Robust Rank-Based Inference Procedures in the Heteroscedastic Linear Model

Sherry L. Dixon
Western Michigan University
ROBUST RANK-BASED INFERENCE PROCEDURES
IN THE HETEROSCEDASTIC LINEAR MODEL

by

Sherry L. Dixon

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Mathematics and Statistics

Western Michigan University
Kalamazoo, Michigan
April 1993

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
This Dissertation considers robust, rank-based inference procedures in the heteroscedastic linear model. Attention is confined to the case in which the error variance is a parametric function of the mean response and a variance parameter $\theta$.

We develop a procedure for estimating the variance parameter $\theta$. Our method is based on the asymptotic linearity result of the scale rank statistic. This linearity result is the rank statistic analogue to that of the pseudolikelihood statistic proposed by Carroll and Ruppert (1982a) for variance parameter estimation. Based on any $\sqrt{N}$-consistent estimate of the regression parameter $\beta$ we show that our Hodges-Lehmann (1963) type estimate of $\theta$ is a $\sqrt{N}$-consistent estimate and obtain its asymptotic distribution when the estimated regression parameter is the "adjusted" rank-based estimate of $\beta$, discussed next.

For estimation of $\beta$ we consider a rank-based estimate which is based on weighted observations. Our weighted $R$-estimate is a generalization of the estimate proposed by Jureckova (1971a) for the homoscedastic case. The
methodology is based on the asymptotic linearity result which generalizes Jureckova's result for the homoscedastic case. Under the assumption that the errors are symmetrically distributed about zero, the regression plane intersects the origin. However, it is well known that a no-intercept model cannot be fit with $R$-estimation. Thus we fit the intercept model and adjust our estimate by projecting the fitted values onto the appropriate space. A result which generalizes the asymptotic linearity results of Tardiff (1985) for the signed-rank statistic is also established. Using the linearity results for the regression statistic and the signed-rank statistic we establish asymptotic equivalence of the "adjusted" estimates based on the true and estimated weights.

Our inferential procedure is asymptotic; hence some consideration to small sample properties is needed. We present the results of a Monte Carlo study which verifies our inference for the situations considered.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Robust rank-based inference procedures in the heteroscedastic linear model

Dixon, Sherry Leigh, Ph.D.
Western Michigan University, 1993
ACKNOWLEDGEMENTS

I wish to express my gratitude to Joseph W. McKean for his continual guidance and support throughout this research. I can never express in words my heart felt gratitude to my Mother for her enduring support, unending encouragement and for providing stability in my life. I am indebted to my Grandmother for instilling the value of education in me since childhood. Her pioneering spirit and determination to succeed have served me as a source of inspiration and will guide me in future endeavors. Finally, I thank my Father, who fostered an atmosphere conducive to mathematical scholarship and conveyed to me his love of mathematics.

Sherry L. Dixon
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii

LIST OF TABLES ............................................................................. v

CHAPTER

I. INTRODUCTION ................................................................. 1
   1.1 Statement of the Problem .............................................. 1
   1.2 Previous Results ........................................................... 3
   1.3 Motivation and Our Solution ........................................ 4
   1.4 Assumptions ................................................................. 8

II. GENERAL LINEARITY RESULTS ................................. 13
   2.1 Introduction ................................................................. 13
   2.2 Linearity Results ........................................................... 15

III. ESTIMATION OF THE REGRESSION PARAMETER $\beta$ .... 37
   3.1 Introduction ................................................................. 37
   3.2 Estimation of $\beta$ in the Intercept Model ................. 42
      3.2.1 Linearity ................................................................. 42
      3.2.2 Estimation ................................................................. 46
   3.3 Estimation of the Intercept ............................................ 48
      3.3.1 Linearity ................................................................. 48
Table of Contents—Continued

CHAPTER

3.3.2 Estimation ......................................................... 56

3.4 Estimation of \( \beta \) in the No-intercept Model ............... 61

3.5 Estimation of the Parameter \( \gamma_L \) ......................... 65

IV. ESTIMATION OF THE VARIANCE PARAMETER \( \theta \) ..... 72

4.1 Introduction .......................................................... 72

4.2 Linearity .............................................................. 78

4.3 Adjustment of the Residuals .................................... 83

4.4 Estimation of \( \theta \) ............................................. 95

4.5 Estimation of the Parameter \( \gamma_s \) ......................... 106

V. IMPLEMENTATION AND MONTE CARLO ................. 111

5.1 Implementation .................................................... 111

5.2 Monte Carlo ....................................................... 112

VI. SUMMARY AND SUGGESTIONS FOR FUTURE RESEARCH ........................................ 118

REFERENCES .......................................................... 120

iv
LIST OF TABLES

1. Monte Carlo MSE Ratio With Respect to the Optimal R-Estimate for Estimates of $\beta_1$ ........................................ 115
2. Monte Carlo Simulation Results for the Rank-based Estimate of $\theta$ .............................................................. 116
3. Monte Carlo Simulation Results for the Adjusted Rank-Based Estimate of $\beta_1$ ....................................................... 117

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
This dissertation is concerned with robust, rank based procedures in the heteroscedastic linear model. A frequent occurrence of such models is when the underlying scale (noise) is a function of the response (signal). For instance, the variance may be an increasing function of the mean of the response. Such models play havoc with the usual inference for linear models since homoscedasticity is a crucial assumption behind either classical or robust analysis of linear models. In this thesis we consider modeling this heteroscedasticity and robust estimation procedures of it. Using these scale estimation procedures, we obtain a complete robust analysis of the underlying linear model including estimation and confidence procedures for the regression coefficients. We obtain asymptotic distribution theory for both our scale and linear model procedures. In order to compute our analysis, we have developed the necessary algorithms. We also explore the small sample properties of our analysis with a Monte Carlo study.

We consider the heteroscedastic linear model

\begin{equation}
    y_i = x_i' \beta + \sigma_i e_i, \quad i = 1, \cdots, N,
\end{equation}
where \( \{x_i\} \) are \( p \times 1 \) design constants, \( \beta \) is a \( p \times 1 \) regression parameter, \( \{e_i\} \) are independent and identically distributed with mean zero, finite variance and unknown symmetric distribution function \( F \), and \( \{\sigma_i\} \) are scaling constants which express possible heteroscedasticity. It will be assumed the scaling constants are form

\[
\sigma_i = \exp[h(x'_i\beta)\theta],
\]

where \( \theta \) is an unknown parameter and \( h \) is smooth and known. The two major problems considered are rank-based estimation of the regression parameter and variance parameter.

The procedures we use in the problem of regression parameter estimation are primarily generalizations of the homoscedastic \( R \)-estimation procedures to the heteroscedastic case. The methodology behind the estimation of \( \theta \) is not such a generalization. The procedure we develop for rank-based estimation of \( \theta \) is motivated by Hajek and Sidak (1967). They examine the use of simple linear rank statistics in testing homoscedasticity versus heteroscedasticity. Our approach to the estimation problem is essentially an extension of their results. Before discussing any specific estimation procedures, we will briefly discuss the iterative scheme used in estimation in the heteroscedastic model. After an initial estimate of \( \beta \) is obtained, an estimate of \( \theta \) is determined which allows us to estimate the weights, and hence obtain an
estimate of $\beta$. Next, a new estimate of $\theta$ is obtained using this new estimate of the regression coefficients, and the process is repeated.

1.2 Previous Results

Generalized weighted least squares is probably the most common method of analyzing the heteroscedastic model. Box and Hill (1974) and Jobson and Fuller (1980) both suggest a type of generalized weighted least squares based on the assumption of normal errors. Their methods use the same iterative scheme discussed in Section 1.1. Carroll and Ruppert (1982b) and Jobson and Fuller (1980) develop the asymptotic theory which establishes asymptotic equivalence of these generalized weighted least squares estimates and the "optimal" weighted least squares estimates for the true $\{\sigma_i\}$. Under the normality assumption the maximum likelihood estimate of $\theta$ has an unbounded influence function and may lack robustness against departures from normality.

In a thorough investigation of the heteroscedastic problem, Carroll and Ruppert (1982a) propose a pseudolikelihood estimate of $\theta$ based on an initial $\sqrt{N}$-consistent estimate of $\beta$. As they show, under reasonable conditions both ordinary least squares estimates and ordinary $M$-estimates satisfy this condition. They then suggest solving

$$\sum_{i=1}^{N} \chi \left( \frac{y_i - x_i' \hat{\beta}}{e^{\theta h(x_i' \hat{\beta})}} \right) h(x_i' \hat{\beta}) = 0,$$
to obtain $\hat{\theta}$, where $\hat{\beta}$ is an initial estimate and $\chi$ is an even function. This generalizes Huber's proposal $2$ estimate for the homoscedastic case. Using their linearity results they prove that their estimate is a $\sqrt{N}$-consistent estimate of $\theta$, but do not establish the asymptotic distribution of their estimate of $\theta$ as we do. Carroll and Ruppert also consider weighted $M$-estimation of the regression parameter $\beta$. Using their asymptotic linearity results for weighted $M$-statistic, they prove that the $M$-estimate of $\beta$ based on estimated weights is asymptotically equivalent to the "optimal" weighted $M$-estimate for the true $\{\sigma_i\}$. The estimated weights are formed with the initial $\sqrt{N}$-consistent estimates of $(\theta, \beta)$. However, their estimate of $\theta$ may not have bounded influence; see Giltinan, Carroll and Ruppert (1986).

Davidian and Carroll (1987) give an overview of methods of variance function estimation. They develop general asymptotic theory for variance function estimation and emphasize the importance of robustness in estimation of the variance parameter. They argue that

"robustness issues are even more important here [in variance function estimation] than in the estimation of a regression function for the mean. The loss of efficiency of the standard method away from the normal distribution is much more rapid than in the regression problem (Davidian and Carroll, 1987, p. 1079)."

1.3 Motivation and Our Solution

The motivation of our rank-based estimate of $\theta$ is based on statis-
tics developed in Hajek and Sidak (1967) for testing homoscedasticity versus heteroscedasticity. Their methodology suggests use of a statistic of the form

\[ S^*(\theta_0) = N^{-1/2} \sum_{i=1}^{N} (h(x'_i \hat{\beta}) - \bar{h}^*) a_s \left( R \left( \frac{y_i - x'_i \hat{\beta}}{\phi_{\theta_0} h(x'_i \hat{\beta})} \right) \right), \]

where \( \hat{\beta} \) is any \( \sqrt{N} \)-consistent estimate of \( \beta \) and \( \bar{h}^* = N^{-1} \sum_{i=1}^{N} h(x'_i \hat{\beta}) \). Further, the scores \( a_s(\cdot) \) are generated by a square integrable function \( \phi_s \), by \( a_s(i) = \phi_s \left( \frac{i}{N+1} \right) \). Under the hypothesis \( H_0: \theta = \theta_0 \) the expectation of the test statistic is zero. This suggests inverting the test statistic by solving \( S^*(\theta) = 0 \) to obtain \( \hat{\theta}^* \). In Sections 4.1 and 4.3 we discuss the assumptions and modifications necessary to obtain a more usable scale rank statistic which is monotone in \( \theta \).

In Section 4.4 we define and discuss the Hodges-Lehmann (1963) type estimate of \( \theta \) which is based on the monotone scale rank statistic. We show that this estimate is \( \sqrt{N} \)-consistent when it is based on any \( \sqrt{N} \)-consistent estimate of \( \beta \). In addition, the asymptotic distribution of this estimate is established when the estimated \( \beta \) is the "adjusted" rank estimate of \( \beta \), discussed later in this section. Finally, in Section 4.5 we discuss estimation of the asymptotic variance of this estimate of \( \theta \). The results we establish are founded on asymptotic linearity of the rank based scale process. This linearity result is the rank statistic analogue to that of the pseudolikelihood statistic established by Carroll and Ruppert (1982a). We develop our linearity result in Section 4.2 using Theorem 7.1 of Carroll and Ruppert (1982a) and the general asymptotic results developed in our Chapter II.
The weighted $R$-estimate of $\beta$ that we use is a generalization of the estimate of Jureckováš (1971a) to the heteroscedastic model. Using the estimated weights $\hat{\sigma}_i = \exp[h(x_i^T\hat{\beta}_0)\hat{\theta}]$ where $(\hat{\theta}, \hat{\beta}_0)$ are initial $\sqrt{N}$-consistent estimates of $(\theta, \beta)$, this estimate is found by solving

$$N^{-1/2} \sum_{i=1}^{N} \left( \frac{x_i}{\hat{\sigma}_i} - \left( \frac{\bar{x}}{\hat{\sigma}} \right) \right) a_L \left( R \left( \frac{y_i - x_i^T\beta}{\hat{\sigma}_i} \right) \right) = 0,$$

to obtain $\hat{\beta}_R$, where

$$\left( \frac{\bar{x}}{\hat{\sigma}} \right) = N^{-1} \sum_{i=1}^{N} \frac{x_i}{\hat{\sigma}_i}.$$

The scores $a_L(\cdot)$ are generated by a non-decreasing, square integrable, odd about $\frac{1}{2}$ function $\phi_L$, by $a_L(i) = \phi_L \left( \frac{i}{N+1} \right)$.

In Section 3.2 we establish asymptotic linearity of the weighted regression statistic and discuss estimation of $\hat{\beta}_R$. We also prove asymptotic equivalence of the estimates based on the true and estimated $\{\sigma_i\}$. The methodology we use to establish this equivalence is similar to that of Akritas (1990), however Akritas considers a random coefficient regression model. Furthermore, he does not use a variance function that depends on the mean; his scaling constants are of the form $\sigma_i = \exp[\theta^T Z_i]$ where $\theta$ is an $r \times 1$ variance function parameter and $Z_i$ is a random vector independent of the errors $e_i$. Although we cannot directly apply the conclusions of his paper, we use a similar methodology in Chapter II to develop general asymptotic theory. Using these results and the linearity results of Carroll and Ruppert (1982a) we establish asymptotic linearity of the rank-based weighted regression statistic.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Under the assumption that the errors are symmetrically distributed about zero, the regression plane for the adjusted variables \( \left( \frac{y_i}{\sigma_i}, \frac{x_i}{\sigma_i} \right) \) does not intersect the origin. However, we cannot force a no-intercept model with rank-based estimation, thus we are actually fitting the model

\[
\frac{y_i}{\sigma_i} = \mu + \frac{x_i}{\sigma_i} \beta + e_i
\]

or equivalently

\[
y_i = \sigma_i \mu + x_i \beta + \sigma_i e_i.
\]

Clearly under this model the mean response is a function of \( \sigma_i \); no preliminary estimate of \( \beta \) could be determined. This is not a problem since in our model we assume \( \mu = 0 \). However, since we cannot force a no-intercept model, we fit an intercept model and make necessary adjustments later.

Using a generalization of the signed-rank statistic suggests solving

\[
N^{-1/2} \sum_{i=1}^{N} a_L^+ \left( R_i \left| \frac{y_i - x_i \hat{\beta}_R}{\hat{\sigma}_i} - \mu \right| \right) sgn \left( \frac{y_i - x_i \hat{\beta}_R}{\hat{\sigma}_i} - \mu \right) = 0,
\]

to obtain \( \hat{\mu}_R \). The scores \( a_L^+ \) are defined by \( a_L^+(i) = \phi_L \left( \frac{\frac{i}{N+1} + 1}{2} \right) \).

In Section 3.3 we establish asymptotic linearity results of the weighted signed-rank statistic. This result is a generalization of the result of Tardiff (1985). We also prove that the estimates of intercept based on the true and estimated \( \{\sigma_i\} \) are asymptotically equivalent. Now that we have fit the intercept model
we form the new estimate of $\beta$ by projecting the fitted values (from the intercept model) to the space spanned by the matrix $\hat{D}$, where $\hat{D} = \left[ \frac{x_{ij}}{\hat{\sigma}_i} \right]$, and solving for the estimate $\hat{\beta}_R$. The estimate and its asymptotic distribution are developed in Section 3.4. In Section 3.5 we discuss estimation of the parameter $\gamma_L$.

In Chapter V an algorithm for implementing the procedure is discussed, and the proposed methods are evaluated using Monte Carlo techniques. Finally, in Chapter VI we discuss possibilities for further research.

1.4 Assumptions

As we pointed out in Section 1.3, the development of asymptotic linearity results for the regression and scale rank statistics are based on Theorem 7.1 of Carroll and Ruppert (1982a) and our Theorem 2.1. Application of Carroll and Ruppert’s Theorem 7.1 to the location problem requires assumptions B1 through B5, F1 through F3, L1 and L2. Assumptions B1 through B5, F1 through F3, S1 and S2 are needed to apply their theorem to the scale process. Development of the asymptotic distribution of the estimate of $\theta$ is dependent upon Assumptions B6, B7 and B8.

First, we consider the “adjusted” design matrix

$$D = [x_{ij}/\sigma_i].$$

We will assume the matrix $[1_N : D]$ is of full rank.
Assumption B1.

\[ N^{-1}[1_N : D]'[1_N : D] = \begin{bmatrix} 1_N \\ \tilde{d} \end{bmatrix} \rightarrow \lambda \text{ where } \tilde{d} = N^{-1}D'1_N, \]

Partition \( \lambda \) compatibly:

\[ \lambda = \begin{bmatrix} 1_N \\ \mu_d \\ \Sigma \end{bmatrix}. \]

Aubuchon (1982) proves that Assumption B1 implies \( N^{-1}D'(I - N^{-1}1_N1_N')D \rightarrow S_1 \), a positive definite matrix. Thus, we will not state this as a separate assumption.

The following assumptions are on the distribution of the errors. The errors \( e_1, \ldots, e_N \) are independently, identically distributed with distribution function \( F \) and density \( f \).

**Assumption F1.** \( F(\cdot) \) has absolutely continuous density \( f(\cdot) \) that is bounded; \( f \) has finite variance,

\[ I(f) = \int_{-\infty}^{\infty} \left[ \frac{f'(y)}{f(y)} \right]^2 f(y) dy < \infty \]

and

\[ I_1(f) = \int_{-\infty}^{\infty} \left[ -1 - y \frac{f'(y)}{f(y)} \right]^2 f(y) dy < \infty. \]

**Assumption F2.** \( f(\cdot) \) is symmetric about 0; that is \( f(y) = f(-y) \).

**Assumption F3.**

\[ \sup_{-\infty < x < \infty} \left| x \frac{f(x)}{f(x)} \right| < \infty. \]

The next four assumptions stated are identical to those of Carroll and Ruppert (1982a).
Assumption B2.

\[ \lim_{N \to \infty} \sup_{1 \leq i \leq N} (\|x_i\| + |h(x'_i \beta)|) N^{-1/2} = 0. \]

Assumption B3.

\[ \sup_{N} \{N^{-1} \sum_{i=1}^{N} (\|x_i\|^2 + |h(x'_i \beta)|^2)\} < \infty. \]

Assumption B4. The \( \sigma_i \) are bounded away from \( 0 \).

Assumption B5. On an open interval \( I \) (possibly infinite) containing all the \( \{x'_i \beta\} \), \( h \) is Lipschitz continuous.

The following three assumptions are necessary for developing the asymptotic distribution of our estimate of \( \theta \).

Assumption B6. On an open interval \( I \) (possibly infinite) containing all the \( \{x'_i \beta\} \), \( h \) is differentiable.

Assumption B7.

\[ N^{-1} \sum_{i=1}^{N} (h(x'_i \beta) - \bar{h}) h'(x'_i \beta) x_i \rightarrow L. \]

Assumption B8. On an open interval \( I \) (possibly infinite) containing all the \( \{x'_i \beta\} \), \( h' \) is uniformly continuous.

The following two assumptions are necessary for the location problem. Assumption L1 is a typical assumption used to establish linearity in the homoscedastic case. Assumption L2 is needed to extend the result to the
heteroscedastic case. We should note that we are using Assumption L2 instead of the complicated smoothness conditions of Carroll and Ruppert (their Assumptions B6 - B8).

**Assumption L1.** \( \phi_L \) is non-decreasing, odd about 1/2 and square integrable.

**Assumption L2.** \( \phi_L \) is bounded with a bounded continuous derivative.

Score generating functions that are square integrable can always be standardized without affecting the properties of the resulting test statistics. Thus, without loss of generality

\[
\int_0^1 \phi_L(u)du = 0, \quad \int_0^1 \phi_L^2(u)du = 1 \quad \text{and} \quad \sum_{i=1}^N a_L(i) = 0.
\]

Define

\[
\phi(u, f) = \frac{-f'(F^{-1}(u))}{f(F^{-1}(u))} \quad \text{and} \quad \gamma_L = \int_0^1 \phi_L(u)\phi(u, f)du.
\]

Assumptions L1, F2 and F1 imply \( \gamma_L > 0 \).

Assumptions S1 and S2 are needed for the scale rank process results. We are using Assumption S2 instead of the complicated smoothness conditions of Carroll and Ruppert (their Assumptions B7 and B8 on \( \chi \)).

**Assumption S1.** \( \phi_s \) is even about 1/2, non-decreasing on \([\frac{1}{2}, 1]\) and square integrable.

**Assumption S2.** \( \phi_s \) is bounded with a bounded continuous derivative.

As with \( \phi_L \), we will standardize \( \phi_s \) so that

\[
\int_0^1 \phi_s(u)du = 0, \quad \int_0^1 \phi_s^2(u)du = 1 \quad \text{and} \quad \sum_{i=1}^N a_s(i) = 0.
\]
Define

\[ \phi_1(u, f) = -1 - F^{-1}(u) \frac{f'(F^{-1}(u))}{f(F^{-1}(u))} \]

and \( \gamma_\alpha = \int_0^1 \phi_\alpha(u) \phi_1(u, f) du. \)

Assumptions S1, F1 and F2 imply \( \gamma_\alpha > 0. \)
CHAPTER II

GENERAL LINEARITY RESULTS

2.1 Introduction

In this chapter we shall establish results needed in Chapters III and IV. The main Theorem (2.1) is applicable to the location and scale rank processes. Basically this theorem acts as a bridge; we will use this theorem and the asymptotic linearity results of Carroll and Ruppert (1982a) to establish asymptotic linearity of the location and scale rank processes. Additionally, some of the lemmas of this chapter will be used in developing asymptotic linearity of the signed-rank statistic.

In order to understand the motivation for the results of the section, consider the scores

\[ a^*(i) = \phi(F(e_{(i)})) \]

and

\[ a(i) = \phi\left(\frac{i}{N+1}\right), \]

where \( e_{(i)} \) denotes the ith ordered residual among \( e_1, \ldots, e_N \). Under conditions on the \( \{c_i\} \), Theorem V1.5a of Hajek and Sidak (1967) implies

\[ \sum_{i=1}^{N}(c_i - \bar{c})\phi(F(e_i)) - \sum_{i=1}^{N}(c_i - \bar{c})\phi\left(\frac{R(e_i)}{N+1}\right) \xrightarrow{P} 0, \]

where \( R(e_i) \) is the rank of \( e_i \) among \( e_1, \ldots, e_N \).  

13
Jureckova (1969) used this result in establishing asymptotic linearity of the regression rank statistic for the homoscedastic case. She applied a contiguity argument to prove

\[
\sum_{i=1}^{N} (c_i - \bar{c}) \phi(F(e_i - \Delta c_i)) - \sum_{i=1}^{N} (c_i - \bar{c}) \phi \left( \frac{R(e_i - \Delta c_i)}{N + 1} \right) \overset{P}{\rightarrow} 0
\]

for fixed \( \Delta \). The remainder of her proof is based on monotonicity of the regression rank statistic in \( \Delta \); clearly this method does not extend to the heteroscedastic model.

The main theorem of this chapter establishes that under the heteroscedastic model, the difference of the rank statistic based on the scores \( a_N(R(\cdot)) \) and the statistic based on the scores \( \phi(F(\cdot)) \) converges in probability to zero uniformly on a compact set.

Many of the results in this chapter are similar to those of Akritas (1990). He developed a similar result for the location rank process. However, Akritas considered a random coefficients regression model with \( \sigma_i = \exp[\theta' Z_i] \), where \( \theta \) is an \( r \times 1 \) vector of unknown coefficients and \( Z_i \) are observable \( r \)-dimensional random vectors independent of the errors \( e_i \). Thus, his results are neither applicable to our location problem nor to our scale problem.

Since the results of this chapter are very general, the notation is quite complex. We postpone our discussion of the practical interpretation of the processes until Chapters III and IV, where we will discuss the interpretations as they relate to the specific problems.
2.2 Linearity Results

We shall use notation similar to Carroll and Ruppert (1982a). For $\Delta_1$ and $\Delta_3$ in $R^p$, $\Delta_2$, $\Delta_4$ and $\Delta_5$ in $R^1$, and $\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5)$, define

$$h_i(\Delta) = h(x_i^t \beta + N^{-1/2} x_i^t \Delta_3)$$

$$\alpha_i^{(2)}(\Delta) = \exp[-h_i(\Delta)N^{-1/2} + (h_i(0) - h_i(\Delta))\theta] - 1$$

and

$$\alpha_i^{(1)}(\Delta) = N^{-1/2} \frac{x_i}{\sigma_i} \Delta_1 + N^{-1/2} \frac{1}{\sigma_i} \Delta_4 + N^{-1/2} \Delta_5(1 + \alpha_i^{(2)}(\Delta))^{-1}.$$  

Next, define the processes

$$F_N(x : \Delta) = N^{-1} \sum_{i=1}^N I[(e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta)) \leq x],$$

and

$$F(x : \Delta) = N^{-1} \sum_{i=1}^N F[x(1 + \alpha_i^{(2)}(\Delta))^{-1} + \alpha_i^{(1)}(\Delta)].$$

In what follows we will also use the notation

$$g_i(e_i : \Delta) = (e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta))$$

and

$$G_i(x : \Delta) = x(1 + \alpha_i^{(2)}(\Delta))^{-1} + \alpha_i^{(1)}(\Delta).$$
Next define

\[ U_N(\Delta : F) = N^{-1/2} \sum_{i=1}^{N} (c_i + \alpha_i^{(3)}(\Delta)) \phi(F[(e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta))]) \]

and

\[ U_N(\Delta : F_N) = N^{-1/2} \sum_{i=1}^{N} (c_i + \alpha_i^{(3)}(\Delta)) \phi(F_N[(e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta)) : \Delta]) \].

