}^{k} g_{D_{ni}}(\cdot|\boldsymbol{x}_0) = \prod_{i=1}^{k} g_{ni}(\cdot)$$

and

$$Q_{nk}(\cdot) = \prod_{i=1}^{k} G_{D_{ni}}(\cdot|\boldsymbol{x}_0) = \prod_{i=1}^{k} G_{ni}(\cdot).$$

Now let $Y^0_{S_1}, \cdot, Y^0_{S_k}$ be i.i.d. r.v.'s with a density and d.f. $g(y)$ and $G(y)$, and let

$$p_{nk}(\cdot) = \prod_{i=1}^{k} g(\cdot) \ \text{ and } \ P_{nk} = \prod_{i=1}^{k} G(\cdot).$$

Thus, q_{nk} is the true conditional density of $(Y_{S_1}, \cdots, Y_{S_k})$ given $(\boldsymbol{X}_{S_1}, \cdots, \boldsymbol{X}_{S_k})$, while p_{nk} and P_{nk} stands for the joint density and distribution of an i.i.d. model.

The following regularity conditions are assumed:

[A1] Let $F(\boldsymbol{x})$, $\boldsymbol{x} \in R^q$ be the joint d.f. of the covariate $\boldsymbol{X}$. We assume that F admits an absolutely continuous (q-variate) density function $f(\boldsymbol{x})$, $\boldsymbol{x} \in R^q$, such that

$$f(\boldsymbol{x}) > 0, \ \text{ for every } \ \boldsymbol{x} \in C,$$

and $f'(\boldsymbol{x}) = (\partial/\partial\boldsymbol{x})(f(\boldsymbol{x}))$ exists in a neighborhood of $\boldsymbol{x}_0$. Moreover, there exist positive and finite numbers ϵ, K_0 and α, such that $||\boldsymbol{x} - \boldsymbol{x}^*|| < \epsilon$ implies that

$$||f'(\boldsymbol{x}) - f'(\boldsymbol{x}^*)|| \leq K_0||\boldsymbol{x} - \boldsymbol{x}^*||.$$

[A2] Conditional distribution of Y_i given $\boldsymbol{X}_i = \boldsymbol{x}_0$, by $g(y|\boldsymbol{x}_0)$ and assume that $G(\cdot)$ has a continuous density $g(y|\boldsymbol{x}_0)$ for almost all $(y, \boldsymbol{x}_0)$. And $(\partial/\partial\boldsymbol{x})g(y|\boldsymbol{x})$ and $(\partial^2/\partial\boldsymbol{x}\partial\boldsymbol{x}')g(y|\boldsymbol{x})$ exist in a neighborhood of $\boldsymbol{x}_0$.

[A3] We assume that there exist $\epsilon > 0$ and $\alpha > 0$ and some Lebesgue measurable functions $u_1^*(y)$ and $u_2^*(y)$, such that for all $r \leq \epsilon$,

$$|(\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)| \leq u_1^*(y)$$

$$|(\partial/\partial r^2)g_r(y|\boldsymbol{x}_0) - \left\{(\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)|_{r=0}\right\}| \leq u_2^*(y).$$

[A4] There exists $\epsilon > 0$ such that

$$I(y|\boldsymbol{x}_0, r) = E[\left\{(\partial/\partial R^2)\log g_R(Y|\boldsymbol{x}_0)\right\}^2 |R = r] < \infty,$$

uniformly in $r : 0 < r < \epsilon$.

Let $\mathcal{L}(A, B)$ be a set of continuous linear transformations from a topological vector space A to another B, and let C be a class of compact subset of A such that every subset consisting of a single point belongs to C. For $G \in A$ and for every $H \in K \in C$, we assume that

$$T(H) = T(G + (H - G)) = T(G) + \int T_1(G; y, \boldsymbol{x}_0)\mathrm{d}[H(y) - G(y)] + R_1(G; G - H),$$

where $G(\cdot) = G(\cdot|\boldsymbol{x}_0)$ and for any sequence $H_k \in K \in C$ such that $||H_k - G|| \to 0$, $|R_1(G; H - G) = o(||H_k - G||)$, uniformly in $H_k \in K$; and $||\cdot||$ stands for the "sup-norm". We may refer to Clarke (1986) and Fernholz (1983) for a detailed discussion in this context. This relates to the so called first order *Hadamard-differentiability* of $T(\cdot)$ at G, and $T_1(G; y, \boldsymbol{x}_0)$ is called the first order compact (or Hadamard) derivative of $T(\cdot)$ at G; it is normalized that

$$\int T_1(G; y, \boldsymbol{x}_0)\mathrm{d}G(y|\boldsymbol{x}_0) = 0 \quad (a.a.\ \boldsymbol{x}_0).$$

Similarly, if we assume that

$$\begin{aligned} T(H) &= T(G) + \int T_1(G; y, \boldsymbol{x}_0)\mathrm{d}[H(y) - G(y)] \\ &\quad +(1/2)\int\int T_2(G; y, y', \boldsymbol{x}_0)\mathrm{d}[H(y) - G(y)][H(y') - G(y')] \\ &\quad R_2(G; H - G), \quad \forall H \in K \in C, \end{aligned}$$

where for any sequence $H_k \in K \in C$ such that $||H_k - G|| \to 0$ and

$$||R_2(G; H_k - G)|| = o(||H_k - G||^2), \text{ uniformly in } H_k \in K,$$

then $T(\cdot)$ is second order Hadamard differentiable at G and $T_2(G; y, y', \boldsymbol{x}_0)$ is the second order compact (or Hadamard) derivative of $T(\cdot)$ at G. We may set

$$T_2(G; y, y', \boldsymbol{x}_0) \equiv T_2(G; y', y, \boldsymbol{x}_0),$$

$$\int T_2(G;y,y',\boldsymbol{x}_0)\mathrm{d}G(y') = 0 = \int T_2(G;y,y',\boldsymbol{x}_0)\mathrm{d}G(y) \quad \text{(a.e.)}.$$

[B1] There exist $\epsilon > 0$, $\alpha > 0$ and $A < \infty$ such that uniformly in $r (0 \le r \le \epsilon)$.

$$|\int T_1(G;y,\boldsymbol{x}_0)\left\{(\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)\right\}\mathrm{d}y| \le A;$$

$$|\int T_1(G;y,\boldsymbol{x}_0)\left\{(\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)\right\}\mathrm{d}y - \int T_1(G;y,\boldsymbol{x}_0)\left\{(\partial/\partial r^2)g_D(y|\boldsymbol{x}_0)|_{r=0}\right\}\mathrm{d}y| \le Ar^\alpha.$$

Lemma 4.10 *(a) under [A1], [A2] and [B1], we have*

$$g_r(y|\boldsymbol{x}_0) = g(y|\boldsymbol{x}_0) + r^2 q(y,\boldsymbol{x}_0) + r^2 R(y,r,\boldsymbol{x}_0),$$

where

$$q(y,\boldsymbol{x}_0) = (\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)\Big|_{r=0},$$

and

$$R(y,r,\boldsymbol{x}_0) = (\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)\Big|_{r^*} - (\partial/\partial r^2)g_r(y|\boldsymbol{x}_0)\Big|_{r=0},$$

for some r^* between 0 and r. Moreover, there exists $\epsilon > 0$ and $M (0 < M < \infty)$ such that uniformly for all $r (0 < r < \epsilon)$ for some $\alpha > 0$

$$\left|\int T_1(G;y,\boldsymbol{x}_0)q(y,\boldsymbol{x}_0)dz\right| \le M \tag{4.18}$$

$$\left|\int T_1(G;y,\boldsymbol{x}_0)R(y,r,\boldsymbol{x}_0)dz\right| \le Mr^\alpha. \tag{4.19}$$

(b) If in addition, we assume [A3], then for $0 < r < \epsilon$

$$\int_{-\infty}^{\infty} q(y;\boldsymbol{x}_0)dy = 0$$

and

$$\int_{-\infty}^{\infty} R(y,r,\boldsymbol{x}_0)dy = 0.$$

Proof: Same as the proof of Lemma 3. ◇

Now, we define

$$L_{nk} = \begin{cases} q_{nk}/p_{nk} & \text{if } p_{nk} > 0 \\ 1 & \text{if } p_{nk} = q_{nk} = 0 \\ 0 & \text{if } p_{nk} = 0 < q_{nk}. \end{cases}$$

Then along with fashon of LeCam's first lemma (see, e.g. Hájek and Sidák (1967), page 203-204), we shall prove the following;

Lemma 4.11 $\log L_{nk}$ *is asymptotically normal* $(-1/2d^2, d^2)$, *so that the densities* q_{nk} *are contiguous to the densities* p_{nk}.

