8-1989

# Measures of Partial Association Based on Rank Estimates

Sudhakar H. Rao
*Western Michigan University*

MEASURES OF PARTIAL ASSOCIATION
BASED ON RANK ESTIMATES

by

Sudhakar H. Rao

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
August 1989

MEASURES OF PARTIAL ASSOCIATION
BASED ON RANK ESTIMATES

Sudhakar H. Rao, Ph.D.

Western Michigan University, 1989

In this work we consider a new robust measure of partial association between two variables $X_1$ and $X_2$, holding a third variable $X_3$ fixed. The motivation for this estimate will be the classical sample Pearson's partial correlation coefficient, commonly denoted by $R_{12.3}$. This estimate is known to be the usual sample correlation coefficient between the ordinary least squares (OLS) residuals in the regression of $X_1$ on $X_3$ and $X_2$ on $X_3$. Being based on OLS residuals, this estimate is sensitive to outliers and heavy tailed error distributions. Further, $R_{12.3}$ is symmetric in the arguments $X_1$ and $X_2$, but there are situations where an asymmetric measure of partial correlation is more appropriate. The proposed measure will be based on rank estimates of regression parameters and is asymmetric in nature allowing the variables $X_1$ and $X_2$ to be treated unequally.

We study the properties of this estimate in comparison to $R_{12.3}$ and Kendall's tau calculated from the residuals which was studied by Randles, Shirahata, Samara and others, in terms of their behavior for contaminated normal distributions, their influence functions and their asymptotic relative efficiencies. We find that the new measure is more robust than $R_{12.3}$ in the above senses.

Order Number 8923552

# Measures of partial association based on rank estimates

Rao, Sudhakar Halady, Ph.D.
Western Michigan University, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

## ACKNOWLEDGEMENTS

I wish to thank Professor Gerald L. Sievers for his advice and guidance during this work . I would like to thank Professor Joseph McKean, Professor Michael Stoline, Professor Paul Eenigenburg and Dr. Thomas Vidmar for serving on my committee. I would like to thank Professor Joseph Buckley for his support in many different ways. I would also like to thank my family and friends for their patience, and Margo Chapman for her help.

Sudhakar H. Rao

TABLE OF CONTENTS

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

LIST OF TABLES .......... xx

CHAPTER

I. INTRODUCTION .......... 1

II. A MEASURE OF PARTIAL ASSOCIATION .......... 7

2.1 Introduction .......... 7

2.2 Properties of $\tau^*$ .......... 9

2.3 Effect of Contaminated Normal Neighborhoods on the Coefficient $\tau^*$ .......... 13

2.4 Pearson's and Kendall's Partial Correlation Coefficients .......... 20

III. A SAMPLE ESTIMATE OF THE PARTIAL CORRELATION MEASURE .......... 25

3.1 Introduction .......... 25

3.2 The Effect of Parameters on the Distribution of $T_n(\hat{\underline{\beta}})$ 28

3.3 Consistency of $T_n(\hat{\underline{\beta}})$ .......... 33

3.4 Asymptotic Null Distribution of $T_n(\hat{\underline{\beta}})$ .......... 40

3.5 The Influence Function of $T_n(\hat{\underline{\beta}})$ .......... 52

3.6 Examples .......... 68

Table of Contents – Continued


CHAPTER

IV. THE ASYMPTOTIC RELATIVE EFFICIENCY ............ 76

4.1 Introduction .................................... 76

4.2 Asymptotic Normality of a General U-Statistic with Estimated Parameters ....................... 79

4.3 Asymptotic Normality of $T_n(\underline{\hat{\beta}})$ under a sequence of alternatives ........................... 82

4.4 Asymptotic Distribution of $R_n$ .................... 106

4.5 Asymptotic Relative Efficiency of $T_n(\underline{\hat{\beta}})$ with respect to $R_n$ ............................... 118

V. A GENERAL CORRELATION PROBLEM .................. 130

5.1 Introduction .................................... 130

5.2 A General Bivariate Distribution ................ 131

5.3 The Performance of $T_n$ as a measure of correlation .................................. 133

BIBLIOGRAPHY .................................................. 137

LIST OF TABLES

Table 2.4.6. .......... 24

Table 3.6.1. .......... 69

Table 3.6.3. .......... 70

Table 3.6.4. .......... 71

Table 3.6.6. .......... 72

Table 3.6.7. .......... 74

Table 3.6.10. .......... 75

Table 4.5.17. .......... 128

Table 4.5.18. .......... 129

LIST OF FIGURES

Figure 3.4.12 .......... 44

Figure 3.6.2 .......... 70

Figure 3.6.5 .......... 72

Figure 3.6.8 .......... 75

Figure 3.6.9 .......... 75

# CHAPTER I

# INTRODUCTION

It is often considered desirable to measure the correlation between two variables say, $X_1$ and $X_2$, controlled for a third variable say, $X_3$, given a sample of observations $(X_{1i}, X_{2i}, X_{3i})'$ for $i = 1, 2, \ldots, n$. Any such measure may be called a partial correlation measure.

The most popular measure of partial correlation is the Pearson's partial correlation coefficient commonly denoted by $R_{12.3}$ given by

$$R_{12.3} = \frac{R_{12} - R_{13}R_{23}}{[(1 - R_{13}^2)(1 - R_{23}^2)]^{\frac{1}{2}}} \tag{1.1}$$

where $R_{ij}$ is the usual sample product moment correlation coefficient between $X_i$ and $X_j, i < j; i, j = 1, 2, 3$ i.e.,

$$R_{ij} = \frac{\sum_{k=1}^{n}(X_{ik} - \overline{X}_i)(X_{jk} - \overline{X}_j)}{\left[\sum_{k=1}^{n}(X_{ik} - \overline{X}_i)^2 \sum_{k=1}^{n}(X_{jk} - \overline{X}_j)^2\right]^{\frac{1}{2}}}. \tag{1.2}$$

This estimate has nice properties in the case of the underlying population is multivariate normal. For example, if $\rho_{12.3}$ is the population analogue of $R_{12.3}$, i.e., $\rho_{12.3}$ is the product moment correlation coefficient in the conditional distribution of $(X_1, X_2)'$ given $X_3$, and if $\rho_{12.3} = 0$, then $\sqrt{n-3}\frac{R_{12.3}}{\sqrt{1-R_{12.3}^2}}$ has a t-distribution with n-3 degrees of freedom. This statistic can be used to test the

hypothesis of conditional independence

(1.3)

$H_0 : X_1$ and $X_2$ are independent conditional on $X_3$ being held constant.

Further, $\sqrt{n}\frac{(R_{12.3}-\rho_{12.3})}{1-\rho_{12.3}^2}$ has an asymptotic normal distribution (see for example, Anderson (1984), Muirhead (1982), Rao (1973) etc. for more details on this).

As pointed out by authors such as Robinson (1981) and Samara (1985), the intuitive appeal to the statistic $R_{12.3}$ arises from the fact that it is nothing but the usual product moment correlation coefficient calculated from the residuals of the 'ordinary least squares' (OLS) fit of the linear model,

$$X_1 = \beta_1 X_3 + e_1 \tag{1.4}$$
$$X_2 = \beta_2 X_3 + e_2$$

where $\beta_1, \beta_2$ are the unknown regression coefficients and $X_3$ is independent of both $e_1$ and $e_2$. i.e., $R_{12.3}$ is the usual product moment correlation coefficient between $Z_{1i}$ and $Z_{2i}$ where $Z_{1i} = X_{1i} - \hat{\beta}_1 X_{3i}, Z_{2i} = X_{2i} - \hat{\beta}_2 X_{3i}, i = 1, 2, \ldots, n$ and $\hat{\beta}_1$ and $\hat{\beta}_2$ are the least squares estimates of $\beta_1$ and $\beta_2$ respectively. Under the linear model (1.4), the hypothesis of (1.3) is equivalent to

$$H_0 : e_1 \text{ and } e_2 \text{ are independent.} \tag{1.5}$$

But $R_{12.3}$ being based on the least squares residuals in (1.4), is sensitive to outliers and heavier tailed distributions and can be inefficient for non-normal distributions.

An alternative nonparametric measure of partial correlation is the Kendall's partial correlation coefficient given by

$$t_{12.3} = \frac{t_{12} - t_{13}t_{23}}{[(1 - t_{13}^2)(1 - t_{23}^2)]^{\frac{1}{2}}} \tag{1.6}$$

where $t_{ij}$ is the usual sample Kendall's tau between the variables $X_i$ and $X_j, i < j; i,j = 1,2,3$ i.e.,

$$t_{ij} = \frac{1}{\binom{n}{2}} \sum_{k<l} sgn(X_{ik} - X_{il}) sgn(X_{jk} - X_{jl}) \tag{1.7}$$

and $sgn(t) = -1, 0, 1$ as $t < 0, = 0,$ or $> 0$.

But unlike $R_{12.3}$, $t_{12.3}$ is not in general the Kendall's tau between the residuals from the model (1.4). Although $t_{12.3}$ has the same mathematical structure as $R_{12.3}$, it is merely a coincidence. Also, this coefficient does not have a population parameter associated with it. As Samara points out, the lack of popularity of $t_{12.3}$ is primarily because of it's mathematical complexities. It is not distribution free for example, and in fact, it is not even asymptotically distribution free. It's asymptotic variance depends on the underlying distribution of $(X_1, X_2, X_3)'$ (see for example, Hettmansperger (1984) and Kendall (1970) for more details).

These considerations lead authors like Shirahata (1977), Randles (1984) and Samara (1985) to consider other measures of partial association. Samara proposed Kendall's tau calculated from the residuals in the model (1.4) as an alternative measure of partial correlation coefficient. He considered the problem of testing the conditional independence and derived the asymptotic relative

efficiency of this test based on this estimator relative to the test based on $R_{12.3}$. He employed the ordinary least squares (OLS), least absolute value (LAV) and Theil's methods to estimate the slope parameters. He showed that this estimate has higher efficiency compared to $R_{12.3}$ for heavy tailed distributions.

Both the measures of partial correlation, namely, the Pearson's partial correlation coefficient and Kendall's tau of the residuals in model (1.4), are symmetric in their arguments or equivalently symmetric in the basic variables $X_1$ and $X_2$. However, there are situations where an asymmetric measure of correlation (partial correlation) in the variables $Z_1$ and $Z_2$ ($X_1$ and $X_2$) is appropriate. Somers (1959) proposed a pair of asymmetric measures of association when the two variables under investigation are ordinal variables. Schollenberger, Agresti, and Wackerly (1979) defined a class of measures of association when one of the variables is on an interval scale and the other on an ordinal scale. Their measures and models are defined in terms of scores for pairs of observations - sign scores for the ordinal variable and distance scores for the interval variable. Raveh (1986) defined a general class of measures of monotone association which includes both Somer's and Agresti, Schollenberger and Wackerly's measures in addition to the classical Goodman-Kruskal's gamma, Kendall's tau and Spearman's rho. Also see for example, Quade (1974) for some other asymmetric measures of partial association. In this paper we propose another asymmetric measure of partial correlation. We noted earlier that both Pearson's and Samara's measures of partial correlation are directly related to regression con-

cepts in interpretation and methodology. This property can be retained in defining a more robust partial correlation coefficient if the correlation measure and the residuals are replaced by more robust choices. This work will explore such a measure using residuals based on rank estimates. Estimates of regression coefficients based on rank statistics have been developed by many authors; in particular, see Jurečková (1971), Jaeckel (1972), McKean and Hettmansperger (1976, 1977) and Sievers (1983) for some of the basic properties and results on their robustness and efficiency. The basic measure of partial association used will be the weighted Kendall's tau statistic of the residuals from the model (1.4). The connection between Kendall's tau statistics and rank regression statistics was mentioned in Sievers (1978). A natural population parameter will be used to allow for a direct meaningful interpretation of sample results.

In Chapter II, we formally define this population parameter $\tau^*$ of partial association in our discussion and show that it shares all the desirable properties of a measure of association. We also study the effect of contaminated normal neighborhoods on this parameter and compare it with the other two measures discussed earlier.

In Chapter III, we define a natural sample estimate of $\tau^*$ and discuss it's small sample and consistency properties. Also we derive it's asymptotic distribution under the hypothesis of conditional independence. We obtain the 'influence curve' (IC) for this estimate, discuss some of it's uses and make a comparison of this IC with that of the other two estimates discussed earlier.

The effect of outliers on this estimate is discussed by means of several numerical examples.

In Chapter IV, we derive an expression for the 'asymptotic relative efficiency' (ARE) of the test of conditional independence based on this estimate relative to the one based on $R_{12.3}$ for a general class of alternatives, originally suggested by Gokhale (1966). This is the class of alternatives given by

$$\begin{aligned} e_1 &= W_1 + \Delta W_3 \\ e_2 &= W_2 + \Delta W_3 \end{aligned} \tag{1.8}$$

where $W_1, W_2$ and $W_3$ are independent random variables and $\Delta$ is some constant. The hypothesis of conditional independence in this case is equivalent to the hypothesis that $\Delta = 0$.

In Chapter V, we look at a general correlation problem rather than the partial correlation problem for a general class of bivariate distributions given by Farlie (1960). This is the class of distribution functions $H(x,y) = F(x)G(y)\{1+\alpha A(F)B(G)\}$ for which F and G are the marginal cdfs and A and B are some functions. For this class we propose our estimate calculated directly from the variables X and Y as an asymmetric measure of correlation and compare it with the maximum likelihood estimate (MLE) of the parameter of association and with the classical Pearson's and Kendall's tau statistics in asymptotic relative efficiency.

# CHAPTER II

# A MEASURE OF PARTIAL ASSOCIATION

## 2.1 Introduction

Consider the random variables $X_1, X_2$ and $X_3$. Assume they have finite expectations, but otherwise their distributions can be quite arbitrary. To measure the partial association between $X_1$ and $X_2$, holding $X_3$ fixed, we shall assume that both $X_1$ and $X_2$ are linearly related to $X_3$ according to the model,

$$X_1 = \beta_1 X_3 + e_1 \tag{2.1.1}$$
$$X_2 = \beta_2 X_3 + e_2,$$

where $\underline{\beta} = (\beta_1, \beta_2)'$ is a vector of unknown parameters and $X_3$ is independent of $(e_1, e_2)'$.

Let F be the univariate cdf of $X_3$ and G be the bivariate cdf of $(e_1, e_2)'$. Then, in this model, the conditional cdf of $(X_1, X_2)'$ given $X_3 = x$ is $G(\underline{y} - \underline{\beta} x)$. This property appears in the multivariate normal model, but here G is not assumed normal. No symmetry or centering assumptions are made on G.

Let $(X_{11}, X_{21}, X_{31})'$ and $(X_{12}, X_{22}, X_{32})'$ be independent, each having the same distribution of $(X_1, X_2, X_3)'$. Suppose $\underline{\beta}_* = (\beta_{*1}, \beta_{*2})'$ is a value of $\underline{\nu} = (\nu_1, \nu_2)'$ that minimizes

$$E|X_{11} - X_{12} - \nu_1(X_{31} - X_{32})| = E|(X_{11} - \nu_1 X_{31}) - (X_{12} - \nu_1 X_{32})| \tag{2.1.2}$$

and

(2.1.3) $E|X_{21} - X_{22} - \nu_2(X_{31} - X_{32})| = E|(X_{21} - \nu_2 X_{31}) - (X_{22} - \nu_2 X_{32})|$

as functions of $\nu_1$ and $\nu_2$. (2.1.2) and (2.1.3) are the population analogues of the dispersion function used in Sievers (1983).

Then define a partial correlation parameter by

(2.1.4)
$$\tau^* = \frac{E(X_{11} - X_{12} - \beta_{*1}(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_{*2}(X_{31} - X_{32}))}{E|X_{11} - X_{12} - \beta_{*1}(X_{31} - X_{32})|}.$$

Note that $E|X_{11} - X_{12} - \nu_1(X_{31} - X_{32})|$ and $E|X_{21} - X_{22} - \nu_2(X_{31} - X_{32})|$ are convex functions of $\nu_1$ and $\nu_2$ respectively and hence the minimum exists in both cases although it may not be unique. In most cases of practical interest $\underline{\beta}_*$ will be unique.

Note that $\tau^*$ is the weighted Kendall's tau between $X_1 - \beta_{*1}X_3$ and $X_2 - \beta_{*2}X_3$. The linear function $\beta_{*i}X_3$ can be viewed as a best linear predictor of $X_i, i = 1, 2$, in the sense of minimizing variation in $X_i - \beta_{*i}X_3$ as measured by the absolute difference of two independent copies. Note that if $z_1$ and $z_2$ are two independent copies of a random variable z, then $E|z_1 - z_2|$ measures the variation in z (a Gini mean difference parameter). Being of first order, this will be less sensitive to contamination and heavy tails in the distribution in comparison to the square function $E(z_1 - z_2)^2 = 2var(z)$ used in the classical approach.

REMARK 2.1.5. *Assume model (2.1.1). Let $G_i$ denote the cdf of $e_i$ and $G_i^*$ denote the cdf of the difference of two independent random variables each having*

*cdf $G_i$ and assume that $G_i^*$ has a unique median, i=1,2. Then $\beta_1$ and $\beta_2$ are the unique values minimizing (2.1.2) and (2.1.3) respectively and*

(2.1.6)
$$\tau^* = \frac{E(X_{11} - X_{12} - \beta_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_2(X_{31} - X_{32}))}{E|X_{11} - X_{12} - \beta_1(X_{31} - X_{32})|}$$
$$= \frac{E(e_{11} - e_{12})sgn(e_{21} - e_{22})}{E|e_{11} - e_{12}|}.$$

PROOF: Under the model (2.1.1), the conditional distribution of $W_i = X_{i1} - X_{i2}$ given $X_{31} - X_{32} = t$ has cdf $G_i^*(w_i - \beta_i t), i = 1, 2$. This distribution has a unique median $\beta_i t$ since $G_i^*$ has unique median by assumption and it's value is 0 from $W_i$ being symmetrically distributed about 0. It is well known that the median minimizes an expected absolute deviation. Thus for each fixed t, $E[|w_i - a_i||t]$ is minimum if $a_i = \beta_i t$ and the result follows.

We shall discuss the properties of $\tau^*$ in section 2.2. In section 2.3, we study the effect of contaminated normal neighborhoods on the coefficient $\tau^*$. In section 2.4, we discuss the properties of $\rho_{12.3}$ and $\tau$ (Kendall's tau calculated from the residuals in (2.1.1)) in comparison to $\tau^*$. Also we derive the expressions for these parameters for a contaminated normal distribution and make a comparison with the corresponding expression for $\tau^*$. Numerical values of these coefficients are calculated for several contaminated normal distributions and compared.

## 2.2 Properties of $\tau^*$

REMARK 2.2.1. $|\tau^*| \leq 1$.

PROOF:

$$|E(X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32}))sgn(X_{21}-X_{22}-\beta_{*2}(X_{31}-X_{32}))|$$

$$\leq E|X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32})|,$$

and hence

$$\begin{aligned}|\tau^*(\underline{\beta}_*)| &= \left|\frac{E(X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32}))sgn(X_{21}-X_{22}-\beta_{*2}(X_{31}-X_{32}))}{E|X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32})|}\right| \\ &\leq \frac{E|X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32})|}{E|X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32})|} \\ &= 1.\end{aligned}$$

REMARK 2.2.2. *If $X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32})$ and $X_{21}-X_{22}-\beta_{*2}(X_{31}-X_{32})$ have the same sign with probability 1, then $\tau^*=+1$.*

PROOF: Let $W_1 = X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32})$ and $W_2 = X_{21}-X_{22}-\beta_{*2}(X_{31}-X_{32})$. The hypothesis implies $sgn(W_1) = sgn(W_2)$ with probability 1. Then the numerator of $\tau^*$ is $EW_1sgnW_2 = EW_1sgnW_1 = E|W_1|$, which is the denominator of $\tau^*$.

REMARK 2.2.3. *If $X_1$ and $X_2$ are conditionally independent on $X_3$, then $\tau^* = 0$.*

PROOF: The hypothesis implies that

$$\begin{aligned}&E[(X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32}))sgn(X_{21}-X_{22} \\ &\qquad -\beta_{*2}(X_{31}-X_{32}))|X_{31}-X_{32}=t] \\ &= E(X_{11}-X_{12}-\beta_{*1}t)Esgn(X_{21}-X_{22}-\beta_{*2}t) \\ &= -\beta_{*1}tEsgn(X_{21}-X_{22}-\beta_{*2}t).\end{aligned}$$

Then the hypothesis implies that the numerator of $\tau^*$ is equal to

$$
\begin{aligned}
&-\beta_{*1}E^{X_{31},X_{32}}\left[(X_{31}-X_{32})Esgn(X_{21}-X_{22}-\beta_{*2}(X_{31}-X_{32}))\right]\\
&=-\beta_{*1}E^{X_{31},X_{32}}\left[(X_{31}-X_{32})\cdot 0\right]\\
&=0.
\end{aligned}
$$

Thus $\tau^* = 0$.

REMARK 2.2.4. *If $(X_1, X_2, X_3)'$ has a multivariate normal distribution, then $\tau^* = \rho_{12.3}$, where $\rho_{12.3}$ is the classical Pearson's partial correlation coefficient between $X_1$ and $X_2$ holding $X_3$ fixed.*

PROOF: If $(X_1, X_2, X_3)'$ has a multivariate normal distribution, then $\rho_{12.3}$ is the Pearson's correlation coefficient in the conditional distribution of $(X_1, X_2)'$ given $X_3$.

Also the above hypothesis implies that $(X_1, X_2, X_3)'$ satisfy model (2.1.1). Hence the conditional distribution of $(X_1, X_2)'$ given $X_3$ is same as the distribution of $(e_1, e_2)'$ excepting for the mean. Thus the correlation coefficient in the conditional distribution of $(X_1, X_2)'$ given $X_3$ is same as the correlation coefficient between $e_1$ and $e_2$. This is the same as the correlation coefficient between the differences $W_1 = e_{11} - e_{12}$ and $W_2 = e_{21} - e_{22}$. Thus $W_1$ and $W_2$ have a bivariate normal distribution with mean $\underline{0} = (0, 0)'$ and correlation coefficient $\rho_{12.3}$.