In this section we prove a very general result applicable to the location and scale problems. Thus, \( c_i \) and \( \alpha_i^{(3)}(\Delta) \) will not be explicitly defined. The proofs will be derived under the following assumptions on \( c_i \) and \( \alpha_i^{(3)}(\Delta) \).

**T1:** \[ \sum_{i=1}^{N} c_{i\ell} = 0, \quad \ell = 1, \ldots, R. \]

**T2:** \[ \sup_{N} \{N^{-1} \sum_{i=1}^{N} \|c_i\|^2\} < \infty \]

and

\[ \lim_{N \to \infty} \sup_{i \leq N} (N^{-1/2} \|c_i\|) = 0. \]

**T3:** \( \alpha_i^{(3)}(0) = 0, \quad i = 1, \ldots, N \) and \( \ell = 1, \ldots, R. \)

**T4:** \( \alpha_i^{(3)}(\Delta) = o(c_i\ell) \) as \( N \to \infty \), for \( \|\Delta\| \leq M, \quad i = 1, \ldots, N \) and \( \ell = 1, \ldots, R. \)

We will also assume:

**T5:** \( \phi \) is bounded on \([0,1]\), with a bounded continuous derivative.
Assumptions B2 through B5, F1 and F3 (Chapter I) are also required.

The following lemma gives properties of $\alpha_i^{(1)}(\Delta)$ and $\alpha_i^{(2)}(\Delta)$ that greatly simplify the proofs of subsequent lemmas and theorems.

Lemma 2.1. Under assumptions B2 through B5

(i) $\alpha_i^{(f)}(0) = 0, \quad \ell = 1, 2 \quad \text{and} \quad i = 1, \cdots, N;$

For each compact set $S$ there exists $K$ such that

(ii) $\|\alpha_i^{(f)}(\Delta) - \alpha_i^{(f)}(\Delta')\| \leq K_i\|\Delta - \Delta'\|K, \quad \ell = 1, 2, \quad \text{and}$

(iii) $\|(1 + \alpha_i^{(2)}(\Delta))^{-1} - (1 + \alpha_i^{(2)}(\Delta'))^{-1}\| \leq K_i\|\Delta - \Delta'\|K, \quad \text{for all} \quad \Delta \quad \text{and} \quad \Delta', \quad i = 1, \cdots, N, \quad \text{where} \quad \{K_i\} \quad \text{is a sequence of}$

positive constants such that

$$\lim_{N \to \infty} \sup_{i \leq N} K_i = 0 \quad \text{and} \quad \sup_{N} \sum_{i=1}^{N} (K_i^2 + N^{-1/2} K_i) = c_1 < \infty.$$ 

Proof. We should note that our $\alpha_i^{(1)}(\Delta)$ and $\alpha_i^{(2)}(\Delta)$ are similar to those used by Carroll and Ruppert (1982a). Since these properties hold for their $\alpha_i^{(1)}(\Delta)$ and $\alpha_i^{(2)}(\Delta)$ we will not prove results (i) and (ii). It can be shown that

$$K_i = N^{-1/2} \left\{ \|x_i\| + \frac{\|x_i\|}{\sigma_i} + |h_i(0)| + \frac{1}{\sigma_i} + 1 \right\}.$$ 

Assumptions B2, B3 and B4 imply the conditions on the $\{K_i\}$ that are stated in this lemma, are satisfied. We prove the third part of the lemma by first considering $\alpha_i^{(2)}(\Delta).$
Define $B_i(\Delta) = -h_i(\Delta)\Delta_2 N^{-1/2} + (h_i(0) - h_i(\Delta))\theta$. The result in part (ii) regarding $\alpha_i^{(2)}(\Delta)$ is proven using the following Taylor expansion,

$$|\alpha_i^{(2)}(\Delta') - \alpha_i^{(2)}(\Delta)| = |[B_i(\Delta) - B_i(\Delta')]e^{A_i(\Delta)}|,$$

where $|A_i(\Delta)| \leq |B_i(\Delta) - B_i(\Delta')|$. Clearly, Assumptions B2 and B5 imply $|B_i(\Delta) - B_i(\Delta')| \to 0$, uniformly in $i = 1, \ldots, N$ and $\|\Delta\| \leq M$. Note that by a Taylor expansion,

$$|(1 + \alpha_i^{(2)}(\Delta))^{-1} - (1 + \alpha_i^{(2)}(\Delta'))^{-1}| = |[B_i(\Delta') - B_i(\Delta)]e^{q_i(\Delta)}|,$$

where $|q_i(\Delta)| \leq |B_i(\Delta) - B_i(\Delta')|$. Thus proving that $|\alpha_i^{(2)}(\Delta) - \alpha_i^{(2)}(\Delta')| \leq K_i \|\Delta - \Delta'\|K$ implies

$$|(1 + \alpha_i^{(2)}(\Delta))^{-1} - (1 + \alpha_i^{(2)}(\Delta'))^{-1}| \leq K_i \|\Delta - \Delta'\|K.$$

**Lemma 2.2.** Under assumptions B2 through B5 and F1, for all $M > 0$ and $\varepsilon > 0$,

$$\lim_{N \to \infty} P\{ \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |g_i(e_i : \Delta) - e_i| \geq \varepsilon \} = 0.$$

**Proof.**

$$g_i(e_i : \Delta) - e_i = e_i \alpha_i^{(2)}(\Delta) - \alpha_i^{(1)}(\Delta)(1 + \alpha_i^{(2)}(\Delta))$$
By Lemma 2.1, this lemma holds if we show:

$$\sup_{\|\Delta\| \leq M} \sup_{i \leq N} |e_i \alpha_i^{(2)}(\Delta)| = o_p(1).$$

Using part ii of Lemma 2.1,

$$\sup_{\|\Delta\| \leq M} \sup_{i \leq N} |e_i \alpha_i^{(2)}(\Delta)| \leq \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |e_i| K_i \|\Delta\| K \leq \sup_{i \leq N} K_i |e_i|M K.$$ 

Since $E[|e|] \leq (E[e^2])^{1/2}$, the fact $F$ has finite second moment implies $E[|e|] < \infty$. The expected value and variance of $K_i|e_i|$ converge to zero, $i = 1, \cdots, N$. Thus $K_i|e_i|$ converges in probability to zero uniformly in $i = 1, \cdots, N$, and the proof is complete.

Lemma 2.3. Under assumptions B2 through B5, and F1, for all $M > 0$ and $\varepsilon > 0$,

$$\lim_{N \to \infty} \mathbb{P}\left\{ \sup_{\|\Delta\| \leq M} \sup_{i,j \leq N} |G_j(g_i(\epsilon_i : \Delta) - \Delta) - g_i(\epsilon_i : \Delta)| \geq \varepsilon \right\} = 0.$$ 

Proof.

$$|G_j(g_i(\epsilon_i : \Delta) - \Delta) - g_i(\epsilon_i : \Delta)|$$

$$= |g_i(\epsilon_i : \Delta)(1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta) - g_i(\epsilon_i : \Delta)|$$

$$\leq |g_i(\epsilon_i : \Delta) - \epsilon_i|(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1|$$

$$+ |\epsilon_i|(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1| + |\alpha_j^{(1)}(\Delta)|$$
Using Lemma 2.1 and an argument as in the proof of Lemma 2.2, the sup of the second component of the sum converges in probability to zero. Convergence in probability to zero of the sup of first and third components of the sum is implied by Lemma 2.1 and Lemma 2.2.

Lemma 2.4. Under Assumptions B2 through B5, F1, and F3, for all \( M > 0 \) there exists a \( c > 0 \) such that

\[
\lim_{N \to \infty} P\{\sup \{N^{1/2}|F(x: \Delta) - F(x)| : -\infty < x < \infty, \|\Delta\| \leq M\} \leq c\} = 1.
\]

Proof.

\[
N^{1/2}|F(x: \Delta) - F(x)| = N^{-1/2} \sum_{j=1}^{N} \{F(x(1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta)) - F(x)\}|
\]

\[
= N^{-1/2} \sum_{j=1}^{N} f(A_j)\{x[(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1] + \alpha_j^{(1)}(\Delta)\}|
\]

\[
\leq N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1| f(A_j)|x - A_j|
\]

\[
+ N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1| f(A_j)A_j|
\]

\[
+ N^{-1/2} \sum_{j=1}^{N} |\alpha_j^{(1)}(\Delta)| |f(A_j)|,
\]

where \( A_j \) is between \([x]\) and \([x(1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta)]\). Convergence in probability to zero of the sup of the first component of the sum is implied
by \( f \) bounded and Lemma 2.1. Using boundedness of \( \sup_{-\infty < x < \infty} |xf(x)| \) and Lemma 2.1, we have the \( \sup \) of the second and third components of the sum bounded in probability. Thus the Lemma is proven.

The following corollary is an obvious consequence of this lemma.

Corollary. Under the assumptions of the lemma, for all \( M > 0 \) and \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P\left\{ \sup_{\|\Delta\| \leq M} \sup_{1 \leq n \leq N} |F(g_i(\Delta) : \Delta) - F(g_i(\Delta))| \geq \varepsilon \right\} = 0.
\]

Since \( F \) is continuous we have the result (Theorem 1) of Shorack (1973), which we shall state as a lemma.

Lemma 2.5. (Shorack) There exists a Brownian Bridge \( B^0 \) such that, perhaps on a different probability space, for all \( \chi > 0 \),

\[
\lim_{N \to \infty} P\left\{ \sup_{-\infty < x < \infty} |N^{1/2}[F_N(x) : \Delta) - F(x : \Delta)] - B^0(F(x))| \geq \varepsilon \right\} = 0
\]

for each fixed \( \Delta \). It is useful to note that by definition, the Brownian Bridge, \( B^0 \), is uniformly continuous on \([0, 1]\).

This brings us to the main theorem of this chapter.

Theorem 2.1. Under assumptions \( B2 \) through \( B5 \), \( F1 \), \( F3 \), and \( T1 \) through \( T5 \), for all \( M > 0 \) and \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P\left\{ \sup_{\|\Delta\| \leq M} \|U_N(\Delta) : F_N) - U_N(\Delta) : F\| \geq \varepsilon \right\} = 0.
\]
Proof. Let $U_{N,\ell}(\Delta : \cdot)$ denote the $\ell$th component of $U_N(\Delta : \cdot)$. The theorem will be established by proving

$$\sup_{\|\Delta\| \leq M} |U_{N,\ell}(\Delta : F) - U_{N,\ell}(\Delta : F)| = o_p(1), \text{ for } \ell = 1, \cdots, R.$$ 

From its definition

$$\sup_{\|\Delta\| \leq M} |U_{N,\ell}(\Delta : F) - U_{N,\ell}(\Delta : F)|$$

$$= \sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{i=1}^N (c_{i\ell} + \alpha_{i\ell}^{(3)}(\Delta))[\phi(F_N(g_i(e_i : \Delta)) - \phi(F(g_i(e_i : \Delta)))]|$$

$$\leq \sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{i=1}^N c_{i\ell}[\phi(F_N(g_i(e_i : \Delta)) - \phi(F(g_i(e_i : \Delta)))]|$$

$$+ \sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{i=1}^N \alpha_{i\ell}^{(3)}(\Delta)[\phi(F_N(g_i(e_i : \Delta)) - \phi(F(g_i(e_i : \Delta)))]|.$$ 

Since the second term of the sum is of smaller order of magnitude than the first term (Assumption T4), it is sufficient to show that the first term converges in probability to zero.

Consider the first term of the sum:

$$\sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{i=1}^N c_{i\ell}[\phi(F_N(g_i(e_i : \Delta)) - \phi(F(g_i(e_i : \Delta)))]|$$

$$\leq \sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{i=1}^N c_{i\ell}[\phi(F_N(g_i(e_i : \Delta)) - \phi(F(g_i(e_i : \Delta)))]|$$

$$+ \sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{i=1}^N \alpha_{i\ell}^{(3)}(\Delta)[\phi(F_N(g_i(e_i : \Delta)) - \phi(F(g_i(e_i : \Delta)))]|$$

$$= Q_1 + Q_2,$$ say.
In what follows we will use the notation $v_i = g_i(e_i : \Delta)$. Using the definition of $F(\cdot : \Delta)$ and the mean value theorem:

$$Q_2 = \sup_{\|\Delta\| \leq M} N^{-1/2} \left| \sum_{i=1}^{N} c_i \phi'(F(e_i)) \right| + N^{-1} \sum_{j=1}^{N} \left| F(v_i (1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta)) - F(v_i) \right|
$$

$$+ \sup_{\|\Delta\| \leq M} N^{-1} \left| \sum_{i=1}^{N} c_i [\phi'(F(e_i)) - \phi'(Z_i)] \right| + N^{-1/2} \sum_{j=1}^{N} \left| F(v_i (1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta)) - F(v_i) \right|.$$

where $Z_i$ is between

$$[F(v_i (1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta))]$$

and $[F(v_i)].$

Now,

$$|Z_i - F(e_i)| \leq |Z_i - F(v_i)| + |F(v_i) - F(e_i)|$$

$$\leq |F(v_i (1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta)) - F(v_i)| + |F(v_i) - F(e_i)|.$$

Lemma 2.2, Lemma 2.3 and uniform continuity of $F$ imply $|Z_i - F(e_i)| \overset{P}{\to} 0$ uniformly in $i = 1, \cdots, N$ and $\|\Delta\| \leq M$. Thus, uniform continuity of $\phi'$ implies

$$\sup_{\|\Delta\| \leq M} |\phi'(F(e_i)) - \phi'(Z_i)| = o_p(1), \quad \text{uniformly in } i = 1, \cdots, N.$$
Now $N^{-1} \sum_{i=1}^{N} |c_{it}| < \infty$ and $N^{1/2} [F(\cdot ; \Delta) - F(\cdot)] = O_p(1)$. Thus the second component of $Q_2$ converges in probability to 0. Using the mean value theorem, the first term of $Q_2$ is less than or equal to:

$$
\sup_{\|\Delta\| \leq M} N^{-3/2} \left| \sum_{i=1}^{N} c_{it} \{ \phi'(F(e_i)) \} \sum_{j=1}^{N} \left[ (1 + \alpha^{(2)}_j(\Delta))^{-1} - 1 \right] (v_i - e_i) f(r_{ij}) \right|
$$

$$
+ \sup_{\|\Delta\| \leq M} N^{-1} \left| \sum_{i=1}^{N} c_{it} \phi'(F(e_i)) N^{-1/2} \right|
\cdot \sum_{j=1}^{N} \left[ e_i((1 + \alpha^{(2)}_j(\Delta))^{-1} - 1) + \alpha^{(1)}_j(\Delta) (f(r_{ij}) - f(e_i)) \right]
$$

$$
+ \sup_{\|\Delta\| \leq M} \left| [N^{-1} \sum_{i=1}^{N} c_{it} \phi'(F(e_i)) f(e_i)e_i] \cdot [N^{-1/2} \sum_{j=1}^{N} ((1 + \alpha^{(2)}_j(\Delta))^{-1} - 1)] \right|
$$

$$
+ \sup_{\|\Delta\| \leq M} \left| [N^{-1} \sum_{i=1}^{N} c_{it} \phi'(F(e_i)) f(e_i)] \cdot [N^{-1/2} \sum_{j=1}^{N} \alpha^{(1)}_j(\Delta)] \right|
$$

$$
= Q_{21} + Q_{22} + Q_{23} + Q_{24},
$$

where $|e_i - r_{ij}| \leq |e_i - v_i| + |v_i - e_i (1 + \alpha^{(2)}_j(\Delta))^{-1} + \alpha^{(1)}_j(\Delta)|$. Thus, Lemma 2.2 and Lemma 2.3 imply $|e_i - r_{ij}| \overset{P}{\to} 0$ uniformly in $i = 1, \ldots, N$, $j = 1, \ldots, N$ and $\|\Delta\| \leq M$.

Consider $Q_{21}$:

$$
Q_{21} \leq (N^{-1} \sum_{i=1}^{N} |c_{it}|) \cdot \sup_{i \leq N} |\phi'(F(e_i))| \cdot \sup_{\|\Delta\| \leq M} \sup_{i,j \leq N} |f(r_{ij})|
$$

$$
\cdot \sup_{\|\Delta\| \leq M} N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha^{(2)}_j(\Delta))^{-1} - 1| \cdot \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |v_i - e_i|.
$$

Now, $\phi'$ and $f$ are bounded and $N^{-1} \sum_{i=1}^{N} |c_{it}| < \infty$. Furthermore,
Lemma 2.2 gives \( \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |u_i - e_i| = o_p(1) \) and Lemma 2.1 implies
\[
\sup_{\|\Delta\| \leq M} N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1| < \infty. \text{ Thus } Q_{21} \xrightarrow{P} 0.
\]

Next consider \( Q_{22} \):
\[
Q_{22} \leq (N^{-1} \sum_{i=1}^{N} |c_{it} e_i|) \sup_{i \leq N} |\phi'(F(e_i))| \cdot \sup_{i,j \leq N} |f(r_{ij}) - f(e_i)|
\]
\[
\cdot \sup_{\|\Delta\| \leq M} N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1|
\]
\[
+ (N^{-1} \sum_{i=1}^{N} |c_{it} e_i|) \sup_{i \leq N} |\phi'(F(e_i))| \cdot \sup_{i,j \leq N} |f(r_{ij}) - f(e_i)|
\]
\[
\cdot \sup_{\|\Delta\| \leq M} N^{-1/2} \sum_{j=1}^{N} |\alpha_j^{(1)}(\Delta)|.
\]

Now, since \( |e_i - r_{ij}| \xrightarrow{P} 0 \) uniformly in \( i = 1, \ldots, N, \ j = 1, \ldots, N \) and \( \|\Delta\| \leq M \), and \( f \) is uniformly continuous we have, \( |f(r_{ij}) - f(e_i)| \xrightarrow{P} 0 \) uniformly in \( i = 1, \ldots, N, \ j = 1, \ldots, N \) and \( \|\Delta\| \leq M \). Assumptions F1 and T2 imply \( N^{-1} \sum_{i=1}^{N} |c_{it} e_i| \) is bounded in probability. These facts and the comments in the demonstration of convergence of \( Q_{21} \) imply convergence of \( Q_{22} \).

Next consider \( Q_{23} \):
\[
Q_{23} \leq \sup_{\|\Delta\| \leq M} \left[ N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1| \right] \cdot |N^{-1} \sum_{i=1}^{N} c_{it} \phi'(F(e_i)) f(e_i) e_i|.
\]

The first term of the product is bounded in probability by Lemma 2.1. In order to show convergence in probability to zero of the second term consider
its expectation and variance.

\[ E[N^{-1} \sum_{i=1}^{N} c_{it} \phi'(F(e_i))f(e_i)e_i] = (N^{-1} \sum_{i=1}^{N} c_{it})E[\phi'(F(e))f(e)e] = 0 \]

and

\[ \text{Var}[N^{-1} \sum_{i=1}^{N} c_{it}\phi'(F(e_i))f(e_i)e_i] \leq N^{-1}(N^{-1} \sum_{i=1}^{N} c_{it}^2)E[(\phi'(F(e))f(e)e)^2] \to 0. \]

The facts \( E[\phi'(F(e))f(e)e] < \infty \) and \( E[(\phi'(F(e))f(e)e)^2] < \infty \) follow from \( \phi' \) and \( |ef(e)| \) bounded, and from the Lebesgue dominated convergence theorem. Thus the above results for the expectation and variance follow from Assumptions T1 and T2. Hence, \( N^{-1} |\sum_{i=1}^{N} c_{it} \phi'(F(e_i))f(e_i)e_i| = o_p(1) \), and \( Q_{23} \to 0 \). Finally consider \( Q_{24} \):

\[ Q_{24} \leq \sup_{\|\Delta\| \leq M} |N^{-1/2} \sum_{j=1}^{N} \alpha_j^{(1)}(\Delta)| \cdot |N^{-1} \sum_{i=1}^{N} c_{it}\phi'(F(e_i))f(e_i)| \]

Using Assumptions F1 and a similar argument as in that showing convergence of \( Q_{23} \), we have: \( Q_{24} \to 0 \).

Now, we have shown that all the components of \( Q_2 \) converge in probability to zero. Thus \( Q_2 \to 0 \).
Now consider $Q_1$:

$$Q_1 = \sup_{\|\Delta\| \leq M} \left| N^{-1/2} \sum_{i=1}^{N} c_{it} \{ \phi(F_N(g_i(\epsilon_i : \Delta)) : \Delta) - \phi(F(g_i(\epsilon_i : \Delta)) : \Delta) \} \right|$$

$$\leq \sup_{\|\Delta\| \leq M} \left| N^{-1/2} \sum_{i=1}^{N} c_{it} \{ [\phi(F_N(v_i)) : \Delta) - \phi(F(v_i) : \Delta)] \right|$$

$$- \left[ \phi(F_N(v_i)) - \phi(F(v_i)) \right] |$$

$$+ \sup_{\|\Delta\| \leq M} \left| N^{-1/2} \sum_{i=1}^{N} c_{it} \{ [\phi(F_N(v_i)) : \Delta) - \phi(F(v_i) : \Delta)] \right|$$

$$- [\phi(F_N(v_i)) - \phi(F(v_i))] |$$

$$+ \left[ \phi(F_N(v_i)) - \phi(F(v_i)) \right] |$$

$$\leq (N^{-1} \sum_{i=1}^{N} c_{it}) \left\{ \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left[ [\phi(F_N(v_i)) : \Delta) - \phi(F(v_i) : \Delta)] \right|$$

$$- [\phi(F_N(v_i)) - \phi(F(v_i))] |$$

$$+ \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left[ [\phi(F_N(v_i)) : \Delta) - \phi(F(v_i) : \Delta)] \right|$$

$$- [\phi(F_N(v_i)) - \phi(F(v_i))] |$$

Assumption T2 and Theorem VI.5a of Hajek and Sidak (1967) give convergence in probability to zero of

$$N^{-1/2} \sum_{i=1}^{N} c_{it} \{ [\phi(F_N(v_i)) : \Delta) - \phi(F(v_i) : \Delta)] \}$$

$$- [\phi(F_N(v_i)) - \phi(F(v_i))] |$$

Since $N^{-1} \sum_{i=1}^{N} c_{it} < \infty$, convergence of the sum is implied by convergence of

$$R_1 = \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left[ [\phi(F_N(v_i) : \Delta)) - \phi(F(v_i) : \Delta)] \right|$$

$$- [\phi(F_N(v_i)) - \phi(F(v_i))] |$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
and

\[ R_2 = \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left\{ \phi(F_N(v_i)) - \phi(F(v_i)) - \left[ \phi(F_N(e_i)) - \phi(F(e_i)) \right] \right\}. \]

Thus, the theorem will be proven if we prove \( R_1 \) and \( R_2 \) converge in probability to zero. Next we will show convergence of \( R_1 \) and \( R_2 \) are implied by convergence of \( S_1 \) and \( S_2 \) respectively, where:

\[ S_1 = \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left\{ [F_N(v_i) - F_N(e_i)] - \phi(F_N(v_i)) - \phi(F_N(e_i)) \right\}. \]

\[ S_2 = \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left\{ [F_N(v_i) - F_N(e_i)] - \phi(F_N(v_i)) - \phi(F_N(e_i)) \right\}. \]

First, we will show convergence of \( R_1 \) is implied by convergence of \( S_1 \). Using the mean value theorem:

\[ R_1 = \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left\{ \phi'(B_{1i}) \left[ F_N(v_i) - F_N(v) \right] - \phi'(B_{2i}) \left[ F_N(v_i) - F(v) \right] \right\} \]

\[ \leq \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left\{ [F_N(v_i) - F_N(v)] - [F_N(v_i) - F(v)] \right\} \]

\[ \cdot \sup_{\|\Delta\| \leq M} \sup_{i \leq N} \phi'(B_{1i}) + \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left\{ [F_N(v_i) - F_N(v)] - [F_N(v_i) - F(v)] \right\} \]

\[ \cdot \sup_{\|\Delta\| \leq M} \sup_{i \leq N} \phi'(B_{1i}) - \phi'(B_{2i}) \]

where \( B_{1i} \) is between...
$F_N(v_i : \Delta)$ and $F(v_i : \Delta)$ and $B_{2i}$ is between $F_N(v_i)$ and $F(v_i)$. Thus, we have

$$|B_{1i} - B_{2i}|$$

$$\leq |B_{1i} - F(v_i : \Delta)| + |F(v_i : \Delta) - F(v_i)| + |F(v_i) - B_{2i}|$$

$$\leq |F_N(v_i : \Delta) - F(v_i : \Delta)| + |F(v_i : \Delta) - F(v_i)|$$

$$+ |F(v_i) - F_N(v_i)| \xrightarrow{P} 0 \text{ uniformly in } i = 1, \ldots, N$$

and $\|\Delta\| \leq M$ by Lemma 2.5 and the corollary to Lemma 2.4. Since $\phi'$ is uniformly continuous, $|\phi'(B_{1i}) - \phi'(B_{2i})| \xrightarrow{P} 0$ uniformly in $i = 1, \ldots, N$ and $\|\Delta\| \leq M$. Thus, convergence of the 2nd component of the sum is implied by Lemma 2.4.

Now, $\phi'$ is bounded, and we conclude that we will have convergence in probability to zero of $R_1$ if we have it for $S_1$. Similarly convergence of $R_2$ is implied by convergence of $S_2$. We will prove convergence of $S_1$ and $S_2$ in the following lemmas.