For this purpose, first note that

$$\log L_{nk} = \sum_{i=1}^{k} \log[g_{ni}(Y_{ni})/g(Y_{ni})]. \tag{4.20}$$

In the next lemma we will show that the summands in (4.20) are uniformly asymptotically negligible (UAN).

Lemma 4.12 *Assume [A1]-[A4]. Then under* p_{nk},

$$\lim_{n\to\infty} \max_{1\le i\le k} P\left[|\{g_{ni}(Y_{ni}/g(Y_{ni})\} - 1| > \epsilon\right] = 0.$$

Proof: By Chebychev's inequality

$$P\left[|\{g_{ni}(Y_{ni}/g(Y_{ni})\} - 1| > \epsilon\right] \le E\left[\{g_{ni}(Y_{ni}/g(Y_{ni})\} - 1\right]^2/\epsilon^2.$$

But

$$E\left[\{g_{ni}(Y_{ni}/g(Y_{ni})\} - 1\right]^2$$

$$= D_{ni}^4[\int \left\{q^2(t, \boldsymbol{x}_0)/g(y)\right\} \mathrm{d}y + \int \left\{\int R^2(y, D_{ni}, \boldsymbol{x}_0)/g(y)\right\} \mathrm{d}y$$

$$+2\int \{q(y, \boldsymbol{x}_0)R(y, D_{ni}, \boldsymbol{x}_0)/g(y)\}\,\mathrm{d}y]$$

$$= o(1). \diamond$$

From this point, we restrict ourselves to the case $\lambda = 4/(q+4)$. At the end of this section, we will indicate how the proof can be modified when $\lambda > 4/(q+4)$.

Lemma 4.13 *Let*

$$W_{nk} = 2\sum_{i=1}^{k}\left\{[g_{ni}(Y_{ni})/g(Y_{ni})]^{1/2} - 1\right\}.$$

Assume [A1]-[A4] hold and the statistics W_{nk} are asymptotically normal $(-1/2d^2, d^2)$ under p_{nk}, then under p_{nk}

$$\lim_{n\to\infty} P\left(\left|\log L_{nk} - W_{nk} + 1/4d^2\right| > \epsilon\right) = 0, \quad \textit{for every } \epsilon > 0.$$

and $\log L_{nk}$ *is asymptotically normal* $(-1/2d^2, d^2)$.

Proof: Since the UAN condition is satisfied by lemma 2.4, the result follows from LeCam's second lemma. ◇

Lemma 4.14 *Assume [A1]-[A4] hold. Then under p_{nk},*

$$E(W_{nk}) \to -(1/4)d^2,$$

where $d^2 = \{80f^4(\boldsymbol{x}_0)\}^{-1}\, t\int_{-\infty}^{\infty}\{q^2(y,\boldsymbol{x}_0)/g(y)\}\,dy$.

Proof: Denote

$$s(y|d^2) = [g_d(y|\boldsymbol{x}_0)]^{1/2}$$

and

$$s(y) = s(y|0) = [g(y\boldsymbol{x}_0)]^{1/2}.$$

So under p_{nk}

$$\begin{aligned} E(W_{nk}) &= 2\sum_{i=1}^{k}\int\left[\left\{s(y|D_{ni}^2)/s(y)\right\} - 1\right]^2 s^2(y)\mathrm{d}y \\ &= \sum_{i=1}^{k}\int\left\{2s(y|D_{ni}^2)s(y) - 2s^2(y)\right\}\mathrm{d}y \end{aligned}$$

$$= -\sum_{i=1}^{k} \int \left(s^2(y) + s^2(y|D_{ni}^2) - 2s(y)s(y|D_{ni}^2) \right) \mathrm{d}y$$

$$= -\sum_{i=1}^{k} D_{ni}^4 \int \left\{ \left(s(y|D_{ni}^2) - s(y) \right)^2 / D_{ni}^2 \right\}^2 \mathrm{d}y$$

$$= -\sum_{i=1}^{k} D_{ni}^4 \int \left\{ (\partial/\partial d^2) s(y|d)\Big|_{d_{ni}^*} \right\}^2 \mathrm{d}y$$

where $0 < d_{ni}^* < D_{ni}$; $i = 1, 2, \cdots, k$.

Note that for some $\epsilon > 0$, by assumption [A4]

$$\int \left\{ (\partial/\partial d^2) s(y|d) \right\}^2 \mathrm{d}y = \frac{1}{2} \int \left\{ (\partial/\partial d^2) \log g_d(y|\boldsymbol{x}_0) \right\}^2 g_d(y|\boldsymbol{x}_0) \mathrm{d}y = \frac{1}{2} I(y|\boldsymbol{x}_0, d) < \infty.$$

uniformly for all $d (0 < d < \epsilon)$. Also,

$$(\partial/\partial d^2) s(y|d)\Big|_{d=0} = (1/2) \left\{ q(y, \boldsymbol{x}_0) / (g(y|\boldsymbol{x}_0))^{1/2} \right\}.$$

So,

$$\lim_{n\to\infty} \int \left\{ (\partial/\partial d^2) s(y|d)\Big|_{d_{ni}^*} \right\}^2 \mathrm{d}y = \int \left\{ (\partial/\partial d^2) s(y|d)\Big|_{d=0} \right\}^2 \mathrm{d}y$$

$$= (1/4) \int \left\{ q^2(y, \boldsymbol{x}_0) / g(y|\boldsymbol{x}_0) \right\} \mathrm{d}y,$$

uniformly i $(= 1, 2, \cdots, k)$.

Finally, let $U_{n1} < \cdots < U_{nn}$ denote the order statistics of a random sample $(U_1, \cdots, U_n)$ of size n from the uniform distribution on $(0, 1)$. Then $D_{ni} = h(U_{ni})$, where $h = F_D^{-1}$. It can cbe seen easily that under assumption [A1] the following is true:

(i) $h(u) = F_D^{-1}(u)$ is defined for $0 < u < \epsilon$ for some $\epsilon > 0$ as a unique solution of $F_D^{-1} = u$.

(ii) $h'(u)$ is continuous at 0.

(iii) $h(0) = h''(0) = 0, h'(0) = \{2f(\boldsymbol{x}_0)\}^{-1}$.

Let F_k^* denote the empirical distribution of $U_1, \cdots, U_k$; then

$$\begin{aligned}
& \max_{1 \le j \le k} \left| U_{nj} - (j/tn^{5/(q+4)}) \right| \\
\le\ & \max_{1 \le j \le k} \left| U_{nj} - (j/tn^{4/(q+4)}) \right| + \max_{1 \le j \le k} \left| (j/(tn^{4/(q+4)}) - (j/tn^{5/(q+4)}) \right| \\
=\ & \max_{1 \le j \le k} \left| U_{nj} - F_k^*(u_{nj}) \right| \ + \ \left(1 - \frac{tn^{4/(q+4)}}{tn^{5/(q+4)}} \right) \\
\le\ & \sup_{0<u<1} |F_k^*(u) - u| + (1 - n^{-1/(q+4)} \\
=\ & O_p(n^{-2/(q+4)}).
\end{aligned}$$