Now the coefficient $\tau^*$ can be written in terms of $W_1$ and $W_2$ as

$$\tau^* = \frac{EW_1 sgnW_2}{E|W_1|}.$$

Write

$$EW_1 sgnW_2 = E^{W_2}\left[E(W_1 sgnW_2|W_2)\right]$$

$$= E^{W_2}\left[sgnW_2 E(W_1|W_2)\right].$$

But $W_1|W_2$ is normally distributed with mean $E(W_1|W_2) = \rho_{12.3}\frac{\sigma_{W_1}}{\sigma_{W_2}}W_2$, where $\sigma^2_{W_1}$ and $\sigma^2_{W_2}$ represent the variances of $W_1$ and $W_2$ respectively. Consequently,

$$EW_1 sgnW_2 = \rho_{12.3}\frac{\sigma_{W_1}}{\sigma_{W_2}}E^{W_2}(W_2 sgnW_2)$$

$$= \rho_{12.3}\frac{\sigma_{W_1}}{\sigma_{W_2}}E|W_2|.$$

Because $W_2 \sim N(0, \sigma^2_{W_2})$, straight forward calculations yield

$$E|W_2| = \sqrt{\frac{2}{\pi}}\sigma_{W_2}.$$

Similarly,

$$E|W_1| = \sqrt{\frac{2}{\pi}}\sigma_{W_1}.$$

Hence

$$\tau^* = \frac{\rho_{12.3}\frac{\sigma_{W_1}}{\sigma_{W_2}}\sqrt{\frac{2}{\pi}}\sigma_{W_2}}{\sqrt{\frac{2}{\pi}}\sigma_{W_1}}$$

$$= \rho_{12.3}.$$

REMARK 2.2.5. *$\tau^*$ is invariant under nonsingular linear transformations of $X_1, X_2$ and $X_3$.*

PROOF: If $X_i$ is replaced by $X_i^* = a_i X_i + b_i X_3 + c_i, a_i > 0, i = 1, 2$ and $X_3$ by $X_3^* = cX_3 + d$, then $\beta_{*1}$ and $\beta_{*2}$ minimizing (2.1.2) and (2.1.3) change to $\overline{\beta}_{*1} = \frac{a_1}{c}\beta_{*1} + \frac{b_1}{c}$ and $\overline{\beta}_{*2} = \frac{a_2}{c}\beta_{*2} + \frac{b_2}{c}$ respectively. Then the coefficient $\tau^*$ based on the transformed variables is given by

$$\tau^*{}_{new} = \frac{E(X^*_{11} - X^*_{12} - \overline{\beta}_{*1}(X^*_{31} - X^*_{32}))sgn(X^*_{21} - X^*_{22} - \overline{\beta}_{*2}(X^*_{31} - X^*_{32}))}{E|X^*_{11} - X^*_{12} - \overline{\beta}_{*1}(X^*_{31} - X^*_{32})|}.$$

But

$$(X_{11}^* - X_{12}^* - \overline{\beta}_{*1}(X_{31}^* - X_{32}^*) = a_1(X_{11} - X_{12}) + b_1(X_{31} - X_{32})$$
$$- (\frac{a_1}{c}\beta_{*1} + \frac{b_1}{c})c(X_{31} - X_{32})$$
$$= a_1(X_{11} - X_{12} - \beta_{*1}(X_{31} - X_{32}).$$

Similarly

$$(X_{21}^* - X_{22}^* - \overline{\beta}_{*2}(X_{31}^* - X_{32}^*) = a_2(X_{21} - X_{22} - \beta_{*2}(X_{31} - X_{32}).$$

Substituting the above in the expression for $\tau^*{}_{new}$, we get

$$\tau^*{}_{new}$$
$$= \frac{a_1 sgn a_2}{|a_1|} \frac{E(X_{11} - X_{12} - \beta_{*1}(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_{*2}(X_{31} - X_{32}))}{E|X_{11} - X_{12} - \beta_{*1}(X_{31} - X_{32})|}$$
$= \tau^*$ since $a_1$ and $a_2$ are positive.

## 2.3 Effect of Contaminated Normal Neighborhoods on the Coefficient $\tau^*$

It is proved in the preceding section that if $(X_1, X_2, X_3)'$ has a multivariate normal distribution, then $\tau^* = \rho_{12.3}$. Now for $0 \leq \epsilon \leq 1$, assume that $(X_1, X_2, X_3)'$ has a cdf given by

$$H_\epsilon = (1 - \epsilon)H_1 + \epsilon H_2 \tag{2.3.1}$$

where $H_1$ and $H_2$ are respectively the cdf's of a

$$N\left[\underline{0}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ & 1 & \rho_{23} \\ & & 1 \end{pmatrix}\right] \quad \text{and} \quad N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho_{12}^*\sigma_1\sigma_2 & \rho_{13}^*\sigma_1\sigma_3 \\ & \sigma_2^2 & \rho_{23}^*\sigma_2\sigma_3 \\ & & \sigma_3^2 \end{pmatrix}\right]$$

distributions.

This corresponds to a distribution that is a mixture of two multivariate normal distributions represented by $H_1$ and $H_2$. For small values of $\epsilon$, most of the observations come from the distribution $H_1$ with occasional observations from the distribution $H_2$. This model is called a contaminated normal model. If the $\mu_i$'s are near zero, this provides a model that is hard to distinguish from a normal model for small to moderate values of $\epsilon$.

Now we shall see how this contaminated normal model affects our partial correlation parameter $\tau^*$. First, let's consider the simple case. Assume model (2.1.1) with $(e_1, e_2)'$ having a cdf $G(\cdot)$, and let

$$\tau^* = \frac{E(e_{11} - e_{12})sgn(e_{21} - e_{22})}{E|e_{11} - e_{12}|}.$$

Now consider the contaminated normal model for $(e_1, e_2)'$ given by the cdf,

$$G_\epsilon = (1 - \epsilon)G_1 + \epsilon G_2$$

where $G_1$ and $G_2$ are respectively a $N(0, 0, 1, 1, \rho)$ and a $N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho^*)$ distribution functions.

First note that, if

$$\text{(2.3.2)} \qquad \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ & \sigma_2^2 \end{pmatrix}\right],$$

then

$$
\begin{aligned}
EW_1 sgnW_2 &= E[E(W_1 sgnW_2|W_2)] \\
&= E[sgnW_2 E(W_1|W_2)] \\
&= E\left[sgnW_2\left(\mu_1 + \frac{\sigma_{12}}{\sigma_2^2}(W_2 - \mu_2)\right)\right] \\
&= \left[\mu_1 - \mu_2\frac{\sigma_{12}}{\sigma_2^2}\right] EsgnW_2 + \frac{\sigma_{12}}{\sigma_2^2}EW_2 sgnW_2 \\
&= \left[\mu_1 - \mu_2\frac{\sigma_{12}}{\sigma_2^2}\right][1 - 2P(W_2 < 0)] + \frac{\sigma_{12}}{\sigma_2^2}E|W_2| \\
&= \left[\mu_1 - \mu_2\frac{\sigma_{12}}{\sigma_2^2}\right]\left[1 - 2\Phi(-\frac{\mu_2}{\sigma_2})\right] + \frac{\sigma_{12}}{\sigma_2^2}E|W_2|
\end{aligned}
$$

where $\Phi$ is the standard normal distribution function.

But,

(2.3.3)

$$
\begin{aligned}
E|W_2| &= \int_{-\infty}^{\infty} |x|\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2}dx \\
&= -\int_{-\infty}^{0} x\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2}dx + \int_{0}^{\infty} x\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2}dx \\
&= -\int_{-\infty}^{-\frac{\mu_2}{\sigma_2}} (y\sigma_2 + \mu_2)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy + \int_{-\frac{\mu_2}{\sigma_2}}^{\infty} (y\sigma_2 + \mu_2)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy \\
&= -\frac{\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu_2}{\sigma_2}} ye^{-\frac{1}{2}y^2}dy - \mu_2\Phi(-\frac{\mu_2}{\sigma_2}) \\
&\quad + \frac{\sigma_2}{\sqrt{2\pi}}\int_{-\frac{\mu_2}{\sigma_2}}^{\infty} ye^{-\frac{1}{2}y^2}dy + \mu_2\left[1 - \Phi(-\frac{\mu_2}{\sigma_2})\right] \\
&= \frac{\sigma_2}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{\mu_2^2}{\sigma_2^2}} - \mu_2\Phi(-\frac{\mu_2}{\sigma_2}) + \frac{\sigma_2}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{\mu_2^2}{\sigma_2^2}} + \mu_2\left[1 - \Phi(-\frac{\mu_2}{\sigma_2})\right] \\
&= \mu_2 - 2\mu_2\Phi(-\frac{\mu_2}{\sigma_2}) + \sqrt{\frac{2}{\pi}}\sigma_2 e^{-\frac{1}{2}\frac{\mu_2^2}{\sigma_2^2}}
\end{aligned}
$$

Substituting for $E|W_2|$ in the expression for $EW_1 sgn W_2$, we get

(2.3.4)

$$\begin{aligned} EW_1 sgn W_2 &= \left[\mu_1 - \mu_2\frac{\sigma_{12}}{\sigma_2^2}\right]\left[1 - 2\Phi(-\frac{\mu_2}{\sigma_2})\right] \\ &\quad + \frac{\sigma_{12}}{\sigma_2^2}\left[\mu_2 - 2\mu_2\Phi(-\frac{\mu_2}{\sigma_2}) + \sqrt{\frac{2}{\pi}}\sigma_2 e^{-\frac{1}{2}\frac{\mu_2^2}{\sigma_2^2}}\right] \\ &= \mu_1 - 2\mu_1\Phi(-\frac{\mu_2}{\sigma_2}) + \sqrt{\frac{2}{\pi}}\frac{\sigma_{12}}{\sigma_2}e^{-\frac{1}{2}\frac{\mu_2^2}{\sigma_2^2}}. \end{aligned}$$

For $i \le j; i,j = 1,2$, let $E_{G_i,G_j}$ represent the expected value induced by two independent measurements on $(e_1, e_2)'$ with one distributed according to $G_i$ and the other distributed according to $G_j$. Then the numerator of $\tau^*$ for the contaminated model $G_\epsilon$ is

$$\begin{aligned} &E(e_{11} - e_{12})sgn(e_{21} - e_{22}) \\ &= (1-\epsilon)^2 E_{G_1,G_1}(e_{11} - e_{12})sgn(e_{21} - e_{22}) \\ &\quad + 2\epsilon(1-\epsilon)E_{G_1,G_2}(e_{11} - e_{12})sgn(e_{21} - e_{22}) \\ &\quad + \epsilon^2 E_{G_2,G_2}(e_{11} - e_{12})sgn(e_{21} - e_{22}), \end{aligned}$$

and the denominator of $\tau^*$ is

$$\begin{aligned} &E|e_{11} - e_{12}| \\ &= (1-\epsilon)^2 E_{G_1,G_1}|e_{11} - e_{12}| \\ &\quad + 2\epsilon(1-\epsilon)E_{G_1,G_2}|e_{11} - e_{12}| \\ &\quad + \epsilon^2 E_{G_2,G_2}|e_{11} - e_{12}|. \end{aligned}$$

Let $\begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix}$ and $\begin{pmatrix} e_{12} \\ e_{22} \end{pmatrix}$ be independent with the common cdf $G_1$. Then

$$\begin{pmatrix} e_{11} - e_{12} \\ e_{21} - e_{22} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 2\rho \\ & 2 \end{pmatrix}\right].$$

Hence by (2.3.2) - (2.3.4),

$$E_{G_1,G_1}|e_{11}-e_{12}| = \frac{2}{\sqrt{\pi}}$$

and

$$E_{G_1,G_1}(e_{11}-e_{12})sgn(e_{21}-e_{22}) = \frac{2}{\sqrt{\pi}}\rho.$$

Now let $\begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix}$ and $\begin{pmatrix} e_{12} \\ e_{22} \end{pmatrix}$ be independent with $\begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix}$ having cdf $G_1$ and $\begin{pmatrix} e_{12} \\ e_{22} \end{pmatrix}$ having cdf $G_2$. Then

$$\begin{pmatrix} e_{11}-e_{12} \\ e_{21}-e_{22} \end{pmatrix} \sim N\left[\begin{pmatrix} -\mu_1 \\ -\mu_2 \end{pmatrix}, \begin{pmatrix} 1+\sigma_1^2 & \rho+\rho^*\sigma_1\sigma_2 \\ & 1+\sigma_2^2 \end{pmatrix}\right].$$

Again by (2.3.2) - (2.3.4),

$$E_{G_1,G_2}|e_{11}-e_{12}|$$
$$= -\mu_1 - 2(-\mu_2)\Phi\left(-\frac{\mu_1}{\sqrt{(1+\sigma_1^2)}}\right) + \sqrt{\frac{2}{\pi}}\sqrt{(1+\sigma_1^2)}e^{-\frac{1}{2}\frac{\mu_1^2}{(1+\sigma_1^2)}}$$

and

$$E_{G_1,G_2}(e_{11}-e_{12})sgn(e_{21}-e_{22})$$
$$= -\mu_1 - 2(-\mu_1)\Phi\left(-\frac{\mu_2}{\sqrt{(1+\sigma_2^2)}}\right) + \sqrt{\frac{2}{\pi}}\frac{(\rho+\rho^*\sigma_1\sigma_2)}{\sqrt{(1+\sigma_2^2)}}e^{-\frac{1}{2}\frac{\mu_2^2}{(1+\sigma_2^2)}}.$$

Finally, let $\begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix}$ and $\begin{pmatrix} e_{12} \\ e_{22} \end{pmatrix}$ be independent with the common cdf $G_2$. Then

$$\begin{pmatrix} e_{11}-e_{12} \\ e_{21}-e_{22} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2\sigma_1^2 & 2\rho^*\sigma_1\sigma_2 \\ & 2\sigma_2^2 \end{pmatrix}\right].$$

Once again by (2.3.2) - (2.3.4),

$$E_{G_2,G_2}|e_{11}-e_{12}| = \frac{2}{\sqrt{\pi}}\sigma_1$$

and

$$E_{G_2,G_2}(e_{11}-e_{12})sgn(e_{21}-e_{22}) = \frac{2}{\sqrt{\pi}}\rho^*\sigma_1.$$

Hence $\tau^*$ for the contaminated normal error model $G_\epsilon$ is,

$$\tau^* = \frac{\tau_1^*}{\tau_2^*} \tag{2.3.5}$$

where

$$\tau_1^* = (1-\epsilon)^2\frac{2}{\sqrt{\pi}}\rho + 2\epsilon(1-\epsilon)[-\mu_1 - 2(-\mu_1)\Phi(-\frac{\mu_2}{\sqrt{(1+\sigma_2^2)}})$$
$$+\sqrt{\frac{2}{\pi}}\frac{(\rho+\rho^*\sigma_1\sigma_2)}{\sqrt{(1+\sigma_2^2)}}e^{-\frac{1}{2}\frac{\mu_2^2}{(1+\sigma_2^2)}}] + \epsilon^2\frac{2}{\sqrt{\pi}}\rho^*\sigma_1$$

and

$$\tau_2^* = (1-\epsilon)^2\frac{2}{\sqrt{\pi}} + 2\epsilon(1-\epsilon)[-\mu_1 - 2(-\mu_1)\Phi(-\frac{\mu_1}{\sqrt{(1+\sigma_1^2)}})$$
$$+\sqrt{\frac{2}{\pi}}\sqrt{(1+\sigma_1^2)}e^{-\frac{1}{2}\frac{\mu_1^2}{(1+\sigma_1^2)}}] + \epsilon^2\frac{2}{\sqrt{\pi}}\sigma_1.$$

Note that in (2.3.5) above, if $\epsilon = 0$ then $\tau^* = \rho$ and if $\epsilon = 1$ then $\tau^* = \rho^*$. These are desired as $\epsilon = 0$ and $\epsilon = 1$ correspond to the case of normal error distribution in which case $\tau^*$ and Pearson's correlation coefficient are the same. Also note that if $\sigma_1, \sigma_2$ tend to infinity then $\tau^* \to \rho^*$. The case of $\sigma_1, \sigma_2$ large (one or both of them) reflects the likelihood of getting outliers. Then $\tau^*$ will weight $\rho^*$ heavily and we could get $\tau^*$ close to $\rho^*$. In this case, if $\epsilon$ is small, $G_1$ is the main part of the distribution with $G_2$ being the nuisance contamination effect, yet $\tau^*$ close to $\rho^*$. Thus contamination could have serious effect on our partial correlation measure if $\rho^*$ is not close to $\rho$.

Note that if $(X_1, X_2, X_3)'$ has cdf $H_\epsilon$ given in (2.3.1), then $(e_1, e_2)'$ has cdf $G_\epsilon$ with $G_1$ and $G_2$ are respectively a $N(0, 0, \xi_1, \xi_2, \zeta)$ and a $N(\gamma_1, \gamma_2, \eta_1, \eta_2, \kappa)$ distribution functions, where

$$\xi_1^2 = 1 - 2\beta_1\rho_{13} + \beta_1^2$$

$$\xi_2^2 = 1 - 2\beta_2\rho_{23} + \beta_2^2$$

$$\zeta = (\rho_{12} - \beta_1\rho_{23} - \beta_2\rho_{13} + \beta_1\beta_2)/\xi_1\xi_2$$

$$\gamma_1 = \mu_1 - \beta_1\mu_3$$

$$\gamma_2 = \mu_2 - \beta_2\mu_3$$

$$\eta_1^2 = \sigma_1^2 - 2\beta_1\rho_{13}^*\sigma_1\sigma_2 + \beta_1^2\sigma_3^2$$

$$\eta_2^2 = \sigma_2^2 - 2\beta_2\rho_{23}^*\sigma_2\sigma_3 + \beta_2^2\sigma_3^2$$

$$\kappa = (\rho_{12}^*\sigma_1\sigma_2 - \beta_1\rho_{23}^*\sigma_2\sigma_3 - \beta_2\rho_{13}^*\sigma_1\sigma_3 + \beta_1\beta_2\sigma_3^2)/\eta_1\eta_2.$$

Then the $\tau^*$ for the contaminated normal model $H_\epsilon$ is given by

$$\tau^* = \frac{\tau_1^*}{\tau_2^*} \tag{2.3.6}$$

where

$$\tau_1^* = (1-\epsilon)^2\frac{2}{\sqrt{\pi}}\zeta\xi_1 + 2\epsilon(1-\epsilon)[-\gamma_1 - 2(-\gamma_1)\Phi(-\frac{\gamma_2}{\sqrt{(\xi_2^2+\eta_2^2)}})$$
$$+\sqrt{\frac{2}{\pi}}\frac{(\zeta\xi_1\xi_2 + \kappa\eta_1\eta_2)}{\sqrt{(\xi_2^2+\eta_2^2)}}e^{-\frac{1}{2}\frac{\gamma_2^2}{(\xi_2^2+\eta_2^2)}}] + \epsilon^2\frac{2}{\sqrt{\pi}}\kappa\eta_1$$

and

$$\tau_2^* = (1-\epsilon)^2\frac{2}{\sqrt{\pi}}\xi_1 + 2\epsilon(1-\epsilon)[-\gamma_1 - 2(-\gamma_1)\Phi(-\frac{\gamma_1}{\sqrt{(\xi_1^2+\eta_1^2)}})$$
$$+\sqrt{\frac{2}{\pi}}\sqrt{(\xi_1^2+\eta_1^2)}e^{-\frac{1}{2}\frac{\gamma_1^2}{(\xi_1^2+\eta_1^2)}}] + \epsilon^2\frac{2}{\sqrt{\pi}}\eta_1.$$

## 2.4 Pearson's and Kendall's Partial Correlation Coefficients

We noted in the proof of Remark 2.2.4 that if $(X_1, X_2, X_3)'$ has a multivariate normal distribution, then Pearson's partial correlation coefficient between $X_1$ and $X_2$ holding $X_3$ fixed is the regular product moment correlation coefficient between $e_1$ and $e_2$ given by

$$(2.4.1) \qquad \rho_{12.3} = \frac{E(e_1 - Ee_1)(e_2 - Ee_2)}{[E(e_1 - Ee_1)^2 E(e_2 - Ee_2)^2]^{\frac{1}{2}}}.$$

Also we showed in the same Remark that $\tau^*$ and $\rho_{12.3}$ are one in the same for this distribution. Although this definition of $\rho_{12.3}$ is primarily given for the multivariate normal distribution, we shall compare $\rho_{12.3}$ with $\tau^*$ when the underlying population is not necessarily normal. It is easy to see that the properties of $\tau^*$ given in Remarks 2.2.1 - 2.2.4 are satisfied by $\rho_{12.3}$ and property in Remark 2.2.5 is given as an exercise in Anderson (1984, p. 54).

We shall now study the effect of contaminated normal errors on the coefficient $\rho_{12.3}$. In this regard we consider the distribution $G_\epsilon, 0 \le \epsilon \le 1$, of the preceding section for the random vector $(e_1, e_2)'$. For this model, it is easy to see that

$$(2.4.2) \qquad \rho_{12.3} = \frac{(1-\epsilon)\rho + \epsilon\rho^*\sigma_1\sigma_2}{[(1-\epsilon+\epsilon\sigma_1^2)(1-\epsilon+\epsilon\sigma_2^2)]^{\frac{1}{2}}}.$$

From (2.4.2), it is clear that $\rho_{12.3} \to \rho^*$ as $\sigma_1 \to \infty$ and $\sigma_2 \to \infty$, indicating the non robust nature of $\rho_{12.3}$ for contaminated normal model. Later we will do a comparative study of the robustness in the above sense of the two parameters

$\tau^*$ and $\rho_{12.3}$ to see if the effect of contamination is more severe for one than for the other.

Again as in the case of $\tau^*$, if $(X_1, X_2, X_3)'$ has a contaminated normal model given in (2.3.1), $\rho_{12.3}$ for this model will be

$$\rho_{12.3} = \frac{(1-\epsilon)\zeta\xi_1\xi_2 + \epsilon\kappa\eta_1\eta_2}{[((1-\epsilon)\xi_1^2 + \epsilon\eta_1^2)((1-\epsilon)\xi_2^2 + \epsilon\eta_2^2)]^{\frac{1}{2}}} \tag{2.4.3}$$

with the notations as in section 2.3.

Now we shall briefly discuss the third measure of partial association namely the Kendall's tau. Under model (2.1.1) this parameter is defined as

$$\tau = Esgn(X_{11}-X_{12}-\beta_{*1}(X_{31}-X_{32}))sgn(X_{21}-X_{22}-\beta_{*2}(X_{31}-X_{32})) \tag{2.4.4}$$

where $\beta_{*1}$ and $\beta_{*2}$ are obtained by minimizing (2.1.2) and (2.1.3).

Clearly this parameter satisfies the properties given in Remarks 2.2.1, 2.2.2 and 2.2.4. In the case the underlying population is normal, then $\tau = Esgn(e_{11} - e_{12})sgn(e_{21} - e_{22})$ which is equal to $\frac{2}{\pi}sin^{-1}\rho$, where $\rho$ is the correlation coefficient between $e_1$ and $e_2$. The last equality follows from the fact that if $(U, V)'$ has a $N(0, 0, 1, 1, \rho)$ distribution, then $Esgn(U)sgn(V) = \frac{2}{\pi}sin^{-1}\rho$ and that if $(U, V)'$ has a $N(0, 0, \sigma_1, \sigma_2, \rho)$ distribution, then $Esgn(U)sgn(V) = Esgn(\frac{U}{\sigma}_1)sgn(\frac{V}{\sigma}_2) = \frac{2}{\pi}sin^{-1}\rho$ is an immediate consequence (for more on this see, for example, Kendall (1970)). But $\rho$ is nothing but $\rho_{12.3}$. Thus unlike $\tau^*, \tau$ is not the same as Pearson's partial correlation coefficient for a normal distribution. Now we shall see how the contaminated normal distribution affects $\tau$.