**Lemma 2.6.** Under Assumptions B2 through B5 and F1, for all $M > 0$ and $\varepsilon > 0$,

$$\lim_{N \to \infty} P\left\{ \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} |[F_N(v_i) - F(v_i)] - [F_N(e_i) - F(e_i)]| \geq \varepsilon \right\} = 0.$$
Proof. We have

\[
\sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \|[F_N(v_i) - F(v_i)] - [F_N(e_i) - F(e_i)]\|
\leq \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |N^{1/2}[F_N(v_i) - F(v_i)] - B^\circ(F(v_i))|
+ \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |N^{1/2}[F_N(e_i) - F(e_i)] - B^\circ(F(e_i))|
+ \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |B^0(F(v_i)) - B^0(F(e_i))|
\leq 2|N^{1/2}[F_N(\cdot) - F(\cdot)] - B^\circ(F(\cdot))| + \sup_{\|\Delta\| \leq M} \sup_{i \leq N} |B^0(F(v_i)) - B^0(F(e_i))|.
\]

Lemma 2.5 implies the first component converges in probability to zero. Convergence of the second component is implied by Lemma 2.2 and uniform continuity of the Brownian Bridge.

Next we need to show convergence of

\[
S_1 = \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \|[F_N(v_i : \Delta) - F(v_i : \Delta)] - [F_N(v_i) - F(v_i)]\|.
\]

First we will prove convergence in probability to zero of \( S_1 \) for \( \Delta \) fixed. Since it is needed later, we prove a more general result.

Lemma 2.7. Under assumption \( F1 \), for fixed \( \Delta^1, \Delta^2, i = 1, \ldots, N \) and all \( \varepsilon > 0 \)

\[
\lim_{N \to \infty} P[N^{-1/2} \sum_{i=1}^{N} [I(e_i \leq G_i(\cdot : \Delta^2_i)) - F(G_i(\cdot : \Delta^2_i))]
- \sum_{i=1}^{N} [I(e_i \leq G_i(\cdot : \Delta^1_i)) - F(G_i(\cdot : \Delta^1_i))]] \geq \varepsilon] = 0.
\]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Proof. Let $L_N(\cdot) = L_{2N}(\cdot) - L_{1N}(\cdot)$, where $L_{2N}(\cdot) = N^{-1/2} \sum_{i=1}^{N} [I(e_i \leq G_i(\cdot : \Delta_i^2)) - F(G_i(\cdot : \Delta_i^2))]$ and $L_{1N}(\cdot) = N^{-1/2} \sum_{i=1}^{N} [I(e_i \leq G_i(\cdot : \Delta_i^1)) - F(G_i(\cdot : \Delta_i^1))]$. First, we will show for a fixed $y$ that $L_N(y) \xrightarrow{P} 0$. Note that

$$L_N(y) = N^{-1/2} \sum_{i=1}^{N} \left[ I(e_i \leq G_i(y : \Delta_i^2)) - I(e_i \leq G_i(y : \Delta_i^1)) \right] - [F(G_i(y : \Delta_i^2)) - F(G_i(y : \Delta_i^1))] = \sum_{i=1}^{N} B_i.$$ 

For a fixed $y$, the $B_i$ are independent random variables with $E(B_i) = 0$, $Var(B_i) = N^{-1} P_i(1 - P_i)$ and $P_i = |F(G_i(y : \Delta_i^2)) - F(G_i(y : \Delta_i^1))|$. Using these properties and the fact that $P(|B_i| \leq 2N^{-1/2}) = 1$, Bernstein's inequality (Uspensky, 1937) implies that for all $\varepsilon > 0$,

$$P\left(\left|\sum_{i=1}^{N} B_i\right| > \varepsilon\right) \leq 2 \exp\left\{-\frac{-\varepsilon^2}{2 \sum_{i=1}^{N} Var(B_i) + \frac{4}{3} N^{-1/2} \varepsilon} \right\}.$$ 

Therefore, $L_N(y)$ converges in probability to 0 for a fixed $y$.

Theorem 1 of Shorack (1973) implies $	ilde{L}_{1N}(F^{-1}(\cdot))$ and $	ilde{L}_{2N}(F^{-1}(\cdot))$ are relatively compact, where $	ilde{L}_{iN}$ equals $L_{iN}$ at the jump points and is linear in between the jumps (for $i = 1, 2$). Thus, $	ilde{L}_N = \tilde{L}_{2N} - \tilde{L}_{1N}$ is also relatively compact. It is also clear that $	ilde{L}_N(F^{-1}(t))$ converges weakly for a
fixed \( t \). Now we have

\[
|\tilde{L}_N(F^{-1}(t))| \leq |L_1N(F^{-1}(t)) - \tilde{L}_1N(F^{-1}(t))|
+ |L_2N(F^{-1}(t)) - \tilde{L}_2N(F^{-1}(t))| + |L_N(F^{-1}(t))|.
\]

Since the first and second components converge in probability to 0 by Theorem 1 of Shorack (1973), and since previously in this Lemma we have proven convergence of the third component, \( \tilde{L}_N(F^{-1}(t)) = o_p(1) \).

Hence, we have established that the finite dimensional distributions of the process \( \tilde{L}_N \) converge weakly to 0, and \( \tilde{L}_N \) is relatively compact.

Thus, by Billingsley (1968) \( \tilde{L}_N(\cdot) \) converges weakly to zero. This leads to the desired result since:

\[
\sup_{0 \leq t \leq 1} |L_N(F^{-1}(t))| \leq \sup_{0 \leq t \leq 1} |L_1N(F^{-1}(t)) - \tilde{L}_1N(F^{-1}(t))|
+ \sup_{0 \leq t \leq 1} |L_2N(F^{-1}(t)) - \tilde{L}_2N(F^{-1}(t))|
+ \sup_{0 \leq t \leq 1} |\tilde{L}_N(F^{-1}(t))| = o_p(1).
\]

Corollary 2. Under Assumption F1, for fixed \( \Delta \) and all \( \varepsilon > 0 \).

\[
\lim_{N \to \infty} P\{\sup_{i \leq N} N^{1/2} |[F_N(v_i : \Delta) - F(v_i : \Delta)] - [F_N(v_i) - F(v_i)]| \geq \varepsilon\} = 0.
\]

This is an obvious consequence of Lemma 2.7

\[
\text{with} \quad \Delta^{2i} = \Delta
\]

\[
\text{and} \quad \Delta^{1i} = 0, \text{ for } i = 1 \cdots N.
\]
Lemma 2.8. Under Assumption F1, F3 and B2 through B5, for all \( M > 0 \) and \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P\left\{ \sup_{\|A\| \leq M} \sup_{i \leq N} N^{1/2} \left[ \|F_N(v_i : \Delta) - F(v_i : \Delta)\| - [F_N(v_i) - F(v_i)] \right] \geq \varepsilon \right\} = 0.
\]

Proof. The previous corollary shows convergence for a fixed \( \Delta \). Now, let \( M > 0 \) and \( \delta > 0 \). Using an argument similar to Boldin (1983) or Bickel (1975), decompose the cube \( C = \{ \Delta : |\Delta| < (1/\delta) + 1)\delta M \) as the union of cubes with vertices on the grid of points \((j_1 \delta M, \ldots, j_p \delta M)\) where \( p = [1/\delta] + 1, \ j_i = 0, \pm 1, \ldots, \pm [1/\delta] + 1, \ |\Delta| = \max \text{of the absolute values of the coordinates of} \ \Delta \) and \([\ ]\) denotes the greatest integer function.

If \( |\Delta| \leq M \), let \( P(\Delta) \) be the lowest vertex of the cube containing \( \Delta \). For fixed \( \delta \),

\[
\sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left[ \|F_N(v_i : \Delta) - F(v_i : \Delta)\| - [F_N(v_i) - F(v_i)] \right] \\ \leq \sup_{\|\Delta\| \leq M} \sup_{i \leq N} N^{1/2} \left[ \|F_N(g_i(e_i : \Delta) : \Delta) - F(g_i(e_i : \Delta) : \Delta)\| \\ - [F_N(g_i(e_i : \Delta) : P(\Delta)) - F(g_i(e_i : \Delta) : P(\Delta))] \right] \\ + \max_{\|\Delta\| \leq M} N^{1/2} \left[ \|F_N(\cdot : P(\Delta)) - F(\cdot : P(\Delta))\| - [F_N(\cdot) - F(\cdot)] \right].
\]

Clearly there are a finite number of cubes in the partition. Thus, Lemma 2.7 implies the second term of the sum is \( o_p(1) \). Now, let \( C_1 \) be any cube of the partition and let \( \Delta^p \) be its lowest vertex. Then the lemma will be proven.
if we establish

$$T_1(\Delta) = \sup_{\Delta \in \mathcal{C}_1} N^{1/2}[F_N(\cdot : \Delta) - F(\cdot : \Delta)] - [F_N(\cdot : \Delta^p) - F(\cdot : \Delta^p)]$$

$$\leq A \delta + o_p(1),$$

where $A$ does not depend on $\delta$, holds uniformly in the cubes of the partition.

Re-expressing the sups, we have

$$T_1(\Delta) \leq N^{1/2}[\sup_{\Delta \in \mathcal{C}_1} F_N(\cdot : \Delta) - \inf_{\Delta \in \mathcal{C}_1} F(\cdot : \Delta)]$$

$$- [F_N(\cdot : \Delta^p) - F(\cdot : \Delta^p)]$$

$$+ N^{1/2}[\inf_{\Delta \in \mathcal{C}_1} F_N(\cdot : \Delta) - \sup_{\Delta \in \mathcal{C}_1} F(\cdot : \Delta)]$$

$$- [F_N(\cdot : \Delta^p) - F(\cdot : \Delta^p)].$$

Due to the structure of $F_N(\cdot : \Delta)$ and $F(\cdot : \Delta)$, we must handle the positive and negative arguments separately. First, consider only positive arguments. Negative arguments are handled in exactly the same way. Define $\Delta^{ui}$ and $\Delta^{Li}$ to be such that

$$G_i(x : \Delta^{ui}) = \sup_{\Delta \in \mathcal{C}_1} \{G_i(x : \Delta)\}$$

and

$$G_i(x : \Delta^{Li}) = \inf_{\Delta \in \mathcal{C}_1} \{G_i(x : \Delta)\} \text{ for } i = 1, \cdots, N.$$ 

Recall $G_i(x : \Delta) = x(1+\alpha_i(2)(\Delta))^{-1}+\alpha_i(1)(\Delta)$ and that $(1+\alpha_i(2)(\Delta))^{-1} > 0$ for $i = 1, \cdots, N$. Thus $\Delta^{ui}$ and $\Delta^{Li}$ do not depend on $x$. Therefore, $T_1(\Delta)$ is less than or equal to
\[ N^{-1/2} \left\{ \sum_{i=1}^{N} \{ [I(e_i \leq G_i(\cdot : \Delta^u)) - F(G_i(\cdot : \Delta^L))] \\ - [F_N(\cdot : \Delta^p) - F(\cdot : \Delta^p)] \right\} + \]
\[ N^{-1/2} \left\{ \sum_{i=1}^{N} \{ [I(e_i \leq G_i(\cdot : \Delta^L)) - F(G_i(\cdot : \Delta^u))] \\ - [F_N(\cdot : \Delta^p) - F(\cdot : \Delta^p)] \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \right\} \r...
for some constant $A$ which does not depend on $\delta$.

Proof. Applying the mean value theorem,

$$\left|N^{-1/2} \sum_{i=1}^{N} [F(G_i(x : \Delta_{ui})) - F(G_i(x : \Delta_{Li}))]\right|$$

$$= \left|N^{-1/2} \sum_{i=1}^{N} \left\{ x[(1 + \alpha_i^{(2)}(\Delta_{ui}))^{-1} - (1 + \alpha_i^{(2)}(\Delta_{Li}))^{-1}]ight. \right.$$

$$+ (\alpha_i^{(1)}(\Delta_{ui}) - \alpha_i^{(1)}(\Delta_{Li})) \right\} f(Z_i)\right|$$

$$\leq N^{-1/2} \sum_{i=1}^{N} \left| (1 + \alpha_i^{(2)}(\Delta_{ui}))^{-1} - (1 + \alpha_i^{(2)}(\Delta_{Li}))^{-1} \right| \cdot |f(Z_i)x|$$

$$+ N^{-1/2} \sum_{i=1}^{N} |\alpha_i^{(1)}(\Delta_{ui}) - \alpha_i^{(1)}(\Delta_{Li})| \cdot |f(Z_i)|,$$

where $Z_i$ is between $G_i(x : \Delta_{ui})$ and $G_i(x : \Delta_{Li})$. We should note that $|f(Z_i)x| \leq |f(Z_i)Z_i| + |f(Z_i)||x - Z_i|$ is bounded in probability since $|f(x)x|$ and $f$ are bounded and $|Z_i - x| \overset{P}{\to} 0$ uniformly in $i = 1, \ldots, N$. Thus by Lemma 2.1 the first component is $\leq A_1\delta$ and the second component is $\leq A_2\delta$ where $A_1$ and $A_2$ do not depend on $\delta$. Thus the lemma is established.

This was the last in the series of lemmas to prove our main theorem. Thus, the theorem is proven.
CHAPTER III

ESTIMATION OF THE REGRESSION PARAMETER $\beta$

3.1 Introduction

Model (1.1) can be expressed equivalently as

\begin{equation}
Y = X\beta + w e,
\end{equation}

where

$Y = (y_1, \cdots, y_N)'$, $e = (e_1, \cdots, e_N)'$, $w$ is the $N \times N$ diagonal matrix whose $i$th diagonal element is $\sigma_i$ and $X = (x_{ij})$ is the matrix with rows $\{x_i'\}$. This model can also be written as

\begin{equation}
w^{-1}Y = D\beta + e, \quad \text{where} \quad D = w^{-1}X.
\end{equation}

Assume the location score function $\phi_L$, is a non-decreasing function on $[0, 1]$, odd about $\frac{1}{2}$ and without loss of generality,

\begin{equation}
\int_0^1 \phi_L(u)du = 0 \quad \text{and} \quad \int_0^1 \phi_L^2(u)du = 1.
\end{equation}

To motivate our estimates, suppose that the $\{\sigma_i\}$ were known. The well known $R$-estimate (see Jureckova, 1971a), $\hat{\beta}_{opt}$, solves

\begin{equation}
N^{-1/2} \sum_{i=1}^N (d_i - \bar{d})a_L(R(\frac{y_i - x_i'\beta}{\sigma_i})) \overset{d}{=} 0,
\end{equation}

37
where $d_i$ is the $i$th row of $D$, $\bar{d} = N^{-1} \sum_{i=1}^{N} d_i$,

$$a_L(i) = \phi_L\left(\frac{i}{N+1}\right), \text{ and } R\left(\frac{y_i - x_i'\beta}{\sigma_i}\right)$$

is the rank of

$$\frac{y_i - x_i'\beta}{\sigma_i} \text{ among } \frac{y_1 - x_1'\beta}{\sigma_1}, \ldots, \frac{y_N - x_N'\beta}{\sigma_N}.$$

Although we want to fit model (3.2), which has no intercept, we cannot force this model with rank-based estimation. In fact, we are actually fitting the model $w^{-1}Y = D\beta + \mu 1_N + e$, where $\mu$ is the intercept parameter and $1_N$ is a $N \times 1$ vector $1_N = (1, \cdots, 1)'$.

We use the Hodges-Lehmann(1963) type estimate of intercept; later we will discuss other options for this intercept problem. The Hodges and Lehmann (1963) type estimate of the intercept solves

$$(3.5) \quad N^{-1/2} \sum_{i=1}^{N} a_L^+ \left(R\left|\frac{y_i - x_i'\hat{\beta}_{opt}}{\sigma_i} - \mu\right|\right)s\text{gn}\left(\frac{y_i - x_i'\hat{\beta}_{opt}}{\sigma_i} - \mu\right) = 0,$$

where $R\left|\frac{y_i - x_i'\hat{\beta}_{opt}}{\sigma_i} - \mu\right|$ is the rank of $\left|\frac{y_1 - x_1'\hat{\beta}_{opt}}{\sigma_1} - \mu\right|, \ldots, \left|\frac{y_N - x_N'\hat{\beta}_{opt}}{\sigma_N} - \mu\right|.$
Further, $a_L^+(i) = \phi_L^+(\frac{i}{n+1})$ and $\phi_L^+(u) = \phi_L(\frac{u+1}{2})$. Denote this unknown "optimal" solution to (3.5) as $\hat{\mu}_{opt}$. See Jureckova (1971b) for a discussion of the homoscedastic case. McKean and Hettmansperger (1978) derived the asymptotic distribution of $(\hat{\mu}_{opt}, \hat{\beta}_{opt})$, where $\hat{\beta}_{opt}$ is Jaekel's (1972) estimate.

Jaekel (1972) considers estimates of $\beta$ derived from dispersion measures which are certain linear combinations of the ordered residuals. He proves that his estimates are asymptotically equivalent to those proposed by Jureckova (1971a). Heiler and Willers (1988) proved that Jureckova's (1971a) results for her estimate of $\beta$ can be established without her "concordance conditions" (Assumptions 3a - 3c) on the rows $d_i$.

Clearly, the assumptions of Jaekel (1972), Heiler and Willers (1988) and McKean and Hettmansperger (1978) are satisfied by our Assumptions F1, L1, L2 and B1. Thus, from McKean and Hettmansperger (1978) we have that

$$\sqrt{N} \left[ \begin{pmatrix} \hat{\mu}_{opt} \\ \hat{\beta}_{opt} \end{pmatrix} - \begin{pmatrix} 0 \\ \beta \end{pmatrix} \right] \overset{D}{\rightarrow} N(0, \gamma_L^{-2} \lambda^{-1}) \quad (3.6)$$

where

$$\lambda = \lim_{N \to \infty} N^{-1}[1_N : D]'[1_N : D], \quad \gamma_L = \int_0^1 \phi(u)\phi(u, f)du$$
and
\[ \phi(u, f) = \frac{-f'(F^{-1}(u))}{f(F^{-1}(u))}. \]

Now, we have discussed fitting an intercept model when we actually have no intercept. So the fitted values \( w^{-1} \hat{Y}_{opt} = D \hat{\beta}_{opt} + 1_N \hat{\mu}_{opt} \) are not in the space spanned by \( D = w^{-1}X \). In order to correct this problem we will adjust the rank based estimate of \( \beta \) by projecting the fitted values \( w^{-1} \hat{Y}_{opt} \) into the space spanned by \( D \). The resulting fitted values are

\[ (3.7) \quad w^{-1} \hat{Y}_{opt} = D(D'D)^{-1}D'[1_N : D] \begin{bmatrix} \hat{\mu}_{opt} \\ \hat{\beta}_{opt} \end{bmatrix}. \]

We could then find the new estimate \( \hat{\beta}_{opt} \) by solving

\[ (3.8) \quad D\beta = w^{-1} \hat{Y}_{opt}. \]

Thus the new adjusted "optimal" estimate is

\[ (3.9) \quad \hat{\beta}_{opt} = [(D'D)^{-1}D'1_N : I] \begin{bmatrix} \hat{\mu}_{opt} \\ \hat{\beta}_{opt} \end{bmatrix}. \]

Clearly Equations 3.6 and 3.9 imply

\[ \sqrt{N}(\hat{\beta}_R - \beta) \xrightarrow{D} N(0, \gamma^{-2}_L \Sigma^{-1}), \]

where \( \Sigma = \lim_{N \to \infty} N^{-1}D'D. \)
We have only considered the case that the \( \{ \sigma_i \} \) are known. In practice the \( \{ \sigma_i \} \) are not known. This is the case we consider next. Suppose we have estimates of \( (\theta, \beta) \) which satisfy \( \sqrt{N}(\hat{\theta} - \theta) = O_p(1) \) and \( \sqrt{N}(\hat{\beta}_0 - \beta) = O_p(1) \). We then form the estimated weights as follows. \( \hat{\sigma}_i = \exp[h(x_i^T \hat{\beta}_0) \hat{\theta}] \). The existence of such estimates are discussed in Chapter I. Define \( \hat{\omega} \) to be the diagonal matrix with this \( i \)th diagonal element \( \hat{\sigma}_i \) and define \( \hat{D} = \hat{\omega}^{-1}X \). Simply substituting \( \{ \hat{\sigma}_i \} \) into Equation 3.4 suggests \( \beta \) can be estimated by solving

\[
N^{-1/2} \sum_{i=1}^{N} (\hat{d}_i - \hat{d}^*) a_L(R(\frac{y_i - x_i^T \hat{\beta}}{\hat{\sigma}_i})) = 0,
\]

where \( \hat{d}_i \) is the \( i \)th row of \( \hat{D} \), \( \hat{d}^* = N^{-1} \sum_{i=1}^{N} \hat{d}_i \). We will denote the solution to (3.10) as \( \hat{\beta}_R \). Just as in the case where true errors are known, we cannot force a no-intercept model. Thus, similarly we must estimate the intercept. So, the Hodges-Lehmann (1963) type intercept, \( \hat{\mu}_R \), solves

\[
N^{-1/2} \sum_{i=1}^{N} a_L^+(R(\frac{y_i - x_i^T \hat{\beta}_R}{\hat{\sigma}_i}) - \mu))\text{sgn}(\frac{y_i - x_i^T \hat{\beta}_R}{\hat{\sigma}_i} - \mu) = 0.
\]

In this case we also need to adjust our estimates to fit the no-intercept model. Using the same arguments as with the true \( \{ \sigma_i \} \), define the new adjusted value of \( \beta \) as

\[
\hat{\beta}_R = [(\hat{D}' \hat{D})^{-1} \hat{D}' 1_N : 1] \begin{bmatrix} \hat{\mu}_R \\ \hat{\beta}_R \end{bmatrix}.
\]
Now, $\hat{\beta}_R$ is the estimate of $\beta$ in the true model (3.2). In Section 2, using the general linearity results in Chapter II and Theorem 7.1 of Carroll and Ruppert (1982a) we establish linearity results for our location rank statistic. Using these results we establish asymptotic equivalence of our weighted R-estimates and those based on the true $\{\sigma_i\}$. Furthermore, properties of the R-estimate, $\hat{\beta}_R$, are discussed. In Section 3 the linearity result for the intercept rank process is established. Utilizing this result, we establish asymptotic equivalence of the Hodges-Lehmann (1963) type estimates of intercept based on both the estimated weights and the true weights. Section 4 is concerned with the adjusted estimate that fits the true no-intercept model. Asymptotic equivalence of $\hat{\beta}_{opt}$ and $\hat{\beta}_R$ is also established in this section, thus establishing the asymptotic distribution of $\hat{\beta}_R$. Using a method similar to McKean and Hettmansperger (1976) a consistent estimate of $\gamma_L$ is developed in Section 5.

3.2 Estimation of $\beta$ in the Intercept Model

3.2.1 Linearity

In this section, linearity results for the location rank statistic is established using the results of Carroll and Ruppert (1982a) and the general results of Chapter II. In order to use the results in Chapter II we need some additional notation.
For $\Delta_1$ and $\Delta_3$ in $\mathbb{R}^P$, $\Delta_2$ in $\mathbb{R}^1$ and $\Delta = (\Delta_1, \Delta_2, \Delta_3)$, define

$$c_i = \frac{x_i}{\sigma_i} - N^{-1} \sum_{j=1}^{N} \frac{x_j}{\sigma_j}, \quad \alpha_i^{(1)}(\Delta) = N^{-1/2} \frac{x'_i}{\sigma_i} \Delta_1$$

$$h_i(\Delta) = h(x'_i \beta + x'_i \Delta_3 N^{-1/2})$$

$$\alpha_i^{(2)}(\Delta) = \exp[-h_i(\Delta) \Delta_2 N^{-1/2} + (h_i(0) - h_i(\Delta))\theta] - 1$$

and

$$\alpha_i^{(3)}(\Delta) = \frac{x_i}{\sigma_i} \alpha_i^{(2)}(\Delta) - N^{-1} \sum_{j=1}^{N} \frac{x_j}{\sigma_j} \alpha_j^{(2)}(\Delta).$$

Next define the processes

$$U_L(\Delta : F_N) = N^{-1/2} \sum_{i=1}^{N} (c_i + \alpha_i^{(3)}(\Delta)) \phi_L(\{e_i - \alpha_i^{(1)}(\Delta)(1 + \alpha_i^{(2)}(\Delta)) : \Delta\})$$

and

$$U_L(\Delta : F) = N^{-1/2} \sum_{i=1}^{N} (c_i + \alpha_i^{(3)}(\Delta)) \phi_L(\{e_i - \alpha_i^{(1)}(\Delta)(1 + \alpha_i^{(2)}(\Delta))\}).$$

We have previously used the notation

$$F_N(x : \Delta) = N^{-1} \sum_{i=1}^{N} I(\{e_i - \alpha_i^{(1)}(\Delta)(1 + \alpha_i^{(2)}(\Delta)) \leq x\})$$

and

$$F(x : \Delta) = N^{-1} \sum_{i=1}^{N} F(x(1 + \alpha_i^{(2)}(\Delta))^{-1} + \alpha_i^{(1)}(\Delta)).$$
Note that \( F_N(x : 0) \) is the empirical distribution function of the errors \( e_i \) and that \( F(x : 0) = F(x) \). Also note that \( \hat{\Delta} = (\sqrt{n}(\hat{\beta}_R - \beta), \sqrt{n}(\hat{\beta}_0 - \beta)) \), where \( \hat{\beta}_R \) is the rank-based estimate of \( \beta \) and \( (\hat{\theta}, \hat{\beta}_0) \) are the initial \( N^{1/2} \)-consistent estimates of \( (\theta, \beta) \). Thus

\[
F_N(x : \hat{\Delta}) = N^{-1} \sum_{i=1}^{N} I(\hat{e}_i \leq x)
\]

is the empirical distribution function of the estimated residuals

\[
\hat{e}_i = \frac{y_i - x_i' \hat{\beta}_R}{e^{\hat{h}(x_i', \hat{\beta}_0)}},
\]

Further, \( N \cdot F_N(\hat{e}_j : \hat{\Delta}) \) is the rank of \( \hat{e}_j \) among \( \hat{e}_1, \cdots, \hat{e}_N \). Now, by substituting the rank based estimate, \( \hat{\beta}_R \), into Equation (3.10), Equation (3.10) can be written as

\[
(3.13) \quad U_L(\hat{\Delta} : F_N) = 0.
\]

We will not actually assume the estimate \( \hat{\beta}_R \) solves this equation exactly, but rather that the LHS of Equation (3.13) is less than twice the infimum over all \( \beta \).