So,

$$U_{nj}^4 = (j/tn^{5/(q+4)})^4 + o_p(n^{-2/(q+4)}), \text{ for all } j.$$

By lemma 3.1 and (iii)

$$\begin{aligned}
\sum_{i=1}^{k} D_{ni}^4 &= \sum_{i=1}^{k} \{h(U_{ni})\}^4 = \sum_{i=1}^{k} \{h'(0)U_{ni} + \{h'(\gamma U_{ni}) - h'(0)\} D_{ni}\}^4 \qquad (4.21) \\
&= \sum_{i=1}^{k} U_{ni}^4 \{h'(0) + o_p(1)\}^4 \\
&= (h'(0))^4 \sum_{i=1}^{4} U_{ni}^4 + o_p(1).
\end{aligned}$$

Substituting (4.3) in (4.21), we have

$$\begin{aligned}
\sum_{i=1}^{k} D_{ni}^4 &= (2f(\boldsymbol{x}_0))^{-4} \sum_{i=1}^{k} (j/tn^{5/(q+4)})^4 + o_p(1) \\
&= (2f(\boldsymbol{x}_0))^{-4} t^{-4} \left[n^{-4/(q+4)} \sum_{i=1}^{tn^{4/(q+4)}} (j/n^{4/(q+4)})^4 \right] + o_p(1) \\
&\to (2f(\boldsymbol{x}_0)t)^{-4} \int_0^t s^4 \mathrm{d}s \\
&= (2f(\boldsymbol{x}_0)t)^{-4} t^5/5
\end{aligned}$$

$$= \left\{80 f^4(\boldsymbol{x}_0)\right\}^4 t.$$

So, we have

$$E(W_{nk}) \rightarrow -(1/4)d^2,$$

where

$$d^2 = \left\{80 f^4(\boldsymbol{x}_0)\right\}^{-1} t \int_{-\infty}^{\infty} \left\{q^2(y, \boldsymbol{x}_0)/g(y)\right\} \mathrm{d}y.$$

$\diamond$

Lemma 4.15 *Define*

$$V_{nk} = \sum_{i=1}^{k} D_{ni}^2 \left\{q(Y_{ni}, \boldsymbol{x}_0)/g(Y_{ni})\right\}.$$

Assume [A1]-[A4]. Then under p_{nk}

$$Var(W_{nk} - V_{nk}) \rightarrow 0.$$

Proof: Under p_{nk}

$$\begin{aligned}
Var(W_{nk} - V_{nk}) &= 4 \sum_{i=1}^{k} Var \left[\frac{s(Y_{ni}|D_{ni}}{s(Y_{ni})} - 1 - (1/2) D_{ni}^2 \frac{q(Y_{ni}, \boldsymbol{x}_0)}{g(Y_{ni})} \right] \\
&\leq 4 \sum_{i=1}^{k} E \left[\frac{s(Y_{ni}|D_{ni} - s(Y_{ni})}{s(Y_{ni})} - (1/2) D_{ni}^2 \frac{q(Y_{ni}, \boldsymbol{x}_0)}{g(Y_{ni})} \right]^2 \\
&\leq 4 \sum_{i=1}^{k} D_{ni}^4 \int \left[\left\{ (s(Y_{ni}|D_{ni}) - s(Y_{ni}))/D_{ni}^2 \right\} \right. \\
&\qquad \left. -(1/2) \left\{ q(y, \boldsymbol{x}_0)/(g(y))^{1/2} \right\} \right]^2 \mathrm{d}y \\
&\leq 4 \sum_{i=1}^{k} D_{ni}^4 \int \left[(\partial/\partial d^2) s(y|d) \Big|_{d_{ni}^*} - (1/2) \left\{ q(y, \boldsymbol{x}_0)/(g(y))^{1/2} \right\} \right]^2 \mathrm{d}y
\end{aligned}$$

where $0 \leq d_{ni}^* \leq D_{ni}$, $i = 1, 2, \cdots, k$. So we conclude that the last integral converges to 0 uniformly for all $i = 1, 2, \cdots, k$. $\diamond$

Proof of the theorem 4.2 : Note that under p_{nk},

$$E(V_{nk}) = \sum_{i=1}^{k} D_{ni}^2 \int \{q(y, \boldsymbol{x}_0)/g(y)\} g(y)\mathrm{d}y = \sum_{i=1}^{k} D_{ni}^2 \int q(y, \boldsymbol{x}_0)\mathrm{d}y = 0,$$

and similarly under p_{nk},

$$\begin{aligned} E(V_{nk}) &= \sum_{i=1}^{k} D_{ni}^4 \int \{q(y, \boldsymbol{x}_0)/g(y)\} \,\mathrm{d}y \\ &= \left\{ \left[80 f^4(\boldsymbol{x}_0)\right]^{-1} t + o_p(1) \right\} \int \{q(y, \boldsymbol{x}_0)/g(y)\} \,\mathrm{d}y \\ &\to d^2. \end{aligned}$$

So, it follows

$$V_{nk} \to N(0, d^2).$$

Now, under p_{nk}

$$E\left[W_{nk} - V_{nk} + (1/4)d^2\right] \to 0.$$

And

$$E[\log L_{nk} - V_{nk} + (1/2)d^2] \to 0.$$

Now, we consider the case when $k = [tn^{\lambda}]$ for $0 < \lambda < 4/(q+4)$. By lemma 3.1, we have $\sum_{i=1}^{k} D_{ni}^4 = o_p(1). So, E(W_{nk}) \to 0$ (i.e., $d^2 = 0$). So it follows easily that under p_{nk}

$$\log L_{nk} \to 0 \text{ in probability.}$$

Hence, the LeCam first lemma holds through the degenerate normal law. This completes the proof.

4.4 Multiple conditional quantile process

In the multiple multivariate setting (i.e., we have $l \times 1$ response vector $\mathbf{Y}_i$, and $q \times 1$ vector of covariates $\mathbf{X}_i$ for i-th subject), then we need a extension of mapping $(\mathbf{X}_i, \mathbf{Y}_i) \to (Z_i, \mathbf{Y}_i)$. Consider the transformation $(\mathbf{X}_i, \mathbf{Y}_i) \to (Z_i, \mathbf{Y}_i)$ where

$D_i = \rho(\mathbf{X}_i, \boldsymbol{x}_0)$. Let $D_{n1} < \cdots < D_{nn}$ be the order statistic corresponding to $D_1, \cdots, D_n$. Let $\mathbf{Y}_{n1}, \cdots, \mathbf{Y}_{nn}$ be the induced order statistics. (i.e., $\mathbf{Y}_{ni} = \mathbf{Y}_j$ if $Z_{ni} = Z_j$, for $j = 1, \cdots, n$.) To define $\boldsymbol{\xi}_{pT}(\mathbf{Y}|\boldsymbol{x}_0)$, we need a multivariate empirical d.f. and corresponding conditional quantile. The i-th element of multivariate empirical d.f. and corresponding $k - NN$ estimator of the conditional quantile are

$$\hat{G}_{nk}^{(i)}(y) = k^{-1} \sum_{j=1}^{k} 1(Y_{nj}^{(i)} \leq y), \ y \in R^1$$

and

$$\hat{\xi}_{nk}^{(i)} = \inf\{y : \hat{G}_{nk}^{(i)}(y) \geq k^{-1}[kp]\},$$

where $1(Y \leq y)$, is 1 if $Y \leq y$, 0 otherwise and $i = 1, \cdots, l$. So $l \times 1$ multivariate $k - NN$ estimator of the conditional quintile is defined by

$$\hat{\boldsymbol{\xi}}_{nk} = (\hat{\xi}_{nk}^{(1)}, \cdots, \hat{\xi}_{nk}^{(l)})'.$$

$\hat{\boldsymbol{\xi}}_{nk}$ has component-wise asymptotic normality. But to derive the joint normality for $\hat{\boldsymbol{\xi}}_{nk}$, we need to consider the multivariate conditional quantile process which remains as a future study. An estimator of $\boldsymbol{\mu}$ can be obtained by the marginal approach: let $x^{(j)}$ be the j-the component of $\boldsymbol{x}$ for $j = 1, \cdots, q$. And the j-th the component of $\hat{\boldsymbol{\mu}}$ can be obtained by

$$\inf_{\mu^{(j)}} \left\{ \sup_{x^{(j)}} \rho(\boldsymbol{m}_{pT}(\cdot|x^{(j)}), \boldsymbol{m}_{pS}(\cdot|x^{(j)} + \mu^{(j)})) \right\}, \text{ for } j = 1, \cdots, q. \qquad (4.22)$$

And $\hat{\boldsymbol{\mu}} = (\hat{\mu}^{(1)}, \cdots, \hat{\mu}^{(r)})$.