Assume that $(e_1, e_2)'$ has the cdf $G_\epsilon$ defined in the previous section. Finding $\tau$ in this case requires finding an expectation of the form $Esgn(U)sgn(V)$ where $(U, V)'$ has a bivariate normal distribution with nonzero mean, which is very messy. To avoid this, we shall consider a slightly simpler contaminated model $G_\epsilon$ in which both $G_1$ and $G_2$ have the same means. There is no loss of generality if we assume that both $G_1$ and $G_2$ have zero means. For this model, straight forward calculations show that,

(2.4.5)

$$\tau = (1-\epsilon)^2 \frac{2}{\pi} sin^{-1}\rho + 2\epsilon(1-\epsilon)\frac{2}{\pi} sin^{-1}\left(\frac{\rho + \rho^* \sigma_1 \sigma_2}{\sqrt{1+\sigma_1^2}\sqrt{1+\sigma_2^2}}\right) + \epsilon^2 \frac{2}{\pi} sin^{-1}\rho^*.$$

It is easy to see that as $\sigma_1, \sigma_2 \to \infty$,

$$\tau \to (1-\epsilon)^2 \frac{2}{\pi} sin^{-1}\rho + 2\epsilon(1-\epsilon)\frac{2}{\pi} sin^{-1}\rho^* + \epsilon^2 \frac{2}{\pi} sin^{-1}\rho^*.$$

It appears that this parameter is not as much affected by outliers as either $\tau^*$ or $\rho_{12.3}$. We will make a formal comparison among $\tau^*, \rho_{12.3}$ and $\tau$ at the end of this section.

If $\sigma_1 = \sigma_2 = 1$, then

$$\tau = (1-\epsilon)^2 \frac{2}{\pi} sin^{-1}\rho + 2\epsilon(1-\epsilon)\frac{2}{\pi} sin^{-1}\left(\frac{\rho + \rho^*}{2}\right) + \epsilon^2 \frac{2}{\pi} sin^{-1}\rho^*.$$

Note that, for this simpler contaminated normal model, $\tau^*$ in (2.3.5) takes a simpler form as follows,

$$\tau^* = \frac{(1-\epsilon)^2\rho + \sqrt{2}\epsilon(1-\epsilon)\left(\frac{\rho+\rho^*\sigma_1\sigma_2}{\sqrt{1+\sigma_2^2}}\right) + \epsilon^2\rho^*\sigma_1}{(1-\epsilon)^2 + \sqrt{2}\epsilon(1-\epsilon)\sqrt{1+\sigma_1^2} + \epsilon^2\sigma_1}$$

but $\rho_{12.3}$ in (2.4.2) remains unchanged and is equal to

$$\rho_{12.3} = \frac{(1-\epsilon)\rho + \epsilon\rho^*\sigma_1\sigma_2}{[(1-\epsilon+\epsilon\sigma_1^2)(1-\epsilon+\epsilon\sigma_2^2)]^{\frac{1}{2}}}.$$

(i) If $\sigma_1, \sigma_2 \to \infty$, then $\tau^* \to \rho^*$ and $\rho_{12.3} \to \rho^*$. As pointed out earlier, this corresponds to the case of outliers in which case both $\tau^*$ and $\rho_{12.3}$ weight $\rho^*$ heavily and are close to $\rho^*$. Thus the outliers can have serious effect on both $\tau^*$ and $\rho_{12.3}$ if $\rho^*$ is not close to $\rho$.

(ii) If $\sigma_1 = \sigma_2 = 1$, then $\tau^* = \rho_{12.3} = (1-\epsilon)\rho+\epsilon\rho^*$. Thus if the contaminating distribution $G_2$, that is, $\epsilon$ is small, has about the same variation as $G_1$, we will not get a deterioration from $\rho$ and thus contamination causes little effect on both $\tau^*$ and $\rho_{12.3}$.

Table 2.4.6 below gives the values of the three coefficients for several contaminated normal distributions.

Table 2.4.6

| $\epsilon$ | $\rho$ | $\rho^*$ | $\sigma_1$ | $\sigma_2$ | $\tau^*$ | $\rho_{12.3}$ | $\sin(\frac{\pi}{2} \cdot \tau)$ |
|---|---|---|---|---|---|---|---|
| 0.05 | 0.1 | 0.2 | 1 | 1 | 0.105 | 0.105 | 0.105 |
| | | | 1 | 3 | 0.106 | 0.106 | 0.106 |
| | | | 3 | 1 | 0.111 | 0.106 | 0.106 |
| | | | 3 | 3 | 0.118 | 0.132 | 0.109 |
| | | 0.8 | 1 | 1 | 0.135 | 0.135 | 0.137 |
| | | | 1 | 3 | 0.145 | 0.182 | 0.149 |
| | | | 3 | 1 | 0.191 | 0.182 | 0.149 |
| | | | 3 | 3 | 0.224 | 0.325 | 0.170 |
| | 0.7 | 0.2 | 1 | 1 | 0.675 | 0.675 | 0.678 |
| | | | 1 | 3 | 0.660 | 0.587 | 0.666 |
| | | | 3 | 1 | 0.619 | 0.587 | 0.666 |
| | | | 3 | 3 | 0.611 | 0.540 | 0.663 |
| | | 0.8 | 1 | 1 | 0.705 | 0.705 | 0.705 |
| | | | 1 | 3 | 0.700 | 0.663 | 0.700 |
| | | | 3 | 1 | 0.700 | 0.663 | 0.700 |
| | | | 3 | 3 | 0.718 | 0.732 | 0.709 |
| 0.10 | 0.1 | 0.2 | 1 | 1 | 0.110 | 0.110 | 0.110 |
| | | | 1 | 3 | 0.111 | 0.112 | 0.111 |
| | | | 3 | 1 | 0.121 | 0.111 | 0.111 |
| | | | 3 | 3 | 0.132 | 0.150 | 0.117 |
| | | 0.8 | 1 | 1 | 0.170 | 0.170 | 0.170 |
| | | | 1 | 3 | 0.190 | 0.246 | 0.196 |
| | | | 3 | 1 | 0.266 | 0.246 | 0.196 |
| | | | 3 | 3 | 0.321 | 0.450 | 0.235 |
| | 0.7 | 0.2 | 1 | 1 | 0.650 | 0.650 | 0.655 |
| | | | 1 | 3 | 0.621 | 0.514 | 0.631 |
| | | | 3 | 1 | 0.555 | 0.514 | 0.631 |
| | | | 3 | 3 | 0.542 | 0.450 | 0.625 |
| | | 0.8 | 1 | 1 | 0.710 | 0.710 | 0.710 |
| | | | 1 | 3 | 0.700 | 0.648 | 0.700 |
| | | | 3 | 1 | 0.700 | 0.648 | 0.700 |
| | | | 3 | 3 | 0.732 | 0.750 | 0.718 |

CHAPTER III

# A SAMPLE ESTIMATE OF THE PARTIAL CORRELATION MEASURE

## 3.1 Introduction

Let $(X_{11}, X_{21}, X_{31})', (X_{12}, X_{22}, X_{32})', \ldots, (X_{1n}, X_{2n}, X_{3n})'$, be n independent measurements on $(X_1, X_2, X_3)'$ and let $(X_1, X_2, X_3)'$ have a distribution function H. To measure the conditional dependence of $X_1$ and $X_2$ holding $X_3$ fixed, we consider the following linear model,

$$
\begin{aligned}
X_{1i} &= \beta_1 X_{3i} + e_{1i} \\
X_{2i} &= \beta_2 X_{3i} + e_{2i}, \quad i = 1, 2, \ldots, n,
\end{aligned} \tag{3.1.1}
$$

where $\beta_1$ and $\beta_2$ are unknown parameters which need to be estimated. Intercept parameters could be added to this model but the procedure here is based on differences and it would cancel out and have no effect.

MODEL ASSUMPTIONS. *The covariate terms $X_{3i}, i = 1, 2, \ldots, n$, are i.i.d. random variables with distribution function $F(\cdot)$, mean $\mu_3$ and variance $\sigma_3^2$. The error terms $(e_{1i}, e_{2i})', i = 1, 2, \ldots, n$, are i.i.d. bivariate random variables with distribution function G. The marginal distribution of $e_{1i}(e_{2i})$ is assumed to have mean $\mu_1(\mu_2)$, variance $\sigma_1^2(\sigma_2^2)$ and distribution function $G_1(\cdot)(G_2(\cdot))$. Further, we assume that $X_{3i}$ is independent of $(e_{1i}, e_{2i})', i = 1, 2, \ldots, n$.*

An estimate of $\tau^*$ can be defined in a natural way as follows. Let $\hat{\beta}_1$ be a value of $b_1$ that minimizes a dispersion measure of the residuals given by,

$$(3.1.2) \qquad D_1(b_1) = \sum_{i<j} |X_{1i} - b_1 X_{3i} - X_{1j} + b_1 X_{3j}|.$$

Similarly let $\hat{\beta}_2$ be a value of $b_2$ that minimizes

$$(3.1.3) \qquad D_2(b_2) = \sum_{i<j} |X_{2i} - b_2 X_{3i} - X_{2j} + b_2 X_{3j}|.$$

Then define an estimate of $\tau^*$ as

(3.1.4)

$$\begin{aligned} T_n(\underline{\hat{\beta}}) &= \frac{\sum_{i<j}(X_{1i} - X_{1j} - \hat{\beta}_1(X_{3i} - X_{3j}))sgn(X_{2i} - X_{2j} - \hat{\beta}_2(X_{3i} - X_{3j}))}{\sum_{i<j}|X_{1i} - X_{1j} - \hat{\beta}_1(X_{3i} - X_{3j})|} \\ &= \frac{\sum_{i<j}(Z_{1i} - Z_{1j})sgn(Z_{2i} - Z_{2j})}{\sum_{i<j}|Z_{1i} - Z_{1j}|}, \end{aligned}$$

where $Z_{1i} = X_{1i} - \hat{\beta}_1 X_{3i}$ and $Z_{2i} = X_{2i} - \hat{\beta}_2 X_{3i}, i = 1, 2, \ldots, n$, are the residuals from the model (3.1.1).

The dispersion functions in (3.1.2) and (3.1.3) are convex, piecewise linear functions of $b_1$ and $b_2$ respectively and as a result there will be points attaining the minimum, although they may not be unique. These are the same dispersion functions used in Sievers (1983) and are algebraically equal to the dispersion function used in Jaeckel (1972) and McKean and Hettmansperger (1976, 1977) when Wilcoxon scores are used. These references point out that the diameter of the set of points attaining the minimum tends to zero asymptotically. Further $\hat{\beta}_1$ and $\hat{\beta}_2$ are the rank estimates of the regression slopes $\beta_1$ and $\beta_2$ and these references contain further results on properties of $\hat{\beta}_1$ and

$\hat{\beta}_2$, computational methods and more. This estimate $T_n(\underline{\hat{\beta}})$ has the following properties: $-1 \leq T_n(\underline{\hat{\beta}}) \leq +1$, $T_n(\underline{\hat{\beta}}) = +1$ if the rank order of the residuals $Z_1 = (Z_{11}, Z_{12}, ..., Z_{1n})'$ is the same as the rank order of the residuals $Z_2 = (Z_{21}, Z_{22}, ..., Z_{2n})'$ and $T_n(\underline{\hat{\beta}})$ is invariant under nonsingular linear transformations of $X_{1i}$, $X_{2i}$ and $X_{3i}$. The estimate $T_n(\underline{\hat{\beta}})$ can be expressed in another form to view it more explicitly as a rank statistic. First note that the numerator of $T_n(\underline{\hat{\beta}})$

$$\begin{aligned}\sum_{i<j}(Z_{1i} - Z_{1j})sgn(Z_{2i} - Z_{2j}) &= \sum_{i=1}^{n} Z_{1i}(2R_{2i} - (n+1)) \\ &= 2\sum_{i=1}^{n}(Z_{1i} - \overline{Z}_1)R_{2i}\end{aligned}$$

where $R_{2i}$ is the rank of $Z_{2i}$ among $Z_{21}, Z_{22}, ..., Z_{2n}$ and $\overline{Z}_1 = \sum_{i=1}^{n} Z_{1i}/n$. Similarly, the denominator of $T_n(\underline{\hat{\beta}})$ can be written as $2\sum_{i=1}^{n}(Z_{1i} - \overline{Z}_1)R_{1i}$, where $R_{1i}$ represents the rank of $Z_{1i}$ among $Z_{11}, Z_{12}, ..., Z_{1n}$. Hence $T_n(\underline{\hat{\beta}})$ can be written as

(3.1.5)
$$\begin{aligned}T_n(\underline{\hat{\beta}})) &= \frac{\sum_{i=1}^{n}(Z_{1i} - \overline{Z}_1)R_{2i}}{\sum_{i=1}^{n}(Z_{1i} - \overline{Z}_1)R_{1i}} \\ &= \frac{Z_{1c}' R_2}{Z_{1c}' R_1}\end{aligned}$$

where $Z_{1c}$ is the vector of the centered residuals $Z_{1i} - \overline{Z}_1$, $R_1 = (R_{11}, R_{12}, ..., R_{1n})'$ and $R_2 = (R_{21}, R_{22}, ..., R_{2n})'$. The formula (3.1.5) suggests an interesting generalization to allow arbitrary scores instead of ranks. It appears that such a statistic would have the same properties as $T_n(\underline{\hat{\beta}})$.

In the following sections we shall discuss other properties of the estimate $T_n(\underline{\hat{\beta}})$. In section 3.2, it will be shown that the distribution of this estimate is free

from the regression parameters $\beta_1$ and $\beta_2$, the location parameters $\mu_1, \mu_2$ and $\mu_3$, and scale parameters $\sigma_1^2, \sigma_2^2$ and $\sigma_3^2$. In section 3.3, we discuss the consistency property of $T_n(\hat{\underline{\beta}})$. In section 3.4, we derive the asymptotic normality of $T_n(\hat{\underline{\beta}})$ under the hypothesis of conditional independence discussed in chapter I. In section 3.5, we will obtain an expression for the influence curve of this estimate and discuss this in comparison with the influence curves of Pearson's partial correlation estimate and Kendall's tau estimate. In section 3.6, we present three numerical examples and discuss the effect of outliers on these estimates.

## 3.2 The Effect of Parameters on the Distribution of $T_n(\hat{\underline{\beta}})$

$T_n(\hat{\underline{\beta}})$ is not a distribution free statistic even under the hypothesis of independence of the "error terms". It's distribution depends on the distribution of $X_{3i}$'s, $e_{1i}$'s and $e_{2i}$'s, $i = 1, 2, ..., n$. To see this, write the residuals given in (3.1.4) in terms of the error terms to obtain

(3.2.1)
$$Z_{1i} = X_{1i} - \hat{\beta}_1 X_{3i}$$
$$= -(\hat{\beta}_1 - \beta_1) X_{3i} + e_{1i},$$

and similarly

$$Z_{2i} = -(\hat{\beta}_2 - \beta_2) X_{3i} + e_{2i}.$$

Thus we can write

(3.2.2)
$$T_n(\hat{\underline{\beta}}) = \frac{\sum_{i<j}(e_{1i} - e_{1j} - (\hat{\beta}_1 - \beta_1)(X_{3i} - X_{3j}))sgn(e_{2i} - e_{2j} - (\hat{\beta}_2 - \beta_2)(X_{3i} - X_{3j}))}{\sum_{i<j} |e_{1i} - e_{1j} - (\hat{\beta}_1 - \beta_1)(X_{3i} - X_{3j})|}.$$

(3.2.2) suggests that the distribution of $T_n(\hat{\underline{\beta}})$ depends on the distribution of $X_{3i}$'s, $e_{1i}$'s and $e_{2i}$'s, $i = 1, 2, ..., n$.

Although the statistic $T_n(\hat{\underline{\beta}})$ is not distribution free, we show in what follows that it's distribution does not depend on the parameters $\beta_1, \beta_2, \mu_1, \mu_2, \mu_3, {\sigma_1}^2, {\sigma_2}^2$, and ${\sigma_3}^2$.

First, if $\mu_1 \neq 0, \mu_2 \neq 0, \mu_3 \neq 0$, consider $e_{1i}^* = e_{1i} - \mu_1, e_{2i}^* = e_{2i} - \mu_2$, and $X_{3i}^* = X_{3i} - \mu_3$. The underlying model (3.1.1) may now be written as

$$
\begin{aligned}
X_{1i} &= (\mu_1 + \beta_1\mu_3) + \beta_1 X_{3i}^* + e_{1i}^* \\
X_{2i} &= (\mu_2 + \beta_2\mu_3) + \beta_2 X_{3i}^* + e_{2i}^*, \quad i = 1, 2, ..., n.
\end{aligned} \tag{3.2.3}
$$

This model (3.2.3) is same as the model (3.1.1) except for the intercept terms. But as noted earlier, the intercept terms have no effect on our estimate and we have $T_n(\hat{\underline{\beta}})$ free of the location parameters $\mu_1, \mu_2$ and $\mu_3$.

To show that the distribution of $T_n(\hat{\underline{\beta}})$ is free from the remaining parameters, we first need to show the following "translation" and "scale" properties of our slope estimates.

Denote $\hat{\beta}_i$ by $\hat{\beta}_i(X_{31}, X_{32}, ..., X_{3n}; X_{i1}, X_{i2}, ..., X_{in}), i = 1, 2$. Then $\hat{\beta}_i$, satisfies the following properties:

"Translation Property".

(3.2.4)

$$\hat{\beta}_i(x_1, x_2, ..., x_n; y_1+cx_1, y_2+cx_2, ..., y_n+cx_n) = \hat{\beta}_i(x_1, x_2, ..., x_n; y_1, y_2, ..., y_n)+c,$$

*$i = 1, 2$, for every $x_1, x_2, ..., x_n, y_1, y_2, ..., y_n$ and $c$.*

PROOF: Note that $\hat{\beta}_1$ is a value of $b_1$ that satisfies

$$\begin{aligned}\min_{b_1}\sum_{i<j}|y_i - y_j - b_1(x_i - x_j)| &= \sum_{i<j}|y_i - y_j - \hat{\beta}_1(x_i - x_j)| \\ &= \sum_{i<j}|(y_i + cx_i) - (y_j + cx_j) - (\hat{\beta}_1 + c)(x_i - x_j)|.\end{aligned}$$

So if $y_i$ is replaced by $y_i + cx_i, i = 1, 2, \ldots, n$, then the new slope estimate is given by $\hat{\beta}_1 + c$, which proves the translation property of $\hat{\beta}_1$. Similarly the translation property of $\hat{\beta}_2$ can be established.

"SCALE PROPERTY".

(3.2.5) $\hat{\beta}_i(ax_1, ax_2, ..., ax_n; by_1, by_2, ..., by_n) = \dfrac{b}{a}\hat{\beta}_i(x_1, x_2, ..., x_n; y_1, y_2, ..., y_n),$

$i = 1, 2,$ *for every* $x_1, x_2, ..., x_n, y_1, y_2, ..., y_n, b$ *and* $a \neq 0$.

PROOF: For $a \neq 0, b \neq 0$,

$$\begin{aligned}\min_{b_1}\sum_{i<j}|y_i - y_j - b_1(x_i - x_j)| &= \sum_{i<j}|y_i - y_j - \hat{\beta}_1(x_i - x_j)| \\ &= \frac{1}{|b|}\sum_{i<j}|by_i - by_j - \frac{b}{a}\hat{\beta}_1(ax_i - ax_j)|,\end{aligned}$$

which shows that when $x_i$ is replaced by $ax_i$ and $y_i$ is replaced by $by_i, i = 1, 2, \ldots, n$, the new slope estimate is given by $\frac{b}{a}\hat{\beta}_1$.

Note that the case b=0 is equivalent to all the $y_i$'s being equal to zero, in which case $\hat{\beta}_1 = 0$. This proves the 'scale' property of $\hat{\beta}_1$. Similarly the 'scale' property of $\hat{\beta}_2$ can be established.

From the 'translation' property (3.2.4), we have,

$$\hat{\beta}_i(X_{31}, X_{32}, ..., X_{3n}; X_{i1}-\beta_i X_{31}, X_{i2}-\beta_i X_{32}, ..., X_{in}-\beta_i X_{3n}) = \hat{\beta}_i-\beta_i, \quad i = 1, 2.$$

Thus $(\hat{\beta}_i - \beta_i), i = 1, 2$ in (3.2.2) may be replaced by

$$\hat{\hat{\beta}}_i = \hat{\beta}_i(X_{31}, X_{32}, ..., X_{3n}; X_{i1} - \beta_i X_{31}, X_{i2} - \beta_i X_{32}, ..., X_{in} - \beta_i X_{3n}), \quad i = 1, 2,$$

without changing the value of $T_n(\underline{\hat{\beta}})$. These new estimators $\hat{\hat{\beta}}_1$ and $\hat{\hat{\beta}}_2$ are the regression estimators obtained by replacing $X_{1i}$ by $X_{1i} - \beta_1 X_{3i}$ and $X_{2i}$ by $X_{2i} - \beta_2 X_{3i}$ respectively in the model

$$X_{1i} = \beta_1 X_{3i} + e_{1i}$$

$$X_{2i} = \beta_2 X_{3i} + e_{2i}, \quad i = 1, 2, \ldots, n.$$

This is equivalent to the slope estimators obtained from the model

$$X_{1i} = e_{1i}$$

$$X_{2i} = e_{2i}, \quad i = 1, 2, \ldots, n,$$

which is the usual model with $\beta_1 = \beta_2 = 0$. Consequently, the estimate $T_n(\underline{\hat{\beta}})$ does not depend on the slope parameters $\beta_1$ and $\beta_2$.

From the 'scale' property (3.2.5), we have,

$$\hat{\beta}_i(\frac{1}{\sigma_3}X_{31}, \frac{1}{\sigma_3}X_{32}, ..., \frac{1}{\sigma_3}X_{3n}; X_{i1}, X_{i2}, ..., X_{in}) = \sigma_3\hat{\beta}_i, \quad i = 1, 2. \tag{3.2.6}$$

Then the residuals can be written as

$$\begin{aligned} Z_{1i} &= X_{1i} - \hat{\beta}_1 X_{3i} \\ &= X_{1i} - (\sigma_3\hat{\beta}_1)\frac{X_{3i}}{\sigma_3} \\ &= X_{1i} - \hat{\beta}_1(\frac{1}{\sigma_3}X_{31}, \frac{1}{\sigma_3}X_{32}, ..., \frac{1}{\sigma_3}X_{3n}; X_{11}, X_{12}, ..., X_{1n})\frac{X_{3i}}{\sigma_3} \end{aligned}$$

and

$$Z_{2i} = X_{2i} - \hat{\beta}_2(\frac{1}{\sigma_3}X_{31}, \frac{1}{\sigma_3}X_{32}, ..., \frac{1}{\sigma_3}X_{3n}; X_{21}, X_{22}, ..., X_{2n})\frac{X_{3i}}{\sigma_3}.$$

From the above it is clear that $X_{3i}$'s may be replaced by their standardized values $\frac{X_{3i}}{\sigma_3}$ without altering the values of $Z_{1i}$ and $Z_{2i}$. Consequently $T_n(\hat{\underline{\beta}})$ does not depend on $\sigma_3$ as can be seen from (3.1.4).