Lemma 3.1. Under Assumptions F1 through F3, L1, L2 and B1 through B5, for all \( M > 0 \) and \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P \left\{ \sup_{\|\Delta\| \leq M} \| U_L(\Delta : F) - U_L(0 : F) + \gamma_L \Delta_1 S \| \geq \varepsilon \right\} = 0
\]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
\[
\gamma_L = \int_0^1 \phi_L(u)\phi(u,f)du, \quad \phi(u,f) = -\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}
\]

and \( S = \lim_{N \to \infty} N^{-1} D'(I - N^{-1}1_N1'_N)D. \)

Proof. Letting

\[
K_i = N^{-1/2} \left\{ N^{-1} \sum_{j=1}^N \|x_j\|^2 + N^{-1} \sum_{j=1}^N \|\frac{x_j}{\sigma_j}\|^2 + N^{-1} \sum_{j=1}^N h_j^2(0) \right. \\
\left. + |h_i(0)| + \|x_i\| + \|\frac{x_i}{\sigma_i}\| \right\}, \quad \phi_i(e,r,s) = \phi_L(F((e-r)(1+s))),
\]

the conditions of Theorem 7.1 of Carroll and Ruppert (1982a) are implied by

the listed assumptions, so the lemma is established.

**Theorem 3.1.** Under the assumptions of Lemma 3.1 for all \( M > 0 \) and \( \varepsilon > 0, \)

\[
\lim_{N \to \infty} P \left\{ \sup_{\|\Delta\| \leq M} \|U_L(\Delta : F_N) - U_L(0 : F_N) + \gamma_L \Delta S\| \geq \varepsilon \right\} = 0.
\]

Proof.

\[
\sup_{\|\Delta\| \leq M} \|U_L(\Delta : F_N) - U_L(0 : F_N) + \gamma_L \Delta S\|
\]

\[
\leq \sup_{\|\Delta\| \leq M} \|U_L(\Delta : F') - U_L(0 : F') + \gamma_L \Delta S\|
\]

\[
+ \sup_{\|\Delta\| \leq M} \|U_L(\Delta : F_N) - U_L(\Delta : F)\| + \|U_L(0 : F_N) - U_L(0 : F)\|
\]

Convergence in probability to zero of the first 2 components of the sum

is implied by Lemma 3.1 and Theorem 2.1 respectively. Under Assumptions
B2 and B3, application of Hajek and Sidak's (1967) Theorem V.1.5a implies the third component of the sum converges in probability to zero. Thus the theorem is proven.

3.2.2 Estimation

This section is concerned with the $R$-estimate of $\beta$ and its properties.

Lemma 3.2. (Hajek and Sidak (1967) Theorem V.1.5 a). Under Assumptions B2 and B3, there exists a $C > 0$ such that

$$\lim_{N \to \infty} P[\|U_L(0 : F_N)\| \leq C] = 1.$$ 

Lemma 3.3. Under the assumptions of Lemma 3.1, for all $\epsilon > 0$,

$$\lim_{N \to \infty} P[\|U_L(\hat{\Delta} : F_N)\| \geq \epsilon] = 0.$$ 

We previously noted that we are not assuming $\hat{\beta}_R$ solves $U_L(\hat{\Delta} : F_N) = 0$, but rather $U_L(\hat{\Delta} : F_N)$ is less than twice its infimum over all $\beta$. The proof of this lemma follows exactly as in Carroll and Ruppert (1982a) as a consequence of Lemma 3.2 and Theorem 3.1.

Lemma 3.4. Under the assumptions of Lemma 3.1, there exists a $C > 0$ such that

$$\lim_{N \to \infty} P[\sqrt{N}\|\hat{\beta}_R - \beta\| \leq C] = 1.$$
Proof. By Lemma 3.3 this lemma holds if we can establish that: for all \( z > 0 \) and \( \epsilon > 0 \) and \( M_1 \), there exists \( M_2 \) satisfying

\[
P\left\{ \inf_{\|\Delta_1\| \geq M_2} \inf_{\|\Delta_2\| \leq M_1} \inf_{\|\Delta_3\| \leq M_1} \|U_L(\Delta : F_N)\| > z \right\} > 1 - \epsilon.
\]

This can be proven by arguments as in Jureckova\'s (1977) proof of her Lemma 5.2.

Theorem 3.2. Under the assumptions of Lemma 3.1, for all \( \epsilon > 0 \),

\[
\lim_{N \to \infty} P\left\{ \sqrt{N} \|\hat{\beta}_R - \hat{\beta}_{opt}\| \geq \epsilon \right\} = 0.
\]

Proof. Theorem 3.1, Lemma 3.3 and Lemma 3.4 give the asymptotic expansion

\[
\sqrt{N} (\hat{\beta}_R - \beta) = [\gamma_L S]^{-1} U_L(0 : F_N) + o_p(1).
\]

Certainly these lemmas and this theorem also apply when the \( \{\sigma_i\} \) are known. Thus we also have the asymptotic expansion

\[
\sqrt{N} (\hat{\beta}_{opt} - \beta) = [\gamma_L S]^{-1} U_L(0 : F_N) + o_p(1).
\]

Combining the two asymptotic expansions proves the theorem.

We will not consider the asymptotic distribution of \( \hat{\beta}_R \) at this time.

Our interest is really in the joint distribution of \( \hat{\beta}_R \) and the estimate of the intercept, which we will consider in Section 3.4.
3.3. Estimation of the Intercept

3.3.1 Linearity

In this section, linearity results for the signed rank statistic are developed using the results of Carroll and Ruppert (1982a) and the general results of Chapter II. In order to apply the linearity results in Chapter II, we need some additional notation.

For $\Delta_1$ and $\Delta_3$ in $\mathbb{R}^p$, $\Delta_2$ and $\Delta_4$ in $\mathbb{R}^1$ and $\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4)$, define

\[
    h_i(\Delta) = h(x_i^T \theta + N^{-1/2} x_i^T \Delta_3)
\]

\[
    \alpha_i^{(2)}(\Delta) = \exp[-h_i(\Delta) \Delta_2 N^{-1/2} + (h_i(0) - h_i(\Delta))\theta] - 1
\]

and

\[
    \alpha_i^{(1)}(\Delta) = N^{-1/2} \frac{x_i^T}{\sigma_i} \Delta_1 + N^{-1/2} \Delta_4 (1 + \alpha_i^{(2)}(\Delta))^{-1}.
\]

Define the process

\[
    U_I(\Delta : F_N) = N^{-1/2} \sum_{i=1}^{N} \text{sgn}[(e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta))] \cdot \varphi_i^T(\Delta)(e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta)).
\]

We should note here that in the notation of Chapter II

\[
    U_I(\Delta : F_N) = N^{-1/2} \sum_{i=1}^{N} \phi_i \left( \frac{1}{2} + \frac{1}{2(N+1)} (F_N(v_i : \Delta) - F_N(-v_i : \Delta)) \right),
\]
where
\[ v_i = (e_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta)), \]
\[ F_N(v_i : \Delta) = N^{-1} \sum_{j=1}^{N} I(v_j \leq v_i), \]
and
\[ F_N(-v_i : \Delta) = N^{-1} \sum_{j=1}^{N} I(v_j \leq -v_i). \]

This expression is a consequence of the fact \( \phi \) is odd about \( \frac{1}{2} \) and \( (1 + \alpha_k^{(2)}(\Delta)) > 0 \) for \( k = 1, \ldots, N \).

Note that \( \hat{\Delta} = (\sqrt{n}(\hat{\beta}_R - \beta), \sqrt{n}(\hat{\theta} - \theta), \sqrt{n}(\hat{\beta}_0 - \beta), \sqrt{n}(\hat{\mu}_R - \mu)), \)
where \( \hat{\mu}_R \) is the rank-based estimate of the intercept, \( \hat{\beta}_R \) is the rank-based estimate of \( \beta \), and \( (\hat{\beta}_0, \hat{\theta}) \) are the initial \( N^{1/2} \)-consistent estimates of \( (\beta, \theta) \). Now, by substituting \( \hat{\mu}_R \) into Equation (3.11) it can be written as:
\[ U_f(\hat{\Delta} : F_N) = 0. \]

**Theorem 3.3.** Under Assumptions F1 through F3, B1 through B5, L1 and L2, for all \( \varepsilon > 0 \) and \( M > 0 \),
\[ \lim_{N \to \infty} P \left\{ \sup_{\|\Delta\| \leq M} |U_f(\Delta : F_N) - U_f(0 : F_N) + \gamma_L \mu_d' \Delta_1 + \gamma_L \Delta_4| \geq \varepsilon \right\} = 0. \]
Proof.

\[
\sup_{\|\Delta\| \leq M} \left| U_I(\Delta : F_N) - U_I(0 : F_N) + \gamma_L \mu'_d \Delta_1 + \gamma_L \Delta_4 \right|
\]

\[
\leq \sup_{\|\Delta\| \leq M} \left| U_I(\Delta : F_N) - N^{-1/2} \right|
\]

\[
\cdot \sum_{i=1}^{N} \phi_L \left( \frac{1}{2} + \frac{N}{2(N+1)} [F(v_i : \Delta) - F(-v_i : \Delta)]\right)
\]

\[
- U_I(0 : F_N) + N^{-1/2} \sum_{i=1}^{N} \phi_L \left( \frac{1}{2} + \frac{N}{2(N+1)} [F(e_i) - F(-e_i)]\right)
\]

\[
+ \sup_{\|\Delta\| \leq M} \left| N^{-1/2} \sum_{i=1}^{N} \phi_L(F(v_i)) - N^{-1/2} \sum_{i=1}^{N} \phi_L(F(e_i)) \right|
\]

\[
+ \gamma_L \mu'_d \Delta_1 + \gamma_L N^{-1} \sum_{i=1}^{N} (1 + \alpha_i^{(2)}(\Delta))^{-1} \Delta_4
\]

\[
+ \sup_{\|\Delta\| \leq M} \left| \gamma_L \Delta_4 (1 - N^{-1} \sum_{i=1}^{N} (1 + \alpha_i^{(2)}(\Delta))^{-1}) \right|
\]

\[
+ |N^{-1/2} \sum_{i=1}^{N} \phi_L(F(e_i)) - N^{-1/2} \sum_{i=1}^{N} \phi_L \left( \frac{1}{2} + \frac{N}{2(N+1)} [F(e_i) - F(-e_i)]\right)|
\]

\[
= P_1 + P_2 + P_3 + P_4 + P_5, \text{ say.}
\]

Under Assumptions F1 through F3, B2 through B5, L1 and L2, Theorem 7.1 of Carroll and Ruppert (1982a) implies \( P_3 \xrightarrow{P} 0 \). Since

\[
N^{-1/2} \sum_{i=1}^{N} [(1 + \alpha_i^{(2)}(\Delta))^{-1} - 1]
\]

is bounded in probability (Lemma 2.1), \( P_4 \xrightarrow{P} 0 \).
Next consider

\[ P_5 = N^{-1/2} \sum_{i=1}^{N} \phi'(A_i)(F(e_i) - \frac{1}{2}) \cdot \frac{1}{N+1}, \]

where \( A_i \) is between \([F(e_i)]\) and \([\frac{1}{2} + \frac{N}{2(N+1)}(2F(e_i)-1)]\). Now, \( F(e_i) - \frac{1}{2} \)
and \( \phi' \) are bounded. Thus \( P_5 \to 0. \)

Consider \( P_1: \)

\[
\sup_{\|A\| \leq M} |U_I(\Delta : F_N) - N^{-1/2} \sum_{i=1}^{N} \phi_L(\frac{1}{2} + \frac{N}{2(N+1)}[F(v_i : \Delta) - F(-v_i : \Delta)])
- U_I(0 : F_N) + N^{-1/2} \sum_{i=1}^{N} \phi_L(\frac{1}{2} + \frac{N}{2(N+1)}[F(e_i) - F(-e_i)])| \\
\leq \sup_{\|A\| \leq M} |N^{-1/2} \sum_{i=1}^{N} \phi'(A_i) \{ [F_N(v_i : \Delta) - F(v_i : \Delta)]
- [F_N(-v_i : \Delta) - F(-v_i : \Delta)] \} \\
- N^{-1/2} \sum_{i=1}^{N} \phi'(B_i) \{ [F_N(e_i) - F(e_i)] - [F_N(-e_i) - F(-e_i)] \} | \frac{N}{2(N+1)} \\
\leq \sup_{\|A\| \leq M} N^{-1} \sum_{i=1}^{N} |\phi'(A_i)| \\
\cdot \left\{ N^{1/2} |[F_N(v_i : \Delta) - F(v_i : \Delta)] - [F_N(e_i) - F(e_i)]| \\
+ N^{1/2} |[F_N(-v_i : \Delta) - F(-v_i : \Delta)] - [F_N(-e_i) - F(-e_i)]| \right\} \cdot \frac{N}{2(N+1)} \\
+ \sup_{\|A\| \leq M} N^{-1} \sum_{i=1}^{N} |\phi'(A_i) - \phi'(B_i)| \\
\cdot N^{1/2} |[F_N(e_i) - F(e_i)] - [F_N(-e_i) - F(-e_i)]| \frac{N}{2(N+1)},
\]

where \( A_i \) is between \([\frac{1}{2} + \frac{N}{2(N+1)}[F_N(v_i : \Delta) - F_N(-v_i : \Delta)])\) and
\[ \left[ \frac{1}{2} + \frac{N}{2(N+1)}[F(v_i : \Delta) - F(-v_i : \Delta)] \right], \text{ and } B_i \text{ is between} \]
\[ \left[ \frac{1}{2} + \frac{N}{2(N+1)}[F_N(e_i) - F_N(-e_i)] \right] \quad \text{and} \quad \left[ \frac{1}{2} + \frac{N}{2(N+1)}[F(e_i) - F(-e_i)] \right]. \]

Thus, \(|A_i - B_i| \overset{p}{\to} 0\) uniformly in \(i = 1, \cdots, N\) and \(\|\Delta\| \leq M\) by Lemma 2.6 and Lemma 2.8. Since \(\phi'\) is uniformly continuous, this implies \(|\phi'(A_i) - \phi'(B_i)| \overset{p}{\to} 0\) uniformly in \(i = 1, \cdots, N\) and \(\|\Delta\| \leq M\). Thus, since

\[ N^{1/2}[F_N(e_i) - F(e_i)] \quad \text{and} \quad N^{1/2}[F_N(-e_i) - F(-e_i)] \]

are bounded in probability, the second component of the sum is \(o_p(1)\). Now, consider the first component of the sum. Lemma 2.6 and Lemma 2.8 imply

\[ N^{1/2}[F_N(v_i : \Delta) - F(v_i : \Delta)] - [F_N(e_i) - F(e_i)] \]

and

\[ N^{1/2}[F_N(-v_i : \Delta) - F(-v_i : \Delta)] - [F_N(-e_i) - F(-e_i)] \]

converge in probability to zero uniformly in \(i = 1, \cdots, N\) and \(\|\Delta\| \leq M\). Since \(\phi'\) is bounded, the first component of the sum converges in probability to zero.

Finally consider \(P_2\):
\[
P_2 = \sup_{\|\Delta\| \leq M} N^{-1/2} \left| \sum_{i=1}^{N} \phi'(F(e_i)) \left\{ [F(v_i : \Delta) - F(-v_i : \Delta)] - [F(v_i) - F(-v_i)] \right\} \right|
\]

\[
- [F(v_i) - F(-v_i)] \left( \frac{N}{2(N+1)} \right)
\]

\[
+ \sup_{\|\Delta\| \leq M} N^{-1} \left| \sum_{i=1}^{N} \phi'(A_i) [F(v_i) - F(-v_i)] \right| \frac{N^{1/2}}{2(N+1)}
\]

\[
+ \sup_{\|\Delta\| \leq M} N^{-1} \left| \sum_{i=1}^{N} (\phi'(A_i) - \phi'(F(e_i))) \right|
\]

\[
\cdot N^{1/2} \left\{ [F(v_i : \Delta) - F(v_i)] - [F(-v_i : \Delta) - F(-v_i)] \right\} \cdot \frac{N}{2(N+1)}
\]

\[
= Q_1 + Q_2 + Q_3, \text{ say,}
\]

where $A_i$ is between

\[
\left[ \frac{1}{2} + \frac{1}{2(N+1)} \right] [F(v_i : \Delta) - F(-v_i : \Delta)] \quad \text{and} \quad \left[ \frac{1}{2} + \frac{1}{2} [F(v_i) - F(-v_i)] \right].
\]

Now, $|A_i - F(e_i)| \overset{P}{\to} 0$ uniformly in $i = 1, \cdots, N$ and $\|\Delta\| \leq M$ by the corollary to Lemma 2.4. Thus $|\phi'(A_i) - \phi'(F(e_i))| \overset{P}{\to} 0$ uniformly in $i = 1, \cdots, N$ and $\|\Delta\| \leq M$ since $\phi'$ is uniformly continuous. Lemma 2.4 implies that

\[
N^{1/2} [F(v_i : \Delta) - F(v_i)] \quad \text{and} \quad N^{1/2} [F(-v_i : \Delta) - F(-v_i)]
\]

are bounded in probability. Thus $Q_3 \overset{P}{\to} 0$. Clearly, $Q_2 \overset{P}{\to} 0$ since $\phi'$ is bounded.
Finally, we must consider $Q_1$:

$$Q_1 = \frac{N^{-1/2}}{2(N+1)} \sum_{i=1}^{N} \left[ \phi'(F(e_i)) \sum_{j=1}^{N} \left\{ F \left[ v_i(1 + \alpha_j^{(2)}(\Delta)^{-1} + \alpha_j^{(1)}(\Delta) \right] - F(v_i) ight] 
- F \left[ -v_i(1 + \alpha_j^{(2)}(\Delta)^{-1} + \alpha_j^{(1)}(\Delta) \right] + F(-v_i) \right] \right]$$

$$= \frac{N^{-1/2}}{2(N+1)} \sum_{i=1}^{N} \left[ \phi'(F(e_i)) \sum_{j=1}^{N} \left\{ f(R_{ij}) \left[ v_i[(1 + \alpha_j^{(2)}(\Delta)^{-1} - 1] + \alpha_j^{(1)}(\Delta) \right] 
- f(T_{ij}) \left[ -v_i[(1 + \alpha_j^{(2)}(\Delta)^{-1} - 1] + \alpha_j^{(1)}(\Delta) \right] \right] \right]$$

$$\leq \frac{1}{2} \left( \frac{1}{N+1} \right) \sum_{i=1}^{N} \left[ \left| \phi'(F(e_i)) \right| |N^{-1/2} \sum_{j=1}^{N} [(1 + \alpha_j^{(2)}(\Delta)^{-1} - 1] \right| \right] \cdot |v_i - e_i| \left| f(R_{ij}) + f(T_{ij}) \right|$$

$$+ \left| \left( \frac{1}{N+1} \right) \sum_{i=1}^{N} \phi'(F(e_i))e_if(e_i) \right| \cdot |N^{-1/2} \sum_{j=1}^{N} [(1 + \alpha_j^{(2)}(\Delta)^{-1} - 1] \right|$$

$$+ \frac{1}{2} \left( \frac{1}{N+1} \right) \sum_{i=1}^{N} \left\{ |e_i| \left| \phi'(F(e_i)) \right| |N^{-1/2} \sum_{j=1}^{N} [(1 + \alpha_j^{(2)}(\Delta)^{-1} - 1] \right| \right.$$
is between \([-v_i(1 + \alpha_j^{(2)}(\Delta))^{-1} + \alpha_j^{(1)}(\Delta)]\) and \([-v_i]\). Thus, by Lemma 2.2 and Lemma 2.3,

\[|R_{ij} - e_i| \xrightarrow{P} 0 \quad \text{and} \quad |T_{ij} - e_i| \xrightarrow{P} 0\]

uniformly in \(i = 1, \cdots, N\), \(j = 1, \cdots, N\) and \(\|\Delta\| \leq M\). Since \(f\) is uniformly continuous,

\[|f(R_{ij}) - f(e_i)| \xrightarrow{P} 0 \quad \text{and} \quad |f(T_{ij}) - f(e_i)| \xrightarrow{P} 0\]

uniformly in \(i = 1, \cdots, N\), \(j = 1, \cdots, N\) and \(\|\Delta\| \leq M\). Now \(N^{-1/2} \sum_{j=1}^{N} |(1 + \alpha_j^{(2)}(\Delta))^{-1} - 1|\) and \(N^{-1/2} \sum_{j=1}^{N} |\alpha_j^{(1)}(\Delta)|\) are bounded in probability. Further, \(\phi'\) and \(f\) are bounded. Thus, by Lemma 2.2 the first term of the sum is \(o_p(1)\). Clearly the fourth term of the sum is \(o_p(1)\). The third term of the sum is \(o_p(1)\) since \(N^{-1} \sum_{i=1}^{N} |e_i|\) is also bounded in probability (\(E[e_i^2] < \infty\)). Convergence of the second component of the sum is implied by convergence of

\[N^{-1} \sum_{i=1}^{N} \phi'(F(e_i))e_i f(e_i)\]

The expectation of this sum is zero, since \(\phi\) is odd about \(\frac{1}{2}\) and \(f\) is symmetric. Now, \(E[\phi'(F(e_i))e f(e)]^2 < \infty\) by the Lebesgue Dominated Convergence Theorem since \(\phi'\) and \(|e f(e)|\) are bounded. Thus, the variance
converges to zero. Thus, \( N^{-1} \sum_{i=1}^{N} \phi'(F(e_i))e_i f(e_i) = o_p(1) \). Hence,
\[ Q_1 \overset{P}{\to} 0, \text{ and the theorem is proven.} \]

### 3.3.2 Estimation

In this section, we shall discuss the Hodges-Lehmann (1963) type estimate of intercept based on estimated weights. The residual process \( U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu) : F_N) \) (as a function of \( \mu \)) is a non-increasing step function which steps down at each

\[
\left[ \frac{y_i - x_i^T \hat{\beta}_R}{\hat{\sigma}_i} + \frac{y_j - x_j^T \hat{\beta}_R}{\hat{\sigma}_j} \right] / 2.
\]

See Bauer (1972).

We define \( U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu) : F_N) \) to be right continuous at the jumps.

Before proceeding with the estimation of \( \mu \) and \( \gamma_L \) we need the limiting distribution of

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N). \]

We need the following obvious corollary of Theorem 3.3.

**Lemma 3.5.** Under the assumptions of Theorem 3.3, for all \( M > 0 \) and all \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P \left[ \sup \left\{ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \Delta^U_i) : F_N) 
- U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \Delta^L_i) : F_N)
+ \gamma_L(\Delta^U_i - \Delta^L_i) : -M \leq \Delta^L_i \leq \Delta^U_i \leq M \right\} \geq \varepsilon \right] = 0.
\]
Lemma 3.6. Under the Assumptions of Theorem 3.3

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \xrightarrow{D} N(0, 1 + \mu'_d S^{-1} \mu_d) \]

Proof. The asymptotic expansion

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) = U_I(0 : F_N) - \gamma_L \mu'_d \hat{\Delta}_1 + o_p(1), \]

follows from Theorem 3.3. From the proof of Theorem 3.2, we have the asymptotic expansion

\[ \hat{\Delta}_1 = \gamma_L^{-1} S^{-1} U_L(0 : F_N) + o_p(1). \]

Combining these two expansions gives:

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) = U_I(0 : F_N) - \mu'_d S^{-1} U_L(0 : F_N) + o_p(1) \]
\[ = (U_I(0 : F_N) - U_I(0 : F)) - \mu'_d S^{-1} (U_L(0 : F_N) - U_L(0 : F)) \]
\[ + U_I(0 : F) - \mu'_d S^{-1} U_L(0 : F) + o_p(1) \]

Now,

\[ U_I(0 : F_N) - U_I(0 : F) = o_p(1) \]

by the proof of Theorem V.1.7 of Hajek and Sidak (1967). Assumptions B2 and B3 imply we can apply Theorem V.1.5a of Hajek and Sidak (1967) to conclude

\[ U_L(0 : F_N) - U_L(0 : F) = o_p(1). \]
Thus we are left with

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) = U_I(0 : F) - \mu'_d S^{-1} U_L(0 : F) + o_p(1). \]

Theorem V.1.2 of Hajek and Sidak (1967) implies \( U_I(0 : F) \overset{D}{\rightarrow} N(0, 1) \). Theorem A.11 of Hettmansperger (1984) implies \( U_L(0 : F) \overset{D}{\rightarrow} N(0, S) \). Since the covariance of \( U_L(0 : F) \) and \( U_I(0 : F) \) is zero,

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \overset{D}{\rightarrow} N(0, 1 + \mu'_d S^{-1} \mu_d). \]

**Corollary.** Under the assumptions of the lemma,

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, a) : F_N) \overset{D}{\rightarrow} N(-a \gamma_L, 1 + \mu'_d S^{-1} \mu_d) \]

for any real \( a \).

**Proof.** Using Lemma 3.5 gives the expansion

\[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, a) : F_N) = U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) - \gamma_L a + o_p(1). \]

Thus the Corollary is established.

Now, we turn our focus to the Hodges and Lehmann (1963) type estimate of the intercept. Define

\[ \hat{\mu}_L = \sup\{\mu : U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu) : F_N) > 0\} \]

\[ \hat{\mu}_U = \inf\{\mu : U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu) : F_N) < 0\} \]
and
\[ \hat{\mu}_R = \frac{\hat{\mu}_U + \hat{\mu}_L}{2}. \]

The following two lemmas are a direct consequence of Hodges and Lehmann (1963) and the corollary of Lemma 3.6.

**Lemma 3.7.** Under the assumptions of Theorem 3.3,
\[ \sqrt{N} \hat{\mu}_R \overset{D}{\to} N(0, \gamma^{-2}_L (1 + \mu_d' S^{-1} \mu_d)). \]

**Lemma 3.8.** Under the assumptions of Theorem 3.3, \( \sqrt{N} \hat{\mu}_U \) and \( \sqrt{N} \hat{\mu}_L \) are bounded in probability.