Chapter 5

Numerical Examples

5.1 Introduction

In this chapter, we demonstrate numerical examples of our work. We use $p = 1/2$, the median of the response. The following is the summary of the paper by The ARIC Investigators (1989). Atherosclerosis Risk in Communities (ARIC) is a prospective investigation of the etiology and natural history of atherosclerosis and the etiology of clinical atherosclerotic disease in four US communities. In each of four US communities (Forsyth County, North Carolina; Jackson, Mississippi; suburbs of Minneapolis Minnesota: and Washington County, Maryland) 4,000 adults aged 45-64 years are examined twice three years apart. Examinations include ultrasound scanning of carotid artery. In this example, we separately studied the black female and the white female groups. The black female data included the participants with second visits from all four communities; the white female data included the participants with second visits from Forsyth County, North Carolina. The response variable is the average carotid artery far wall thickness(WT) scanned by Beta-mode ultrasound equipment in mm. The explanatory variable is Body Mass Index (BMI) which is defined by weight/height^2 (Kg/m^2). The group variable for black females is the hypertension medicine history which is one if the participant has taken hypertension lowering medication in past two weeks, zero otherwise. The group variable for white females is the smoking status which is zero if the participant is (was) not a smoker, one if the participant is a current or former smoker. The main interest is to test the bioequivalence of the carotid artery far wall thickness between the two groups and to estimate the relative potency.

5.2 ARIC Black females

Since the group variable is the use of the hypertension lowering medication for black females, the participants were classified as hypertensive or not. In turn, the use of the hypertensive medicine is determined by the exhibition of hypertension symptoms. The black female data has sample size 623 and 735 for the hypertension and the normotension groups, respectively.

5.2.1 Inferences at a given point of x

Table 5.1 shows the results of estimating the conditional median carotid artery far wall thickness conditioned on the mean BMI of the blacks (30.162 Kg/m^2). All the variance estimators are calculated by the bootstrap method of 200 iterations. The δNN estimator is shown to be asymptotically unbiased, but it has a much larger variance than the NN estimator does. The 95 % confidence interval of the difference between two medians of the δNN estimators includes zero; therefore, we cannot say that the conditional medians are different between the hypertension and the normotension groups by 95% significance level. The next two lines show the inference on the ratio. The lower and upper bounds of the 95% C.I. of the ratio by Fieller's theorem are not in the range of $\pm$ 20 %. From the viewpoint of bioequivalence, the bioequivalence is not acceptable based on the C.I. of the ratio. Estimated variances

Table 5.1: Estimated conditional median carotid artery far wall thickness (mm) conditioned on BMI=30.162 Kg/m^2 (black females)

		Normotension	Hypertension
	n	623	735
δNN	$k_n{}^1 = n^{4/5-2\delta}$	90	101
	$\hat{\xi}_1$	0.62965	0.66521
	$\hat{\sigma}$	0.01553	0.02406
	95% C.I.	(0.59921,0.66009)	(0.61805,0.71237)
	95% C.I. of the diff.	(-0.02057,0.09169)	
	ratio	1.06	
	Fieller's C.I.(95%)	(0.53,1.96)	
NN	$k_n = n^{4/5}$	172	196
	$\hat{\xi}_2$	0.64124	0.68388
	$\hat{\sigma}$	0.006596	0.006399
NBC	$\hat{\xi}_2 - \hat{\xi}_1$	0.01159	0.01867
	bias correction on n	0.00958	0.01207
	95% C.I.	(0.61878,0.64454)	(0.65927,0.68435)
	95% C.I. of the diff.	(0.02214,0.05816)	
	ratio	1.06	
	Fieller's C.I.(95%)	(0.75,1.52)	
KBC	bias correction on k_n	0.01870	0.01098
	95% C.I.	(0.60961,0.63547)	(0.66036,0.68544)
	95% C.I. of the diff.	(0.03235,0.06837)	
	ratio	1.08	
	Fieller's C.I.(95%)	(0.76,1.55)	

Table 5.2: Jackknife bias correction on k_n sample (black females)

p	Normotension $k_n = n^{4/5}$=172	Hypertension $k_n = n^{4/5}$= 196
0.90	0.00000	0.00000
0.80	0.01981	-0.00133
0.70	0.01527	-0.02294
0.60	0.01811	-0.00075
0.50	0.01823	0.01404
0.40	0.01589	0.02483
0.30	0.02055	0.02292
0.20	0.01424	0.02173
0.10	0.02751	0.02896
mean	0.01870	0.01098

from the NN method are much smaller, but the median estimators are theoretically biased. The differences between the δNN and the NN estimator for the hypertension and the normotension groups are 0.01159 and 0.01867. Jackknife bias corrections (KBC (bias correction on the k_n samples) and NBC (bias correction on the whole n data)) were employed. Tables 5.2 and 5.3 show the estimated bias correction in terms of p (sub-sampling rate). There is no standard way to choose a particular p, so we tried several possible cases and used the mean for our correction. Using the KBC method, we do not need to sort the data for every p; however, using the NBC method, the data does need to be sorted for every subsample, since the subsamples are changing for each different p. Therefore, computationally KBC is less intensive than NBC, but the estimated bias from the NBC approach looks more stable than the bias from KBC. The 95% C.I.'s of the differences between the two medians do not include zero for both bias correction methods. Therefore, the results of the NN

Table 5.3: Jackknife bias correction on the whole n data (black females)

		Normotension n =623			Hypertensi on n =735		
# of gps	p	pn	k_{pn}	bias	pn	k_{pn}	bias
10	0.1000	62.3	27	0.02517	72.3	31	0.01740
9	0.1111	69.2	30	0.00402	80.3	33	0.01586
8	0.1250	77.9	33	0.01081	90.3	37	0.00702
7	0.1429	89.0	36	0.00757	103.3	41	0.02390
6	0.1667	103.8	41	0.00449	120.5	46	0.00586
5	0.2000	124.6	48	0.01069	144.6	53	0.00793
4	0.2500	155.8	57	0.00432	180.8	64	0.00655
mean				0.00958			0.01207

method with bias correction show that the median carotid artery far wall thickness of hypertensive patients is significantly thicker than that of normotensive people after adjusting for BMI. The Fieller's C.I.'s of both the δNN and the NN methods are somewhat larger than $\pm$ 20%; the bioequivalence is not acceptable from the results. If we increase the sample size, we may be able to conclude the bioequivalence.

5.2.2 Inferences for the processes

To test the bioequivalence between the two groups and to estimate the relative potency, we assume the bioequivalence model (parallelism of the two response curves) which is:

$$\text{Ho}: \xi^T(x) - \xi^S(x+\mu) = 0 \text{ for all } x \in \mathcal{X}. \tag{5.1}$$

We need to check the parallelism before doing any statistical inferences based on the assumption (5.1). Figure 5.1 shows the estimated conditional median curves over