From (3.2.2), we can see that $T_n(\hat{\underline{\beta}})$ involves quantities of the form $e_{1i} - (\hat{\beta}_1 - \beta_1)X_{3i}$ and $e_{2i} - (\hat{\beta}_2 - \beta_2)X_{3i}$. Further from (3.2.4),

$$\hat{\beta}_i(X_{31}, X_{32}, ..., X_{3n}; e_{i1}, e_{i2}, ..., e_{in}) = \hat{\beta}_i - \beta_i, \quad i = 1, 2.$$

Thus $T_n(\hat{\underline{\beta}})$ may be written in terms of the residual estimates of the form

$$R_{1i} = e_{1i} - \hat{\beta}_1(X_{31}, X_{32}, ..., X_{3n}; e_{11}, e_{12}, ..., e_{1n})X_{3i}$$

and

$$R_{2i} = e_{2i} - \hat{\beta}_2(X_{31}, X_{32}, ..., X_{3n}; e_{21}, e_{22}, ..., e_{2n})X_{3i}$$

as

$$T_n(\hat{\underline{\beta}}) = \frac{\sum_{i<j}(R_{1i} - R_{1j})sgn(R_{2i} - R_{2j})}{\sum_{i<j}|R_{1i} - R_{1j}|}.$$

Applying 'scale' property

$$\begin{aligned}&\hat{\beta}_i(X_{31}, X_{32}, ..., X_{3n}; \frac{1}{\sigma_i}e_{i1}, \frac{1}{\sigma_i}e_{i2}, ..., \frac{1}{\sigma_i}e_{in})\\ &= \frac{1}{\sigma_i}\hat{\beta}_i(X_{31}, X_{32}, ..., X_{3n}; e_{i1}, e_{i2}, ..., e_{in}), \quad i = 1, 2.\end{aligned}$$

Thus if we replace the error terms $e_{1i}(e_{2i})$ by their standardized forms $\frac{e_{1i}}{\sigma_1}(\frac{e_{2i}}{\sigma_2})$, the residual estimates $R_{1i}(R_{2i})$ will be transformed to $\frac{R_{1i}}{\sigma_1}(\frac{R_{2i}}{\sigma_2}), i = 1, 2, \ldots, n$, so that

$$\begin{aligned}T_n(\hat{\underline{\beta}}) &= \frac{\sum_{i<j}(\frac{R_{1i}}{\sigma_1} - \frac{R_{1j}}{\sigma_1})sgn(\frac{R_{2i}}{\sigma_2} - \frac{R_{2j}}{\sigma_2})}{\sum_{i<j}|\frac{R_{1i}}{\sigma_1} - \frac{R_{1j}}{\sigma_1}|}\\ &= \frac{\sum_{i<j}(R_{1i} - R_{1j})sgn(R_{2i} - R_{2j})}{\sum_{i<j}|R_{1i} - R_{1j}|} \quad \text{since } \sigma_1 \text{ and } \sigma_2 \text{ are positive.}\end{aligned}$$

Thus the distribution of $T_n(\underline{\hat{\beta}})$ is free of $\sigma_1^2$ and $\sigma_2^2$.

So far in this section we have shown that the distribution of $T_n(\underline{\hat{\beta}})$ is free of

(i) the intercept parameters and the location parameters $\mu_1, \mu_2$ and $\mu_3$.

(ii) the intercept parameters, location parameters $\mu_1, \mu_2$ and $\mu_3$ and the slope parameters $\beta_1$ and $\beta_2$, because $\hat{\beta}_1$ and $\hat{\beta}_2$ satisfy the 'translation' property.

(iii) the intercept parameters, the location parameters $\mu_1, \mu_2$ and $\mu_3$, the slope parameters $\beta_1$ and $\beta_2$ and the scale parameters $\sigma_1^2, \sigma_2^2$ and $\sigma_3^2$ since $\hat{\beta}_1$ and $\hat{\beta}_2$ satisfy both the 'translation' and 'scale' properties.

REMARK 3.2.7. *In addition to our method of estimation of slope parameters $\beta_1$ and $\beta_2$, other commonly used methods such as OLS(ordinary least squares) and LAV(least absolute value) methods also have these "translation" and "scale" properties.*

## 3.3 Consistency of $T_n(\underline{\hat{\beta}})$

In this section we show that $T_n(\underline{\hat{\beta}})$ is consistent for $\tau^*(\underline{\beta})$ under model (3.1.1) with some additional regularity conditions:

CONDITION 3.3.1. *The cdf $G_i$ has an absolutely continuous density function $g_i$ with*

$$\int \left(\frac{g_i'}{g_i}\right)^2 g_i < \infty, \quad i = 1, 2.$$

CONDITION 3.3.2. *The difference of two independent random variables with cdf $G_i$ has cdf $G_i^*$ and density $g_i^*$ which is continuous at 0, $g_i^*(0) > 0, i = 1,2$.*

CONDITION 3.3.3. *The random variable $X_3$ has a positive variance ${\sigma_3}^2$.*

CONDITION 3.3.4. *There exists a positive $\delta$ such that $E(X_3 - \mu_3)^{4+\delta} < \infty$.*

LEMMA 3.3.5 (SIEVERS). *Assume model (3.1.1) and conditions 3.3.1 - 3.3.4. Then*

$$\sqrt{n}(\hat{\beta}_i - \beta_i) \Rightarrow N(0, \frac{1}{12{\gamma_i}^2{\sigma_3}^2}).$$

*where*

$$\gamma_i = \int {g_i}^2, \quad i = 1,2.$$

*and $\Rightarrow$ stands for convergence in distribution.*

REMARK 3.3.6. *It follows from Lemma 3.3.5 that $\hat{\beta}_i \to \beta_i$ in probability, $i = 1,2$.*

Let's introduce some additional notations. Define for $\underline{\nu} = (\nu_1, \nu_2)'$

$$U_n^{(1)}(\underline{\nu}) = \frac{1}{\binom{n}{2}} \sum_{i<j} (X_{1i} - X_{1j} - \nu_1(X_{3i} - X_{3j}))sgn(X_{2i} - X_{2j} - \nu_2(X_{3i} - X_{3j}))$$

$$U_n^{(2)}(\underline{\nu}) = \frac{1}{\binom{n}{2}} \sum_{i<j} |X_{1i} - X_{1j} - \nu_1(X_{3i} - X_{3j})|$$

$$\theta^{(1)}(\underline{\nu}) = E(X_{11} - X_{12} - \nu_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \nu_2(X_{31} - X_{32}))$$

$$\theta^{(2)}(\underline{\nu}) = E|X_{11} - X_{12} - \nu_1(X_{31} - X_{32})|$$

Note that $U_n^{(1)}(\underline{\nu})$ and $U_n^{(2)}(\underline{\nu})$ are U-Statistics with means $\theta^{(1)}(\underline{\nu})$ and $\theta^{(2)}(\underline{\nu})$ respectively. Also note that

$$T_n(\underline{\hat{\beta}}) = \frac{U_n^{(1)}(\underline{\hat{\beta}})}{U_n^{(2)}(\underline{\hat{\beta}})}$$

and

$$\tau^*(\underline{\beta}) = \frac{\theta^{(1)}(\underline{\beta})}{\theta^{(2)}(\underline{\beta})}.$$

The following theorem gives the consistency property of $T_n(\hat{\underline{\beta}})$.

THEOREM 3.3.7. *Assume model (3.1.1) and conditions 3.3.1 - 3.3.4. Then* $T_n(\hat{\underline{\beta}}) \to \tau^*(\underline{\beta})$ *in probability.*

PROOF: We show that $U_n^{(i)}(\hat{\underline{\beta}}) \to \theta^{(i)}(\underline{\beta})$, in probability, i=1,2, from which the theorem follows by an application of Slutsky's theorem. First consider,

$$\begin{aligned}
&|U_n^{(2)}(\hat{\underline{\beta}}) - U_n^{(2)}(\underline{\beta})| \\
&= \frac{1}{\binom{n}{2}} \Big| \sum_{i<j} [|X_{1i} - X_{1j} - \hat{\beta}_1(X_{3i} - X_{3j})| - |X_{1i} - X_{1j} - \beta_1(X_{3i} - X_{3j})|] \Big| \\
&\le \frac{1}{\binom{n}{2}} \sum_{i<j} \big| |X_{1i} - X_{1j} - \hat{\beta}_1(X_{3i} - X_{3j})| - |X_{1i} - X_{1j} - \beta_1(X_{3i} - X_{3j})| \big| \\
&\le \frac{1}{\binom{n}{2}} \sum_{i<j} |(\hat{\beta}_1 - \beta_1)(X_{3i} - X_{3j})| \\
&= |\hat{\beta}_1 - \beta_1| \frac{1}{\binom{n}{2}} \sum_{i<j} |(X_{3i} - X_{3j})|.
\end{aligned}$$

By Remark 3.3.6, $|\hat{\beta}_1 - \beta_1| \to 0$ in probability and hence the right hand side of the preceding equation converges to zero in probability. But $U_n^{(2)}(\underline{\beta})$ is a U-Statistic converging in probability to $\theta^{(2)}(\underline{\beta})$. It follows that $U_n^{(2)}(\hat{\underline{\beta}}) \to \theta^{(2)}(\underline{\beta})$ in probability.

Now Consider

$$
\begin{aligned}
&U_n^{(1)}(\hat{\underline{\beta}}) - U_n^{(1)}(\underline{\beta}) \\
&= \frac{1}{\binom{n}{2}} \sum_{i<j} [(X_{1i} - X_{1j} - \hat{\beta}_1(X_{3i} - X_{3j}))sgn(X_{2i} - X_{2j} - \hat{\beta}_2(X_{3i} - X_{3j})) \\
&\qquad - (X_{1i} - X_{1j} - \beta_1(X_{3i} - X_{3j}))sgn(X_{2i} - X_{2j} - \beta_2(X_{3i} - X_{3j}))] \\
&= -(\hat{\beta}_1 - \beta_1)\frac{1}{\binom{n}{2}} \sum_{i<j} (X_{3i} - X_{3j})sgn(X_{2i} - X_{2j} - \hat{\beta}_2(X_{3i} - X_{3j})) \\
&\qquad + \frac{1}{\binom{n}{2}} \sum_{i<j} (X_{1i} - X_{1j} - \beta_1(X_{3i} - X_{3j}))[sgn(X_{2i} - X_{2j} - \hat{\beta}_2(X_{3i} - X_{3j})) \\
&\qquad\quad - sgn(X_{2i} - X_{2j} - \beta_2(X_{3i} - X_{3j}))]
\end{aligned}
$$

The first term on the right hand side of the preceding equation in absolute value is bounded by

$$
|\hat{\beta}_1 - \beta_1| \frac{1}{\binom{n}{2}} \sum_{i<j} |X_{3i} - X_{3j}|
$$

and the latter converges to zero in probability.

For fixed value of $\underline{\beta} = (\beta_1, \beta_2)'$, denote by

(3.3.8)

$$
\begin{aligned}
\varphi(\hat{\beta}_2) = \frac{1}{\binom{n}{2}} \sum_{i<j} &(X_{1i} - X_{1j} - \beta_1(X_{3i} - X_{3j}))[sgn(X_{2i} - X_{2j} - \hat{\beta}_2(X_{3i} - X_{3j})) \\
&- sgn(X_{2i} - X_{2j} - \beta_2(X_{3i} - X_{3j}))]
\end{aligned}
$$

Let $\Delta > 0$ belong to a bounded interval and consider

(3.3.9)

$$
\varphi(\beta_2 + \frac{\Delta}{\sqrt{n}}) = \frac{1}{\binom{n}{2}} \sum_{i<j} (e_{1i} - e_{1j})[sgn(e_{2i} - e_{2j} - \frac{\Delta}{\sqrt{n}}(X_{3i} - X_{3j})) - sgn(e_{2i} - e_{2j})]
$$

But with probability 1,

$$sgn(e_{2i}-e_{2j}-\frac{\Delta}{\sqrt{n}}(X_{3i}-X_{3j}))-sgn(e_{2i}-e_{2j})$$
$$=\begin{cases} 2, & \text{if } (e_{2i}-e_{2j})<0<(e_{2i}-e_{2j}-\frac{\Delta}{\sqrt{n}}(X_{3i}-X_{3j})) \\ -2, & \text{if } (e_{2i}-e_{2j}-\frac{\Delta}{\sqrt{n}}(X_{3i}-X_{3j}))<0<(e_{2i}-e_{2j}) \\ 0, & \text{if otherwise.} \end{cases}$$
$$=\begin{cases} 2, & \text{if } \frac{\Delta}{\sqrt{n}}(X_{3i}-X_{3j})<(e_{2i}-e_{2j})<0 \\ -2, & \text{if } 0<(e_{2i}-e_{2j})<\frac{\Delta}{\sqrt{n}}(X_{3i}-X_{3j}) \\ 0, & \text{if otherwise.} \end{cases}$$

and thus with probability 1,

$$|sgn(e_{2i}-e_{2j}-\frac{\Delta}{\sqrt{n}}(X_{3i}-X_{3j}))-sgn(e_{2i}-e_{2j})|$$
$$\leq 2I\{-\frac{\Delta}{\sqrt{n}}|X_{3i}-X_{3j}|<(e_{2i}-e_{2j})<\frac{\Delta}{\sqrt{n}}|X_{3i}-X_{3j}|\}$$
$$=2I\{-\frac{\Delta}{\sqrt{n}}|X_{3i}-\overline{X}_3+\overline{X}_3-X_{3j}|<(e_{2i}-e_{2j})<\frac{\Delta}{\sqrt{n}}|X_{3i}-\overline{X}_3+\overline{X}_3-X_{3j}|\}$$
$$\leq 2I\{-2\Delta\max_{1\leq i\leq n}\frac{|X_{3i}-\overline{X}_3|}{\sqrt{n}}<(e_{2i}-e_{2j})<2\Delta\max_{1\leq i\leq n}\frac{|X_{3i}-\overline{X}_3|}{\sqrt{n}}\}$$
$$=2I\{|e_{2i}-e_{2j}|<2\Delta\max_{1\leq i\leq n}\frac{|X_{3i}-\overline{X}_3|}{\sqrt{n}}\}.$$

Substituting in (3.3.9),

$$|\varphi(\beta_2+\frac{\Delta}{\sqrt{n}})|\leq 2\frac{1}{\binom{n}{2}}\sum_{i<j}|e_{1i}-e_{1j})|I\{|e_{2i}-e_{2j}|<2\Delta\max_{1\leq i\leq n}\frac{|X_{3i}-\overline{X}_3|}{\sqrt{n}}\}.$$

From the above it follows that

(3.3.10)

$$\sup_{|\Delta|\leq M}|\varphi(\beta_2+\frac{\Delta}{\sqrt{n}})|\leq 2\frac{1}{\binom{n}{2}}\sum_{i<j}|e_{1i}-e_{1j})|I\{|e_{2i}-e_{2j}|<2M\max_{1\leq i\leq n}\frac{|X_{3i}-\overline{X}_3|}{\sqrt{n}}\}.$$

The right hand side of (3.3.10) is a U-Statistic $U_n$ with symmetric kernel,

$$h_n(X_3,\underline{e}_1,\underline{e}_2;\underline{\beta})=2|e_{11}-e_{12}|I\{|e_{21}-e_{22}|<2M\max_{1\leq i\leq n}\frac{|X_{3i}-\overline{X}_3|}{\sqrt{n}}\}.$$

Note that

$$
\begin{aligned}
EU_n &= Eh_n(X_3, \underset{\sim}{e}_1, \underset{\sim}{e}_2; \underset{\sim}{\beta}) \\
&= 2E[|e_{11} - e_{12})| I\{|e_{21} - e_{22}| < 2M \max_{1 \le i \le n} \frac{|X_{3i} - \overline{X}_3|}{\sqrt{n}}\}] \\
&\le 2[E(e_{11} - e_{12})^2 P(|e_{21} - e_{22}| < 2M \max_{1 \le i \le n} \frac{|X_{3i} - \overline{X}_3|}{\sqrt{n}})]^{\frac{1}{2}}.
\end{aligned}
$$

Since $E(e_{11} - e_{12})^2 = 2{\sigma_1}^2 < \infty$ and because of conditions 3.3.3 and 3.3.4, $\max_{1 \le i \le n} \frac{|X_{3i} - \overline{X}_3|}{\sqrt{n}} \to 0$ a.e., we have

$$EU_n \to 0, \quad \text{as n} \to \infty.$$

Also,

$$E{h_n}^2(X_3, \underset{\sim}{e}_1, \underset{\sim}{e}_2; \underset{\sim}{\beta}) \le 4E(e_{11} - e_{12})^2 = 8{\sigma_1}^2 < \infty, \quad \text{for every n.}$$

Then by results such as Lemma 5.3.11 of Randles and Wolfe (1979),

$$Var(U_n) \to 0, \quad \text{as n} \to \infty.$$

Hence $U_n$ converges to zero in probability.

Then by (3.3.10),

(3.3.11)
$$\sup_{|\Delta| \le M} |\varphi(\beta_2 + \frac{\Delta}{\sqrt{n}})| \to 0, \quad \text{in probability.}$$

From Lemma 3.3.5, we have

(3.3.12)
$$\sqrt{n}(\hat{\beta}_2 - \beta_2) = O_p(1), \quad \text{as n} \to \infty.$$

Let $\epsilon$ and $\delta$ be arbitrary constants. By 3.3.12, there exists a positive real number M such that for all n,

(3.3.13)
$$P(\sqrt{n}|\hat{\beta}_2 - \beta_2| > M) \le \frac{\delta}{2}.$$

Consider now

(3.3.14)

$$
\begin{aligned}
&P(|\varphi(\hat{\beta}_2)| > \epsilon) \\
&= P(|\varphi(\beta_2 + \frac{\hat{\Delta}}{\sqrt{n}}) > \epsilon) \quad \text{where } \hat{\Delta} = \sqrt{n}(\hat{\beta}_2 - \beta_2) \\
&= P(|\varphi(\beta_2 + \frac{\hat{\Delta}}{\sqrt{n}}) > \epsilon, |\hat{\Delta}| \le M) \\
&\quad + P(|\varphi(\beta_2 + \frac{\hat{\Delta}}{\sqrt{n}}) > \epsilon, |\hat{\Delta}| > M) \\
&\le P(|\varphi(\beta_2 + \frac{\hat{\Delta}}{\sqrt{n}}) > \epsilon, |\hat{\Delta}| \le M) + P(|\hat{\Delta}| > M) \\
&\le P(\sup_{|\Delta| \le M} |\varphi(\beta_2 + \frac{\Delta}{\sqrt{n}}) > \epsilon) + P(\sqrt{n}|\hat{\beta}_2 - \beta_2| > M) \\
&\le P(\sup_{|\Delta| \le M} |\varphi(\beta_2 + \frac{\Delta}{\sqrt{n}}) > \epsilon) + \frac{\delta}{2}.
\end{aligned}
$$

From (3.3.11), there is a positive integer N such that for all $n > N$,

$$P(\sup_{|\Delta| \le M} |\varphi(\beta_2 + \frac{\Delta}{\sqrt{n}}) > \epsilon) < \frac{\delta}{2}.$$

Now from (3.3.14), we have for all $n > N$,

$$P(|\varphi(\hat{\beta}_2)| > \epsilon) < \delta,$$

which establishes that $\varphi(\hat{\beta}_2)$ converges to zero in probability.

Thus we have proved that

$$U_n^{(1)}(\underline{\hat{\beta}}) - U_n^{(1)}(\underline{\beta}) \to 0, \quad \text{in probability.}$$

But $U_n^{(1)}(\underline{\beta})$ is a U-Statistic converging in probability to $\theta^{(1)}(\underline{\beta})$, consequently

$$U_n^{(1)}(\underline{\hat{\beta}}) \to \theta^{(1)}(\underline{\beta}), \quad \text{in probability.}$$

By Slutsky's theorem, it follows that $T_n(\underline{\hat{\beta}}) \to \tau^*(\underline{\beta})$ in probability and the proof of the theorem is complete.

## 3.4 Asymptotic Null Distribution of $T_n(\hat{\underline{\beta}})$

In this section we derive the asymptotic normality of $T_n(\hat{\underline{\beta}})$ under the null hypothesis $H_0$ of independence of errors in model (3.1.1). With the notations as in the previous sections, we note that $T_n(\hat{\underline{\beta}})$ is the ratio of two U-statistics $U_n^{(1)}$ and $U_n^{(2)}$ with estimated slope parameters. We shall establish the null asymptotic normality of $T_n(\hat{\underline{\beta}})$ by first deriving the null asymptotic normality of $U_n^{(1)}(\hat{\underline{\beta}})$ and then by an application of Slutsky's theorem. The asymptotic normality of $U_n^{(1)}(\hat{\underline{\beta}})$ under $H_0$ is a direct result of a theorem by Randles (1982), which gives the asymptotic normality of a U-statistic involving estimated parameters. To verify the conditions of Randles' theorem, we need the following assumptions.

CONDITION 3.4.1. *$e_1(e_2)$ is a continuous random variable with a bounded continuous density function, and a finite variance.*

CONDITION 3.4.2. *$X_3$ is a continuous random variable and there exists a $\delta > 0$ such that $E(X_3 - \mu_3)^{4+\delta} < \infty$.*

Now we shall state Randles conditions and theorem.

Let $X_1, X_2, \ldots, X_n$, denote a random sample from some population. Let $h(x_1, x_2, \ldots, x_r : \nu)$ denote a symmetric kernel of degree r with expected value

$$\theta(\nu) = E_\lambda[h(X_1, X_2, \ldots, X_r; \nu)],$$

where $\lambda$ denotes a parameter value, and $\nu$ is, in general, a mathematical variable.

The corresponding U-statistic is

$$U_n(\nu) = \frac{1}{\binom{n}{r}} \sum_{\underline{\alpha} \in A^*} h(X_{\alpha_1}, X_{\alpha_2}, \dots, X_{\alpha_r}; \nu),$$

where $A^*$ denotes the collection of all subsets of size r from the set of integers $\{1, ..., n\}$. Then the conditions of Randles' theorem are as follows.