**Lemma 3.9.** Under the assumptions of Theorem 3.3, for all \( \varepsilon > 0 \),
\[ \lim_{N \to \infty} P \{ |U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu_L) : F_N)| \geq \varepsilon \} = 0 \]
and
\[ \lim_{N \to \infty} P \{ |U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu_U) : F_N)| \geq \varepsilon \} = 0. \]

**Proof.** We will begin by proving the second statement of the lemma. First, we should note that by the definition of \( \hat{\mu}_U \), \( U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_L) : F_N) < 0 \). Let \( \varepsilon > 0 \) be given and let \( z = \frac{\varepsilon}{2\gamma_L} \). By Lemma 3.8 \( \sqrt{N} \hat{\mu}_U \) and \( \sqrt{N} \hat{\mu}_U - z \) are bounded in probability. Thus,
\[ A_N = \{ \sqrt{N} |\hat{\mu}_U| \leq C \text{ and } \sqrt{N} |\hat{\mu}_U - z N^{-1/2}| \leq C \} \]
converges in probability to 1 for some $C > 0$. Note that $z > 0$ implies $U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U - z) : F_N) > 0$. Now,

\[
P[|U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U) : F_N)| \geq \varepsilon] = P[-U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U) : F_N) \geq \varepsilon] \\
\leq P[U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U - z) : F_N) \\
- U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U) : F_N) \geq \varepsilon] \\
\leq P(A_N^c) + P(A_N \text{ and } U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U - z) : F_N) \\
- U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U) : F_N) - \frac{\varepsilon}{2} \geq \frac{\varepsilon}{2}]
\]

Since $(\sqrt{N} \hat{\mu}_U - \sqrt{N} \hat{\mu}_U + z)\gamma_L = \frac{\varepsilon}{2}$, Lemma 3.5 ensures the second term converges in probability to 0. Thus, $U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_U) : F_N) = o_p(1)$.

A similar argument could be used to show

\[
U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_L) : F_N) = o_p(1).
\]

The result we need is an obvious corollary:

**Corollary.** Under the assumptions of the lemma, for all $\varepsilon > 0$,

\[
\lim_{N \to \infty} P\{|U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \hat{\Delta}_4) : F_N)| \geq \varepsilon\} = 0,
\]

where $\hat{\Delta}_4 = \sqrt{N} \hat{\mu}_R$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Theorem 3.4. Under the conditions of Theorem 3.3, for all $\varepsilon > 0$

$$\lim_{N \to \infty} P\{\sqrt{N}\vert \hat{\mu}_{\text{opt}} - \hat{\mu}_R \vert \geq \varepsilon \} = 0.$$ 

Proof. Theorem 3.3, Lemma 3.7 and the corollary to Lemma 3.9 give the asymptotic expansion

$$\sqrt{N} \hat{\mu}_R = -\mu'_d \sqrt{N}(\hat{\beta}_R - \beta) + \gamma^{-1}_L U_I(0 : F_N) + o_p(1).$$

Clearly these lemmas and this theorem also are applicable when the $\{\sigma_i\}$ are known. Thus, we also have the asymptotic expansion

$$\sqrt{N} \hat{\mu}_{\text{opt}} = -\mu'_d \sqrt{N}(\hat{\beta}_{\text{opt}} - \beta) + \gamma^{-1}_L U_I(0 : F_N) + o_p(1).$$

Combining the two asymptotic expansions proves the theorem.

3.4 Estimation of $\beta$ in the No-intercept Model

In this section we obtain the properties of the estimate of $\beta$ in the true, no-intercept model.

Lemma 3.10. Under Assumptions F1 through F3, B1 through B5, L1 and L2,

$$\begin{bmatrix} \sqrt{N} \hat{\mu}_R \\ \sqrt{N}(\hat{\beta}_R - \beta) \end{bmatrix} \overset{D}{\to} N(0, \gamma^{-2}_L \lambda^{-1}).$$
Proof.

\[
\begin{bmatrix}
\sqrt{N} \hat{\mu}_R \\
\sqrt{N} (\hat{\beta}_R - \beta)
\end{bmatrix} = \begin{bmatrix}
\sqrt{N} (\hat{\mu}_R - \hat{\mu}_{opt}) \\
\sqrt{N} (\hat{\beta}_R - \hat{\beta}_{opt})
\end{bmatrix} + \begin{bmatrix}
\sqrt{N} \hat{\mu}_{opt} \\
\sqrt{N} (\hat{\beta}_{opt} - \beta)
\end{bmatrix}
\]

The first term of the sum is \( o_p(1) \) by Theorem 3.2 and Theorem 3.4. McKean and Hettmansperger (1978) prove that

\[
\begin{bmatrix}
\sqrt{N} \hat{\mu}_{opt} \\
\sqrt{N} (\hat{\beta}_{opt} - \beta)
\end{bmatrix} \overset{D}{\rightarrow} N(0, \gamma_L^{-2} \lambda^{-1}).
\]

Thus, the lemma is proven.

Lemma 3.11. Under assumptions F1 through F3, B1 through B5, L1 and L2,

\[
\sqrt{N}(\hat{\beta}_{opt} - \beta) \overset{D}{\rightarrow} N(0, \gamma_L^{-2} \Sigma^{-1}).
\]

Proof.

\[
\sqrt{N}(\hat{\beta}_{opt} - \beta) = [(D' D)^{-1} D' 1_N : 1] \begin{bmatrix}
\sqrt{N} \hat{\mu}_{opt} \\
\sqrt{N} (\hat{\beta}_{opt} - \beta)
\end{bmatrix}
\]

Using the identity

\[
([1 : D]' [1 : D])^{-1} = \begin{pmatrix}
N^{-1} & 0 \\
0 & 0
\end{pmatrix} + \begin{pmatrix}
-N^{-1} 1_N D \\
D' (I - N^{-1} 1_N 1_N') D
\end{pmatrix}
\]

and the limiting distribution result of McKean and Hettmansperger (1978) cited in Lemma 3.10 proves this theorem.
Theorem 3.5. Under Assumptions F1 through F3, B1 through B5, L1 and L2, for all $\varepsilon > 0$,

$$\lim_{N \to \infty} P \left\{ \sqrt{N} \left| \hat{\beta}_{opt} - \hat{\beta}_{R} \right| \geq \varepsilon \right\} = 0.$$

Proof.

$$\sqrt{N}(\hat{\beta}_{opt} - \hat{\beta}_{R})$$

$$= \sqrt{N} \left\{ \beta_{opt} + (D'D)^{-1}D'1N \mu_{opt} - \beta_R - (D' \hat{D})^{-1} \hat{D}'1N \tilde{\mu}_R \right\}$$

$$= \sqrt{N}(\hat{\beta}_{opt} - \hat{\beta}_{R}) + (N^{-1}D'D)^{-1}D'1N^{-1} \sqrt{N}(\mu_{opt} - \tilde{\mu}_R)$$

$$+ \sqrt{N} \tilde{\mu}_R[(N^{-1}D'D)^{-1}D'1N^{-1} - (N^{-1}D' \hat{D})^{-1} \hat{D}'1N^{-1}]$$

The first term of the sum is $o_p(1)$ by Theorem 3.2. Assumptions B1 implies $(N^{-1}D'D)^{-1}$ and $D'1N^{-1}$ converge. Thus, Theorem 3.4 implies the second term of the sum is $o_p(1)$. Since Lemma 3.7 implies $\sqrt{N} \tilde{\mu}_R$ is bounded in probability, the theorem will be proven if we show

$$(N^{-1}D'D)^{-1}D'1N^{-1} - (N^{-1}D' \hat{D})^{-1} \hat{D}'1N^{-1}$$

converges in probability to zero. We will proceed by examining the components of this term.

$$D'1N^{-1} - \hat{D}'1N^{-1} = -N^{-1} \sum_{i=1}^{N} \frac{x_i}{\sigma_i} \alpha_i^{(2)}(\hat{\Delta})$$

$$\leq N^{-1} \sum_{i=1}^{N} \| \frac{x_i}{\sigma_i} \| |\alpha_i^{(2)}(\hat{\Delta})|.$$
Now, \( N^{-1} \sum_{i=1}^{N} \| \frac{x_i}{\sigma_i} \| \) is bounded (Assumption B3 and B4) and \( \alpha_i^{(2)}(\hat{\Lambda}) = o_p(1) \). Thus \( \hat{D}'_1N^{-1} - \hat{D}'_1N^{-1} = o_p(1) \). Now, consider the \((j,k)\)th element of \([N^{-1}D'D - N^{-1}\hat{D}'\hat{D}]\).

\[
[N^{-1}D'D - N^{-1}\hat{D}'\hat{D}]_{jk} = N^{-1} \sum_{i=1}^{N} \frac{x_{ij}}{\sigma_i} \frac{x_{ik}}{\sigma_i} [1 - (1 + \alpha_i^{(2)}(\hat{\Lambda}))^2].
\]

\( N^{-1} \sum_{i=1}^{N} \frac{x_{ij}}{\sigma_i} \frac{x_{ik}}{\sigma_i} \) is bounded (Assumption B3 and B4) and \( \alpha_i^{(2)}(\hat{\Lambda}) = o_p(1) \). Thus, \((N^{-1}D'D) - (N^{-1}\hat{D}'\hat{D}) = o_p(1)\) and consequently

\[
(N^{-1}D'D)^{-1} - (N^{-1}\hat{D}'\hat{D})^{-1} = o_p(1).
\]

By Assumption B1

\[
N^{-1}D'D \rightarrow \Sigma \quad \text{and} \quad N^{-1}D'1_N \rightarrow \mu_d.
\]

Combining these results gives

\[
(N^{-1}D'D)^{-1}D'1_NN^{-1} - (N^{-1}\hat{D}'\hat{D})^{-1}\hat{D}'1_NN^{-1} = o_p(1).
\]

Thus the theorem is proven.
Corollary. Under the assumptions of Theorem 3.5,

\[ \sqrt{N}(\hat{\beta}_R - \beta) \xrightarrow{D} N(0, \gamma_L^{-2} \Sigma^{-1}) \]

This is an obvious consequence of Theorem 3.5 and Lemma 3.11.

3.5 Estimation of the Parameter \( \gamma_L \)

In this section, we discuss an estimation of \( \gamma_L \) based on the methods of Lehmann (1963) and McKean and Hettmansperger (1976). Essentially, the results of McKean and Hettmansperger are extended to the case that the weights are estimated.

For a given \( \delta, 0 < \delta < 1 \), define

\[ \hat{\mu}_S = \sup \{ \mu : U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu) : F_N)\hat{\sigma}_L^{-1} > -Z_{\delta/2} \} \]

and

\[ \hat{\mu}_M = \inf \{ \mu : U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \mu) : F_N)\hat{\sigma}_L^{-1} < Z_{\delta/2} \}, \]

where

\[ \hat{\sigma}_L = \{1 + N^{-1}1_N'\hat{D}(N^{-1}\hat{D}'(I - N^{-1}1_N1_N')\hat{D})^{-1}\hat{D}'1_NN^{-1}\}, \]

\[ 1 - \Phi(Z_{\delta/2}) = \delta/2 \quad \text{and} \quad \Phi \quad \text{is the cumulative distribution function of a standard normal.} \]

Define the estimate of \( \gamma_L \):

\[ \hat{\gamma}_L = 2 Z_{\delta/2} \hat{\sigma}_L(\hat{\mu}_S - \hat{\mu}_M)^{-1} N^{-1/2}. \]

Now, we must prove \( \hat{\gamma}_L \) is a consistent estimate of \( \gamma_L \).
Lemma 3.12. Under Assumptions B2 through B5, for all $\varepsilon > 0$,

$$\lim_{N \to \infty} P \left\{ \| N^{-1} N' \hat{D} (N^{-1} \hat{D}' (I - N^{-1} 1_N 1_N') \hat{D})^{-1} \hat{D}' 1_N N^{-1} - \mu_d' S^{-1} \mu_d \| \geq \varepsilon \right\} = 0.$$

Proof. We should first note that in the proof of Theorem 3.5 we demonstrated

$$D'1_N N^{-1} - \hat{D}'1_N N^{-1} = o_p(1) \quad \text{and} \quad N^{-1} D'D - N^{-1} \hat{D}' \hat{D} = o_p(1).$$

Also note, by Assumption B1

$$D'1_N N^{-1} \to \mu_d, \quad N^{-1} D'D \to \Sigma$$

and

$$N^{-1} D'(I - N^{-1} 1_N 1_N')D \to S.$$

These facts clearly imply

$$\hat{D}'1_N N^{-1} \xrightarrow{p} \mu_d \quad \text{and} \quad N^{-1} \hat{D}' \hat{D} \xrightarrow{p} \Sigma.$$

Thus, we only need consider

$$N^{-1} \hat{D}'(I - N^{-1} 1_N 1_N')\hat{D} = N^{-1} D'(I - N^{-1} 1_N 1_N')D$$

$$+ \left[ N^{-1} \hat{D}' \hat{D} - N^{-1} D'D \right]$$

$$- \left[ N^{-1} \hat{D}'1_N 1_N' \hat{D} N^{-1} - N^{-1} D'1_N 1_N' DN^{-1} \right].$$

Clearly the second and third components are $o_p(1)$ by the above comments.

Now, since $N^{-1} D'(I - N^{-1} 1_N 1_N')D \to S$, we have

$$N^{-1} \hat{D}'(I - N^{-1} 1_N 1_N')\hat{D} \xrightarrow{p} S.$$ Combining these results proves the lemma.
Lemma 3.13. Under the assumptions of Theorem 3.3, and there exists a \( C > 0 \) such that

\[
\lim_{N \to \infty} P[\sqrt{N} |\hat{\mu}_S| < C] = 1 \quad \text{and} \quad \lim_{N \to \infty} P[\sqrt{N} |\hat{\mu}_M| < C] = 1.
\]

We will first prove \( \sqrt{N} \hat{\mu}_S \) is bounded above in probability. Let \( \varepsilon > 0 \) be given. Lemma 3.6 and 3.12 imply there exists a \( C > 0 \) such that if \( N \) is large enough \( P[U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \hat{\sigma}_L^{-1} > C] < \frac{\varepsilon}{3} \). Define \( C' = \gamma_L^{-1} \sigma_L (C + \varepsilon + Z_{\delta/2}) \). Note that by the definition of \( \hat{\mu}_S \)

\[
\sqrt{N} \hat{\mu}_S > C' \implies U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \hat{\sigma}_L^{-1} \geq -Z_{\delta/2}.
\]

Thus

\[
P\left[\sqrt{N} \hat{\mu}_S > C'\right] \leq P \left[U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \hat{\sigma}_L^{-1} \geq Z_{\delta/2}\right]
\]

\[
\leq P \left[ |U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \hat{\sigma}_L^{-1} - U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \hat{\sigma}_L^{-1} + C' \gamma_L \sigma_L^{-1} | \geq \varepsilon \right]
\]

\[
+ P \left[ U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \hat{\sigma}_L^{-1} - C' \gamma_L \sigma_L^{-1} \geq -Z_{\delta/2} - \varepsilon \right]
\]

\[
\leq P \left[ \hat{\sigma}_L^{-1} |U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, C') : F_N) - U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \right] C' \gamma_L \sigma_L^{-1} \geq \varepsilon /2\right]
\]

\[
+ P \left[ |C' \gamma_L | \sigma_L^{-1} - \hat{\sigma}_L^{-1} | \geq \varepsilon /2\right]
\]

\[
+ P \left[ \hat{\sigma}_L^{-1} U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N) \geq C \right].
\]
The first component of the sum can be made smaller than $\varepsilon/3$ for sufficiently large $N$ by Lemma 3.5 and Lemma 3.12. The second term can also be made less than $\varepsilon/3$ by Lemma 3.12 and since $C'$ and $\gamma_L$ are constants. The third term is shown to be less than $\varepsilon/3$ for sufficiently large $N$ above.

A similar argument could be made to show $\sqrt{N} \hat{\mu}_M$ is bounded below in probability. Then, using the fact $\hat{\mu}_M \leq \hat{\mu}_S$, the lemma is proven.

Lemma 3.14. Under the assumptions of Theorem 3.3, for all $\varepsilon > 0$

$$\lim_{N \to \infty} P[|Z_{\delta/2} - U_1((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N)\sigma_L^{-1} | \geq \varepsilon] = 0$$

and

$$\lim_{N \to \infty} P[|Z_{\delta/2} + U_1((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_S) : F_N)\sigma_L^{-1} | \geq \varepsilon] = 0.$$ 

Proof. We will begin by proving the first statement of the lemma. First note by right continuity $Z_{\delta/2} - U_1((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N)\sigma_L^{-1} \geq 0$. Let $\varepsilon > 0$ be given and define $z = \frac{\varepsilon}{2\gamma_L} \sigma_L$. By Lemma 3.13, $\sqrt{N} \hat{\mu}_M$ and $\sqrt{N} \hat{\mu}_M - z$ are bounded in probability. Thus,

$$A_N = \{|\sqrt{N}\hat{\mu}_M| \leq C \text{ and } \sqrt{N}|\hat{\mu}_M - zN^{-1/2}| \leq C\}$$
converges in probability to 1 for some $C > 0$. Now,

$$P[|Z_{t/2} - U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N)\hat{\sigma}_L^{-1}| \geq \varepsilon]$$

$$= P[Z_{t/2} - U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N)\hat{\sigma}_L^{-1} \geq \varepsilon]$$

$$\leq P\left[-U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M - z) : F_N)\hat{\sigma}_L^{-1}ight]$$

$$\leq P(A_N^c) + P\left[U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M - z) : F_N)\hat{\sigma}_L^{-1} - \varepsilon/2 \geq \varepsilon/2 \text{ and } A_N\right]$$

$$\leq P(A_N^c) + P\left\{U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M - z) : F_N)\right\}$$

$$- U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N) - \gamma_L|\hat{\sigma}_L^{-1} - \varepsilon/4 \text{ and } A_N\right\}$$

$$+ P\left\{\gamma_L|\hat{\sigma}_L^{-1} - \hat{\sigma}_L^{-1}| \geq \varepsilon/2 \text{ and } A_N\right\}.$$ 

Hence, by the same arguments as in the previous lemma

$$\lim_{N \to \infty} P[|Z_{t/2} - U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N)\hat{\sigma}_L^{-1}| \geq \varepsilon] = 0.$$ 

The second statement of the lemma could be similarly proven.

Theorem 3.6. Under the assumptions of Theorem 3.3, for all $\varepsilon > 0$,

$$\lim_{N \to \infty} P[|\gamma_L - \hat{\gamma}_L| \geq \varepsilon] = 0.$$ 

Proof. Lemmas 3.5 and 3.13 imply

$$\lim_{N \to \infty} P[|U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_S) : F_N)$$

$$- U_l((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N} \hat{\mu}_M) : F_N) + N^{1/2}(\hat{\mu}_S - \hat{\mu}_M)\gamma_L| \geq \varepsilon] = 0.$$
Combining this result with Lemmas 3.12 and Lemma 3.14 results in:

\[
\lim_{N \to \infty} P[| -2 Z_{\delta/2} + N^{1/2}(\hat{\mu}_S - \hat{\mu}_M)\gamma_L \hat{\sigma}_L^{-1}| \geq \varepsilon] = 0.
\]

Now, by the definition of \( \gamma_L \) we know

\[
N^{1/2}(\hat{\mu}_S - \hat{\mu}_M)\hat{\sigma}_L^{-1} = 2 Z_{\delta/2} \gamma_L^{-1}.
\]

Substituting this into the previous probability statement results in:

\[
\lim_{N \to \infty} P[|2 Z_{\delta/2}| |1 - \gamma_L \hat{\gamma}_L^{-1}| \geq \varepsilon] = 0.
\]

Thus the theorem is proven.

One potential problem with this method of estimating \( \gamma_L \) is the case that \( U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N}\mu) : F_N) \hat{\sigma}_L^{-1} \) does not attain values above \( Z_{\delta/2} \) and/or values below \(-Z_{\delta/2}\). In this case it is preferable to using the following alternative method.

For a given \( \delta, 0 < \delta < 1 \), define

\[
\hat{\mu}_S^{(0)} = \sup \left\{ \mu : U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N}\mu) : F_N) > -Z_{\delta/2} \right\}
\]

and

\[
\hat{\mu}_M^{(0)} = \inf \left\{ \mu : U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, \sqrt{N}\mu) : F_N) < Z_{\delta/2} \right\}.
\]

Define the estimate of \( \gamma_L \):

\[
\hat{\gamma}_0 = 2Z_{\delta/2}(\hat{\mu}_S^{(0)} - \hat{\mu}_M^{(0)})^{-1}N^{-1/2}.
\]
Lemma 3.15. Under the assumptions of Theorem 3.3, for all $\varepsilon > 0$,

$$\lim_{N \to \infty} P[|\hat{\gamma}_0 - \gamma_0| \geq \varepsilon] = 0.$$ 

Proof. The proof follows exactly as that for the consistency of $\hat{\gamma}_L$ in McKean and Hettmansperger (1976) as a consequence of Lemma 3.5 and Lemma 3.6.

Although consistency of $\hat{\gamma}_0$ is established, there is little control of the bandwidth since $U_I((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3, 0) : F_N)$ does not converge to a standard normal random variable. Hence it is preferable to use the first method presented if it is possible.
CHAPTER IV

ESTIMATION OF THE VARIANCE PARAMETER \( \theta \)

4.1 Introduction

In this chapter we will discuss estimation of the scale parameter \( \theta \). To motivate our estimates, we will consider the scale rank tests discussed in Hajek and Sidak (1967).

Suppose \( e_1, \cdots, e_N \) are independent random variables from a population with likelihood \( \prod_{i=1}^{N} f(e_i \epsilon^{-(c_i - \bar{c})\epsilon}) \epsilon^{-(c_i - \bar{c})\epsilon} \), \( f \) absolutely continuous and \( \int_{-\infty}^{\infty} |xf'(x)|dx < \infty \). Hajek and Sidak (1967) consider testing hypotheses of the form \( H_0 : \epsilon = \theta \), versus \( H_A : \epsilon > \theta \). First we define the families of densities \( p_0 = \prod_{i=1}^{N} f(e_i) \) and \( q_\epsilon = \prod_{i=1}^{N} f(e_i \epsilon^{-(c_i - \bar{c})\epsilon}) \epsilon^{-(c_i - \bar{c})\epsilon} \). Then the hypothesis can equivalently be stated as testing that the joint distribution of \( (e_1, \cdots, e_N) \) belongs to \( p_0 \) against the alternative that it is an element of \( \{q_\epsilon : \epsilon > \theta \} \). They proved that there exists a sequence of scores \( a_N(\cdot) \) to every scale alternative \( q_\epsilon \) so that the critical region \( \sum_{i=1}^{N} (e_i - \bar{e})a_N(R(e_i)) \geq k \) gives the locally most powerful rank test for testing homoscedasticity \( (p_0) \) against \( \{q_\epsilon : \epsilon > \theta \} \) at the respective level.

They prove that this optimal sequence of scores, \( a_N(1), \cdots, a_N(N) \),
are generated by the score function

\[ \phi_1(u, f) = -1 - F^{-1}(u) \frac{f'(F^{-1}(u))}{f(F^{-1}(u))}. \]

Further, let \( u^{(1)}_N < \cdots < u^{(N)}_N \) be an ordered sample from a uniform distribution on \([0, 1]\). Also suppose \( e^{(1)}_N < \cdots < e^{(N)}_N \) is an ordered sample from a population with density \( f \). The sequence of scores \( a_N(1), \ldots, a_N(N) \) is given by

\[ a_N(i) = E[\phi_1(u^{(i)}_N, f)] = E \left\{ -1 - \frac{e^{(i)}_N f'(e^{(i)}_N)}{f(e^{(i)}_N)} \right\}. \]

Before defining our estimates we will consider several specific examples of scale rank tests.

Example 1. In the case that the errors are normally distributed, the sequence of scores that give the locally most powerful rank test for the hypothesis is given by

\[ a(i) = E[\phi_1(u^{(i)}_N, f)] \text{ where } \phi_1(u, f) = \left( \Phi^{-1}(u) \right)^2 - 1 \]

and \( \Phi \) is the standard normal distribution function.

The Klotz test is based on the approximate optimal scores for the normal distribution. It employs the statistic

\[ K = \sum_{i=1}^{N} (c_i - \bar{c}) \left[ \Phi^{-1} \left( \frac{R_i}{N + 1} \right) \right]^2, \]
where $R_i$ is the rank of $e_i$ among $e_1, \ldots, e_N$. This is a scale analogue of the van der Waerden test and is asymptotically optimum for the normal distribution.

Example 2. The Ansari-Bradley statistic,

$$ A = \sum_{i=1}^{N} (c_i - \bar{c}) \left[ \frac{1}{2} (N + 1) - |R_i - \frac{1}{2} (N + 1)| \right], $$

employs the approximate optimal scores, $\phi(u) = 2|2u - 1| - 1$, for the problem associated with the density $f(x) = \frac{1}{2} (1 + |x|)^{-2}$.

Example 3. Mood's scale statistic,

$$ M = \sum_{i=1}^{N} (c_i - \bar{c}) \left[ R_i - \frac{(N + 1)}{2} \right]^2 $$

is derived from scores of the form $\phi(u) = (u - \frac{1}{2})^2$.

One common property of these three examples is that the score functions used are even about $\frac{1}{2}$. In fact, under symmetry of $f$, the optimal score function for testing $H_0$ versus $\{ \varepsilon \in [0, \infty) : \varepsilon \geq \theta \}$ is even about $\frac{1}{2}$.