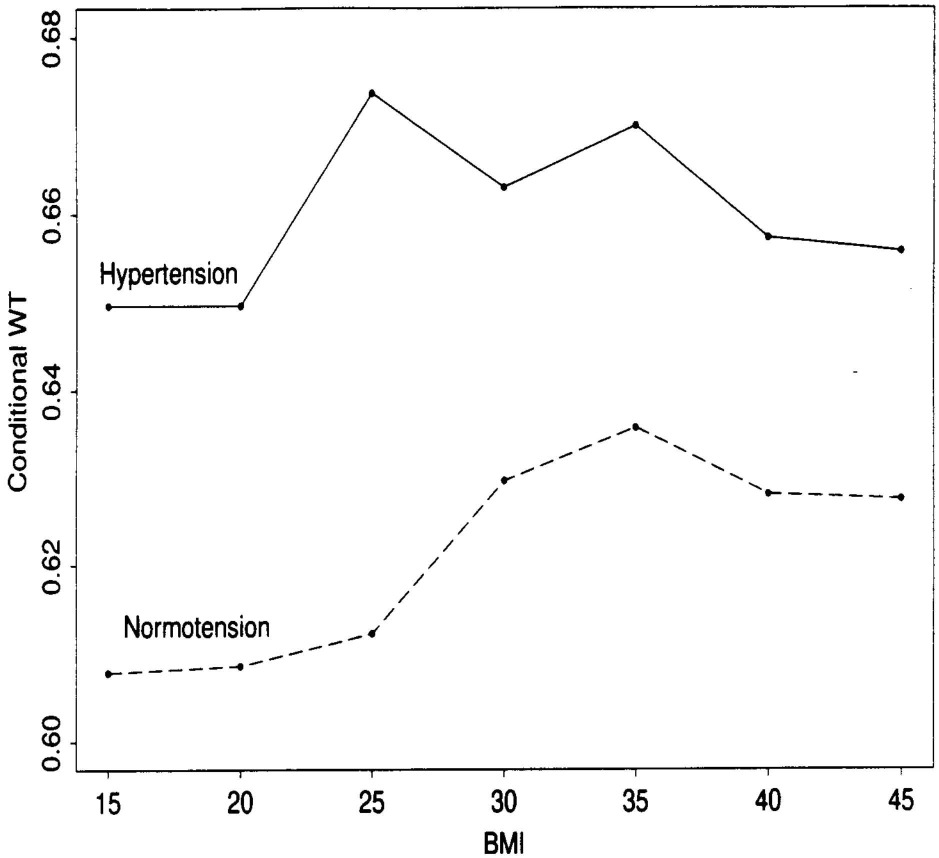

Figure 5.1: Estimated conditional median carotid artery far wall thickness (mm) conditioned on BMI (Kg/m^2) for two groups, $\mu = 0$ (black females)

BMI. and we decided that there is no violation of the parallelism. WT curve of the hypertension group is greater than that of the normotension group for all BMI's. To estimate μ, we discussed the mini-max rule in chapter 1. Figure 5.2 shows

$$\sup_{x \in \mathcal{X}} |\xi_n^T(x) - \xi_n^S(x + \mu)| \tag{5.2}$$

as a function of μ. From the figure, we estimate μ is 10 which minimizes (5.2). i.e. the WT curve of hypertension patients is shifted to the left by 10 units of BMI compared to the WT curve of the normotension group.

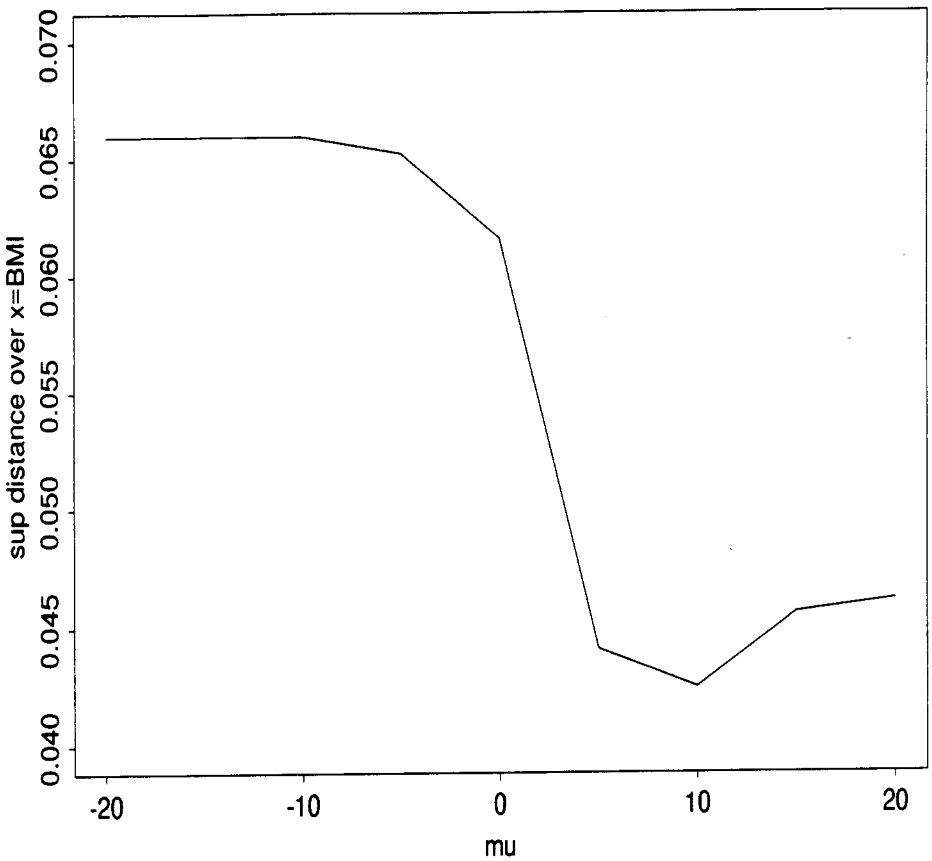

Figure 5.2: $\sup_{x \in \mathcal{X}} |\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x + \mu)|$ vs. μ (black females)

To test (5.1) when $\mu = 10$, as we discussed in section 2.6,

$$k_n^{1/2}[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x}) - \hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x} + \boldsymbol{\mu})] \xrightarrow{\mathcal{D}} N_m(0, \Sigma),$$

for $\boldsymbol{x} = (x_1, \cdots, x_m)'$ as $n \to \infty$. We reject Ho if

$$Q = k_n \left[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x}) - \hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x} + \boldsymbol{\mu})\right]' \hat{\Sigma}_n^{-1} \left[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x}) - \hat{\boldsymbol{\xi}}_n^S(\boldsymbol{x} + \boldsymbol{\mu})\right]' > \chi^2(m, 1 - \alpha). \qquad (5.3)$$

where $\hat{\Sigma}_n$ is the variance matrix of $\left[\hat{\boldsymbol{\xi}}_n^T(\boldsymbol{x}) - \hat{\boldsymbol{\xi}}^S(\boldsymbol{x} + \boldsymbol{\mu})\right]$ which is asymptotically a diagonal matrix. where $\chi^2(m, 1 - \alpha)$ is the upper α percentile from χ^2 distribution

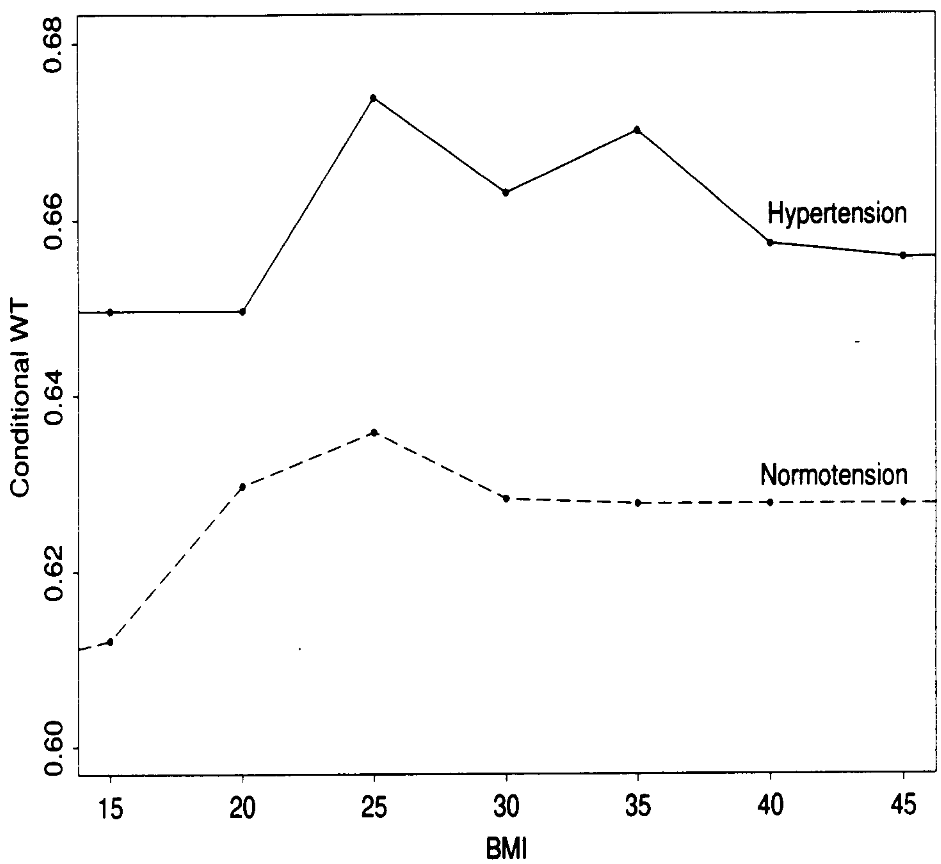

Figure 5.3: Estimated conditional median carotid artery far wall thickness (mm) conditioned on BMI (Kg/m^2) for two groups, $\mu = 10$ (estimated relative potency) (black females)