CONDITION 3.4.3. *Suppose*

$$n^{\frac{1}{2}}(\hat{\lambda} - \lambda) = O_p(1) \quad \textit{as} \quad n \to \infty.$$

CONDITION 3.4.4. *Suppose there is a neighborhood of $\lambda$, say $K(\lambda)$, and a constant $K_1 > 0$ such that if $\nu \in K(\lambda)$ and $D(\nu, d)$ is a sphere centered at $\nu$ with radius d satisfying $D(\nu, d) \subset K(\lambda)$, then,*

$$E[\sup_{\nu' \in D(\nu,d)} |h(X_1, X_2, \dots, X_r; \nu') - h(X_1, X_2, \dots, X_r; \nu)|] \leq K_1 d,$$

*and*

$$\lim_{d \to 0} E[\sup_{\nu' \in D(\nu,d)} |h(X_1, X_2, \dots, X_r; \nu') - h(X_1, X_2, \dots, X_r; \nu)|^2] = 0$$

CONDITION 3.4.5. *$\theta(\nu)$ has a zero differential at $\nu = \lambda$.*

CONDITION 3.4.6.

$$n^{\frac{1}{2}}[U_n(\lambda) - \theta(\lambda)] \Rightarrow N(0, \tau^2), \quad \textit{as} \quad n \to \infty,$$

*where*

$$\tau^2 = r^2 Var(E[h(X_1, X_2, \dots, X_r; \lambda | X_1)]) > 0.$$

Then Randles theorem is

THEOREM 3.4.7 (RANDLES). *Under conditions 3.4.3-3.4.6,*

$$n^{\frac{1}{2}}[U_n(\hat{\lambda}) - \theta(\lambda)] \Rightarrow N(0, \tau^2), \quad as \quad n \to \infty.$$

Now consider the U-statistic,

$$\begin{aligned} U_n^{(1)}(\underline{\nu}) &= \frac{1}{\binom{n}{2}} \sum_{i<j} (X_{1i} - X_{1j} - \nu_1(X_{3i} - X_{3j})) sgn(X_{2i} - X_{2j} - \nu_2(X_{3i} - X_{3j})) \\ &= \frac{1}{\binom{n}{2}} \sum_{i<j} (e_{1i} - e_{1j} - (\nu_1 - \beta_1)(X_{3i} - X_{3j})) sgn(e_{2i} - e_{2j} \\ &\quad - (\nu_2 - \beta_2)(X_{3i} - X_{3j})) \end{aligned}$$

where the mathematical variable $\underline{\nu} = (\nu_1, \nu_2)'$ replaces the estimator $\hat{\underline{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$. The corresponding kernel of degree 2 is

(3.4.8)

$$h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}) = (e_{11} - e_{12} - (\nu_1 - \beta_1)(X_{31} - X_{32})) sgn(e_{21} - e_{22} - (\nu_2 - \beta_2)(X_{31} - X_{32}))$$

with $\underline{S}_i = (X_{3i}, e_{1i}, e_{2i})', i = 1, 2$. Let

(3.4.9)

$$\theta^{(1)}(\underline{\nu}) = Eh^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})$$

$$= E_{\underline{\beta}}(e_{11} - e_{12} - (\nu_1 - \beta_1)(X_{31} - X_{32})) sgn(e_{21} - e_{22} - (\nu_2 - \beta_2)(X_{31} - X_{32}))$$

Now we shall state the main theorem of this section.

THEOREM 3.4.10. *Under conditions 3.4.1 and 3.4.2 and under $H_0$,*

$$n^{\frac{1}{2}} U_n^{(1)}(\hat{\underline{\beta}}) \Rightarrow N(0, \frac{4}{3}\sigma_1{}^2), \quad as \quad n \to \infty.$$

PROOF: This follows from Theorem 3.4.7 (Randles) once we verify conditions 3.4.3 - 3.4.6.

*Condition* 3.4.3 : Conditions 3.4.1 and 3.4.2 imply conditions of Lemma 3.3.5 of section 3.3 and hence both $\sqrt{n}(\hat{\beta}_1 - \beta_1)$ and $\sqrt{n}(\hat{\beta}_2 - \beta_2)$ converge in distribution to normal random variables. This implies that $\sqrt{n}(\hat{\underline{\beta}} - \underline{\beta}) = O_p(1)$ as $n \to \infty$ and hence condition 3.4.3.

*Condition* 3.4.4 : For $\underline{\nu}' = (\nu_1', \nu_2')'$, we consider the following:

$$h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})$$

$$= (e_{11} - e_{12} - (\nu_1' - \beta_1)(X_{31} - X_{32}))sgn(e_{21} - e_{22} - (\nu_2' - \beta_2)(X_{31} - X_{32}))$$

$$- (e_{11} - e_{12} - (\nu_1 - \beta_1)(X_{31} - X_{32}))sgn(e_{21} - e_{22} - (\nu_2 - \beta_2)(X_{31} - X_{32}))$$

Denoting $e_{11} - e_{12}$ by $T_1, e_{21} - e_{22}$ by $T_2$ and $X_{31} - X_{32}$ by S, we have,

$$h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})$$

$$= (T_1 - (\nu_1' - \beta_1)S)sgn(T_2 - (\nu_2' - \beta_2)S)$$

$$- (T_1 - (\nu_1 - \beta_1)S)sgn(T_2 - (\nu_2 - \beta_2)S)$$

$$= (T_1 - (\nu_1' - \beta_1)S)[sgn(T_2 - (\nu_2' - \beta_2)S) - sgn(T_2 - (\nu_2 - \beta_2)S)]$$

$$- (\nu_1' - \nu_1)S \cdot sgn(T_2 - (\nu_2 - \beta_2)S)$$

Hence

(3.4.11)

$$|h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})|$$

$$\leq |(T_1 - (\nu_1' - \beta_1)S)[sgn(T_2 - (\nu_2' - \beta_2)S) - sgn(T_2 - (\nu_2 - \beta_2)S)]|$$

$$+ |\nu_1' - \nu_1||S|.$$

Consider the first term on the right hand side of (3.4.11),

$$|(T_1 - (\nu_1' - \beta_1)S)[sgn(T_2 - (\nu_2' - \beta_2)S) - sgn(T_2 - (\nu_2 - \beta_2)S)]|$$

$$\leq 2|T_1 - (\nu_1' - \beta_1)S|I_A$$

where A is the set of $(S, T_2)$ between the lines $T_2 = (\nu_2{}' - \beta_2)S$ and $T_2 = (\nu_2 - \beta_2)S$. (see figure 3.4.12).

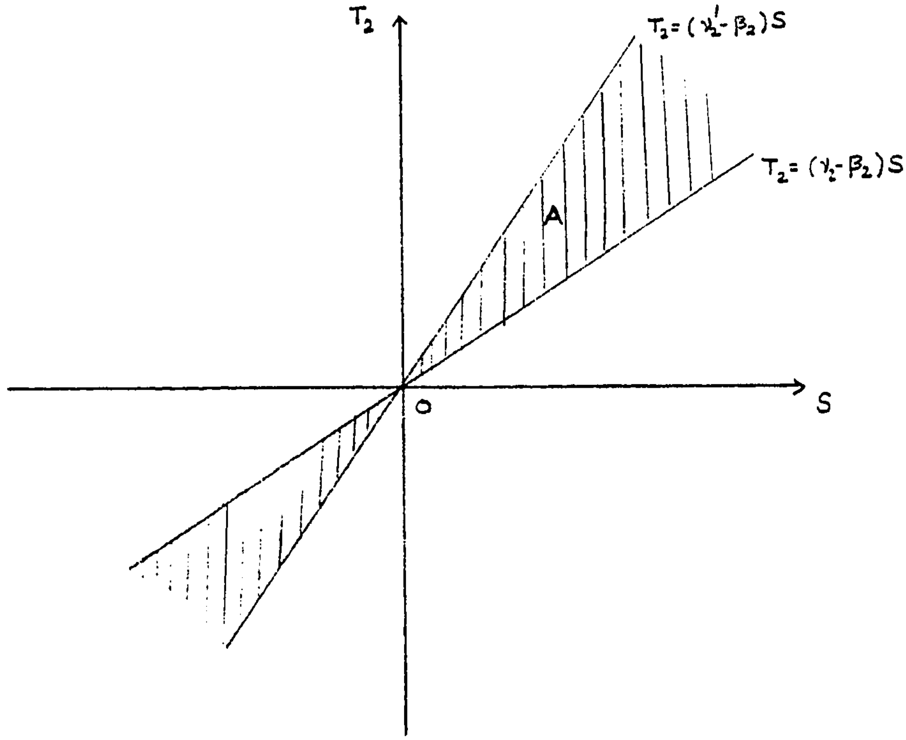


Figure 3.4.12

Note that in figure 3.4.12 above, $[sgn(T_2 - (\nu_2{}' - \beta_2)S) - sgn(T_2 - (\nu_2 - \beta_2)S)]$ is zero outside the band and equal to 2 within the band. The positions of the lines may be interchanged and one or both the lines can have negative slopes. If B represents the set of $(S, T_2)$ between the lines $(max(\nu_2, \nu_2{}') - \beta_2)S$ and $(min(\nu_2, \nu_2{}') - \beta_2)S$, then $A \subset B$ and

$$|(T_1 - (\nu_1{}' - \beta_1)S)[sgn(T_2 - (\nu_2{}' - \beta_2)S) - sgn(T_2 - (\nu_2 - \beta_2)S)]|$$
$$\leq 2|T_1 - (\nu_1{}' - \beta_1)S|I_B.$$

From (3.4.11), it follows that

(3.4.13)

$$|h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu}') - h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})| \le 2|(T_1 - ({\nu_1}' - \beta_1)S)|I_B + |{\nu_1}' - \nu_1||S|.$$

Therefore,

$$\begin{aligned}
&E[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu}') - h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|] \\
&\le 2E[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |(T_1 - ({\nu_1}' - \beta_1)S)|I_B] + E[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |{\nu_1}' - \nu_1||S|] \\
&\le 2E[I_C \sup_{\underline{\nu}'\in D(\underline{\nu},d)} |(T_1 - ({\nu_1}' - \beta_1)S)|] + dE|S|,
\end{aligned}$$

where C is the set of $(S, T_2)$ obtained by replacing $max(\nu_2, {\nu_2}')$ by $\nu_2 + d$ and $min(\nu_2, {\nu_2}')$ by $\nu_2 - d$. Thus,

$$\begin{aligned}
&E[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu}') - h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|] \\
&\le 2E|T_1|I_C + 2\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |{\nu_1}' - \beta_1|E[|S|I_C] + dE|S|.
\end{aligned} \tag{3.4.14}$$

Now consider

$$E[|S|I_C] = E[E(|S|I_C|S)] = E[|S|E(I_C|S)].$$

But

$$\begin{aligned}
E(I_C|S) &= P[(\nu_2 - d - \beta_2)S < T_2 < (\nu_2 + d - \beta_2)S, S > 0] \\
&\quad + P[(\nu_2 + d - \beta_2)S < T_2 < (\nu_2 - d - \beta_2)S, S < 0] \\
&= P[(\nu_2 - d - \beta_2) < \frac{T_2}{S} < (\nu_2 + d - \beta_2), S > 0] \\
&\quad + P[(\nu_2 - d - \beta_2) < \frac{T_2}{S} < (\nu_2 + d - \beta_2), S < 0] \\
&\le P[(\nu_2 - d - \beta_2) < \frac{T_2}{S} < (\nu_2 + d - \beta_2)] \\
&\quad + P[(\nu_2 - d - \beta_2) < \frac{T_2}{S} < (\nu_2 + d - \beta_2)] \\
&= 2P[(\nu_2 - d - \beta_2) < \frac{T_2}{S} < (\nu_2 + d - \beta_2)].
\end{aligned}$$

By conditions 3.4.1 and 3.4.2, the random variable $\frac{T_2}{S} = \frac{e_{21}-e_{22}}{X_{31}-X_{32}}$ has a distribution function $L_2(\cdot)$ and a density $l_2(\cdot)$ which is bounded by $B_2$ and continuous, so that

$$\begin{aligned} &P[(\nu_2 - d - \beta_2) < \frac{T_2}{S} < (\nu_2 + d - \beta_2)] \\ &= L_2(\nu_2 + d - \beta_2) - L_2(\nu_2 - d - \beta_2) \\ &= 2dl_2(\xi) \quad \text{for some} \quad \xi \in (\nu_2 - d - \beta_2, \nu_2 + d - \beta_2) \\ &\leq 2dB_2. \end{aligned}$$

Thus,

$$E(I_C|S) \leq 4dB_2$$

and hence

(3.4.15) $$E[|S|I_C] \leq 4dB_2E|S|.$$

Next consider

$$E[|T_1|I_C] = E[E(|T_1|I_C|T_1)] = E[|T_1|E(I_C|T_1)]$$

But, on lines similar to the calculation of $E(I_C|S)$, we have $E(I_C|T_1) \leq 4dB_2$ so that

(3.4.16) $$E[|T_1|I_C] \leq 4dB_2E|T_1|.$$

Thus from (3.4.14) we have

$$\begin{aligned} &E[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})|] \\ &\leq 8dB_2E|T_1| + 8dB_2E|S| \cdot \sup_{\underline{\nu}' \in D(\underline{\nu},d)} |\nu_1{}' - \beta_1| + dE|S|. \end{aligned}$$

Thus for any neighborhood $K(\underline{\beta})$ of $\underline{\beta}$, with diameter say m, for $\underline{\nu} \in K(\underline{\beta})$ and $D(\underline{\nu}, d) \subset K(\underline{\beta})$,

$$
\begin{aligned}
&E[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})|] \\
&\leq 8dB_2E|T_1| + 8dB_2mE|S| + dE|S|. \\
&= K_1 d
\end{aligned}
$$

where $K_1 = 8B_2E|T_1| + 8B_2mE|S| + E|S|$, is positive and finite. This verifies the first part of condition 3.4.4.

To verify the second part of condition 3.4.4., we first note that from (3.4.13),

$$
\begin{aligned}
&|h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})|^2 \\
&\leq (2|(T_1 - (\nu_1{}' - \beta_1)S)|I_B + |\nu_1{}' - \nu_1||S|)^2 \\
&\leq (2|T_1|I_C + 2|\nu_1{}' - \beta_1||S|I_C + |\nu_1{}' - \nu_1||S|)^2 \\
&= 4T_1{}^2 I_C + 4(\nu_1{}' - \beta_1)^2 S^2 I_C \\
&\qquad + (\nu_1{}' - \nu_1)^2 S^2 + 8|\nu_1{}' - \beta_1||T_1||S|I_C \\
&\qquad + 4|\nu_1{}' - \nu_1||T_1||S|I_C + 4|\nu_1{}' - \beta_1||\nu_1{}' - \nu_1|S^2 I_C
\end{aligned}
$$

By a similar argument as in the first part,

$$
\begin{aligned}
&E[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu}') - h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\nu})|^2] \\
&\leq 16dB_2ET_1{}^2 + 16dB_2m^2ES^2 \\
&\qquad + d^2ES^2 + 32dB_2mE|T_1|E|S| \\
&\qquad + 4dE|T_1|E|S| + 4mdES^2
\end{aligned}
$$

Thus,

$$\lim_{d\to 0} E[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu}') - h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|^2] = 0,$$

which is the second part of condition 3.4.4.

*Condition* 3.4.5 : Consider

$$\theta^{(1)}(\underline{\nu}) = Eh^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})$$

$$= E[E(h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|X_{31},X_{32})]$$

Denoting $(\nu_1-\beta_1)(X_{31}-X_{32})$ by $B_1(\underline{\nu})$ and $(\nu_2-\beta_2)(X_{31}-X_{32})$ by $B_2(\underline{\nu})$, consider the inner conditional expectation in the above expression for the fixed values of $X_{31}=x_{31}$ and $X_{32}=x_{32}$,

$$E(h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|X_{31}=x_{31},X_{32}=x_{32})$$

$$= E(e_{11}-e_{12}-b_1(\underline{\nu}))sgn(e_{21}-e_{22}-b_2(\underline{\nu}))$$

where $b_1(\underline{\nu}) = (\nu_1-\beta_1)(x_{31}-x_{32})$ and $b_2(\underline{\nu}) = (\nu_2-\beta_2)(x_{31}-x_{32})$.

Under the null hypothesis, $(e_{11}-e_{12})$ is independent of $(e_{21}-e_{22})$ and thus the above expectation can be written as the product of two expectations.

Thus,

$$E(h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|X_{31}=x_{31},X_{32}=x_{32})$$

$$= E(e_{11}-e_{12}-b_1(\underline{\nu}))\cdot Esgn(e_{21}-e_{22}-b_2(\underline{\nu}))$$

$$= -b_1(\underline{\nu})[1-2P(e_{21}-e_{22}<b_2(\underline{\nu}))].$$

Recall that $e_{21}-e_{22}$ has a cdf $G_2{}^*$, density $g_2{}^*$, continuous at 0, $g_2{}^*(0)>0$. Hence

$$E(h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|X_{31}=x_{31},X_{32}=x_{32}) = -b_1(\underline{\nu})[1-2G_2{}^*(b_2(\underline{\nu}))].$$

Thus

$$E(h^{(1)}(\underline{S}_1,\underline{S}_2;\underline{\nu})|X_{31},X_{32}) = -B_1(\underline{\nu})[1-2G_2{}^*(B_2(\underline{\nu}))]$$

and

$$\theta^{(1)}(\underline{\nu}) = -E^{X_{31},X_{32}} B_1(\underline{\nu})[1 - 2G_2{}^*(B_2(\underline{\nu}))]$$

where $E^{X_{31},X_{32}}$ represents the expectation induced by the joint distribution of $X_{31}$ and $X_{32}$.

Note that

$$E^{X_{31},X_{32}} B_1(\underline{\nu}) = (\nu_1 - \beta_1)E(X_{31} - X_{32}) = 0,$$

so that

(3.4.17)
$$\theta^{(1)}(\underline{\nu}) = 2E^{X_{31},X_{32}} B_1(\underline{\nu}) G_2{}^*(B_2(\underline{\nu})).$$

We shall use the following result from Randles and Wolfe (1979) in order to pass the differentiation inside the integral.

Result 3.4.18 (Differentiation of an integral). *Let F(.) denote the cdf of a continuous random variable. For the function k(x,t), suppose that the following conditions are satisfied:*

*(i) for each fixed t in $(a,b)$, $\frac{\partial k(x,t)}{\partial t}$ exists for all but at most a countable number of x-values,*

*(ii) for each fixed x, $\frac{\partial k(x,t)}{\partial t}$ exists and satisfies*

$$\left|\frac{\partial k(x,t)}{\partial t}\right| \leq k_*(x),$$

*for all but a countable number of t-values in (a,b), where $k_*(.)$ is such that*

$$\int k_*(x)dF(x) < \infty,$$

*and (iii) k(.) is an absolutely continuous function of t. Then*

$$\frac{d}{dt}\int k(x,t)dF(x) = \int \frac{\partial k(x,t)}{\partial t}dF(x)$$

*for all t in (a,b).*

Clearly because of Conditions 3.4.1 and 3.4.2, differentiation with respect to $\nu_i, i = 1,2$, may be passed inside the integral in (3.4.17) giving

$$\frac{\partial \theta^{(1)}(\underline{\nu})}{\partial \nu_1} = 2E^{X_{31},X_{32}}(X_{31} - X_{32})G_2{}^*(B_2(\underline{\nu}))$$

and

$$\frac{\partial \theta^{(1)}(\underline{\nu})}{\partial \nu_2} = 2E^{X_{31},X_{32}}B_1(\underline{\nu})g_2{}^*(B_2(\underline{\nu}))(X_{31} - X_{32})$$

Since $B_i(\underline{\beta}) = 0, i = 1,2$, and $G_2{}^*$, being symmetric about zero and hence $G_2^*(0) = \frac{1}{2}$, we have

$$\left.\frac{\partial \theta^{(1)}(\underline{\nu})}{\partial \nu_1}\right|_{\underline{\nu}=\underline{\beta}} = E(X_{31} - X_{32}) = 0$$

and

$$\left.\frac{\partial \theta^{(1)}(\underline{\nu})}{\partial \nu_2}\right|_{\underline{\nu}=\underline{\beta}} = 0.$$

Hence $\theta^{(1)}(\underline{\nu})$ has a zero differential at $\underline{\nu} = \underline{\beta}$ and this verifies condition 3.4.5.

*Condition* 3.4.6 : This condition is verified as a direct result of U-Statistic theorems, since

$$U_n^{(1)}(\underline{\beta}) = \frac{1}{\binom{n}{2}}\sum_{i<j}(e_{1i} - e_{1j})sgn(e_{2i} - e_{2j})$$

is a U-Statistic based on i.i.d. observations $(e_{11}, e_{21})', (e_{12}, e_{22})', \ldots, (e_{1n}, e_{2n})'$.

Under $H_0$, the mean of the above U-Statistics is

$$\begin{aligned}\theta^{(1)}(\underline{\beta}) &= E(e_{11} - e_{12})sgn(e_{21} - e_{22}) \\ &= E(e_{11} - e_{12})Esgn(e_{21} - e_{22}) \\ &= 0.\end{aligned}$$

The null asymptotic variance of the above U-Statistics is

$$\tau_1{}^2 = 4Var_{H_0}(E_{H_0}[h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\beta})|\underline{S}_1]).$$

Consider now for the fixed value $\underline{s}_1 = (x_3, y_1, y_2)$ of $\underline{S}_1$,

$$\begin{aligned}&E_{H_0}[h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\beta})|\underline{S}_1 = \underline{s}_1] \\ &= E_{H_0}[(y_1 - e_{12})sgn(y_2 - e_{22})] \\ &= E(y_1 - e_{12})Esgn(y_2 - e_{22}) \\ &= (y_1 - Ee_{12})(2G_2(y_2) - 1) \\ &= (y_1 - Ee_{11})(2G_2(y_2) - 1).\end{aligned}$$

Thus

$$E_{H_0}[h^{(1)}(\underline{S}_1, \underline{S}_2; \underline{\beta})|\underline{S}_1] = (e_{11} - Ee_{11})(2G_2(e_{21}) - 1)$$

and hence

$$\begin{aligned}\tau_1{}^2 &= 4Var_{H_0}[(e_{11} - Ee_{11})(2G_2(e_{21}) - 1)] \\ &= 4E(e_{11} - Ee_{11})^2 E(2G_2(e_{21}) - 1)^2 \\ &= 16\sigma_1{}^2 E(G_2(e_{21}) - \frac{1}{2})^2 \\ &= 16\sigma_1{}^2 \cdot \frac{1}{12} \\ &= \frac{4}{3}\sigma_1{}^2.\end{aligned}$$

Thus we have by a standard U-Statistic theorem, under $H_0$,

$$n^{\frac{1}{2}} U_n^{(1)}(\underline{\beta}) \Rightarrow N(0, \frac{4}{3}{\sigma_1}^2) \quad \text{as} \quad n \to \infty,$$

which verifies condition 3.4.6 and this completes the proof of theorem 3.4.10.

Now we shall state the last theorem of this section which establishes the null asymptotic normality of $T_n(\hat{\underline{\beta}})$.

THEOREM 3.4.19. *Under conditions 3.4.1 and 3.4.2, and under* $H_0$,

$$n^{\frac{1}{2}} T_n(\hat{\underline{\beta}}) \Rightarrow N(0, \frac{4}{3} \cdot \frac{{\sigma_1}^2}{(E|e_{11} - e_{12}|)^2}) \quad as \quad n \to \infty.$$

PROOF: Conditions 3.4.1 and 3.4.2 imply condition 3.3.1 - 3.3.4 of section 3.3 and from Theorem 3.3.7, under the latter conditions,

$$U_n^{(2)}(\hat{\underline{\beta}}) \to \theta^{(2)}(\underline{\beta}) \quad \text{in probability.}$$

But

$$\theta^{(2)}(\underline{\beta}) = E|e_{11} - e_{12}|.$$

This together with Theorem 3.4.10 and an application of Slutsky's theorem establishes Theorem 3.4.19.

## 3.5 The Influence Function of $T_n(\hat{\underline{\beta}})$

The influence curve measures the influence on the population characteristic such as $\tau^*$, being estimated, of a contaminating point mass in the underlying

distribution. The definition of the influence curve for an estimator depends on being able to represent the estimator as a functional evaluated at the empirical cumulative distribution function. Then as pointed out by Hampel (1974), the influence curve is essentially the first derivative of the estimator, viewed as functional at some distribution. This influence curve can be used to study several local robustness properties and also in obtaining the asymptotic variance of the estimator.