For the two-sample scale problem, Bauer (1972) suggests using a score function which satisfies

$$ a_N(k) \leq a_N(k + 1) \quad \text{for} \quad k \leq k_0 \quad \text{and} $$

$$ a_N(k) \geq a_N(k + 1) \quad \text{for} \quad k > k_0. $$
Considering these scale statistics indicates selecting a scale score function which is either non-increasing on \([0, \frac{1}{2}]\) or non-decreasing on \([0, \frac{1}{2}]\). However, for our proofs, it is necessary to select one. Thus, we define the scale score function as follows. Assume the scale score function, \(\phi_s\), is even about \(\frac{1}{2}\), non-decreasing on \([\frac{1}{2}, 1]\) and set \(a_s(i) = \phi_s\left(\frac{i}{N+1}\right)\). Without loss of generality assume

\[
\int_0^1 \phi_s(u)du = 0 \quad \text{and} \quad \int_0^1 \phi_s^2(u)du = 1.
\]

To motivate our estimate, suppose \(\beta\) is known. For our model \(e_i = \frac{y_i-x_i^T\beta}{\sigma_i}\) and \(\sigma_i = \exp[h(x_i^T\beta)\theta]\). Thus Hajek and Sidak’s (1967) previously stated result suggests using a statistic of the form

\[
S(\varepsilon) = N^{-1/2} \sum_{i=1}^{N} (h(x_i^T\beta) - \overline{h})a_s(R(\frac{y_i-x_i^T\beta}{e^{x_i^T\theta}})),
\]

where \(\overline{h} = N^{-1} \sum_{i=1}^{N} h(x_i^T\beta)\). Under \(H_0\), the expectation of \(S(\varepsilon)\) is zero. This suggests inverting the test statistic and estimating \(\varepsilon\) by solving \(S(\varepsilon) = 0\).

Clearly we are interested in the distribution of the test statistic under the alternative. Now, if \(I_1(f) < \infty\), the fact \(H_N = N^{-1} \sum_{i=1}^{N} (h(x_i^T\beta) - \overline{h})^2 < \infty\) (Assumption B3) implies \(q_{\varepsilon}\) is a contiguous alternative to \(p_0\). Thus, Theorem VI 2.4 of Hajek and Sidak (1967) implies that under \(H_0\)

\[
S(\alpha) \xrightarrow{D} N(-\sqrt{N}(a - \theta)H_N\gamma_s, H_N), \quad \text{where} \quad \gamma_s = \int_0^1 \phi_s(u)\phi_1(u, f)du
\]
and \( \phi_1(u, f) = -1 - F^{-1}(u) \frac{f'(F^{-1}(u))}{f(F^{-1}(u))} \). For the moment assume \( S(a) \) is monotone in \( a \).

Define
\[
\hat{\theta}^*_L = \inf \{ a : S(a) < 0 \}
\]
and
\[
\hat{\theta}^*_U = \sup \{ a : S(a) > 0 \}.
\]
The Hodges-Lehmann type estimate of \( \theta \) is \( \hat{\theta}^* = \frac{\hat{\theta}^*_L + \hat{\theta}^*_U}{2} \). Following the method of Hodges and Lehmann (1963) we can get the following results:

(a) The distribution of \( \hat{\theta}^*_L \) and \( \hat{\theta}^*_U \) is continuous, since the distribution of \( e_1, \ldots, e_N \) is continuous.

(b) \( P[S(a) < 0] \leq P[\hat{\theta}^* < a] \leq P[S(a) \leq 0] \) for any real \( a \).

(c) \( \lim_{N \to \infty} P_\theta[\sqrt{N}(\hat{\theta}^* - \theta) \leq a] = \lim_{N \to \infty} P_\theta[S(\theta + a/\sqrt{N}) \leq 0] \)

Combining part c and the limiting distribution result of \( S(a) \) imply that \( \sqrt{N}(\hat{\theta}^* - \theta) \) is asymptotically normal

\[
(0, H_N^{-1} \gamma_2^2).
\]

A necessary condition for monotonicity in \( \varepsilon \) of the function the RHS of Equation (4.2) it that the median of \( y_1 - x'_1 \beta, \ldots, y_N - x'_N \beta \) is zero (see Lemma 4.6). Although symmetry of \( f \) implies the population medians of \( y_1 - x'_1 \beta, \ldots, y_N - x'_N \beta \) are zero, this does not imply the median of the sample is zero. In order to accommodate for this problem, we will make an adjustment; define \( \tilde{\alpha}^* = \text{median} \{ y_i - x'_i \beta \} \). Then the median of the "adjusted" residuals \( y_1 - x'_1 \beta - \tilde{\alpha}^*, \ldots, y_N - x'_N \beta - \tilde{\alpha}^* \) is zero.
From our linearity results, we demonstrate that making this adjustment to the residuals does not affect the limiting distribution of the estimate of $\theta$. Thus if $\beta$ is known, our rank-based estimate of $\theta$ solves

\[
N^{-1/2} \sum_{i=1}^{N} (h(x_i') - \bar{h}) a \left( \frac{y_i - x_i' \hat{\beta} - \hat{\alpha}^*}{e^{\beta h(x_i')}} \right) = 0.
\]

We denote the unknown, "optimal" Hodges-Lehmann (1963) type estimate as $\hat{\theta}_{opt}$. In Section 4.4, we will show that the function on the LHS of (4.3) is a non-increasing step function in $\theta$.

We have discussed the "optimal" case, where the true $\beta$ is known. However, in practice, $\beta$ is not known. Thus, we will utilize an estimate $\hat{\beta}$ which satisfies $\sqrt{N}(\hat{\beta} - \beta) = O_p(1)$. Substituting $\hat{\beta}$ into Equation 4.3 suggests estimating $\theta$ by solving

\[
N^{-1/2} \sum_{i=1}^{N} (h(x_i') - \bar{h}) a \left( \frac{y_i - x_i' \hat{\beta} - \hat{\alpha}}{e^{\beta h(x_i')}} \right) = 0,
\]

where $\bar{h}^* = N^{-1} \sum_{i=1}^{N} h(x_i')$ and $\hat{\alpha} = \text{median} \{y_i - x_i' \hat{\beta}\}$.

We establish in Section 4.3 that $\hat{\alpha}$ is a $N^{1/2}$-consistent estimate of $\alpha$. Using the general linearity results in Chapter II and Theorem 7.1 of Carroll and Ruppert (1982a), we establish linearity results for our scale rank statistic in Section 4.2. Furthermore, in Section 4.4 we establish that the function on the LHS of Equation 4.4 is a non-increasing step function in $\theta$.

Thus, define $\hat{\theta}$ to be the Hodges-Lehmann (1963) type estimate of $\theta$ based
on Equation 4.4. In Section 4.4, we prove that if \((\hat{\alpha}, \hat{\beta})\) are \(N^{1/2}\)-consistent estimates of \((0, \beta)\), then the Hodges-Lehmann (1963) type estimate of \(\theta\) is a \(N^{1/2}\)-consistent estimate. We also establish asymptotic normality of the Hodges-Lehmann type estimate \(\theta\) for the case that the estimate of \(\beta\) is the adjusted rank-based estimate. In this case

\[
\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{D} N(0, H_N^{-1} \gamma_s^{-2} + \theta^2 \gamma_s^{-2} L' \Sigma^{-1} L H_N^{-2}).
\]

Although we establish the limiting distribution of our estimate, it is not asymptotically equivalent to the "optimal" estimate with the true \(\beta\). Using the same methodology, it can be shown that

\[
\sqrt{N}(\hat{\theta}_{opt} - \theta) \xrightarrow{D} N(0, H_N^{-1} \gamma_s^{-2}).
\]

Using a method similar to McKean and Hettmansperger (1976), a consistent estimate of \(\gamma_s\) is developed in Section 5.

4.2 Linearity

In order to use the results in Chapter II we need some additional notation. For \(\Delta_1\) in \(R^p\), \(\Delta_2\) and \(\Delta_3\) in \(R^1\), and \(\Delta = (\Delta_1, \Delta_2, \Delta_3)\), define
\[ h_i(\Delta) = h(x_i'\beta + x_i' \Delta_1 N^{-1/2}) \]

\[ c_i = h_i(0) - N^{-1} \sum_{j=1}^{N} h_j(0), \]

\[ \alpha_i^{(1)}(\Delta) = N^{-1/2} \frac{x_i' \Delta_1}{\sigma_i} + N^{-1/2} \frac{1}{\sigma_i} \Delta_3 \]

\[ \alpha_i^{(2)}(\Delta) = \exp[-h_i(\Delta)\Delta_2 N^{-1/2} + (h_i(0) - h_i(\Delta))\theta] - 1 + N^{-1/2} \frac{1}{\sigma_i} \Delta_3 \]

\[ \alpha_i^{(3)}(\Delta) = [h_i(\Delta) - h_i(0)] - N^{-1} \sum_{j=1}^{N} [h_j(\Delta) - h_j(0)]. \]

Next define the processes

\[ U_{\Delta}(\Delta : F_N) = N^{-1/2} \sum_{i=1}^{N} (c_i + \alpha_i^{(3)}(\Delta))\phi_s \left( F_N[(c_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta)) : \Delta]\right) \]

and

\[ U_{s}(\Delta : F) = N^{-1/2} \sum_{i=1}^{N} (c_i + \alpha_i^{(3)}(\Delta))\phi_s \left( F[(c_i - \alpha_i^{(1)}(\Delta))(1 + \alpha_i^{(2)}(\Delta))]ight). \]

\( F_N(x : \Delta) \) and \( F(x : \Delta) \) were defined in Chapter II. We noted in Section 3.2.1 that \( F_N(x : 0) \) is the empirical distribution function of the errors \( e_i \) and that \( F(x : 0) = F(x) \). Note that \( \tilde{\Delta} = (\sqrt{N}(\hat{\beta} - \beta), \sqrt{N}(\hat{\theta} - \theta), \sqrt{N}(\hat{\alpha})) \), where \( \hat{\beta} \) is a \( N^{1/2} \)-consistent estimate of \( \beta \), \( \hat{\alpha} = \text{median} \{y_i - x_i'\hat{\beta}\} \) and \( \hat{\theta} \) is the rank based estimate of \( \theta \) obtained by solving Equation (4.4). Thus \( F_N(X; \tilde{\Delta}) \) is the empirical distribution.
function of the estimated residuals \( \hat{e}_i = \frac{y_i - x_i' \hat{\beta} - \hat{\alpha}}{e^{\hat{h}(x_i')}} \). Substituting the rank based estimate, \( \hat{\theta} \), into equation (4.4), it can be written as

\[
U_3(\hat{\Delta} : F_N) = 0.
\]

**Lemma 4.1.** Under Assumptions F1 through F3, S1, S2 and B1 through B5, for all \( M > 0 \) and \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P \left\{ \sup_{\|\Delta\| \leq M} \| U_3(\Delta : F) - U_3(0 : F) \| + \gamma_s N^{-1/2} \sum_{i=1}^{N} (h_i(0) - \bar{h}) \alpha_i^{(2)}(\Delta) \| \geq \varepsilon \right\} = 0,
\]

where

\[
\gamma_s = \int_0^1 \phi_s(u) \phi_1(u, f) du, \phi_1(u, f) = -1 - F^{-1}(u) \frac{f'(F^{-1}(u))}{f(F^{-1}(u))}
\]

and \( \bar{h} = N^{-1} \sum_{i=1}^{N} h_i(0) \).

**Proof.** Letting

\[
K_i = N^{-1/2} \left\{ \| x_i \| + \| h_i(0) \| + \frac{1}{\sigma_i} + N^{-1} \sum_{j=1}^{N} \| h_j^2(0) \| + N^{-1} \sum_{j=1}^{N} \| x_j \|^2 \right\},
\]

the conditions of Theorem 7.1 of Carroll and Ruppert (1982a) are implied by the listed assumptions, so the lemma is established.
Corollary. Under the assumptions of the lemma, for all \( \varepsilon > 0 \) and \( M > 0 \),

\[
\lim_{N \to \infty} P \left\{ \sup_{\|\Delta\| \leq M} \left\| U_s(\Delta : F') - U_s(0 : F') + \gamma_s H_N \Delta_2 - \gamma_s G_N(\Delta) \right\| \geq \varepsilon \right\} = 0,
\]

where \( H_N = N^{-1} \sum_{i=1}^{N} (h_i(0) - \bar{h})^2 \) and

\[
G_N(\Delta) = N^{-1/2} \sum_{i=1}^{N} (h_i(0) - h_i(\Delta))(h_i(0) - \bar{h}) \theta.
\]

Proof. Clearly the corollary will be proven if we establish

\[
\sup_{\|\Delta\| \leq M} \left\| N^{-1/2} \sum_{i=1}^{N} (h_i(0) - \bar{h}) \alpha_i^{(2)}(\Delta) - H_N \Delta_2 + G_N(\Delta) \right\| = o_p(1).
\]

Define \( B_i(\Delta) = -h_i(\Delta) \Delta_2 N^{-1/2} + (h_i(0) - h_i(\Delta)) \theta \). We noted in the proof of Lemma 2.1 that \( B_i(\Delta) \xrightarrow{P} 0 \) uniformly in \( i = 1, \ldots, N \) and \( \|\Delta\| \leq M \). By a Taylor series expansion,

\[
\alpha_i^{(2)}(\Delta) = B_i(\Delta) e^{A_i(\Delta)}, \quad \text{where} \quad \|A_i(\Delta)\| \leq |B_i(\Delta)|.
\]

Now,
\[ N^{-1/2} \sum_{i=1}^{N} (h_i(0) - \overline{h_i}) \alpha_i^{(2)}(\Delta) \]
\[
= N^{-1/2} \sum_{i=1}^{N} (h_i(0) - \overline{h_i}) \left[ -h_i(0) \Delta_2 N^{-1/2} + (h_i(0) - h_i(\Delta)) \theta \right. \\
+ \left. (h_i(0) - h_i(\Delta)) \Delta_2 N^{-1/2} \right] e^{A_i(\Delta)} \\
= [ -H_N \Delta_2 + G_N(\Delta) ] e^{A_i(\Delta)} \\
+ N^{-1} \sum_{i=1}^{N} (h_i(0) - \overline{h_i})(h_i(0) - h_i(\Delta)) \Delta_2 e^{A_i(\Delta)}. \]

By Assumptions B3 and B5,

\[ N^{-1} \sum_{i=1}^{N} (h_i(0) - \overline{h_i})(h_i(0) - h_i(\Delta)) \xrightarrow{P} 0 \quad \text{uniformly in } \|\Delta\| \leq M. \]

Therefore, since \(|A_i(\Delta)| \xrightarrow{P} 0\) uniformly in \(i = 1, \ldots, N\) and \(\|\Delta\| \leq M\), the corollary is established.

Theorem 4.1. Under the assumptions of Lemma 4.1, for all \(M > 0\) and \(\varepsilon > 0\),

\[
\lim_{N \to \infty} P \left\{ \sup_{\|\Delta\| \leq M} \|U_s(\Delta : F_N) - U_s(0 : F_N) \right. \\
+ \gamma_s \Delta_2 H_N - \gamma_s G_N(\Delta) \| \geq \varepsilon = 0. 
\]
Proof.

\[
\sup_{\|\Delta\| \leq M} \| U_s(\Delta : F_N) - U_s(0 : F_N) + \gamma_S \Delta_2 H_N - \gamma_S G_N(\Delta) \|
\]

\[
\leq \sup_{\|\Delta\| \leq M} \| U_s(\Delta : F) - U_s(0 : F) + \gamma_S \Delta_2 H_N - \gamma_S G_N(\Delta) \|
\]

\[
+ \sup_{\|\Delta\| \leq M} \| U_s(\Delta : F_N) - U_s(\Delta : F) \| + \| U_s(0 : F_N) - U_s(0 : F) \|
\]

Convergence of the first 2 components of the sum is implied by the corollary to Lemma 4.1 and Theorem 2.1 respectively. Assumptions B2 and B3 imply

\[
\frac{\max_{1 \leq i \leq N} (h_i(0) - \bar{h})^2 / N}{N^{-1} \sum_{j=1}^{N} (h_j(0) - \bar{h})^2} \longrightarrow 0.
\]

Therefore, the proof of Theorem V 1.5a of Hajek and Sidak (1967) yields:

\[
U_s(0 : F_N) - U_s(0 : F) \overset{P}{\rightarrow} 0.
\]

Thus the lemma is established.

4.3 Adjustment of the Residuals

Assume \( \hat{\beta} \) is a \( N^{1/2} \)-consistent estimate of \( \beta \). In this section, we establish \( \hat{\alpha} = \text{median } \{ y_i - x_i' \hat{\beta} \} \) is a \( N^{1/2} \)-consistent estimate of zero using an argument similar to Aubuchon (1982). In Section 4.1 we discussed the need of the estimate \( \hat{\alpha} \); it is necessary to obtain monotonicity of the
process \( U_s(\Delta : F_N) \). The \( N^{1/2} \)-consistency is required to apply the linearity result (Theorem 4.1) that we need. Define

\[
G_N(a : \Delta) = N^{-1} \sum_{i=1}^{N} I \left( e_i - \frac{x_i'}{\sqrt{N}} \Delta - \frac{1}{\sqrt{N}} \frac{a}{\sigma_i} \leq 0 \right).
\]

Note that

\[
G_N(a : \sqrt{N}(\hat{\beta} - \beta)) = N^{-1} \sum_{i=1}^{N} I \left( \frac{y_i - x_i' \hat{\beta}}{\sigma_i} - \frac{a}{\sqrt{N}} \leq 0 \right)
= N^{-1} \sum_{i=1}^{N} I \left( y_i - x_i' \hat{\beta} - \frac{a}{\sqrt{N}} \leq 0 \right).
\]

Now, \( G_N(a : \Delta) \) is a regression process with sign scores. Thus for \( \Delta \) fixed, \( G_N(a : \Delta) \) is a non-decreasing function of \( a \). Then the estimate, \( \hat{a} \), solves

\[
G_N(\sqrt{N} \hat{a} : \sqrt{N}(\hat{\beta} - \beta)) = 1/2.
\]

There are no ties in the residuals, with probability 1. Thus, for even \( N \), \( \hat{a} \) solves

\[
G_N(\sqrt{N} \hat{a} : \sqrt{N}(\hat{\beta} - \beta)) = 1/2,
\]

and for odd \( N \), \( \hat{a} \) solves

\[
G_N(\sqrt{N} \hat{a} : \sqrt{N}(\hat{\beta} - \beta)) = \frac{N + 1}{2N}.
\]

We proceed by a series of lemmas to prove \( N^{1/2} \)-consistency of \( \hat{a} \).
Lemma 4.2. Suppose $B_2, B_3, B_4$ and $F_1$ hold. For all $a \in R$ and all $\epsilon > 0$,

$$
\lim_{N \to \infty} P \left\{ N^{1/2} \left| G_N(a : \sqrt{N}(\hat{\beta} - \beta)) - G_N(a : 0) \right| - N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N}\sigma_i} \right) \frac{x_i}{\sigma_i} (\hat{\beta} - \beta) \geq \epsilon \right\} = 0.
$$

Proof. First we will prove

$$
N^{1/2} \left[ G_N(a : \Delta) - G_N(a : 0) - N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N}\sigma_i} \right) \frac{x_i}{\sqrt{N}\sigma_i} \Delta \right] = o_p(1)
$$

for fixed $\Delta$. We will prove this lemma by establishing that the second moment of the quantity converges to zero.
\[
E \left[ N^{1/2} \left( G_N(a : \Delta) - G_N(a : 0) \right) \right] \\
= -N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N} \sigma_i} \frac{x_i'}{\sqrt{N} \sigma_i} \Delta \right)^2 \\
= N^{-1} E \left[ \sum_{i=1}^{N} \left\{ I(e_i - \frac{x_i'}{\sqrt{N} \sigma_i} \Delta - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) - I(e_i - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) - f \left( \frac{a}{\sqrt{N} \sigma_i} \frac{x_i'}{\sqrt{N} \sigma_i} \Delta \right) \right\}^2 \right] \\
= N^{-1} \sum_{i=1}^{N} E \left\{ I(e_i - \frac{x_i'}{\sqrt{N} \sigma_i} \Delta - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) - I(e_i - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) - f \left( \frac{a}{\sqrt{N} \sigma_i} \frac{x_i'}{\sqrt{N} \sigma_i} \Delta \right) \right\} \\
+ N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} E \left\{ I \left[ e_i - \frac{x_i'}{\sqrt{N} \sigma_i} \Delta - \frac{1}{\sqrt{N} \sigma_i} a \leq 0 \right] - I \left( e_j - \frac{x_j'}{\sqrt{N} \sigma_j} \Delta - \frac{1}{\sqrt{N} \sigma_j} a \leq 0 \right) - f \left( \frac{a}{\sqrt{N} \sigma_i} \frac{x_i'}{\sqrt{N} \sigma_i} \Delta \right) \right\} \\
= P_1 + P_2, \quad \text{say.}
\]
First consider $P_1$:

$$P_1 = N^{-1} \sum_{i=1}^{N} E \left\{ \left[I(e_i - \frac{x'_i}{\sqrt{N} \sigma_i} \Delta - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) - I(e_i - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) \right]^2 \right. $$

$$- 2 \left[I(e_i - \frac{x'_i}{\sqrt{N} \sigma_i} \Delta - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) - I(e_i - \frac{1}{\sqrt{N} \sigma_i} a \leq 0) \right]$$

$$\cdot \left\{ f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \frac{x'_i}{\sqrt{N} \sigma_i} \Delta \right\} + f^2 \left( \frac{a}{\sqrt{N} \sigma_i} \right) \left( \frac{x'_i}{\sqrt{N} \sigma_i} \Delta \right)^2$$

$$= N^{-1} \sum_{i=1}^{N} \left\{ |F \left( \frac{x'_i}{\sqrt{N} \sigma_i} \Delta + \frac{1}{\sqrt{N} \sigma_i} a \right) - F \left( \frac{1}{\sqrt{N} \sigma_i} a \right) | $$

$$- 2 \left[F \left( \frac{x'_i}{\sqrt{N} \sigma_i} \Delta + \frac{1}{\sqrt{N} \sigma_i} a \right) - F \left( \frac{1}{\sqrt{N} \sigma_i} a \right) \right]$$

$$\cdot f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \frac{x'_i}{\sqrt{N} \sigma_i} \Delta + f^2 \left( \frac{a}{\sqrt{N} \sigma_i} \right) \left( \frac{x'_i}{\sqrt{N} \sigma_i} \Delta \right)^2 \right\}.$$

Now

$$\left| \frac{x'_i}{\sqrt{N} \sigma_i} \Delta \right| \leq \left\| \frac{x'_i}{\sqrt{N} \sigma_i} \right\| \| \Delta \| \rightarrow 0$$

uniformly in $i = 1, \cdots, N$ since $\{\sigma_i\}$ are bounded away from zero and

$$\sup_{i \leq N} N^{-1/2} \|x_i\| \rightarrow 0.$$ Since $F$ is uniformly continuous, this implies

$$|F \left( \frac{x'_i}{\sqrt{N} \sigma_i} \Delta + \frac{1}{\sqrt{N} \sigma_i} a \right) - F \left( \frac{1}{\sqrt{N} \sigma_i} a \right) | \rightarrow 0$$

uniformly in $i = 1, \cdots, N$. Further, since $f$ is bounded, each of the summands converge uniformly. Thus $P_1 \rightarrow 0$. Next consider $P_2$:
\[ P_2 = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ \left[ F \left( \frac{x_i}{\sqrt{N} \sigma_i} \Delta + \frac{1}{\sqrt{N} \sigma_i} a \right) - F \left( \frac{1}{\sqrt{N} \sigma_i} a \right) \right] \right. \\
- f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \frac{x_i}{\sqrt{N} \sigma_i} \Delta \\
\left. \times \left[ F \left( \frac{x_j}{\sqrt{N} \sigma_j} \Delta + \frac{1}{\sqrt{N} \sigma_j} a \right) - F \left( \frac{1}{\sqrt{N} \sigma_j} a \right) - f \left( \frac{a}{\sqrt{N} \sigma_j} \right) \frac{x_j}{\sqrt{N} \sigma_j} \Delta \right] \right\} \]
\[ = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ \left[ \left( f(B_i) - f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \right) \frac{x_i}{\sqrt{N} \sigma_i} \Delta \right] \right. \\
\left. \times \left[ \left( f(B_j) - f \left( \frac{a}{\sqrt{N} \sigma_j} \right) \right) \frac{x_j}{\sqrt{N} \sigma_j} \Delta \right] \right\} \]
\[ \leq \left( N^{-1} \sum_{i=1}^{N} \frac{x_i}{\sigma_i} \Delta \right)^2 \sup_{i \leq N} \left| f(B_i) - f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \right|, \]

where

\[ |B_i - \frac{1}{\sqrt{N} \sigma_i} a| \leq \left| \frac{x_i}{\sqrt{N} \sigma_i} \Delta \right| \longrightarrow 0 \] uniformly in \( i = 1, \cdots, N \).

So uniform continuity of \( f \) implies

\[ |f(B_i) - f \left( \frac{a}{\sqrt{N} \sigma_i} \right) | \longrightarrow 0 \] uniformly in \( i = 1, \cdots, N \).

Now, \( N^{-1} \sum_{i=1}^{N} \frac{x_i}{\sigma_i} \Delta \leq N^{-1} \sum_{i=1}^{N} \frac{x_i}{\sigma_i} \| \Delta \| \) is bounded, since the \( \{ \sigma_i \} \) are bounded away from zero and \( N^{-1} \sum_{i=1}^{N} \| x_i \|^2 < \infty. \) Thus \( P_2 \longrightarrow 0, \) and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
we have established:

\[ \frac{1}{N^{1/2}} \left[ G_N(a : \Delta) - G_N(a : 0) - N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \frac{x_i}{\sqrt{N} \sigma_i} | \Delta | \right] = o_p(1). \]

Note that

\[ \|N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \frac{x_i}{\sigma_i} \| \leq \sup_{i \leq N} f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \left( N^{-1} \sum_{i=1}^{N} \|x_i\| \right) \sup \frac{1}{\sigma_i} \]

is bounded. Thus Lemma 1.3c of Aubuchon (1982) can be applied to yield

\[ \sup_{\|\Delta\| \leq M} N^{1/2} |G_N(a : \Delta) - G_N(a : 0) - N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N} \sigma_i} \right) \frac{x_i}{\sqrt{N} \sigma_i} | \Delta | = o_p(1). \]

Thus the lemma is proven, since \( \sqrt{N} (\hat{\beta} - \beta) \) is bounded in probability.