with m d.f. Therefore, we employ the following

$$Q_d = \sum_{i=1}^{m} (\hat{\xi}_n^T(x_i) - \hat{\xi}_n^S(x+\mu))^2 \tilde{\sigma}_{ii}^{-1}, \tag{5.4}$$

where

$$\tilde{\sigma}_{ii} = \hat{Var}(\xi_n^T(x_i) - \xi_n^S(x_i+\mu)).$$

We use $\boldsymbol{x}' = (15, 20, 25, 30, 35, 40, 45)$ with 200 bootstrap iterations to estimate $\hat{\Sigma}_n$

and $\tilde{\sigma}_{ii}$. We have $Q = 20.36$ with p-value 0.005 and $Q_d = 18.87$ with $p = 0.009$. Therefore, we reject the hypothesis (5.1) which means the difference between the two WT curves are statistically significant. The difference $Q - Q_d = 1.30$ can be interpreted as the increase of the χ^2 due to the non-zero off-diagonal elements in $\hat{\Sigma}_n$. Since $Q - Q_d$ is small, we can say that the independence assumption of the conditional quantile over $x \in \mathcal{X}$ is not violated.

The difference of the mean WT's between the two groups is 0.04974. We subtract the difference from the hypertension group and apply the test again. The values $Q = 5.88$ with $p = 0.55$ and $Q_d = 5.56$ with $p = 0.59$ were obtained. Our conclusions are as follows: 1) the WT curve of the hypertension group is significantly higher than that of the normotension group for the entire range of BMI, 2) but hypertension effect-adjusted WT curves are equivalent for the two groups; WT curves have the same shape for the two groups.

5.3 ARIC White females

Using white females, we compared the two conditional median WT curves for the smokers and the non-smokers conditioned on BMI and evaluated the bioequivalence between the two curves.

5.3.1 Statistical inferences

Since Tables 5.4 to 5.6 show the similar results as the black females, we omit the details of interpretations for these tables. From Figure 5.4, we estimate μ as zero, and Figure 5.5 shows the median curves for smokers and non-smokers.

We notice two things from Figure 5.5 . First, the curves cross, clearly, the parallel assumption is violated where BMI is approximately 17. Furthermore, the

Table 5.4: Estimated conditional median carotid artery far wall thickness (mm) conditioned on BMI=25.173 Kg/m^2 (white females)

		Non Smoker	Smoker
	n	651	679
δNN	$k_n = n^{4/5-2\delta}$	93	96
	$\hat{\xi}_1$	0.61823	0.63371
	$\hat{\sigma}$	0.01694	0.01315
	95% C.I.	(0.58503,0.65143)	(0.60793,0.65947)
	95% C.I. of the diff.	(-0.02655,0.05751)	
	ratio	1.03	
	Fieller's C.I.(95%)	(0.59,1.88)	
NN	$k_n = n^{4/5}$	178	184
	$\hat{\xi}_2$	0.61518	0.62432
	$\hat{\sigma}$	0.004833	0.004388
KBC	$\hat{\xi}_2 - \hat{\xi}_1$	-0.00305	-0.00939
	bias correction on k_n	-0.00485	-0.00713
	95%C.I.	(0.61056,0.62950)	(0.62285,0.64005)
	95% C.I. of the diff.	(-0.00137,0.02422)	
	ratio	1.02	
	Fieller's C.I.(95%)	(0.75,1.39)	
NBC	bias correction on n	-0.00865	-0.00430
	95%C.I.	(0.61436,0.63300)	(0.62002,0.63722)
	95% C.I. of the diff.	(-0.00800,0.01759)	
	ratio	1.01	
	Fieller's C.I.(95%)	(0.74,1.37)	

Table 5.5: Jackknife bias correction on k_n sample (white females)

p	Non Smoker $k_n = n^{4/5}$=178	Smoker $k_n = n^{4/5}$= 184
0.95	0.0000	-0.01036
0.90	0.00000	-0.00532
0.80	0.00000	-0.00281
0.70	-0.00571	-0.00198
0.60	-0.00477	-0.00158
0.50	-0.00407	-0.02070
mean	-0.00485	-0.00713

Table 5.6: Jackknife bias correction on the whole n data (white females)

		Non Smoker n =651			Smoker n =679		
# of gps	p	pn	k_{pn}	bias	pn	k_{pn}	bias
10	0.1000	65.1	28	-0.01040	67.9	29	-0.00519
9	0.1111	72.3	31	-0.01210	75.4	32	-0.00359
8	0.1250	81.4	34	-0.00993	84.8	35	-0.00698
7	0.1429	93.0	38	-0.00549	97.0	39	-0.00337
6	0.1667	108.5	42	-0.00697	113.5	44	0.00193
5	0.2000	130.2	49	-0.00675	135.8	51	-0.00389
4	0.2500	162.8	59	-0.00891	169.8	61	-0.00902
mean				-0.00958			-0.01207

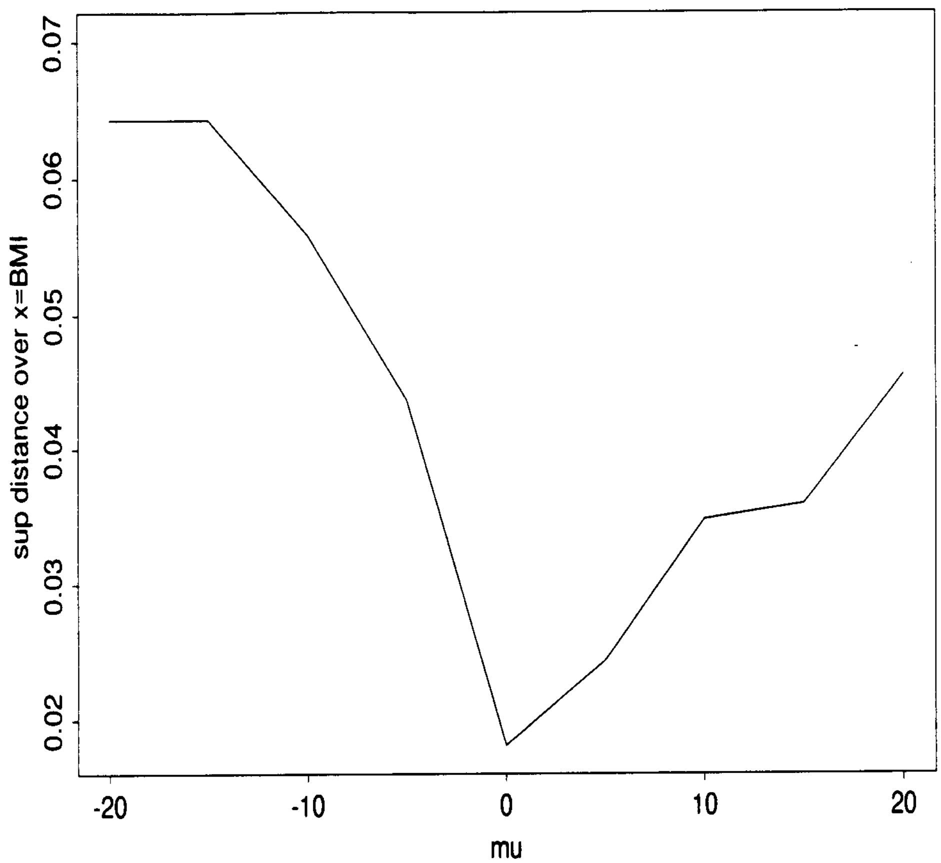

Figure 5.4: $\sup_{x \in \mathcal{X}} |\hat{\xi}_n^T(x) - \hat{\xi}_n^S(x+\mu)|$ vs. μ (white females)

curves are flat for large values of BMI, say, greater than 40. These two issues are related to the k_n samples for given BMI; BMI has a bell-shape distribution. Therefore, the data points are dense in the middle of the range and sparse at both ends. To calculate the conditional median conditioned on BMI=10 for non-smokers, the range of k_n sample is (16.6,21.1). To calculate the conditional median at BMI=10, we use the data where the BMI range does not cover the point of BMI=10. Consequently, we have too few data points at these extreme points, and the inferences at these points reflect the trend for the mid-ranged data points. Table 5.7 shows the BMI