DEFINITION OF INFLUENCE CURVE IN THE GENERAL CASE. *Let $\Omega$ be a complete separable metric space, let $T$ be a vector valued mapping from a subset of the probability measures on $\Omega$ into the k-dimensional Euclidean space $R^k$, and let $F$ be in the domain of $T$. Let $\delta_\omega$ denote the atomic probability measure concentrated at any given $\omega \in \Omega$. Then the vector-valued "influence curve" $IC_{T,F}(\cdot)$ of (the "estimator") $T$ at (the "underlying probability distribution") $F$ is defined by*

$$IC_{T,F}(\omega) = \lim_{\epsilon \downarrow 0} \frac{T[(1-\epsilon)F + \epsilon\delta_\omega] - T(F)}{\epsilon}$$

*if this limit is defined for every $\omega \in \Omega$.*

From the point of view of robustness it is desirable to have estimates with bounded influence curves. This means that a single observation cannot have an arbitrarily large effect on the estimate. Another desirable property, perhaps better reflected in the influence curve, is continuity. Jumps or discontinuities in the influence curves indicate local instability at the jump point. For example,

an estimate with such a influence curve may be adversely effected by round off error at these points.

In this section, we shall derive the influence curve of our estimate $T_n(\hat{\underline{\beta}})$ of the partial correlation parameter $\tau^*$. We shall assume the following mild regularity conditions in order to establish the results of this section.

CONDITION 3.5.1. $e_1(e_2)$ *is a continuous random variable with bounded continuous density function and a finite variance.*

CONDITION 3.5.2. $X_3$ *is a continuous random variable with a finite first moment.*

We derive the influence curve of $T_n(\hat{\underline{\beta}})$ in two steps. In the first step we assume that the parameter values $\beta_1$ and $\beta_2$ are known constants and obtain the influence curve of $T_n(\underline{\beta})$. Then we shall extend it to the case where $\beta_1$ and $\beta_2$ are estimated by $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively.

Let $\underline{\beta} = (\beta_1, \beta_2)'$ and $H(\cdot)$ represent the joint distribution of $(X_1, X_2, X_3)'$. Then the coefficient $\tau^*$ in (2.1.4) being estimated by $T_n(\underline{\beta})$ can be written in the functional form as

(3.5.3)

$$T(\beta_1, \beta_2, H) = \frac{E_H(X_{11} - X_{12} - \beta_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_2(X_{31} - X_{32}))}{E_H|X_{11} - X_{12} - \beta_1(X_{31} - X_{32})|}$$

where $E_H$ represents the expectation induced by two independent copies of $(X_1, X_2, X_3)'$ with cdf H.

Let $\underline{x} = (x_1, x_2, x_3)'$ be a fixed point in $R^3$ and $\delta_{\underline{x}}$ is the point mass at $\underline{x}$. Then for $\underline{y} = (y_1, y_2, y_3)'$ in $R^3$, denote by

$$H_t(\underline{y}) = (1-t)H(\underline{y}) + t\delta_{\underline{x}}(\underline{y}). \tag{3.5.4}$$

Then by definition, the influence curve of $T_n(\underline{\beta})$ denoted by $IC(T, H, \underline{x})$ is given by

$$IC(T, H, \underline{x}) = \lim_{t \downarrow 0} \frac{T(\beta_1, \beta_2, H_t) - T(\beta_1, \beta_2, H)}{t}. \tag{3.5.5}$$

Note that the numerator of $T(\beta_1, \beta_2, H_t)$ is equal to

$$(1-t)^2 E_H(X_{11} - X_{12} - \beta_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_2(X_{31} - X_{32}))$$
$$+ 2t(1-t)E_H(X_{11} - x_1 - \beta_1(X_{31} - x_3))sgn(X_{21} - x_2 - \beta_2(X_{31} - x_3))$$

and the denominator of $T(\beta_1, \beta_2, H_t)$ is equal to

$$(1-t)^2 E_H|X_{11} - X_{12} - \beta_1(X_{31} - X_{32})|$$
$$+ 2t(1-t)E_H|X_{11} - x_1 - \beta_1(X_{31} - x_3)|.$$

For simplicity, let a,b, $r(\underline{x})$ and $s(\underline{x})$ denote respectively the quantities $E_H(X_{11} - X_{12} - \beta_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_2(X_{31} - X_{32}))$, $E_H|X_{11} - X_{12} - \beta_1(X_{31} - X_{32})|$, $E_H(X_{11} - x_1 - \beta_1(X_{31} - x_3))sgn(X_{21} - x_2 - \beta_2(X_{31} - x_3))$ and $E_H|X_{11} - x_1 - \beta_1(X_{31} - x_3)|$. Noting that $T(\beta_1, \beta_2, H) = \frac{a}{b}$, we have

$$\begin{aligned}\frac{T(\beta_1, \beta_2, H_t) - T(\beta_1, \beta_2, H)}{t} &= \frac{1}{t}\left[\frac{(1-t)^2 a + 2t(1-t)r(\underline{x})}{(1-t)^2 b + 2t(1-t)s(\underline{x})} - \frac{a}{b}\right] \\ &= \frac{2[r(\underline{x})b - s(\underline{x})a]}{t[(1-t)b + 2ts(\underline{x})]}.\end{aligned}$$

The right hand side of the above expression converges to $2[\frac{r(\underline{x})}{b} - \frac{s(\underline{x})}{b} \cdot \frac{a}{b}]$ as $t \to 0$, and this limit is the influence curve of $T_n(\underline{\beta})$.

Thus,

$$
\begin{aligned}
&IC(T,H,\underline{x}) \\
&= 2\left[\frac{E_H(X_{11}-x_1-\beta_1(X_{31}-x_3))sgn(X_{21}-x_1-\beta_2(X_{31}-x_3))}{E_H|X_{11}-X_{12}-\beta_1(X_{31}-X_{32})|}\right. \\
&\qquad \left.-\frac{E_H|X_{11}-x_1-\beta_1(X_{31}-x_3)|}{E_H|X_{11}-X_{12}-\beta_1(X_{31}-X_{32})|}\cdot T(H)\right],
\end{aligned}
$$

where

$$T(H)=\frac{E_H(X_{11}-X_{12}-\beta_1(X_{31}-X_{32}))sgn(X_{21}-X_{22}-\beta_2(X_{31}-X_{32}))}{E_H|X_{11}-X_{12}-\beta_1(X_{31}-X_{32})|}.$$

Note that $IC(T,H,\underline{x})$ can be written alternatively as,

(3.5.6)
$$
\begin{aligned}
&IC(T,H,\underline{x}) \\
&= 2\left[\frac{E_G(e_{11}-(x_1-\beta_1x_3))sgn(e_{21}-(x_2-\beta_2x_3))}{E_G|e_{11}-e_{12}|}\right. \\
&\qquad \left.-\frac{E_G|e_{11}-(x_1-\beta_1x_3)|}{E_G|e_{11}-e_{12}|}\cdot T(H)\right],
\end{aligned}
$$

where

$$T(H)=\frac{E_G(e_{11}-e_{12})sgn(e_{21}-e_{22})}{E_G|e_{11}-e_{12}|}.$$

Since $\beta_1$ and $\beta_2$ are typically unknown and are estimated, it is meaningful to characterize the influence of the contaminated point mass in the underlying distribution on the values of estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ of $\beta_1$ and $\beta_2$ respectively and in turn on the estimate $T_n(\hat{\underline{\beta}})$. In what follows we shall show that the influence curve for $T_n(\hat{\underline{\beta}})$ is unaffected by the influence of $\hat{\beta}_1$ and $\hat{\beta}_2$.

The parameters $\beta_1$ and $\beta_2$ are defined as solutions to the functional equations

$$\beta_1 = T_1(H) \quad \beta_2 = T_2(H)$$

for which $E_H|X_{11}-X_{12}-T_1(H)(X_{31}-X_{32})|$ and $E_H|X_{21}-X_{22}-T_2(H)(X_{31}-X_{32})|$ are minimum. Now let

$$\psi(\nu_1,\nu_2,t)=T(\nu_1,\nu_2,H_t).$$

Denote by,

$$\frac{\partial\psi(\nu_1,\nu_2,t)}{\partial\nu_1}=\psi_1(\nu_1,\nu_2,t)$$
$$\frac{\partial\psi(\nu_1,\nu_2,t)}{\partial\nu_2}=\psi_2(\nu_1,\nu_2,t)$$
$$\frac{\partial\psi(\nu_1,\nu_2,t)}{\partial t}=\psi_3(\nu_1,\nu_2,t)$$

provided these derivatives exist.

Note that

$$\psi(\beta_1,\beta_2,0)=T(\beta_1,\beta_2,H)=\tau^*$$

$$\psi_3(\beta_1,\beta_2,0)=IC(T,H,\underline{x}),$$

the influence curve of $T_n$ without using estimates of $\beta_1$ and $\beta_2$.

Now let

$$\phi(t)=\psi(T_1(H_t),T_2(H_t),t).$$

Suppose further, $\frac{d}{dt}T_1(H_t)$ and $\frac{d}{dt}T_2(H_t)$ exist, then by chain rule (see for example, Lees (1971)),

$$\begin{aligned}\frac{d}{dt}\phi(t)=\ &\psi_1(T_1(H_t),T_2(H_t),t)\cdot\frac{d}{dt}T_1(H_t)\\&+\psi_2(T_1(H_t),T_2(H_t),t)\cdot\frac{d}{dt}T_2(H_t)\\&+\psi_3(T_1(H_t),T_2(H_t),t).\end{aligned}$$

Note that $\frac{d}{dt}\phi(t)|_{t=0}$ is the influence curve of $T_n$ with estimated parameters $\hat{\beta}_1$ and $\hat{\beta}_2$. We shall denote this by $IC[T(T_1,T_2),H,\underline{x}]$. Also note that $\frac{d}{dt}T_1(H_t)$ and

$\frac{d}{dt}T_2(H_t)$ evaluated at t=0 are the influence curves of $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively. We shall denote these quantities by $IC(T_1, H, \underline{x})$ and $IC(T_2, H, \underline{x})$ respectively.

Substituting for t=0 in the preceding equation and with the notations introduced above, we have

(3.5.7)
$$\begin{aligned} & IC[T(T_1, T_2), H, \underline{x}] \\ &= \psi_1(T_1(H), T_2(H), 0) \cdot IC(T_1, H, \underline{x}) \\ &\quad + \psi_2(T_1(H), T_2(H), 0) \cdot IC(T_2, H, \underline{x}) \\ &\quad + \psi_3(T_1(H), T_2(H), 0) \\ &= \psi_1(\beta_1, \beta_2, 0) \cdot IC(T_1, H, \underline{x}) \\ &\quad + \psi_2(\beta_1, \beta_2, 0) \cdot IC(T_2, H, \underline{x}) \\ &\quad + IC(T, H, \underline{x}). \end{aligned}$$

Here

$$\begin{aligned} &\psi(\nu_1, \nu_2, t) \\ &= \frac{E_{H_t}(X_{11} - X_{12} - \nu_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \nu_2(X_{31} - X_{32}))}{E_{H_t}|X_{11} - X_{12} - \nu_1(X_{31} - X_{32})|} \\ &= \frac{N(\nu_1, \nu_2, t)}{D(\nu_1, \nu_2, t)} \quad \text{say,} \end{aligned}$$

Then

$$\begin{aligned} \psi_1(\nu_1, \nu_2, t) &= \frac{\partial \psi(\nu_1, \nu_2, t)}{\partial \nu_1} \\ &= \frac{D(\nu_1, \nu_2, t)\frac{\partial N(\nu_1, \nu_2, t)}{\partial \nu_1} - N(\nu_1, \nu_2, t)\frac{\partial D(\nu_1, \nu_2, t)}{\partial \nu_1}}{D^2(\nu_1, \nu_2, t)}. \end{aligned}$$

Consider

$$
\begin{aligned}
&D(\nu_1, \nu_2, t) \\
&= E_{H_t}|X_{11} - X_{12} - \nu_1(X_{31} - X_{32})| \\
&= (1-t)^2 E_H|X_{11} - X_{12} - \nu_1(X_{31} - X_{32})| \\
&\qquad + 2t(1-t)E_H|X_{11} - x_1 - \nu_1(X_{31} - x_3)| \\
&= (1-t)^2 E_H(X_{11} - X_{12} - \nu_1(X_{31} - X_{32}))sgn(X_{11} - X_{12} - \nu_1(X_{31} - X_{32})) \\
&\qquad + 2t(1-t)E_H(X_{11} - x_1 - \nu_1(X_{31} - x_3)sgn(X_{11} - x_1 - \nu_1(X_{31} - x_3)).
\end{aligned}
$$

By conditions 3.5.1 and 3.5.2 and Result 3.4.18 (Differentiation of an integral), we can pass the differential inside the expectation yielding the differential of $D(\nu_1, \nu_2, t)$ with respect to $\nu_1$ to be

$$
\begin{aligned}
&\frac{\partial D(\nu_1, \nu_2, t)}{\partial \nu_1} \\
&= -(1-t)^2 E_H(X_{31} - X_{32})sgn(X_{11} - X_{12} - \nu_1(X_{31} - X_{32})) \\
&\qquad - 2t(1-t)E_H(X_{31} - x_3)sgn(X_{11} - x_1 - \nu_1(X_{31} - x_3)).
\end{aligned}
$$

Thus

$$
\begin{aligned}
\left.\frac{\partial D(\nu_1, \nu_2, t)}{\partial \nu_1}\right|_{(\beta_1, \beta_2, 0)} &= -E(X_{31} - X_{32})sgn(X_{21} - X_{22} - \beta_1(X_{31} - X_{32})) \\
&= -E(X_{31} - X_{32})sgn(e_{11} - e_{12}) \\
&= 0 \quad \text{since } X_3 \text{ and } e_1 \text{ are independent.}
\end{aligned}
$$

Now

$$
\begin{aligned}
&N(\nu_1, \nu_2, t) \\
&= E_{H_t}(X_{11} - X_{12} - \nu_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \nu_2(X_{31} - X_{32})) \\
&= (1-t)^2 E_H(X_{11} - X_{12} - \nu_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \nu_2(X_{31} - X_{32})) \\
&\qquad + 2t(1-t)E_H(X_{11} - x_1 - \nu_1(X_{31} - x_3))sgn(X_{21} - x_2 - \nu_2(X_{31} - x_3)).
\end{aligned}
$$

Again because of conditions 3.5.1 and 3.5.2, we can interchange the order of differentiation and taking expectation yielding,

$$\frac{\partial N(\nu_1,\nu_2,t)}{\partial \nu_1}$$
$$= -(1-t)^2 E_H(X_{31}-X_{32})sgn(X_{21}-X_{22}-\nu_2(X_{31}-X_{32}))$$
$$-2t(1-t)E_H(X_{31}-x_3)sgn(X_{21}-x_2-\nu_2(X_{31}-x_3))$$

and hence

$$\left.\frac{\partial N(\nu_1,\nu_2,t)}{\partial \nu_1}\right|_{(\beta_1,\beta_2,0)} = -E(X_{31}-X_{32})sgn(X_{21}-X_{22}-\beta_2(X_{31}-X_{32}))$$
$$= -E(X_{31}-X_{32})sgn(e_{21}-e_{22})$$
$$= 0.$$

Thus

$$\psi_1(\beta_1,\beta_2,0) = 0.$$

Also since $D(\nu_1,\nu_2,t)$ is free from $\nu_2$, clearly

$$\frac{\partial D(\nu_1,\nu_2,t)}{\partial \nu_2} = 0$$

and again by conditions 3.5.1 and 3.5.2

$$\frac{\partial N(\nu_1,\nu_2,t)}{\partial \nu_2} = 0.$$

Thus,

$$\psi_2(\beta_1,\beta_2,0) = 0.$$

Substituting for $\psi_1(\beta_1,\beta_2,0)$ and $\psi_2(\beta_1,\beta_2,0)$ in (3.5.7), we get

$$IC[T(T_1,T_2),H,\underline{x}] = IC(T,H,\underline{x})$$

proving our claim tat the influence curve of $T_n$ with estimated parameters is unaffected by the influence of the estimated parameters.

REMARK 3.5.8.

It is to be noted from (3.5.6) that the estimate $T_n(\hat{\underline{\beta}})$ has unlimited influence in the variables $x_1$ and $x_3$ through the differences $x_1 - \beta_1 x_3$. Hence the influence function is unbounded, and an extreme value of the residual $z_1 = x_1 - \hat{\beta}_1 x_3$ will have a large impact on the estimate. However, the influence is bounded relative to large or extreme residual $z_2 = x_2 - \hat{\beta}_2 x_3$.

REMARK 3.5.9.

A relation between IC and the asymptotic variance of the estimate can be obtained from the following equation,

$$\sqrt{n}(T_n - T(H)) \cong \frac{1}{\sqrt{n}} \sum_{i=1}^{n} IC(T, H, \underline{X}_i) + \text{remainder},$$

for sufficiently large n, where $\underline{X}_i = (X_{1i}, X_{2i}, X_{3i})', i = 1, 2, \ldots, n$, are i.i.d. according to $H(\cdot)$. (See for example, Hampel, Rousseeuw, Ronchetti, and Stahel (1985) for more on this).

The first term on the right rand side of the above equation is asymptotically normal by the central limit theorem. Thus if the remainder becomes negligible for $n \to \infty$, which according to Hampel is true in most cases, $T_n$ itself is asymptotically normal. That is $\sqrt{n}(T_n - T(H))$ converges in distribution to $N(0, V(T, H))$, where the asymptotic variance equals

$$V(T, H) = \int IC^2(T, H, \underline{x}) dH(\underline{x}). \tag{3.5.10}$$

Hampel also points out that when one has calculated the formal asymptotic variance of a certain estimator by means of (3.5.10), it is usually easier to

verify the asymptotic normality in another way instead of trying to assess the necessary regularity conditions to make this approach rigorous.

REMARK 3.5.11.

As an application of the use of IC in evaluating the asymptotic variance of an estimator remarked above, we shall consider the asymptotic null distribution of $T_n(\hat{\underline{\beta}})$.

From (3.5.6), the IC for $T_n(\hat{\underline{\beta}})$ in the null case is

$$\begin{aligned} IC_0(T,H,\underline{x}) &= 2\frac{E(e_{11}-(x_1-\beta_1x_3))Esgn(e_{21}-(x_2-\beta_2x_3))}{E|e_{11}-e_{12}|} \\ &= 2\frac{[Ee_{11}-(x_1-\beta_1x_3)][2G_2(x_2-\beta_2x_3)-1]}{E|e_{11}-e_{12}|} \end{aligned}$$

Thus the asymptotic null variance of $T_n(\hat{\underline{\beta}})$ given by (3.5.10) is

$$\begin{aligned} &V_0(T,H) \\ &= \frac{4}{(E|e_{11}-e_{12}|)^2}\int [Ee_{11}-(x_1-\beta_1x_3)]^2[2G_2(x_2-\beta_2x_3)-1]^2dH(x_1,x_2,x_3) \\ &= \frac{4}{(E|e_{11}-e_{12}|)^2}\int [Ee_{11}-y_1]^2[2G_2(y_2)-1]^2dG_1(y_1)dG_2(y_2) \\ &= \frac{16}{(E|e_{11}-e_{12}|)^2}\int [Ee_{11}-y_1]^2dG_1(y_1)\int\left[G_2(y_2)-\frac{1}{2}\right]^2dG_2(y_2) \\ &= \frac{16}{(E|e_{11}-e_{12}|)^2}\cdot\sigma_1^2\cdot\frac{1}{12} \\ &= \frac{4}{3}\frac{\sigma_1^2}{(E|e_{11}-e_{12}|)^2}, \end{aligned}$$

which is the asymptotic variance of $T_n(\hat{\underline{\beta}})$ given in Theorem 3.4.19.

REMARK 3.5.12.

An alternative method of obtaining the influence curve of $T_n(\underline{\beta})$ is by treating $T_n(\underline{\beta})$ as the ratio of two U-statistics $U_n^{(1)}(\underline{\beta})$ and $U_n^{(2)}(\underline{\beta})$ and using the formula for influence curve of the ratio of two estimators.

Let $T = \frac{T_1}{T_2}$ where $T, T_1$ and $T_2$ are respectively the functionals representing $T_n(\underline{\beta}), U_n^{(1)}(\underline{\beta})$ and $U_n^{(2)}(\underline{\beta})$, and let $\tau^* = \frac{\tau_1^*}{\tau_2^*}$ where $\tau^*, \tau_1^*$, and $\tau_2^*$ are the parameters estimated by $T_n(\underline{\beta}), U_1^{(1)}(\underline{\beta})$ and $U_n^{(2)}(\underline{\beta})$ respectively. Then

$$\begin{aligned}\frac{d}{dt}T(H_t) &= \frac{T_2(H_t)\frac{d}{dt}T_1(H_t) - T_1(H_t)\frac{d}{dt}T_2(H_t)}{[T_2(H_t)]^2} \\ &= \frac{\frac{d}{dt}T_1(H_t)}{T_2(H_t)} - \frac{T_1(H_t)}{T_2(H_t)} \cdot \frac{\frac{d}{dt}T_2(H_t)}{T_2(H_t)}.\end{aligned}$$

Setting t=0 in the above equation,

$$IC(T, H, \underline{x}) = \frac{1}{T_2(H)}[IC(T_1, H, \underline{x}) - T(H)IC(T_2, H, \underline{x})]. \tag{3.5.13}$$

We have

$$\begin{aligned}&T_1(H_t) \\ &= (1-t)^2 E_H(X_{11} - X_{12} - \beta_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_2(X_{31} - X_{32})) \\ &\quad + 2t(1-t)E_H(X_{11} - x_1 - \beta_1(X_{31} - x_3))sgn(X_{21} - x_2 - \beta_2(X_{31} - x_3))\end{aligned}$$

and

$$\begin{aligned}&IC(T_1, H, \underline{x}) \\ &= \frac{d}{dt}T_1(H_t)\Big|_{t=0} \\ &= -2E_H(X_{11} - X_{12} - \beta_1(X_{31} - X_{32}))sgn(X_{21} - X_{22} - \beta_2(X_{31} - X_{32})) \\ &\quad + 2E_H(X_{11} - x_1 - \beta_1(X_{31} - x_3))sgn(X_{21} - x_2 - \beta_2(X_{31} - x_3)) \\ &= 2[E_G(e_{11} - (x_1 - \beta_1 x_3))sgn(e_{21} - (x_2 - \beta_2 x_3)) - T_1(H)].\end{aligned}$$

Similarly,

$$IC(T_2, H, \underline{x}) = 2[E_G|e_{11} - (x_1 - \beta_1 x_3)| - T_2(H)].$$

Thus by (3.5.13),

$$
\begin{aligned}
&IC(T,H,\underline{x})\\
&= \frac{2}{T_2(H)}[E_G(e_{11}-(x_1-\beta_1 x_3))sgn(e_{21}-(x_2-\beta_2 x_3)) - T_1(H)\\
&\qquad - T(H)(E_G|e_{11}-(x_1-\beta_1 x_3)| - T_2(H))]\\
&= \frac{2}{T_2(H)}[E_G(e_{11}-(x_1-\beta_1 x_3))sgn(e_{21}-(x_2-\beta_2 x_3))\\
&\qquad - T(H)\cdot(E_G|e_{11}-(x_1-\beta_1 x_3)|]
\end{aligned}
$$

which is same as the expression (3.5.6), derived directly.