Lemma 4.3. Under Assumptions B4 and F1, for all \( a \in \mathbb{R} \) and all \( \varepsilon > 0 \),

\[ \lim_{N \to \infty} P \left\{ N^{1/2} |G_N(a : 0) - G_N(0 : 0) - f(0) \frac{a}{\sqrt{N}} N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i} | \geq \varepsilon \right\} = 0. \]

Proof. We will establish this lemma by proving that the second moment
converges to zero.

\[ N^{-1} E \left[ \sum_{i=1}^{N} \left\{ I \left( e_i - \frac{a}{\sqrt{N\sigma_i}} \leq 0 \right) - I(e_i \leq 0) - f(0) \frac{a}{\sqrt{N\sigma_i}} \right\}^2 \right] \]

\[ = N^{-1} \sum_{i=1}^{N} E \left\{ \left[ I \left( e_i - \frac{a}{\sqrt{N\sigma_i}} \leq 0 \right) - I(e_i \leq 0) - f(0) \frac{a}{\sqrt{N\sigma_i}} \right]^2 \right\} \]

\[ + N^{-1} \sum_{i=1}^{N} \sum_{j=1 \, i \neq j}^{N} E \left\{ \left[ I \left( e_i - \frac{a}{\sqrt{N\sigma_i}} \leq 0 \right) - I(e_i \leq 0) - f(0) \frac{a}{\sqrt{N\sigma_i}} \right] \cdot \left[ I \left( e_j - \frac{a}{\sqrt{N\sigma_j}} \leq 0 \right) - I(e_j \leq 0) - f(0) \frac{a}{\sqrt{N\sigma_j}} \right] \right\} \]

\[ = Q_1 + Q_2, \text{ say.} \]

First consider \( Q_1 \):

\[ Q_1 = N^{-1} \sum_{i=1}^{N} \left\{ |F \left( \frac{a}{\sqrt{N\sigma_i}} \right) - F(0)| - 2 \left( F \left( \frac{a}{\sqrt{N\sigma_i}} \right) - F(0) \right) f(0) \frac{a}{\sqrt{N\sigma_i}} + \frac{f^2(0)a^2}{N\sigma_i^2} \right\} \]

Now, the \( \{\sigma_i\} \) are bounded away from zero, \( F \) is uniformly continuous and \( f \) is bounded. Thus each of the summands converge uniformly to zero and consequently \( Q_1 \to 0 \).
Next consider $Q_2$:

$$Q_2 = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ \left[ F\left( \frac{a}{\sqrt{N} \sigma_i} \right) - F(0) - \frac{f(0)a}{\sqrt{N} \sigma_i} \right] 
\times \left[ F\left( \frac{a}{\sqrt{N} \sigma_j} \right) - F(0) - \frac{f(0)a}{\sqrt{N} \sigma_j} \right] \right\}$$

$$= N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{a}{\sqrt{N} \sigma_i} (f(A_i) - f(0)) \frac{a}{\sqrt{N} \sigma_j} (f(A_j) - f(0))$$

$$\leq \sup_{i \leq N} \frac{1}{\sigma_i} \sup_{i \leq N} |f(A_i) - f(0)| |a|,$$

where $|A_i| \leq \frac{a}{\sqrt{N} \sigma_i}$. Since the $\{\sigma_i\}$ are bounded away from zero, $|A_i| \rightarrow 0$ uniformly in $i = 1, \cdots, N$. Thus, uniform continuity of $f$ implies

$$|f(A_i) - f(0)| \rightarrow 0 \quad \text{uniformly in} \quad i = 1, \cdots, N.$$ 

Thus $Q_2 \rightarrow 0$ and the lemma is proven.

Lemma 4.4. Under Assumptions B2, B3, B4 and F1, for all $\epsilon > 0$ and all $a \in \mathbb{R}$,

$$\lim_{N \rightarrow \infty} P \left\{ N^{1/2} |G_N(a : \sqrt{N}(\hat{\beta} - \beta)) - G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - f(0) \frac{a}{\sqrt{N}} N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i} | \geq \epsilon \right\} = 0.$$
Proof.

\[ N^{1/2}|G_N(a : \sqrt{N}(\hat{\beta} - \beta)) - G_N(0, \sqrt{N}(\hat{\beta} - \beta)) - f(0)\frac{a}{\sqrt{N}N^{-1}} \sum_{i=1}^{N} \frac{1}{\sigma_i} | \]

\[ \leq N^{1/2}|G_N(a : \sqrt{N}(\hat{\beta} - \beta)) - G_N(a : 0) - N^{-1} \sum_{i=1}^{N} f \left( \frac{a}{\sqrt{N}\sigma_i} \right) \frac{x_i}{\sigma_i}(\hat{\beta} - \beta) | \]

\[ + N^{1/2}|G_N(a : 0) - G_N(0, 0) - f(0)\frac{a}{\sqrt{N}}N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i} | \]

\[ + N^{1/2}|G_N(0, \sqrt{N}(\hat{\beta} - \beta)) - G_N(0, 0) - N^{-1} \sum_{i=1}^{N} f(0)\frac{x_i}{\sigma_i}(\hat{\beta} - \beta) | \]

\[ + N^{1/2}|N^{-1} \sum_{i=1}^{N} \left( f \left( \frac{a}{\sqrt{N}\sigma_i} \right) - f(0) \right) \frac{x_i}{\sigma_i}(\hat{\beta} - \beta) | . \]

We can apply Lemma 4.2 to the first and third terms on the RHS of the inequality. Lemma 4.3 applies to the second term. Now the fourth term is bounded by

\[ N^{-1} \sum_{i=1}^{N} \left\| x_i \right\| \frac{1}{\sigma_i} \left( f \left( \frac{a}{\sqrt{N}\sigma_i} \right) - f(0) \right) \left\| \sqrt{N}(\hat{\beta} - \beta) \right\| . \]

Since \( N^{-1} \sum_{i=1}^{N} \left\| x_i \right\|^2 < \infty \), the \( \{\sigma_i\} \) are bounded away from 0, \( \sqrt{N}(\hat{\beta} - \beta) = O_p(1) \) and \( f \) is uniformly continuous, this term is \( o_p(1) \). Thus the lemma is established.

Lemma 4.5. Under Assumptions B2, B3, B4 and F1,

\[ N^{1/2} \left[ G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - 1/2 \right] \text{ is bounded in probability.} \]
Proof.

\[ N^{1/2}|G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - 1/2| \leq N^{1/2}|G_N(0,0) - 1/2| \]

\[ + N^{1/2}|G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - G_N(0 : 0) - N^{-1} \sum_{i=1}^{N} \frac{x_i^T}{\sigma_i}(\hat{\beta} - \beta)f(0)| \]

\[ + N^{-1} \sum_{i=1}^{N} ||x_i|| \frac{1}{\sigma_i}||\sqrt{N}(\hat{\beta} - \beta)||. \]

Note that \( N \cdot G_N(0,0) \) is a binomial random variable. Thus,

\( N^{1/2}[G_N(0,0) - 1/2] \) is bounded in probability, since it converges in distribution to a normal random variable. The second component of the sum is \( o_p(1) \) by Lemma 4.2. The third component is clearly bounded in probability since \( N^{-1} \sum_{i=1}^{N} ||x_i||^2 < \infty \), \( \{\sigma_i\} \) are bounded away from 0 and \( \sqrt{N}(\hat{\beta} - \beta) = O_p(1) \). Thus the lemma is established.

Theorem 4.2. Under Assumptions B2, B3, B4 and F1, there exists a \( c > 0 \) such that

\[ \lim_{N \to \infty} P \left[ \sqrt{N} |\hat{\alpha}| \leq c \right] = 1. \]

Proof. First we show that \( \sqrt{N} \hat{\alpha} \) is bounded above in probability. Let \( \varepsilon > 0 \) be given. By Lemma 4.5 there exists a \( M < 0 \) such that

\[ P \left\{ N^{1/2} \left[ G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - 1/2 \right] < M \right\} < \frac{\varepsilon}{2} \]

if \( N \) is large enough. Define \( M^* = \frac{(\varepsilon - M)B}{f(0)} \), where \( B \) is the bound on the \( \{\sigma_i\} \) specified by Assumption B4. If \( \sqrt{N} \hat{\alpha} > M^* \) then \( G_N(M^* : \)
\[\sqrt{N}(\hat{\beta} - \beta) \leq 1/2.\] Thus,

\[
P(\sqrt{N}\hat{\alpha} > M^*) \leq P(\sqrt{N}(G_N(M^* : \sqrt{N}(\hat{\beta} - \beta)) - 1/2) \leq 0)
\]

\[
= P\left(\sqrt{N}\left[G_N(M^* : \sqrt{N}(\hat{\beta} - \beta)) - G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) + \frac{f(0)M^*}{\sqrt{N}} \left(N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i}\right) \right] - 1/2 \right) \leq 0\right)
\]

\[
\leq P\left(\sqrt{N}|G_N(M^* : \sqrt{N}(\hat{\beta} - \beta)) - G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - f(0)\frac{M^*}{\sqrt{N}} \left(N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i}\right) | \geq \epsilon \right)
\]

\[
+ P\left(\sqrt{N}\left[G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) + \frac{f(0)M^*}{\sqrt{N}} \left(N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i}\right) \right] - 1/2 \right) \leq \epsilon \right).
\]

The first term of the sum can be made less than \(\frac{\epsilon}{2}\) for large \(N\) by Lemma 4.4. Consider the second term:

\[
P\left[\sqrt{N}\left[G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - 1/2\right] \leq \epsilon - (\epsilon - M)B \left(N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i}\right)\right]
\]

\[
\leq P\left[\sqrt{N}\left[G_N(0 : \sqrt{N}(\hat{\beta} - \beta)) - 1/2\right] \leq M\right],
\]

Since \(0 < B N^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i} \leq 1\) and \((\epsilon - M) > 0\). Thus, by the initial selection of \(M\), this term can be made less than \(\epsilon/2\) for \(N\) large enough.

The proof that \(\sqrt{N}\hat{\alpha}\) is bounded below in probability follows similarly.
4.4. Estimation of $\theta$

Assume $\hat{\Delta} = (\sqrt{N}(\hat{\beta} - \beta), 0, \sqrt{N}(\hat{\alpha}))$ where $(\hat{\beta}, \hat{\alpha})$ are $N^{1/2}$-consistent estimates of $(\beta, \alpha)$. In this section we show that the Hodges-Lehmann type estimate of $\theta$ is $\sqrt{N}$-consistent when it is based on any $\sqrt{N}$-consistent estimate of $\beta$. In addition, the asymptotic distribution of this estimate is established when the estimate of $\beta$ is the "adjusted" rank estimate.

Lemma 4.6. Suppose $\sqrt{N}(\hat{\beta} - \beta)$ is bounded in probability and $\hat{\alpha}$ is the median of $y_1 - x_1'\hat{\beta}, \cdots, y_N - x_N'\hat{\beta}$. Then under Assumption S1, $U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N)$ is a non-increasing function of $\theta$, which steps down at

$$\frac{\ln[(y_i - x_i'\hat{\beta} - \hat{\alpha})/(y_j - x_j'\hat{\beta} - \hat{\alpha})]}{h(x_i'\hat{\beta}) - h(x_j'\hat{\beta})}$$

when

$$(y_i - x_i'\hat{\beta} - \hat{\alpha})(y_j - x_j'\hat{\beta} - \hat{\alpha}) > 0$$

and

$$h(x_i'\hat{\beta}) \neq h(x_j'\hat{\beta}).$$

Proof. For simplicity in notation define $Z_i = y_i - x_i'\hat{\beta} - \hat{\alpha}$ and $c_i = h(x_i'\hat{\beta})$, $i = 1, \cdots, N$, and $\bar{c} = N^{-1}\sum_{i=1}^N c_i$. Define the process

$$s(\theta) = N^{-1/2}\sum_{t=1}^N (c_t - \bar{c})a_s(R(Z_t e^{\theta} c_t)).$$
Note that by the definition of $\hat{\alpha}$, the median of $Z_1, \cdots, Z_N$ is zero. Suppose $\theta_1 < \theta_2$ and $s(\theta)$ changes value between $\theta_1$ and $\theta_2$. Then

$$Z_i e^{-\theta_1 c_i} < Z_j e^{-\theta_1 c_j} \quad \text{and} \quad Z_i e^{-\theta_2 c_i} > Z_j e^{-\theta_2 c_j},$$

for some $i$ and $j$. Clearly this could only occur if $Z_i$ and $Z_j$ are the same sign. Now, suppose ranking $Z_1 e^{-\theta_1 c_1}, \cdots, Z_N e^{-\theta_1 c_N}$ results in

$$R(Z_i e^{-\theta_1 c_i}) = K \quad \text{and} \quad R(Z_j e^{-\theta_1 c_j}) = K + 1.$$ 

Furthermore, ranking $Z_1 e^{-\theta_2 c_1}, \cdots, Z_N e^{-\theta_2 c_N}$ results in

$$R(Z_i e^{-\theta_2 c_i}) = K + 1 \quad \text{and} \quad R(Z_j e^{-\theta_2 c_j}) = K.$$ 

Thus,

$$s(\theta_2) - s(\theta_1) = N^{-1/2}(c_i - \bar{c})a(K + 1) + N^{-1/2}(c_j - \bar{c})a(K) - N^{-1/2}(c_i - \bar{c})a(K) - N^{-1/2}(c_j - \bar{c})a(K + 1) = N^{-1/2}(c_i - c_j)(a(K + 1) - a(K)).$$

Now, we must consider two cases.

**Case I: $Z_i > 0$ and $Z_j > 0$**

Since the median of $Z_1, \cdots, Z_N$ is zero, this implies $K > N/2$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Now, $\phi_s$ is non-decreasing on $[1/2,1]$. Thus $a(K+1) \geq a(K)$. The initial selection of $i$ and $j$ imply

$$e^{-\theta_2(c_j - c_i)} < \frac{Z_i}{Z_j} < e^{-\theta_1(c_j - c_i)}.$$ 

This equation implies $(c_i - c_j) < 0$ and $\theta_1 < \frac{\ell n(Z_i/Z_j)}{c_i - c_j} < \theta_2$. Thus $(c_i - c_j)(a(K+1) - a(K)) \leq 0$.

Case II: $Z_i < 0$ and $Z_j < 0$

Since the median of $Z_1, \ldots, Z_N$ is zero, this implies $K \leq N/2$.

Now, $\phi_s$ is non-increasing on $[0,1/2]$. Thus $a(K+1) \leq a(K)$.

This initial selection of $i$ and $j$ imply

$$e^{-\theta_1(c_j - c_i)} < \frac{Z_i}{Z_j} < e^{-\theta_2(c_j - c_i)}.$$ 

The equation implies $(c_i - c_j) > 0$ and $\theta_1 < \frac{\ell n(Z_i/Z_j)}{c_i - c_j} < \theta_2$.

Thus $(c_i - c_j)(a(K+1) - a(K)) \leq 0$. Therefore the lemma is established.

Lemma 4.7. Under the assumptions of Theorem 4.1, there exists $c > 0$ such that

$$\lim_{n \to \infty} P \left[ |U_s((\hat{A}_1,0,\hat{A}_3) : F_n) | \leq c \right] = 1.$$

Proof. Theorem 4.1 implies

$$U_s((\hat{A}_1,0,\hat{A}_3) : F_n) - U_s(0 : F_n) - \gamma_s G_N(\hat{A}) = o_p(1),$$
where

\[ G_N(\hat{\Delta}) = N^{-1/2} \theta \sum_{i=1}^{N} (h(x'_i \beta) - \bar{h})(h(x'_i \beta) - h(x'_i \hat{\beta})). \]

Since \( h \) is Lipschitz continuous,

\[ ||G_N(\hat{\Delta})|| \leq N^{-1/2} \sum_{i=1}^{N} |h(x'_i \beta) - \bar{h})| \frac{|x'_i|}{\sqrt{N}} \sqrt{N(\hat{\beta} - \beta)} |\theta| \]

\[ \leq \left( N^{-1} \sum_{i=1}^{N} (h(x'_i \beta) - \bar{h})^2 \right)^{1/2} \left( N^{-1} \sum_{i=1}^{N} |x_i|^2 \right)^{1/2} ||\sqrt{N}(\hat{\beta} - \beta)|| |\theta|. \]

Now,

\[ N^{-1} \sum_{i=1}^{N} (h(x'_i \beta) - \bar{h})^2 < \infty \text{ and } N^{-1} \sum_{i=1}^{N} |x_i|^2 < \infty \]

by Assumption B3. Further \( \sqrt{N}(\hat{\beta} - \beta) = O_p(1) \). Thus \( G_N(\hat{\Delta}) \) is bounded in probability. Theorem V1.5a of Hajek and Sidak (1967) implies \( U_3(0:F_N) = O_p(1) \), and the proof of the lemma is complete. The following lemma is an obvious consequence of Theorem 4.1.

**Lemma 4.8.** Under the assumptions of Theorem 4.1, for all \( M > 0 \) and all \( \varepsilon > 0 \),

\[ \lim_{N \to \infty} P \left[ \sup \left\{ |U_3((\hat{\Delta}_1, \Delta_2^U, \hat{\Delta}_3) : F_N) - U_3((\hat{\Delta}_1, \Delta_2^L, \hat{\Delta}_3) : F_N) + \gamma_2 H_N(\Delta_2^U - \Delta_2^L) : -M \leq \Delta_2^L \leq \Delta_2^U \leq M \right\} \geq \varepsilon \right] = 0. \]
Define
\[ \hat{\theta}_L = \sup \left\{ \theta : U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) > 0 \right\} \]
\[ \hat{\theta}_U = \inf \left\{ \theta : U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) < 0 \right\} \]
and \[ \hat{\theta} = \frac{\hat{\theta}_L + \hat{\theta}_U}{2}. \]

Lemma 4.9. Under the assumptions of Theorem 4.1, there exists a \( c > 0 \) such that

\[ \lim_{N \to \infty} P \left[ N^{1/2} |\hat{\theta} - \theta| < c \right] = 1 \quad \text{and} \quad \lim_{N \to \infty} P \left[ N^{1/2} |\hat{\theta} - \theta| < c \right] = 1. \]

Proof. We will prove \( \sqrt{N} (\hat{\theta}_L - \theta) \) is bounded below in probability. Let \( \varepsilon > 0 \) be given. Lemma 4.7 implies there exists a \( c < 0 \) such that if \( N \) is large enough \( P \left[ U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) < c \right] < \frac{\varepsilon}{2} \). Now define \( c' = \gamma_s^{-1} H_N^{-1}(c - \varepsilon) \).

By the definition of \( \hat{\theta}_L \), \( \sqrt{N} (\hat{\theta}_L - \theta) < c' \) implies \( U_s((\hat{\Delta}_1, c', \hat{\Delta}_3) : F_N) \leq 0 \).

Thus,
\[
P \left[ \sqrt{N} (\hat{\theta}_L - \theta) < c' \right] \leq P \left[ U_s((\hat{\Delta}_1, c', \hat{\Delta}_3) : F_N) \leq 0 \right] \]
\[
\leq P \left[ U_s((\hat{\Delta}_1, c', \hat{\Delta}_3) : F_N) - U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) + \gamma_s H_N c' \geq \varepsilon \right] \]
\[
+ P \left[ U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) - \gamma_s H_N c' \leq \varepsilon \right]. \]

The first term of the sum is less than \( \frac{\varepsilon}{2} \) for \( N \) sufficiently large by Lemma 4.8. Note that by the definition of \( c' \) the third term can be written as
\[
P \left[ U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) \leq c \right], \]
which is less than \( \frac{\varepsilon}{2} \) for \( N \) large enough. So, we have proven there exists a \( c' < 0 \) such that \( \lim_{N \to \infty} P \left[ \sqrt{N} (\hat{\theta}_L - \theta) < c' \right] = 0 \). Similarly, an argument could be given to show \( \sqrt{N} (\hat{\theta}_U - \theta) \) is bounded above in probability. These two results and the fact \( \hat{\theta}_L \leq \hat{\theta}_U \) prove the lemma.

Lemma 4.10. Under the assumptions of Theorem 4.1, for all \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P \left\{ |U_s((\hat{\Delta}_1, \sqrt{N}(\hat{\theta}_L - \theta), \hat{\Delta}_3) : F_N)| \geq \varepsilon \right\} = 0
\]

and

\[
\lim_{N \to \infty} P \left\{ |U_s((\hat{\Delta}_1, \sqrt{N}(\hat{\theta}_U - \theta), \hat{\Delta}_3) : F_N)| \geq \varepsilon \right\} = 0.
\]

Proof. Follows exactly as the proof of Lemma 3.9 as a consequence of Lemma 4.6, Lemma 4.7, Lemma 4.8 and Lemma 4.9.

The result we need is an obvious corollary of this lemma.

Corollary. Under the assumptions of the lemma, for all \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P \left\{ |U_s((\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_3) : F_N)| \geq \varepsilon \right\} = 0,
\]

where \( \hat{\Delta}_2 = \sqrt{N}(\hat{\theta} - \theta) \).

Next, we need to obtain the limiting distribution of \( \sqrt{N}(\hat{\theta} - \theta) \). Lemma 4.9, Lemma 4.10 and Theorem 4.1 give the asymptotic expansion

\[
\sqrt{N}(\hat{\theta} - \theta) = \gamma_s^{-1} H_N^{-1} [U_s(0 : F_N) + \gamma_s G_N(\hat{\Delta})].
\]
Theorem V1.5a of Hajek and Sidak implies $U_s(0 : F_N) \rightarrow^P N(0, H_N)$ under conditions B2 and B3. However at this point we only know $G_N(\hat{\Delta})$ is bounded in probability (proof of Lemma 4.7). To establish the limiting distribution of $\sqrt{N}(\hat{\theta} - \theta)$, we need several more lemmas.

Lemma 4.11. Under Assumptions B3, B6 and B7, for all $\varepsilon > 0$,

$$\lim_{N \to \infty} P \left[ \| G_N(\hat{\Delta}) - L \sqrt{N}(\hat{\beta} - \beta) \| \geq \varepsilon \right] = 0,$$

where

$$L = \lim_{N \to \infty} N^{-1} \sum_{i=1}^{N} (h(x_i^r \beta) - \overline{h}) h'(x_i^r \beta) x_i.' $$

Proof.

$$\| G_N(\hat{\Delta}) - L \sqrt{N}(\hat{\beta} - \beta) \|$$

$$\leq \| G_N(\hat{\Delta}) - N^{-1} \sum_{i=1}^{N} (h(x_i^r \beta) - \overline{h}) h'(x_i^r \beta) x_i' \sqrt{N}(\hat{\beta} - \beta) \|$$

$$+ \| N^{-1} \sum_{i=1}^{N} (h(x_i^r \beta) - \overline{h}) h'(x_i^r \beta) x_i' - L \| \sqrt{N}(\hat{\beta} - \beta) \| \cdot |\theta|.$$
Now, since $h$ is differentiable

$$\left| \frac{h(x_i'\hat{\beta}) - h(x_i'\beta)}{x_i'\hat{\beta} - x_i'\beta} - h'(x_i'\beta) \right| \to 0 \quad \text{uniformly in } i = 1, \ldots, N.$$ 

Further

$$N^{-1} \sum_{i=1}^{N} \|x_i\| |h(x_i'\beta) - \bar{h}|$$

$$\leq \left( N^{-1} \sum_{i=1}^{N} \|x_i\|^2 \right)^{1/2} \left( N^{-1} \sum_{i=1}^{N} (h(x_i'\beta) - \bar{h})^2 \right)^{1/2} < \infty$$

by Assumption B3. Thus the lemma is proven.

We should note that $G_N(\Delta) = G_N(\Delta_1, \Delta_2, \Delta_3)$ depends on $\Delta_1$ but not $\Delta_2$ or $\Delta_3$. Hence, Lemma 4.11 clearly implies that the distribution of $G_N(\Delta)$ depends upon the distribution of $\hat{\Delta}_1 = \sqrt{N}(\hat{\beta} - \beta)$. In initial estimation of $\theta$, we only know $\sqrt{N}(\hat{\beta} - \beta) = O_p(1)$. However, in subsequent iterations the limiting distribution of $\sqrt{N}(\hat{\beta}_R - \beta)$ is known. Thus the following limiting distribution arguments are not valid for the initial estimate of $\theta$.