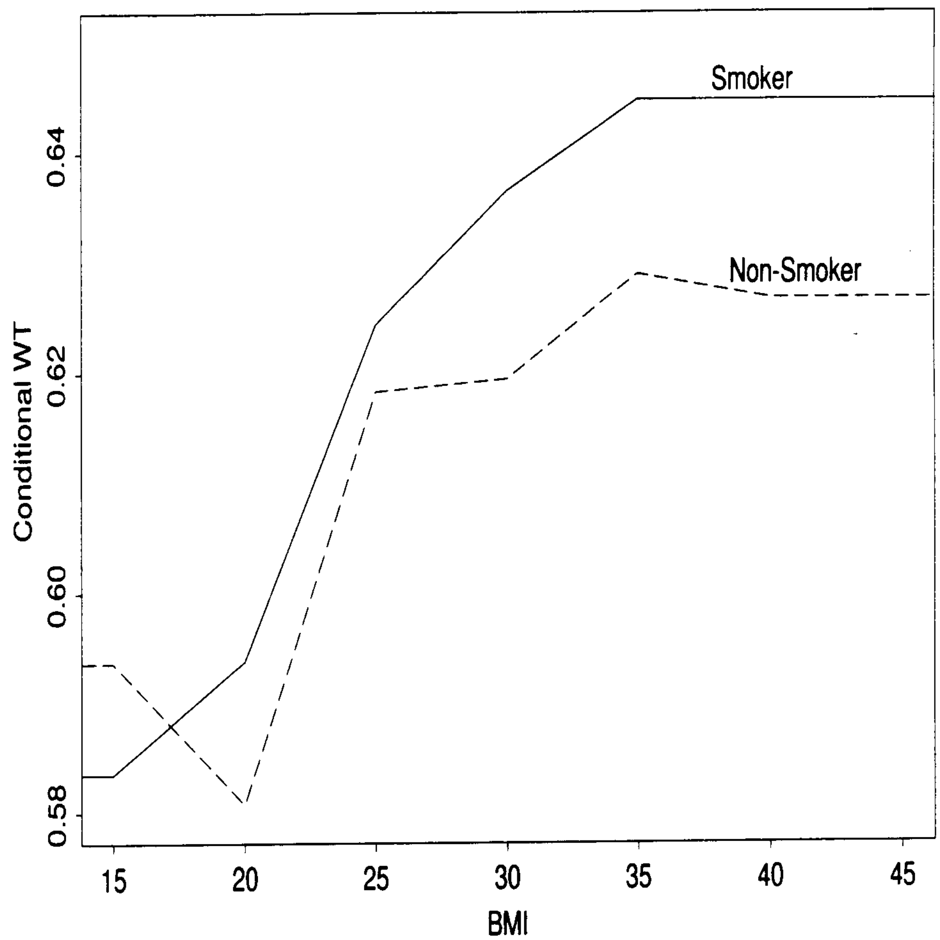

Figure 5.5: Estimated conditional median carotid artery far wall thickness (mm) conditioned on BMI (Kg/m^2) for two groups: $\mu = 0$ (white females)

ranges for k_n samples at each given point of BMI. And Figure 5.6 is the diagram for the Table 5.7. Ideally, the ranges should be exclusive, which is true in the middle of the range. If the ranges are exclusive, then the asymptotic independence theory will work well. In Figure 5.6, BMI ranges for BMI=10 and 15 and the same and the range for BMI=40, 45 and 50 are the same. We conclude that we cannot make statistical inferences for those points due to the lack of data. We restrict our analysis for BMI bigger than or equal to 20 and less than or equal to 35.

From Figure 5.5, the WT curve of the smoking group is higher than that of the non-smoking group. The difference is small for thin people, but the difference for heavy people is much larger. Moreover, WT increases as BMI increases for the two groups. Figure 5.4 shows the $\sup_{x\in\mathcal{X}} |\hat{\xi}^T(x) - \hat{\xi}^S(x+\mu)|$ vs. μ. Our estimator of the relative potency is zero. We use $\boldsymbol{x} = (20, 22.5, 25, 27.5, 30, 32.5, 35)$ for the test of the fundamental assumption, and obtain $Q = 2.74$ with $p = 0.91$ and $Q_d = 2.73$ with $p = 0.91$.

Our conclusions of the 2nd example are these: 1) WT increases as BMI increases for the smokers and the non-smokers, 2) WT of the smoking group is higher than that of the non-smoking group, 3) the difference of WT between the two groups is larger for heavy people than for thin people, 4) but the difference in WT between the smokers and non-smokers is not statistically significant.

Table 5.7: BMI range of k_n sample at each point of BMI (white females)

given BMI($= x_0$)	Non Smoker	Smoker
10	(16.6,21.1)	(13.7,20.7)
15	(16.6,21.1)	(13.7,20.7)
20	(18.6,21.4)	(19.0,21.1)
22.5	(21.8,23.2)	(21.9,23.1)
25	(24.3,25.7)	(24.2,25.9)
30	(28.1,31.9)	(28.0,32.0)
35	(30.0,39.5)	(29.3,40.1)
40	(30.2,48.3)	(29.5,43.6)
45	(30.2,48.3)	(29.5,43.6)
50	(30.2,48.3)	(29.5,43.6)

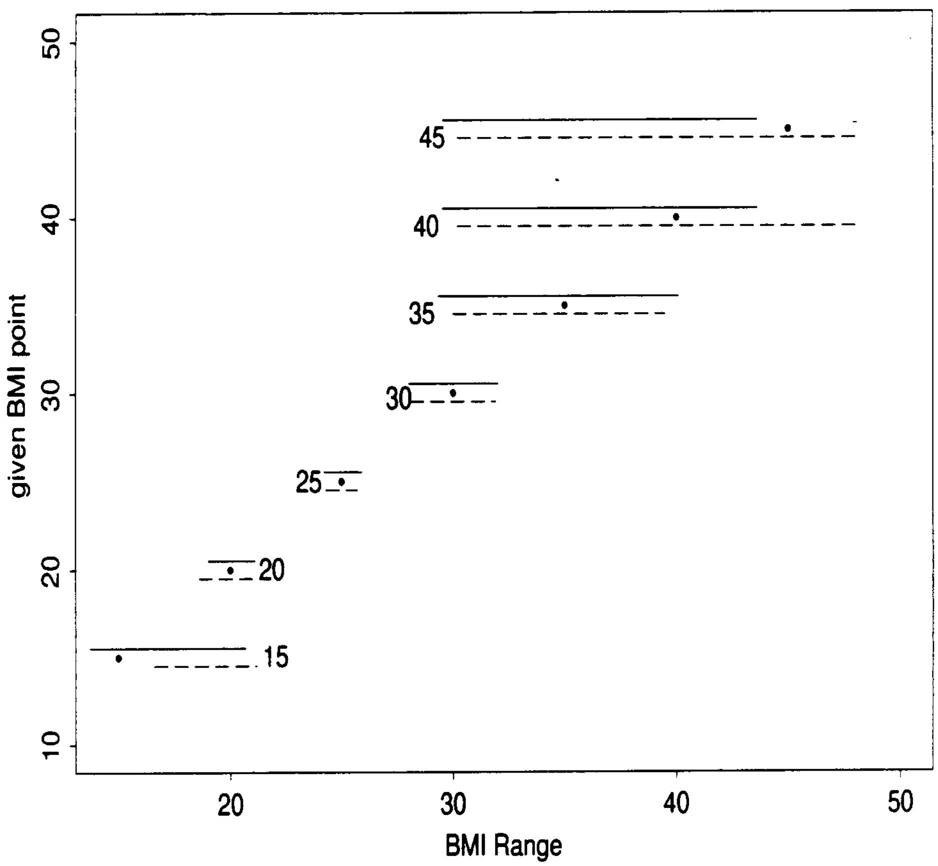

Figure 5.6: BMI range of k_n NN sample at each point of BMI (white females): solid line: smoker, dotted line: non-smoker