THE INFLUENCE FUNCTION OF $R_{12.3}$.

First note that the coefficient $\rho_{12.3}$ given in (2.4.1) can be written in an alternative form as

$$
\rho_{12.3} = \frac{E(e_{11}-e_{12})(e_{21}-e_{22})}{[E(e_{11}-e_{12})^2 E(e_{21}-e_{22})^2]^{\frac{1}{2}}} \tag{3.5.14}
$$

where $(e_{11}, e_{21})'$ and $(e_{12}, e_{22})'$ are two independent copies of $(e_1, e_2)'$.

Let's assume that $\beta_1$ and $\beta_2$ are known parameter values and consider $H_t$ given in (3.5.4) for the joint distribution of $(X_1, X_2, X_3)'$. Then the coefficient $\rho_{12.3}$ in (3.5.14) being estimated by $R_{12.3}$ can be written in the functional form as

$$
T(\beta_1,\beta_2,H_t) = \frac{T_{12}(\beta_1,\beta_2,H_t)}{[T_{11}(\beta_1,\beta_2,H_t)T_{22}(\beta_1,\beta_2,H_t)]^{\frac{1}{2}}} \tag{3.5.15}
$$

where

(3.5.16)

$$
T_{ij}(\beta_1,\beta_2,H_t) = E_{H_t}(X_{i1}-X_{i2}-\beta_i(X_{31}-X_{32}))(X_{j1}-X_{j2}-\beta_j(X_{31}-X_{32}))
$$

$i \le j; i,j = 1,2.$

By similar calculations as in Remark 3.5.12, we can write the influence function of $R_{12.3}$, suppressing $\beta_1$ and $\beta_2$ as,

$$IC(T,H,\underline{x}) \tag{3.5.17}$$
$$= \frac{\left[IC(T_{12},H,\underline{x}) - \frac{T_{12}(H)}{2T_{11}(H)} IC(T_{11},H,\underline{x}) - \frac{T_{12}(H)}{2T_{22}(H)} IC(T_{22},H,\underline{x})\right]}{[T_{11}(H)T_{22}(H)]^{\frac{1}{2}}}$$

Now

(3.5.18)
$$\begin{aligned}
&IC(T_{ij},H,\underline{x}) \\
&= \frac{d}{dt}T_{ij}(H_t)\Big|_{t=0} \\
&= \frac{d}{dt}[(1-t)^2 E_H(X_{i1} - X_{i2} - \beta_i(X_{31} - X_{32}))(X_{j1} - X_{j2} - \beta_j(X_{31} - X_{32})) \\
&\qquad + 2t(1-t)E_H(X_{i1} - x_i - \beta_i(X_{31} - x_3))(X_{j1} - x_j - \beta_j(X_{31} - x_3))]|_{t=0} \\
&= 2E_H(X_{i1} - X_{i2} - \beta_i(X_{31} - X_{32}))(X_{j1} - X_{j2} - \beta_j(X_{31} - X_{32})) \\
&\qquad + 2E_H(X_{i1} - x_i - \beta_i(X_{31} - x_3))(X_{j1} - x_j - \beta_j(X_{31} - x_3)) \\
&= 2E_G(e_{i1} - e_{i2})(e_{j1} - e_{j2}) + 2E_G(e_{i1} - (x_i - \beta_i x_3))(e_{j1} - (x_j - \beta_j x_3))
\end{aligned}$$

and

$$T_{ij}(H) = E_G(e_{i1} - e_{i2})(e_{j1} - e_{j2}). \tag{3.5.19}$$

Substituting (3.5.18) and (3.5.19) in (3.5.17), we get

(3.5.20)
$$\begin{aligned}
IC(T,H,\underline{x}) = {} & 2\frac{E_G(e_{11} - (x_1 - \beta_1 x_3))(e_{21} - (x_2 - \beta_2 x_3))}{[E_G(e_{11} - e_{12})^2 E_G(e_{21} - e_{22})^2]^{\frac{1}{2}}} \\
& - T(H) \cdot \frac{E_G(e_{11} - (x_1 - \beta_1 x_3))^2}{E_G(e_{11} - e_{12})^2} \\
& - T(H) \cdot \frac{E_G(e_{21} - (x_2 - \beta_2 x_3))^2}{E_G(e_{21} - e_{22})^2}
\end{aligned}$$

In what follows we shall show that the influence function of $R_{12.3}$ is unaffected by the influence of the contaminated point mass on the estimates of $\beta_1$ and $\beta_2$. From (3.5.7) we can establish the above claim by showing that

$$\left.\frac{\partial\psi(\nu_1,\nu_2,t)}{\partial\nu_i}\right|_{(\beta_1,\beta_2,0)} = 0, \quad i = 1,2, \tag{3.5.21}$$

where

$$\psi(\nu_1,\nu_2,t) = T(\nu_1,\nu_2,H_t).$$

In order to show (3.5.21), it suffices to show that for $i \leq j; i,j = 1,2,$

$$\left.\frac{\partial T_{ij}(\nu_1,\nu_2,H_t)}{\partial\nu_1}\right|_{(\beta_1,\beta_2,0)} = 0 \quad \text{and} \quad \left.\frac{\partial T_{ij}(\nu_1,\nu_2,H_t)}{\partial\nu_2}\right|_{(\beta_1,\beta_2,0)} = 0.$$

Because of conditions 3.5.1 and 3.5.2, we can pass the differentiation inside the expectation in each of the above cases. And because of the independence of $X_3$ and $(e_1, e_2)$, this yields

$$\left.\frac{\partial T_{ij}(\nu_1,\nu_2,H_t)}{\partial\nu_1}\right|_{(\beta_1,\beta_2,0)} = 0$$

and

$$\left.\frac{\partial T_{ij}(\nu_1,\nu_2,H_t)}{\partial\nu_2}\right|_{(\beta_1,\beta_2,0)} = 0.$$

This establishes our claim.

It is clear from (3.5.20) that $R_{12.3}$ has unbounded influence in all three variables $x_1, x_2$ and $x_3$ through the squares and cross products of the differences $x_1 - \beta_1 x_3$ and $x_2 - \beta_2 x_3$. Consequently, an extreme value of either of the residuals $z_1 = x_1 - \hat{\beta}_1 x_3$ and $z_2 = x_2 - \hat{\beta}_2 x_3$ will have a large impact on the estimate in contrast to $T_n$ where it is unbounded relative to only $z_1$ and bounded relative to $z_2$.

THE INFLUENCE FUNCTION OF KENDALL'S TAU $t$.

Assuming that $\beta_1$ and $\beta_2$ are known, we can write the coefficient $\tau$ given in (2.4.4) when the underlying distribution is $H_t$ is given by

(3.5.22)

$$T(\beta_1,\beta_2,H_t) = E_{H_t} sgn(X_{11}-X_{12}-\beta_1(X_{31}-X_{32}))sgn(X_{21}-X_{22}-\beta_2(X_{31}-X_{32})).$$

Then the influence curve of t, Kendall's tau calculated from the errors is given by

(3.5.23)

$$\begin{aligned}
&IC(T,H,\underline{x}) \\
&= \frac{d}{dt}T(\beta_1,\beta_2,H_t)\Big|_{t=0} \\
&= \frac{d}{dt}[(1-t)^2 E_G sgn(e_{11}-e_{12})sgn(e_{21}-e_{22}) \\
&\qquad + 2t(1-t)E_G sgn(e_{11}-(x_1-\beta_1 x_3))sgn(e_{21}-(x_2-\beta_2 x_3))]|_{t=0} \\
&= -2E_G sgn(e_{11}-e_{12})sgn(e_{21}-e_{22}) \\
&\qquad + 2E_G sgn(e_{11}-(x_1-\beta_1 x_3))sgn(e_{21}-(x_2-\beta_2 x_3)) \\
&= 2[E_G sgn(e_{11}-(x_1-\beta_1 x_3))sgn(e_{21}-(x_2-\beta_2 x_3)) - T(H)]
\end{aligned}$$

As before,let $\psi(\nu_1,\nu_2,t) = T(\nu_1,\nu_2,H_t)$. Then by conditions 3.5.1 and 3.5.2, we can pass the differentiation inside the expectation yielding, $\frac{\partial\psi(\nu_1,\nu_2,t)}{\partial\nu_1} = 0$ and $\frac{\partial\psi(\nu_1,\nu_2,t)}{\partial\nu_2} = 0$. Again by (3.5.7), it follows that the influence function of 't' is unaffected by the influence of the estimated parameters and the influence function is given by (3.5.23).

It is clear from (3.5.23) that the influence function of Kendall's tau is bounded in all the variables $x_1, x_2$ and $x_3$. Thus large or extreme residuals do

not have any impact on this estimate, exhibiting maximum robustness in this sense among the three estimates under consideration.

## 3.6 Examples

In this section we present three simple examples illustrating the effect of outliers on our estimate of partial correlation, $T_n$ as compared to Pearson's and Kendall's correlation estimates.

EXAMPLE 1.

Table 3.6.1 below represents a sample of 15 independent measurements on 3 normal variables $X_3, e_1$ and $e_2$ where $X_3 \sim N(0,1), e_1 \sim N(0,1), e_2 \sim N(0,2)$ and $X_3$ is independent of both $e_1$ and $e_2$. Minitab statistical package is used in generating these measurements and in the following work. A plot of $e_2$ vs $e_1$ is shown in Figure 3.6.2. Recalling that the distribution of our estimate $T_n$, as well as the other estimates under consideration are free of the regression constants involved in the underlying linear models and of the location and scale parameters of $X_3, e_1$ and $e_2$, the following models are obtained:

$$X_{1i} = X_{3i} + e_{1i}$$

$$X_{2i} = X_{3i} + e_{2i}, \quad i = 1,2,...,15.$$

From these models the residual pairs $(Z_{1i}, Z_{2i}), i = 1,2,...,15$, are obtained in two ways: (i) by the ordinary least squares (OLS) method, and (ii) by the rank method. From the rank residuals we calculated our estimate $T_n$ which is denoted here by $Tn(1)$ and $Tn(2)$ which is same as $Tn(1)$ with the roles of $Z_{1i}$ and $Z_{2i}$

being interchanged. Also calculated from these rank residuals is the Kendall's tau estimate. Recall that this is an estimate of $\frac{2}{\pi} sin^{-1}(\rho)$ in the case of normal population, where $\rho$ is the Pearson's correlation coefficient between $e_1$ and $e_2$. Thus we get an estimate of $\rho$ based on Kendall's tau as $sin(\frac{\pi}{2} \cdot \text{Kendall's tau})$. In Table 3.6.3, the values in the column titled Kendall's refer to these transformed values. Pearson's correlation coefficient between the OLS residuals is calculated. These values are given in row 0 of Table 3.6.3. Rows 1-8 in Table 3.6.3 represent the values of the same statistics calculated when one of the $e_1$ or $e_2$ values is replaced by an outlier point indicated by the same row number in Figure 3.6.2. We see that the $Tn(1)$ and $Tn(2)$ values tend to be less sensitive to outliers than Pearson's correlation coefficient.

Table 3.6.1

| ROW | x3 | e1 | e2 | x1 | x2 |
|---|---|---|---|---|---|
| 1 | -0.92656 | -0.80939 | 0.16478 | -1.73595 | -0.76178 |
| 2 | -1.11506 | 0.78476 | -0.21557 | -0.33030 | -1.33063 |
| 3 | -0.37152 | -0.79275 | -2.36557 | -1.16427 | -2.73709 |
| 4 | -1.63823 | 0.40147 | -0.90313 | -1.23676 | -2.54136 |
| 5 | 1.26228 | 1.35693 | -0.26810 | 2.61921 | 0.99418 |
| 6 | 0.40651 | 0.27768 | 0.18866 | 0.68419 | 0.59517 |
| 7 | -0.85982 | -0.01276 | -1.24084 | -0.87258 | -2.10066 |
| 8 | 0.78136 | -0.26396 | -0.95317 | 0.51740 | -0.17181 |
| 9 | 1.13864 | 0.11984 | -0.76874 | 1.25848 | 0.36990 |
| 10 | -0.44667 | 0.54138 | 0.89386 | 0.09471 | 0.44719 |
| 11 | -1.32507 | -0.17719 | -0.09322 | -1.50226 | -1.41829 |
| 12 | 2.18358 | 0.30060 | -0.11163 | 2.48418 | 2.07195 |
| 13 | 0.21522 | -1.98136 | -1.74780 | -1.76614 | -1.53258 |
| 14 | 1.26396 | -0.41944 | -2.16185 | 0.84452 | -0.89789 |
| 15 | -0.98156 | -0.36664 | 1.19816 | -1.34820 | 0.21660 |

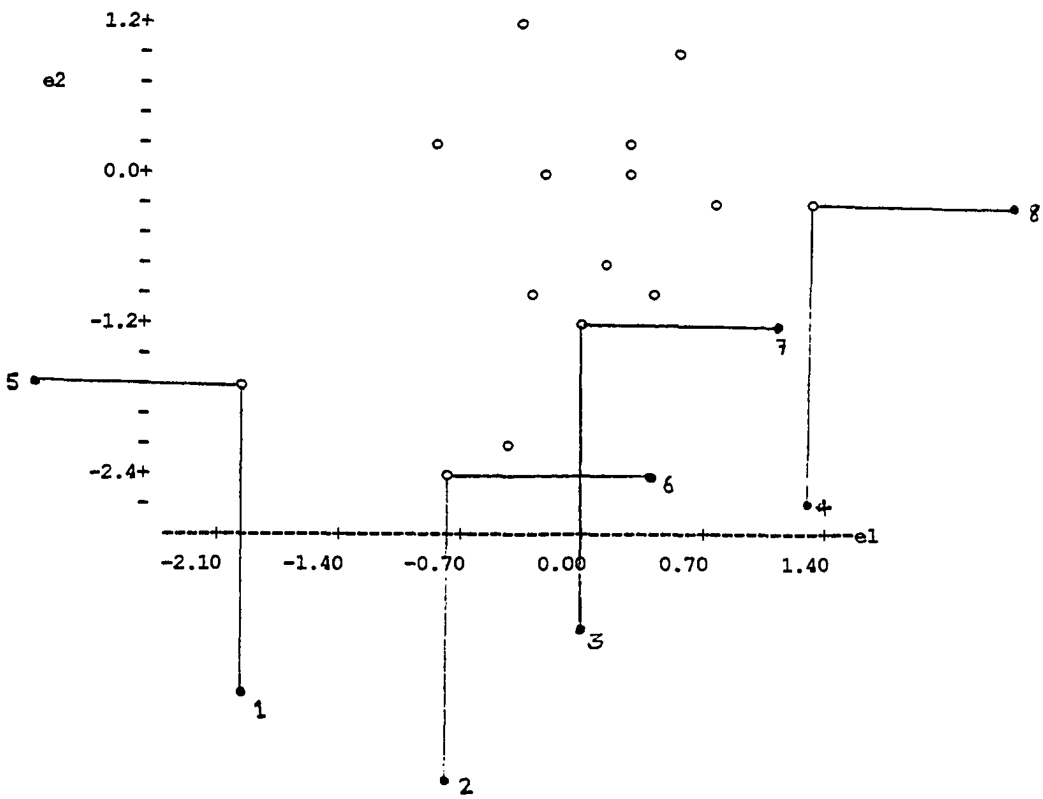


Figure 3.6.2

Table 3.6.3

| Row | corr(e1,e2) | Tn(1) | Tn(2) | Kendall's | Pearson's |
|---|---|---|---|---|---|
| 0 | 0.430 | 0.447 | 0.503 | 0.474 | 0.464 |
| 1 | 0.613 | 0.506 | 0.601 | 0.526 | 0.646 |
| 2 | 0.424 | 0.449 | 0.567 | 0.500 | 0.442 |
| 3 | 0.353 | 0.375 | 0.399 | 0.420 | 0.366 |
| 4 | 0.168 | 0.177 | 0.269 | 0.223 | 0.215 |
| 5 | 0.420 | 0.510 | 0.516 | 0.500 | 0.438 |
| 6 | 0.108 | 0.137 | 0.089 | 0.134 | 0.121 |
| 7 | 0.256 | 0.223 | 0.259 | 0.252 | 0.256 |
| 8 | 0.333 | 0.409 | 0.503 | 0.474 | 0.393 |

EXAMPLE 2.

In this example, $X_{3i}, i = 1,2,...,10$, are generated from a $N(0,1)$ distribution as in Example 1, but $e_{1i}$ and $e_{2i}, i = 1,2,...,10$, are some arbitrary errors. This data is presented in Table 3.6.4 below. Figure 3.6.5 is the plot of $e_2$ vs $e_1$. We do exactly similar calculations as in Example 1 except that we do not transform the values of Kendall's tau in order to get an estimate of $\rho$, since the exact relation in the population is unknown in this case. The values of the 4 statistics are calculated for the data in Table 3.6.4 and when one of the $e_1$ or $e_2$ values is replaced by an outlier. The results are presented in Table 3.6.6. Here. once again we see that $Tn(1)$ and $Tn(2)$ are less sensitive to the outliers than both Kendall's tau and Pearson's correlation coefficients.

Table 3.6.4

| RoW | x3 | e1 | e2 | x1 | x2 |
|---|---|---|---|---|---|
| 1 | -0.92656 | -1.50 | -1.4 | -2.42656 | -2.32656 |
| 2 | -1.11506 | -1.30 | -0.2 | -2.41506 | -1.31506 |
| 3 | -0.37152 | -0.40 | 0.1 | -0.77152 | -0.27152 |
| 4 | -1.63823 | -0.30 | -0.1 | -1.93823 | -1.73823 |
| 5 | 1.26228 | 0.00 | 1.0 | 1.26228 | 2.26228 |
| 6 | 0.40651 | 0.20 | -0.4 | 0.60651 | 0.00651 |
| 7 | -0.85982 | 0.25 | -1.2 | -0.60982 | -2.05982 |
| 8 | 0.78136 | 0.80 | 1.3 | 1.58136 | 2.08136 |
| 9 | 1.13864 | 1.40 | -0.3 | 2.53864 | 0.83864 |
| 10 | -0.44667 | 1.60 | 0.3 | 1.15333 | -0.14667 |

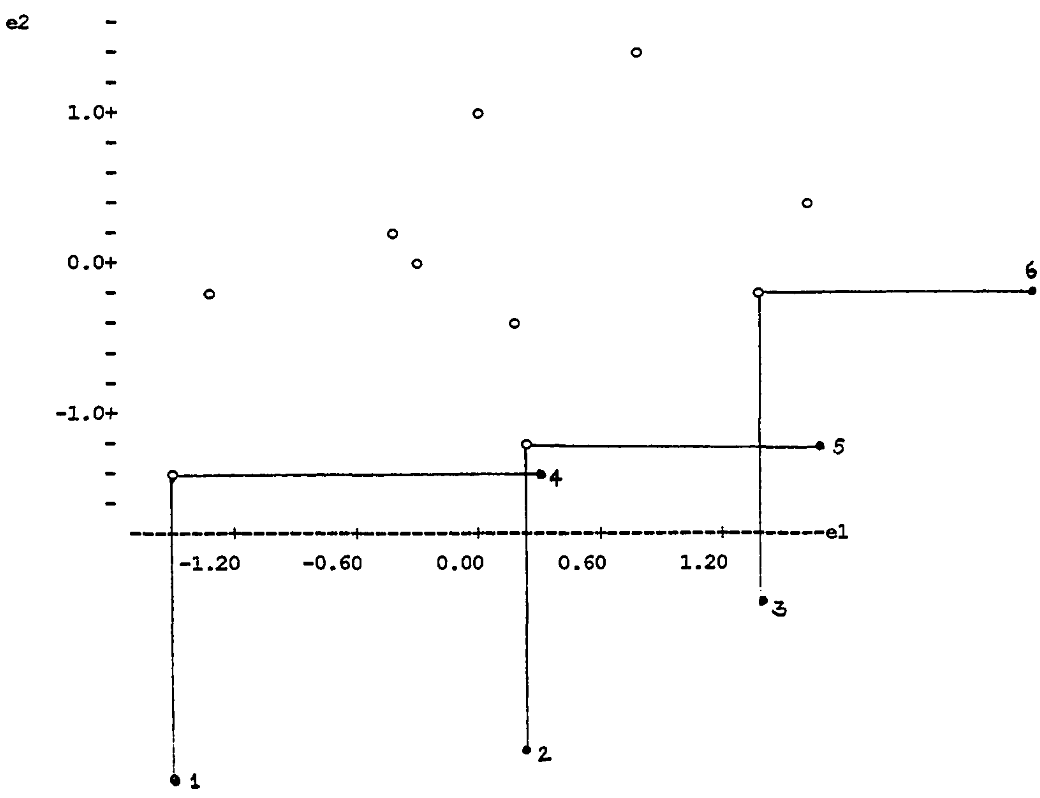


Figure 3.6.5

Table 3.6.6

| Row | corr(e1,e2) | Tn(1) | Tn(2) | Kendall's | Pearson's |
|---|---|---|---|---|---|
| 0 | 0.404 | 0.310 | 0.046 | 0.022 | 0.161 |
| 1 | 0.510 | 0.310 | 0.300 | 0.022 | 0.334 |
| 2 | 0.273 | 0.238 | -0.117 | 0.067 | 0.009 |
| 3 | 0.125 | 0.143 | -0.174 | -0.022 | -0.004 |
| 4 | 0.179 | 0.016 | -0.164 | -0.067 | -0.119 |
| 5 | 0.168 | 0.117 | -0.018 | 0.067 | -0.046 |
| 6 | 0.284 | 0.065 | -0.101 | -0.067 | -0.047 |

EXAMPLE 3.

This example is taken from Steel and Torrie (1960). Table 3.6.7 gives percentages of Nitrogen, ($X_1$), Chlorine, ($X_2$), Potassium, ($X_3$) and log of Leaf burn in seconds, ($Y$), for 30 samples of tobacco taken from farmers' fields. $X_1, X_2$ and $X_3$ are independent variables and $Y$ is a response.

We shall calculate our estimate of partial correlation coefficient between $Y$ and $X_2$ holding $X_1$ fixed, as well as other estimates discussed before. For this we first obtain the regression equations of $Y$ on $X_1$ and $X_2$ on $X_1$ and obtain the residual pairs $(Z_{1i}, Z_{2i}), i = 1, 2, ..., 30$, both by the OLS method and by the rank method. As in examples 1 and 2, rank residuals are used to calculate $Tn(1), Tn(2)$ and the Kendall's tau and OLS residuals for Pearson's partial correlation coefficient. Figures 3.6.8 and 3.6.9 are the plots of these residuals in the two cases. The values of the four statistics calculated are given in row 0 of Table 3.6.10. Once again the effect on these statistics when one of the values of $Y$ or $X_2$ is an outlier in the data set is studied by calculating these statistics in each of these cases. Row 1 of Table 3.6.10 contains values of the four statistics when the 1st measurement on $Y$ 0.34, is being replaced by 3.4 and Row 2 corresponds to 18th measurement on $Y$ 0.51, is being replaced by 5.1. Row 3 contains values of the four statistics when the 15th measurement on $X_2$ 9.2, is being replaced by 9.2 and Row 4 corresponds to the 21st measurement on $X_2$ 1.79, is being replaced by 17.9. We see that $Tn(1)$ is less sensitive to outliers in $X_2$ and $Tn(2)$ is less sensitive to outliers in $Y$. Both of these statistics seem

to be giving more consistent results than Pearson's in each one of the cases.