Lemma 4.12. Suppose $\hat{\beta}_R$ is the “adjusted” rank based estimate of $\beta$ discussed in Chapter III. Also suppose $\hat{\alpha}_R = \text{med} \{y_i - x_i'\hat{\beta}_R\}$. Then under Assumptions F1 through F3, B1 through B7, S1, S2, L1 and L2,

$$U_s(\sqrt{N}(\hat{\beta}_R - \beta), 0, \sqrt{N}\hat{\alpha}_R) : F_N \overset{D}{\rightarrow} \text{Normal} (0, H_N + \gamma_L^{-2}\gamma_S^2\Sigma^{-1}L'\theta^2).$$
Proof. Define $\hat{\Delta} = \left(\sqrt{N}(\hat{\beta}_R - \beta), 0, \sqrt{N}\hat{\alpha}_R\right)$. Note that $\sqrt{N}(\hat{\beta}_R - \beta)$ and $\sqrt{N}\hat{\alpha}_R$ are bounded in probability by the corollary to Theorem 3.5 and Theorem 4.2 respectively. Thus Theorem 4.1 gives the asymptotic representation

$$U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) = U_s(0 : F_N) + \gamma_s G_N(\hat{\Delta}) + o_p(1).$$

If $U_s(0 : F_N)$ and $G_N(\hat{\Delta})$ are asymptotically independent, then Lemma 4.11, the corollary to Theorem 3.5 and Theorem V.1.5a of Hajek and Sidak (1967) would prove this lemma. However, since this is difficult to verify by inspection, we will proceed with a proof. Using Lemma 4.11 results in

$$U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) = U_s(0 : F_N) + \gamma_s L\sqrt{N}(\hat{\beta}_R - \beta)\theta + o_p(1).$$

Recall by the definition of $\hat{\beta}_R$,

$$\sqrt{N}(\hat{\beta}_R - \beta) = \sqrt{N}(\hat{\beta}_R - \beta) + (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}'1_N\sqrt{N}\hat{\mu}_R,$$

where $\hat{\beta}_R$ and $\hat{\mu}_R$ are defined in Chapter III. Note that

$$\sqrt{N}(\hat{\beta}_R - \beta) = \gamma_L^{-1} S^{-1} U_L(0 : F_N) + o_p(1).$$

and

$$\sqrt{N}\hat{\mu}_R = \gamma_L^{-1} U_I(0 : F_N) - \mu_d' \sqrt{N}(\hat{\beta}_R - \beta) + o_p(1)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
from the proofs of Theorem 3.2 and 3.4 respectively. Substituting the latter
two expansions into the expression for \( \sqrt{N}(\hat{\beta}_R - \beta) \) results in:

\[
\sqrt{N}(\hat{\beta}_R - \beta) = (I - (\hat{D}'\hat{D})^{-1}\hat{D}'1_N\mu_d')\gamma_L^{-1} S^{-1} U_L(0 : F_N)
\]

\[
+ (\hat{D}'\hat{D})^{-1}\hat{D}'1_N \gamma_L^{-1} U_I(0 : F_N) + o_p(1).
\]

Substituting this expansion into the initial expansion for \( U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) \) implies,

\[
U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N)
\]

\[
= U_s(0 : F_N) + \gamma_s L(I - (\hat{D}'\hat{D})^{-1}\hat{D}'1_N\mu_d')\gamma_L^{-1} S^{-1} \theta U_L(0 : F_N)
\]

\[
+ \gamma_s L(\hat{D}'\hat{D})^{-1}\hat{D}'1_N \gamma_L^{-1} \theta U_I(0 : F_N) + o_p(1)
\]

\[
= U_s(0 : F) + \gamma_s L(I - (\hat{D}'\hat{D})^{-1}\hat{D}'1_N\mu_d')\gamma_L^{-1} S^{-1} \theta U_L(0 : F)
\]

\[
+ \gamma_s L(\hat{D}'\hat{D})^{-1}\hat{D}'1_N \gamma_L^{-1} \theta U_I(0 : F)
\]

\[
+ [U_s(0 : F_N) - U_s(0 : F)]
\]

\[
+ \gamma_s L(I - (\hat{D}'\hat{D})^{-1}\hat{D}'1_N\mu_d')\gamma_L^{-1} \theta S^{-1} [U_L(0 : F_N) - U_L(0 : F)]
\]

\[
+ \gamma_s L(\hat{D}'\hat{D})^{-1}\hat{D}'1_N \gamma_L^{-1} \theta (U_I(0 : F_N) - U_I(0 : F))
\]

Now, since \( \phi_L \) is odd about 1/2, \( \phi_s \) is even about 1/2 and \( f \) is
symmetric, the first three components of the sum are independent. The fourth
and fifth components are \( o_p(1) \) by Theorem V1.5a of Hajek and Sidak (1967).
Convergence in probability to 0 of the sixth component is implied by Theorem
V1.7 of Hajek and Sidak (1967). Theorem V1.2 of Hajek and Sidak implies

\[
U_I(0, F) \overset{D}{\rightarrow} N(0, 1) \quad \text{and} \quad U_s(0 : F) \overset{D}{\rightarrow} N(0, H_N).
\]
Theorem A.11 of Hettmansperger (1984) implies

\[ U_L(0 : F) \xrightarrow{D} N(0,S). \]

Combining these limiting distribution results,

\[ U_s((\hat{A}_1,0,\hat{A}_3) : F_N) \xrightarrow{D} N(0, H_N + \gamma_L^{-2} \gamma_s^2 L \Sigma^{-1} L' \theta^2), \]

where

\[ \Sigma = \lim_{N \to \infty} N^{-1}(D'D). \]

Theorem 4.3. Under the assumptions of Lemma 4.12,

\[ \sqrt{N}(\hat{\theta} - \theta) \xrightarrow{D} N(0, \gamma_s^{-2} H_N^{-1} + H_N^{-2} \gamma_L^{-2} \theta^2 L \Sigma^{-1} L'). \]

Proof. Lemma 4.8, Lemma 4.10 and Theorem 4.1 imply

\[ \sqrt{N}(\hat{\theta} - \theta) = \gamma_s^{-1} H_N^{-1} [U_s(0 : F_N) + \gamma_s G_N(\hat{A}_1, \hat{A}_2, \hat{A}_3)] + o_p(1). \]

We noted in the proof of Lemma 4.12

\[ U_s((\hat{A}_1,0,\hat{A}_3) : F_N) = U_s(0 : F_N) + \gamma_s G_N(\hat{A}_1, 0, \hat{A}_3) + o_p(1). \]

Now, \( G_N(\Delta) \) does not depend on \( \Delta_2 \) or \( \Delta_3 \). Thus combining these results,
\[
\sqrt{N}(\hat{\theta} - \theta) = \gamma^{-1}H^{-1}_N U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) + o_p(1).
\]

Thus, by Lemma 4.12

\[
\sqrt{N}(\hat{\theta} - \theta) \overset{D}{\Rightarrow} N(0, \gamma^{-2}_s H^{-1}_N + H^{-2}_N \gamma^{-2}_l \theta^2 L \Sigma^{-1} L').
\]

4.5 Estimation of the Parameter \(\gamma_s\)

In this section we define a consistent estimate of \(\gamma_s\) based on the method of Lehmann (1963) and McKean and Hettmansperger (1976). First, we consider initial estimation of \(\gamma_s\) based on \(\hat{\Delta}_1 = \sqrt{N}(\hat{\beta} - \beta)\), \(\hat{\mu} = \text{median}\{y_i - x'_i\hat{\beta}\}\), and \(\hat{\Delta}_3 = \sqrt{N}\hat{\mu}\) where \(\hat{\beta}\) is a \(N^{1/2}\)-consistent estimate of \(\beta\).

For a given \(\delta, 0 < \delta < 1\), define

\[
\bar{\delta}_M^{(0)} = \sup\{\theta : U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) > -Z_{\delta/2}\}
\]

\[
\bar{\delta}_I^{(0)} = \inf\{\theta : U_s((\hat{\Delta}_1, 0, \hat{\Delta}_3) : F_N) < Z_{\delta/2}\},
\]

where \(\Phi\) is the distribution function of a standard normal random variable with \(\Phi(Z_{\delta/2}) = \delta/2\). Define the initial estimate of \(\gamma_s\),

\[
\hat{\gamma}_{os} = 2Z_{\delta/2} H^{-1}_N(\bar{\delta}_M^{(0)} - \bar{\delta}_I^{(0)}),
\]
where
\[
\hat{H}_N = N^{-1} \sum_{i=1}^{N} (h(x_i^\hat{\beta}) - \overline{h}^*)^2
\]
and
\[
\overline{h}^* = N^{-1} \sum_{i=1}^{N} h(x_i^\hat{\beta}).
\]

Lemma 4.13. Under the assumptions of Theorem 4.1 and Assumption B8, for all \( \varepsilon > 0 \),
\[
\lim_{N \to \infty} P[|\hat{\gamma}_o - \gamma_s| \geq \varepsilon] = 0.
\]

Proof. This proof follows exactly as that for the consistency of \( \hat{\gamma}_L \) in McKean and Hettmansperger (1976) as a consequence of Lemma 4.7 and Lemma 4.8. Although consistency of the estimate can be established without the asymptotic variance of the residual process, there is no control of the bandwidth. For subsequent iterations,
\[
U_I((\hat{\Delta}_1^*, 0, \hat{\Delta}_3^*): F_N) \overset{D}{\to} N(0, H_N + \gamma_L^{-2}\gamma_s^2\Sigma^{-1}L'L^2),
\]
where \( \hat{\Delta}_1^* = \sqrt{N}(\hat{\beta}_R - \beta) \), \( \hat{\alpha}_R = \text{median}\{y_i - x_i^\hat{\beta}_R\} \) and \( \hat{\Delta}_3^* = \sqrt{N}(\hat{\alpha}_R - \alpha) \). In order to control the bandwidth, we need a consistent estimate of the asymptotic variance of \( U_I((\hat{\Delta}_1^*, 0, \hat{\Delta}_3^*): F_N) \) to standardize the statistic.
Lemma 4.14. Assume $F_1$ through $F_3$, $B_1$ through $B_8$, $S_1$, $S_2$, $L_1$ and $L_2$ hold. Suppose $\tilde{\gamma}_0 \overset{p}{\to} \gamma_s, \sqrt{N}(\hat{\beta} - \beta) = O_p(1)$ and $\sqrt{N}(\hat{\theta} - \theta) = O_p(1)$. Then for all $\varepsilon > 0$,

$$\lim_{N \to \infty} P \left[ \| \hat{H}_N + \tilde{\gamma}_0^{-2} \tilde{\gamma}_0 \hat{L} \hat{\Sigma}^{-1} \hat{L}' \hat{\theta}^2 \| \geq \varepsilon \right] = 0,$$

where

$$\hat{H}_N = N^{-1} \sum_{i=1}^{N} (h(x'_i \hat{\beta}) - \bar{h}^*)^2, \quad \bar{h}^* = N^{-1} \sum_{i=1}^{N} h(x'_i \hat{\beta})$$

$$\hat{\Sigma}^{-1} = (N^{-1} \hat{\Phi}' \hat{\Phi})^{-1}$$

and

$$\hat{L} = N^{-1} \sum_{i=1}^{N} (h(x'_i \hat{\beta}) - \bar{h}^*) h'(x'_i \hat{\beta}) x'_i.$$

Proof. We noted in Lemma 3.12 $\hat{\Sigma} \overset{p}{\to} \Sigma$. Further $\tilde{\gamma}_L \overset{p}{\to} \gamma_L$ by Theorem 3.6. Also, Lemma 4.7 implies $\hat{\theta} \overset{p}{\to} \theta$. Assumptions $B_3$ and $B_5$ imply $\hat{H}_N \overset{p}{\to} H_N$. Thus, the only component left to consider is $\hat{L}$,

$$|\hat{L} - L| \leq |N^{-1} \sum_{i=1}^{N} (h(x'_i \hat{\beta}) - \bar{h}^*) h'(x'_i \hat{\beta}) x'_i - N^{-1} \sum_{i=1}^{N} (h(x'_i \beta) - \bar{h}) h'_i(x'_i \beta) x'_i|$$

$$+ |N^{-1} \sum_{i=1}^{N} (h(x'_i \beta) - \bar{h}) h'(x'_i \beta) x'_i - L|.$$

Thus, by the initial definition of $L$ the second component of the sum converges to 0. Now, Assumptions $B_5$ and $B_8$ establish the lemma.
Define

\[ \sigma_s = \left\{ H_N + \gamma_L^{-2} L \Sigma^{-1} L' \theta^2 \right\}^{1/2} \]

and \( \sigma_s = \left\{ \hat{H}_N + \gamma_L^{-2} \hat{L} \hat{\Sigma}^{-1} \hat{L}' \hat{\theta}^2 \right\}^{1/2} \).

For a given \( \delta, 0 < \delta < 1 \), define

\[
\hat{\theta}_M = \sup \{ \theta : U_s((\hat{\Delta}_1^*, 0, \hat{\Delta}_3^*) : F_N) \hat{\sigma}_s^{-1} > -Z_{\delta/2} \} \\
\hat{\theta}_I = \inf \{ \theta : U_s((\hat{\Delta}_1^*, 0, \hat{\Delta}_3^*) : F_N) \hat{\sigma}_s^{-1} < Z_{\delta/2} \}.
\]

Next, define the estimate of \( \gamma_s \)

\[ \hat{\gamma}_s = 2Z_{\delta/2} \hat{\sigma}_s \hat{H}_N^{-1} (\hat{\theta}_M - \hat{\theta}_I)^{-1} N^{-1/2} \]

In practice, it is possible that \( U_s((\hat{\Delta}_1^*, 0, \hat{\Delta}_3^*) : F_N) \hat{\sigma}_s^{-1} \) does not attain values above \( Z_{\delta/2} \) and/or values below \(-Z_{\delta/2} \). In this case, it is preferable to use the estimate previously discussed in this section.

Lemma 4.15. Under the assumptions of Lemma 4.14, there exists a \( c > 0 \) such that

\[ \lim_{N \to \infty} P \left[ \sqrt{N} |\hat{\theta}_M - \theta| \leq c \right] = 1 \quad \text{and} \quad \lim_{N \to \infty} P \left[ \sqrt{N} |\hat{\theta}_I - \theta| \leq c \right] = 1. \]

Proof. The proof of this lemma follows as in the proof of Lemma 3.13.

Lemma 4.16. Under the assumptions of Lemma 4.14, for all \( \varepsilon > 0 \),

\[ \lim_{N \to \infty} P \left[ |Z_{\delta/2} - U_s((\hat{\Delta}_1, \sqrt{N}(\hat{\theta}_I - \theta), \hat{\Delta}_3) : F_N) \hat{\sigma}_s^{-1}| \geq \varepsilon \right] = 0 \]
and

\[
\lim_{N \to \infty} P \left[ \left| Z_{\delta/2} + U_s( (\hat{\Delta}_1, \sqrt{N}(\hat{\theta}_M - \theta), \hat{\Delta}_3) : F_N) \hat{\sigma}^{-1}_S \right| \geq \varepsilon \right] = 0.
\]

Proof. Follows exactly as the proof of Lemma 3.14.

**Theorem 4.3.** Under the assumptions of Lemma 4.14, for all \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} P \left[ |\hat{\gamma}_s - \gamma_s| \geq \varepsilon \right] = 0.
\]

Proof. Follows exactly as the proof of Theorem 3.6.
CHAPTER V

IMPLEMENTATION AND MONTE CARLO

5.1 Implementation

In this section we outline and discuss an algorithm for obtaining the rank-based estimates of $\beta$ and $\theta$. We construct the estimates as follows.

Step 1. Obtain $\hat{\beta}_0$, an initial $\sqrt{N}$-consistent estimate of $\beta$.

Step 2. Calculate $\hat{\sigma}_0 = \text{median} \{ y_i - x_i^T \hat{\beta}_0 \}$ and solve Equation (4.4) to obtain $\hat{\theta}$. Calculate $\hat{\gamma}_s$, the estimate of $\gamma_s$ discussed in Section 4.5. Also form the estimated $\{ \hat{\sigma}_i \}$, $\hat{\sigma}_i = \exp[h(x_i^T \hat{\beta}_0)\hat{\theta}]$.

Step 3. Solve Equation (3.10) for $\hat{\beta}_R$ and then solve Equation (3.11) for $\hat{\mu}_R$. Calculate the estimate $\hat{\gamma}_L$, discussed in Section 3.5.

Step 4. Calculate the adjusted estimate $\hat{\beta}_R$ given by Equation (3.12).

Step 5. Repeat Steps 2, 3 and 4 with the preliminary estimate replaced by the current estimate of $\beta$.

For Step 1, ordinary least squares estimates or ordinary $M$-estimates could be used; see Carroll and Ruppert (1982a). Jackel (1972) showed that his estimate of $\beta$ is asymptotically equivalent to that of Jureckova (1971a). Thus, it is not necessary to distinguish between the two in computational methods. Detailed analysis of the RGLM procedure needed to obtain $\hat{\beta}_R$ in

111

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Step 3 can be found in Hettmansperger (1984).

Caution should be used in the estimation of $\gamma_L$ due to the need to standardize the residual process with an estimate of its asymptotic standard deviation. The alternative method discussed in Section 3.5 based on the unstandardized residual process may be preferable in some cases. However, in this case we have little knowledge about the bandwidth. Since $\gamma_s$ is estimated by a similar method, the same type of adjustments may be needed in its estimation.

5.2 Monte Carlo

The purposes of the following Monte Carlo study are to verify the inferences presented in this thesis and compare our procedures with other established methods. This study investigates the empirical confidence of the confidence intervals for the variance parameter $\theta$ and the regression coefficients. We also compare our adjusted rank-based estimate of $\beta$ with several other methods.

For the regression problem, the Wilcoxon scores generated by $\phi_L(u) = \sqrt{12} \left( u - \frac{1}{2} \right)$ are used. The scale problem score generating function is given by $\phi_s(u) = \sqrt{180} \left( (u - \frac{1}{2})^2 - \frac{1}{2} \right)$. This is a standardized version of the scores used in Mood's test. The model we considered was simple linear regression, given by

$$y_i = \beta_0 + \beta_1 x_i + \sigma_i e_i, \ i = 1, \ldots, N.$$
In the study, the \( \{x_i\} \) were from a uniform distribution on \([-2, 2]\) and \( \sigma_i \) are given by
\[
\sigma_i = (1 + |x_i'\beta|)^\theta.
\]
The experiments were composed of 1000 simulations with \( N = 40, \beta_0 = 4.0 \) and \( \beta_1 = 2.0 \). The following distributions were used; the normal; the contaminated normal \( CN(.10, 64) \), i.e. 10% contamination and the ratio of the variance of the contaminated part to the non-contaminated part at 64; and the contaminated normal \( CN(.2, 16) \).

The normal variates were generated as discussed in Marsaglia and Bray (1964) using uniform variates which were generated by a portable FORTRAN generator written by Kahaner, Moler and Nash (1989). R-Estimates of \( \beta \) and \( \mu \) were obtained from the Robust General Linear Model Package developed by Kappenga, McKean and Vidmar (1988). The other estimates were obtained using similar methods. We used the unweighted Huber’s (1981) proposal estimate as the initial estimate of \( \beta \) with \( \psi \) given by \( \psi(x) = \max\{-2.0, \min(x, 2.0)\} \). Three iterations were used in our procedures since little was gained with additional iterations. Preliminary studies suggested using the \( M \)-estimate as the initial estimate of \( \beta \), since under departures from normality, many iterations were needed if the OLS estimate is used. The band width parameter \( \delta = .05 \) was used to construct \( \gamma_L \). The estimate of \( \gamma_L \) was corrected by the standard degree of freedom correction suggested by McKean and Sheather (1991), that is, the form of the estimates.
used was $\hat{\gamma}_L = \sqrt{\frac{N-4}{N}} \hat{\gamma}_L$. For the data considered, we were not able to use the first estimate discussed in Section 3.5 based on the standardized residual process. The second method discussed in Section 3.5 was used. A confidence parameter of $\delta = .10$ was used to construct $\hat{\gamma}_s$.

In Table 1, different methods of estimation of $\beta$ are examined. The comparison of methods is based on $\frac{1}{1000} \sum_{i=1}^{1000} (\hat{\beta}_{1i} - 2.0)^2$, the mean square error. The values listed are ratios of mean square error for estimating $\beta_1$, the ratio being with respect to the "optimal" $R$-estimate if the true $\{\sigma_i\}$ were known. The unweighted $R$-estimate of $\beta_1$ is the usual $R$-estimate; see Jackel (1972). In the case that $\theta = 0$, the optimal rank-based estimate is the same as the unweighted $R$-estimate. It is important to note that the assumptions on the design matrix $[1 : D]$, which are specified in Section 1.4, are not satisfied when $\theta = 0$. We considered this case because it is important for practical application of the methods presented in this thesis. It is important to note that our estimate never has MSE more than 20% larger than the "optimal" rank-based estimate. The presence of outliers in the sample inflates the MSE of the "optimal" weighted least squares estimate to as high as 484% that of the "optimal" rank-based estimate. The rank-based estimates are more or less comparable for the three distributions considered.
Table 1

Monte Carlo MSE Ratio With Respect to the Optimal
\( R \)-Estimate for Estimates of \( \beta_1 \)

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Normal</th>
<th>( CN(.1, 64) )</th>
<th>( CN(.2, 16) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \theta = 0 )</td>
<td>( \theta = .75 )</td>
<td>( \theta = 0 )</td>
</tr>
<tr>
<td>Unweighted LSE</td>
<td>.93</td>
<td>1.46</td>
<td>4.84</td>
</tr>
<tr>
<td>“Optimal” WLSE known weights</td>
<td>.93</td>
<td>.84</td>
<td>4.84</td>
</tr>
<tr>
<td>Unweighted ( R )-estimate</td>
<td>1.00</td>
<td>1.44</td>
<td>1.00</td>
</tr>
<tr>
<td>Our weighted ( R )-estimate</td>
<td>1.09</td>
<td>1.11</td>
<td>1.10</td>
</tr>
</tbody>
</table>

For interval estimation of \( \theta \), we use the estimate of the asymptotic variance of \( \hat{\theta} \)

\[
N^{-1} \left[ \hat{\gamma}_S \hat{H}_N^{-1} + \hat{\gamma}_L \hat{S}^{-1} L^T \right]
\]

and normal critical values. The components of this estimate were defined in Sections 4.4 and 4.5.

Table 2 displays the empirical confidence coefficients at the nominal 95% for our confidence interval of \( \theta \). We also included the mean, standard error and MSE of \( \hat{\theta} \). Confidence intervals for \( \theta \) based on its asymptotic
distribution appear to have stable confidence coefficients. None of the empirical confidence coefficients is less than the two simulation standard errors from the nominal .95 confidence. The results for the three distributions are more or less comparable. The estimates of \( \theta \) seem a little high when \( \theta = .75 \). We should note that the asymptotic variance of \( \hat{\theta} \) is non-decreasing in \( \theta \). Hence, the standard error is higher for \( \theta = .75 \) than \( \theta = 0.0 \).

Table 2
Monte Carlo Simulation Results for the Rank-based Estimate of \( \theta \)

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Normal</th>
<th>( CN(.1, 64) )</th>
<th>( CN(.2, 16) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\theta} )</td>
<td>.020</td>
<td>.861</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>.841</td>
<td>.013</td>
<td>.865</td>
</tr>
<tr>
<td>Std. Err.</td>
<td>.268</td>
<td>.367</td>
<td>.381</td>
</tr>
<tr>
<td></td>
<td>.505</td>
<td>.398</td>
<td>.655</td>
</tr>
<tr>
<td>(MSE)(^{1/2})</td>
<td>.269</td>
<td>.385</td>
<td>.381</td>
</tr>
<tr>
<td></td>
<td>.513</td>
<td>.398</td>
<td>.665</td>
</tr>
<tr>
<td>Nominal Empirical Confidence</td>
<td>.952</td>
<td>.968</td>
<td>.952</td>
</tr>
<tr>
<td></td>
<td>.959</td>
<td>.941</td>
<td>.955</td>
</tr>
</tbody>
</table>

For interval estimation of \( \beta \), we use the estimate of the asymptotic variance of \( \hat{\beta}_R \)

\[
N^{-1} \left[ \hat{\Sigma}^{-1} \right]
\]

and normal critical values. The components of this estimate were defined in Sections 3.4 and 3.5.
Table 3 displays the empirical confidence coefficients at the nominal .95 for our confidence interval of $\beta_1$ under different error distributions. We also included the mean, standard error and MSE of $\hat{\beta}_1$. Confidence intervals for $\beta_1$, based on its asymptotic distribution, appear to have quite stable confidence coefficients. One empirical confidence is more than two simulation standard errors from the nominal .95 confidence; it is marked with an (*). The results for the three distributions are more or less comparable.

**Table 3**
Monte Carlo Simulation Results for the Adjusted Rank-Based Estimate of $\beta_1$

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Normal</th>
<th>$CN(.1, 64)$</th>
<th>$CN(.2, 16)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_1$</td>
<td>$\theta = 0$</td>
<td>$\theta = .75$</td>
<td>$\theta = 0$</td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>2.000</td>
<td>2.014</td>
<td>1.995</td>
</tr>
<tr>
<td>Std. Err.</td>
<td>0.152</td>
<td>0.453</td>
<td>0.178</td>
</tr>
<tr>
<td>$(MSE)^{1/2}$</td>
<td>0.152</td>
<td>0.453</td>
<td>0.178</td>
</tr>
<tr>
<td>Nominal Empirical Confidence</td>
<td>0.938</td>
<td>0.918*</td>
<td>0.945</td>
</tr>
</tbody>
</table>

The results from this study tend to support the applicability of the methods presented in this thesis. Clearly, further studies of these robust analyses are needed.
CHAPTER VI

SUMMARY AND SUGGESTIONS FOR FUTURE RESEARCH

In Chapter II, we developed asymptotic results that were used in obtaining asymptotic linearity of the regression, signed-rank and scale statistics. The assumption that $\phi$ is continuous with a uniformly continuous bounded derivative was necessary in obtaining the results of Chapter II. Although the symmetry assumption is not necessary for the results in Chapter II, it is needed to apply Theorem 7.1 of Carroll and Ruppert (1982a), which is necessary in our three major asymptotic linearity results. Symmetry is not an assumption listed in their theorem. However, it is difficult to meet their assumptions without symmetry. Perhaps the asymptotic linearity for the regression and scale rank statistics can be obtained for other score generating functions or without the symmetry assumption.

The "adjusted" rank-based estimate of $\beta$ was developed in Chapter III. We showed that this estimate is asymptotically normal. Further we established consistency of the heteroscedastic analogue of McKean and Hettmansperger's (1976) estimate of $\gamma_L$, hence, a consistent estimate of the asymptotic variance of $\hat{\beta}_R$ was established. It may be of interest to develop a window estimate of $\gamma_L$ that does not depend upon symmetry, which is similar to that of Koul, McKean and Sievers (1987) for the homoscedastic
linear model. Aubuchon (1982) developed a method of intercept estimation which does not depend on symmetry. Perhaps his method could be extended to the heteroscedastic model. In Chapter IV, we showed that our rank-based estimate of $\theta$ is asymptotically normal and developed a consistent estimate of $\gamma_s$, thus establishing a consistent estimate of the asymptotic variance of $\hat{\theta}$. The Monte Carlo studies presented in Chapter V were preliminary in nature. The study was small, but seems to indicate that our estimates of $\beta$ and $\theta$ work well with contaminated normal errors. Further studies are needed to establish robustness properties of these estimates.

In this thesis we have assumed that $h$ is known, even though in practice this may not be the case. Further studies are needed to examine robustness against misspecification of $h$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
BIBLIOGRAPHY