Chapter 6

Summary and further research

In this paper, we propose a conditional quantile process by using the k-NN estimator. We have shown that our process is asymptotically a Gaussian process which is fully characterized by pointwise asymptotic normality and covariance structures for only shrinking neighborhood of fixed points. A Kolmogorov type estimator of the relative potency was proposed. We have the same asymptotic normality of the conditional quantile under the small measurement errors. The difference of the quantiles between the two cdf's (one without measurement error and one with measurement error) is the order of $O(||\epsilon||^2)$ under the uniform measurement error ($U(-\epsilon, \epsilon)$). We expanded our results to the multivariate case by contiguity arguments. Finally, we applied our method to ARIC data.

We have several questions which may become future research topics. There may be more than two responses in which we need to consider the correlation structure among the responses. We may want to estimate the effect of measurement errors on the quantile which was shown to be the order of $O(||\epsilon||^2)$. To test the fundamental assumption, we can try to obtain the distribution of $\sup_{x \in \mathcal{X}} |W^T(x) - W^S(x + \mu)|$ which would make our test of the fundamental assumption of the bioequivalence models very simple. The Kolmogorov type estimator of the relative potency may possess some appealing properties such as asymptotic normality or consistency. One interesting question is how to select k-NN samples for the processes. In this paper, we select x's by the same increment over the range of interest. We may select exclusive samples to guarantee independence among them, or we may allow some dependency by selecting not totally exclusive samples.

Bibliography

Bahadur, R. R. (1966). A note on quantiles in large samples, *The Annals of Mathematical Statistics* **37**: 577–580.

Bhattacharya, P. K. and Gangopadhyay, A. K. (1988). Kernel and nearest neighbor estimation of a conditional quantile, *Technical report series of the intercollege division of statistics Technical Report 104*, University of California at Davis.

Bhattacharya, P. K. and Mack, Y. P. (1987). Weak convergence of k-NN density and regression estimators with varying k and applications, *The Annals of Statistics* **15**: 976–994.

Bhattachatya, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics, *The Annals of Statistics* **2**: 1034–1039.

Box, G. E. P. (1950). Problems in the analysis of growth and wear curves, *Biometrics* **6**: 362–389.

Clarke, B. R. (1986). Nonsmooth analysis and Frechet differentiability of M-functionals, *Prob. Theor. Rel. Fields* **73**: 197–209.

Cole, J. W. C. and Grizzle, J. E. (1966). Application of multivariate analysis of variance to repeated measurements experiments, *Biometrics* **22**: 810–828.

Fernholz, L. T. (1983). *Von Mises calculu for statistical functions* , *Lecture notes in statistics*, 19 edn, New York: Springer.

Fieller, E. C. (1940). The biological standardization of insulin, *Journal of the Royal Statistical Society, Supplementary 7* pp. 1–64.

Fieller, E. C. (1944). A fundamental formula in the statistics of bioassay , and some applications. *Q. J. Pharm. Pharmaco.* pp. 117–123.

Finney, D. J. (1951). Two new uses of the Behrens-Fisher distribution, *Journal of the Royal Statistical Society, Ser.* **B 12**: 296–300.

Finney, D. J. (1964). *Statistical Methods in Biological Assay*, 2nd edn. London: Charles Griffin.

Gangopadhyay, A. K. and Sen, P. K. (1992). Contiguity in nonparametric estimation of a conditional functional, *in* A. Saleh (ed.), *Nonparametric Statistics and related topics*, Amsterdam: North Holland, pp. 141–162.

Gangopadhyay, A. K. and Sen, P. K. (1993). Contiguity in Bahadur-type representation of a conditional quantile and application in conditional quantile process. *in* J. K. G. et al. (ed.), *Statistics and Probability, a Raghu Rai Bahadur Festschrift*, New Delhi: Wiley Eastern, pp. 219–232.

Gikhman, I. I. and Skorokhod, A. V. (1969). *Introduction to the theory of random processes*, W.B. Saunders, Philadelphia.

Govindarajulu, Z. (1988). *Statistical Techniques in Bioassay*, New York: Karger.

Grizzle, J. E. (1965). The two-period change-over design and its use in clinical trials. *Biometrics* **21**: 467–480.

Grizzle, J. E. and Allen, D. M. (1969). Analysis of grwoth dose-response curves. *Biometrics* **25**: 357–381.

Hájek, J. and Sidák, Z. (1967). *Theory of Rank Statistics*, New York: Academic Press.

Jones, D. A. and Karson, M. J. (1972). On the use of confidence regions to test hypotheses. *Journal of Quality Technology* **4**: 156–158.

Metzler, C. M. (1974). Bioavailability-a problem in bioequivalence. *Biometrics* **30**: 309–317.

Rao, P. V. and Little, R. C. (1976). An estimator of relative potency, *Communications in Statistics A* **5**: 183–189.

Rao, P. V., Schuster, E. F. and Little, R. C. (1975). Estimation of shift and center of symmetry based on Kolmogorov-Smirnov statistics, *The Annals of Statistics* **3**: 862–873.

Sen, P. K. (1963). On the estimation of relative potency in dilute (-direct) assays by distribution-free methods, *Biometrics* **19**: 532–552.

Sen, P. K. (1981). *Sequential Nonparametrics*, New York: John Wiley & Sons.

Sen, P. K. (1994). Regression quantiles in nonparametric regression, *Nonparametric Statistics* **3**: 237–253.

Sen, P. K. and Gangopadhyay, A. K. (1990). Bootstrap confidence intervals for conditional quantile functions, *Sankhyã : The Indian Journal of Statistics A* **52**: 346–363.

Sen, P. K. and Singer, J. M. (1993). *Large Smaple Methods in Statistics*. New York: Chapman & Hall.

Shorack, G. R. (1966). Graphical procedures for using distribution-free methods in the estimation of relative potency in dilution (-direct) assays. *Biometrics* **22**: 610–619.

Snee, R. D. (1972). On the analysis of response curve data, *Technometrics* **14**: 47–62.

Stute, W. (1986). Conditional empirical processes, *The Annals of Statistics* **5**: 595–645.

The ARIC Investigators (1989). The Atherosclerosis Risk in Comminities Study: Design and Objectives, *American Journal of Epidemiology* **129**: 687–702.

Wallenstein, S. and Fisher, A. C. (1977). The analysis of two repeated measurements design with appication to clinical trials, *Biometrics* **33**: 261–269.

Westlake, W. J. (1972). Use of confidence intervals in analysis of comparative bioavailability trials, *Juurnal of Pharmaceutical Sciences* **62**: 1340–1341.

Westlake, W. J. (1974). The use of balanced incomplete block design in comparative bioavailability trials, *Biometrics* **30**: 319–327.

Westlake, W. J. (1979). Statistical aspects of comparative bioavailability trials, *Biometrics* **35**: 273–280.

Yates, F. (1939). An apparent inconsistency arising from tests of significance based on fiducial distributions of unknown parameters, *Proc. Camb. Phil. Soc.* **35**: 579–591.