Table 3.6.7

| ROW | X1 | X2 | X3 | Y |
|---|---|---|---|---|
| 1 | 3.05 | 1.45 | 5.67 | 0.34 |
| 2 | 4.22 | 1.35 | 4.86 | 0.11 |
| 3 | 3.34 | 0.26 | 4.19 | 0.38 |
| 4 | 3.77 | 0.23 | 4.42 | 0.68 |
| 5 | 3.52 | 1.10 | 3.17 | 0.18 |
| 6 | 3.54 | 0.76 | 2.76 | 0.00 |
| 7 | 3.74 | 1.59 | 3.81 | 0.08 |
| 8 | 3.78 | 0.39 | 3.23 | 0.11 |
| 9 | 2.92 | 0.39 | 5.44 | 1.53 |
| 10 | 3.10 | 0.64 | 6.16 | 0.77 |
| 11 | 2.86 | 0.82 | 5.48 | 1.17 |
| 12 | 2.78 | 0.64 | 4.62 | 1.01 |
| 13 | 2.22 | 0.85 | 4.49 | 0.89 |
| 14 | 2.67 | 0.90 | 5.59 | 1.40 |
| 15 | 3.12 | 0.92 | 5.86 | 1.05 |
| 16 | 3.03 | 0.97 | 6.60 | 1.15 |
| 17 | 2.45 | 0.18 | 4.51 | 1.49 |
| 18 | 4.12 | 0.62 | 5.31 | 0.51 |
| 19 | 4.61 | 0.51 | 5.16 | 0.18 |
| 20 | 3.94 | 0.45 | 4.45 | 0.34 |
| 21 | 4.12 | 1.79 | 6.17 | 0.36 |
| 22 | 2.93 | 0.25 | 3.38 | 0.89 |
| 23 | 2.66 | 0.31 | 3.51 | 0.91 |
| 24 | 3.17 | 0.20 | 3.08 | 0.92 |
| 25 | 2.79 | 0.24 | 3.98 | 1.35 |
| 26 | 2.61 | 0.20 | 3.64 | 1.33 |
| 27 | 3.74 | 2.27 | 6.50 | 0.23 |
| 28 | 3.13 | 1.48 | 4.28 | 0.26 |
| 29 | 3.49 | 0.25 | 4.71 | 0.73 |
| 30 | 2.94 | 2.22 | 4.58 | 0.23 |

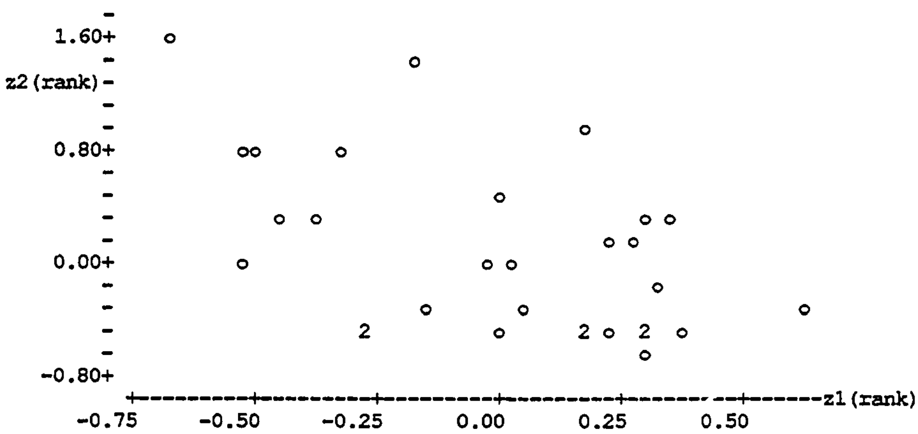


Figure 3.6.8

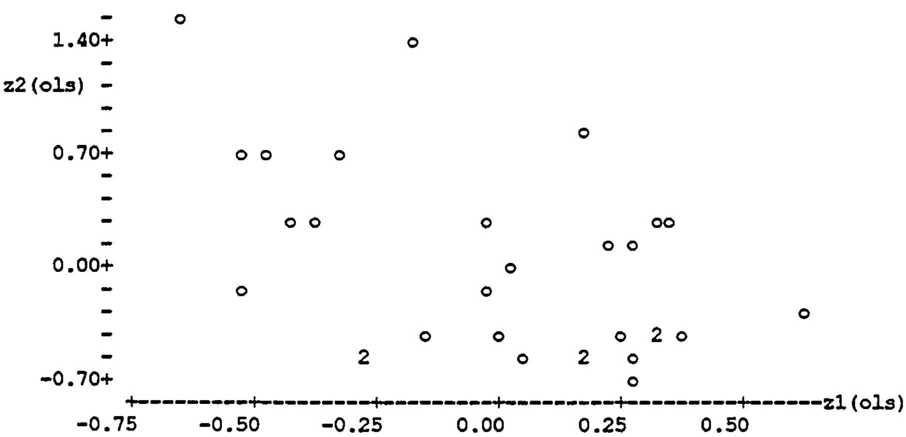


Figure 3.6.9

Table 3.6.10

| Row | Tn(1) | Tn(2) | Kendall's | Pearson's |
|---|---|---|---|---|
| 0 | -0.465 | -0.498 | -0.264 | -0.513 |
| 1 | -0.087 | -0.360 | -0.182 | -0.085 |
| 2 | -0.305 | -0.518 | -0.269 | -0.291 |
| 3 | -0.421 | -0.066 | -0.241 | -0.044 |
| 4 | -0.444 | -0.038 | -0.251 | -0.006 |

CHAPTER IV

# THE ASYMPTOTIC RELATIVE EFFICIENCY

## 4.1 Introduction

We study in this chapter how the two measures of association namely $T_n$ and $R_n$ compare in efficiency when used to test the hypothesis of independence of the errors. The efficiency measure used here will be the asymptotic relative efficiency in the neighborhood of independence as defined by Pitman (1948), later generalized by Noether (1955) and other authors.

It would be obviously desirable if the performances of these coefficients could be studied for a fairly general class of distributions in order to gain insight into the relationship between the coefficients. A general class of distributions has been considered by Konijn (1958), namely the class derived from linear combinations of two independent variables. This is the class of alternatives under which X and Y are given by

$$X = \lambda_1 U + \lambda_2 V, \qquad Y = \lambda_3 U + \lambda_4 V \tag{4.1.1}$$

where the $\lambda_i$'s, $i = 1, 2, 3, 4$ are real numbers, U, V are independently distributed. The hypothesis of independence of X and Y for this model states that $\lambda_1 = \lambda_4 = 1, \lambda_2 = \lambda_3 = 0$. A similar class of alternatives suggested by

Bhuchongkul (1964) are

$$X = (1-\theta)U + \theta Z, \qquad Y = (1-\theta)V + \theta Z \tag{4.1.2}$$

where $0 \leq \theta \leq 1$ and U,V,and Z are independent random variables. The hypothesis of independence to be tested is that $\theta = 0$.

A third model of dependence between two variables which includes Bhuchongkul's model as a special case, was given by Gokhale (1966) and was recommended by Há jek and Šidák (1967) for parametrizing the class of alternatives to the hypothesis of independence. His model is given by

$$X = U + \Delta Z_1, \qquad Y = V + \Delta Z_2 \tag{4.1.3}$$

where U, V and $Z = (Z_1, Z_2)$ are independent random variables. The hypothesis of independence corresponds to $\Delta = 0$, the alternative being $\Delta \neq 0$. This model can be useful, for example, in factor analysis, where test scores are expressed as linear combinations of individual factors and interdependent factors.

Farlie (1960) suggested a model of dependence in terms of the joint distribution of X and Y. The form of the bivariate distribution function proposed by Farlie is an extension of an idea of Morgenstern; see for example, Gumbel (1958). Morgenstern (1956) proposed $F(x)G(y)\{1+\alpha(1-F(x))(1-G(y))\}$ as a bivariate distribution function having F(x) and G(y) as marginal distribution functions and Gumbel noted that this differs from the bivariate normal distribution if $F(x) = \Phi(x)$ and $G(y) = \Phi(y)$, where $\Phi$ is the normal error

function. Farlie suggested that

$$H(x,y) = F(x)G(y)\{1 + \alpha A(F)B(G)\} \tag{4.1.4}$$

where the functions A(F) and B(G) are not completely arbitrary but, are bounded and have bounded first differential coefficients with respect to their arguments, is a suitably general class of bivariate distribution functions for which the marginal distribution functions are F(x) and G(y).

So we are faced with the problem of specifying an appropriate class of alternatives which is sufficiently wide to encompass a large variety of situations and is mathematically manageable. In our setting, this problem is further complicated by the presence of estimated slope parameters which induce dependence among the residual pairs $(Z_{1i}, Z_{2i}), i = 1, 2, \dots, n$. To attain maximum generality and at the same time keep our investigations mathematically manageable, we consider Gokhale's model (4.1.3) with the same variable in place of both $Z_1$ and $Z_2$. This model is same as the one with $Z_1$ and $Z_2$ being different can be seen as follows.

Let $X = U + \Delta Z_1, \quad Y = V + \Delta Z_2$ where U, V and $(Z_1, Z_2)$ are independent. Then there exist random variables $T_1, T_2$, and $T_3$ such that

(i) $T_1, T_2, T_3$ are independent and

(ii) $Z_1 = T_1 + T_3$ and $Z_2 = T_2 + T_3$. Then

$$X = (U + \Delta T_1) + \Delta T_3$$

$$Y = (V + \Delta T_2) + \Delta T_3$$

and $U + \Delta T_1$, $V + \Delta T_2$ and $T_3$ are independent which is similar to (4.1.3).

This class of alternatives is constructed as follows: Let

$$e_{1i} = W_{1i} + \Delta W_{3i}$$

(4.1.5)

$$e_{2i} = W_{2i} + \Delta W_{3i} \quad i = 1, 2, \ldots, n$$

where $\{W_{1i}\}, \{W_{2i}\}$ and $\{W_{3i}\}, i = 1, 2, \ldots, n$, are three independent random samples of random variables. The hypothesis that $e_{1i}$'s and $e_{2i}$'s are independent is equivalent to the hypothesis $\Delta = 0$ versus $\Delta \neq 0$. In order to obtain the Pitman asymptotic relative efficiency(ARE), we will further suppose that $\Delta_n$ is a sequence of parameters converging to the null hypothesis value. i.e., $\lim_{n\to\infty} \Delta_n = 0$. Section 4.2 discusses the asymptotic normality of a general U-Statistic with estimated parameters under a sequence of alternatives, given a set of sufficient conditions. In section 4.3, the asymptotic normality of $T_n$ under a sequence of alternatives converging to the null distribution is established. In section 4.4, we derive the asymptotic distribution of the Pearson's partial correlation coefficient $R_{12.3}$ which we shall denote by $R_n$ henceforth. The Pitman asymptotic relative efficiency of $T_n$ with estimated parameters relative to $R_n$ is discussed in section 4.5 and the calculated values of $ARE(T_n, R_n)$ for several underlying distributions is also given.

## 4.2 Asymptotic Normality of a General U-Statistic with Estimated Parameters

The main result of this section follows by an application of a result due to

Samara (1985) which is an extension of theorem by Randles (1982). Randles' theorem was slightly modified by Samara to apply to the more general case where the U-statistic, $U_n$, and it's moments are functions of the sample size, n, through the observations $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$, whose distribution in turn depends on n perhaps through a sequence of parameters $\Delta_n$.

Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$, denote a random sample from some distribution with distribution function $F_n(\cdot)$, possibly changing as n changes, and let $h(x_1, x_2, \ldots, x_r; \underline{\nu})$ denote a symmetric kernel of degree r with expected value

$$\theta_n(\underline{\nu}) = E_{\underline{\beta}}[h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu})], \tag{4.2.1}$$

where $\underline{\beta}$ denotes a p-dimensional parameter value, and $\underline{\nu}$ is, in general, a mathematical variable. Both the kernel and its expected value may depend on $\underline{\nu}$, and on n through $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$. The correspondin U-statistic is then

$$U_n(\underline{\nu}) = \frac{1}{\binom{n}{r}} \sum_{\underline{\alpha} \in A} h(X_{\alpha_1:n}, X_{\alpha_2:n}, \ldots, X_{\alpha_r:n}; \underline{\nu}) \tag{4.2.2}$$

where A denotes the collection of all subsets of size r from the set of integers $\{1, ..., n\}$. The main result of this section which is due to Samara, gives the asymptotic normality of $n^{\frac{1}{2}}[U_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})]$, where $\hat{\underline{\beta}}$ is an estimator of the parameter $\underline{\beta}$. Now we shall state the conditions needed for Samara's theorem.

CONDITION 4.2.3. *Suppose*

$$n^{\frac{1}{2}}(\hat{\underline{\beta}} - \underline{\beta}) = O_p(1) \quad as \quad n \to \infty.$$

CONDITION 4.2.4. *Suppose there is a neighborhood of $\underline{\beta}$, say $K(\underline{\beta})$, and a positive constant $K_1$ such that if $\underline{\nu} \in K(\underline{\beta})$ and $D(\underline{\nu}, d)$ is a sphere centered at $\underline{\nu}$ with radius d satisfying $D(\underline{\nu}, d) \subset K(\underline{\beta})$, then for every n,*

$$E[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu}') - h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu})|] \leq K_1 d,$$

*and*

$$\lim_{d \to 0} E[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu}') - h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu})|^2] = 0$$

*uniformly in n. That is, for $\epsilon' > 0$ and every n, there exists a positive constant $D'$ such that for $0 < d < D'$ and $D(\underline{\nu}, d) \subset K(\underline{\beta})$,*

$$E[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu}') - h(X_{1:n}, X_{2:n}, \ldots, X_{r:n}; \underline{\nu})|^2] < \epsilon'.$$

The key step in establishing the main result of this section is given in the following theorem.

THEOREM 4.2.5(SAMARA). *Under conditions 4.2.3 and 4.2.4,*

$$n^{\frac{1}{2}}[U_n(\hat{\underline{\beta}}) - \theta_n(\hat{\underline{\beta}}) - U_n(\underline{\beta}) - \theta_n(\underline{\beta})] \to 0, \quad \textit{in probability as} \quad n \to \infty.$$

The main result is stated in the following theorem.

THEOREM 4.2.6(SAMARA). *Suppose that $\theta_n(\underline{\nu})$ is uniformly(in n) differentiable at $\underline{\nu} = \underline{\beta}$ and that this differential is zero. Suppose further that the conditions of Theorem 4.2.5 are satisfied. If in addition,*

$$n^{\frac{1}{2}}[U_n(\underline{\beta}) - \theta_n(\underline{\beta})] \Rightarrow N(0, \tau^2), \quad \textit{as} \quad n \to \infty,$$

*for some* $\tau^2 > 0$, *then*

$$n^{\frac{1}{2}}[U_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] \Rightarrow N(0, \tau^2), \quad \text{as} \quad n \to \infty.$$

The steps involved in the proof of the preceding theorem are as follows. Write

$$\begin{aligned} n^{\frac{1}{2}}[U_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] = & \; n^{\frac{1}{2}}[U_n(\hat{\underline{\beta}}) - \theta_n(\hat{\underline{\beta}}) - U_n(\underline{\beta}) - \theta_n(\underline{\beta})] \\ & + n^{\frac{1}{2}}[\theta_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] \\ & + n^{\frac{1}{2}}[U_n(\underline{\beta}) - \theta_n(\underline{\beta})] \end{aligned}$$

By conditions 4.2.3 and 4.2.4 and Theorem 4.2.5,

$$n^{\frac{1}{2}}[U_n(\hat{\underline{\beta}}) - \theta_n(\hat{\underline{\beta}}) - U_n(\underline{\beta}) - \theta_n(\underline{\beta})] \to 0, \quad \text{in probability as} \quad n \to \infty.$$

Further because of conditions 4.2.3 and uniform differentiability of $\theta_n(\underline{\nu})$ with the derivative at $\underline{\beta}$ being 0,

$$n^{\frac{1}{2}}[\theta_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] \to 0, \quad \text{in probability as} \quad n \to \infty.$$

Then the result follows by Slutsky's theorem.

## 4.3 Asymptotic Normality of $T_n(\hat{\underline{\beta}})$ under a sequence of alternatives

In this section, we derive the asymptotic normality of $T_n(\hat{\underline{\beta}})$ under a sequence of alternatives in the class given by (4.1.4). This is the class in which

$$e_{1i} = W_{1i} + \Delta W_{3i}$$

$$e_{2i} = W_{2i} + \Delta W_{3i}, \quad i = 1, 2, \dots, n,$$

where $\{W_{1i}\}, \{W_{2i}\}$ and $\{W_{3i}\}, i = 1, 2, \ldots, n$, are three independent random samples of random variables. The hypothesis of independence of $e_{1i}$ and $e_{2i}, i = 1, 2, \ldots, n$, is equivalent to the hypothesis $\Delta = 0$. We suppose that $\{\Delta_n\}$ is a sequence of parameter values converging to the null hypothesis value. i.e., $\lim_{n\to\infty} \Delta_n = 0$. In order to establish the asymptotic normality of $T_n(\hat{\underline{\beta}})$, we need the following assumptions.

CONDITION 4.3.1. $\{W_{1i}\}, \{W_{2i}\}$ *and* $\{W_{3i}\}$, $i = 1, \ldots, n$, *are three independent random samples of random variables with absolutely continuous distribution functions* $G_1(\cdot), G_2(\cdot)$ *and* $G_3(\cdot)$, *respectively. Also assume that* $W_{1i}, W_{2i}$ *and* $W_{3i}$ *have finite third moments.*

CONDITION 4.3.2. *The variables* $T_i = W_{i1} - W_{i2}$ *have distribution functions* $F_i(\cdot)$ *and bounded and continuous density functions* $f_i(\cdot), i = 1, 2, 3$.

CONDITION 4.3.3. $X_3$ *is a continuous random variable and there exists a* $\delta > 0$ *such that* $E(X_3 - \mu_3)^{4+\delta} < \infty$.

Let $\underline{S}_{1:n}, \underline{S}_{2:n}, \ldots, \underline{S}_{n:n}$ denote a random sample from a trivariate distribution with distribution function $H_n(\cdot,\cdot,\cdot)$ depending on n, where

$$\underline{S}_{i:n} = \begin{pmatrix} X_{3i} \\ W_{1i} + \Delta_n W_{3i} \\ W_{2i} + \Delta_n W_{3i} \end{pmatrix}. \tag{4.3.4}$$

For $\underline{\nu} = (\nu_1, \nu_2)'$, let

$$\begin{aligned} U_n^{(1)}(\underline{\nu}) = \frac{1}{\binom{n}{2}} \sum_{i<j} &(W_{1i} - W_{1j} + \Delta_n(W_{3i} - W_{3j}) \\ &- (\nu_1 - \beta_1)(X_{3i} - X_{3j}))sgn(W_{2i} - W_{2j} + \Delta_n(W_{3i} - W_{3j}) \\ &- (\nu_2 - \beta_2)(X_{3i} - X_{3j})) \end{aligned}$$

and

$$U_n^{(2)}(\underline{\nu}) = \frac{1}{\binom{n}{2}} \sum_{i<j} |W_{1i} - W_{1j} + \Delta_n(W_{3i} - W_{3j}) - (\nu_1 - \beta_1)(X_{3i} - X_{3j})|.$$

Then $U_n^{(1)}(\underline{\nu})$ and $U_n^{(2)}(\underline{\nu})$ are U-statistics with symmetric kernels of degree 2 given by

$$\begin{aligned} h^{(1)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}) = (W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) \\ - (\nu_1 - \beta_1)(X_{31} - X_{32}))sgn(W_{21} - W_{22} + \Delta_n(W_{31} - W_{32}) \\ - (\nu_2 - \beta_2)(X_{31} - X_{32})) \end{aligned}$$

and

$$\begin{aligned} h^{(2)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}) &= |W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) - (\nu_1 - \beta_1)(X_{31} - X_{32})| \\ &= (W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) \\ &\quad - (\nu_1 - \beta_1)(X_{31} - X_{32}))sgn(W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) \\ &\quad - (\nu_1 - \beta_1)(X_{31} - X_{32})) \end{aligned}$$

respectively.

Note that

$$T_n(\underline{\hat{\beta}}) = \frac{U_n^{(1)}(\underline{\hat{\beta}})}{U_n^{(2)}(\underline{\hat{\beta}})}.$$

Let $\theta_n^{(i)}(\underline{\nu}) = E_{\underline{\beta}} h^{(i)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})$ represent the mean of the U-statistic $U_n^{(i)}(\underline{\nu})$ and let

$$\theta_n(\underline{\nu}) = \frac{\theta_n^{(1)}(\underline{\nu})}{\theta_n^{(2)}(\underline{\nu})}.$$

We shall show that $\sqrt{n}(T_n(\underline{\hat{\beta}}) - \theta_n(\underline{\beta}))$ is asymptotically normal. To this end we first show that $\sqrt{n}(U_n^{(i)}(\underline{\hat{\beta}}) - \theta_n^{(i)}(\underline{\beta}))$ is asymptotically normal for i=1,2. We

use standard U-statistic theorems such as Theorem 5.3.10 of Randles and Wolfe (1979) to show the asymptotic normality of $\sqrt{n}(U_n^{(i)}(\underline{\beta}) - \theta_n^{(i)}(\underline{\beta})), i = 1, 2$ and then Theorem 4.2.6 for the case of $\underline{\beta}$ being estimated by $\hat{\underline{\beta}}$.

Since $\Delta_n$ converges to zero, assume without loss of generality that $|\Delta_n| \leq \Delta$ for some $\Delta > 0$, and all n. We now verify the conditions of Theorem 5.3.10 of Randles and Wolfe (1979). First consider the U-statistic $U_n^{(1)}(\underline{\beta})$ with kernel of degree 2 given by

$$h^{(1)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta}) = (W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}))sgn(W_{21} - W_{22} + \Delta_n(W_{31} - W_{32})).$$

Then

(i)

$$\begin{aligned} Eh^{(1)^2}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta}) &= E(T_1 + \Delta_n T_3)^2 \\ &\leq ET_1^2 + \Delta^2 ET_3^2 \\ &< \infty \quad \text{by condition 4.3.1.} \end{aligned}$$

(ii)

$$\begin{aligned} \theta_n^{(1)}(\underline{\beta}) &= Eh^{(1)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta}) \\ &= E(T_1 + \Delta_n T_3)sgn(T_2 + \Delta_n T_3) \\ &= E^{T_3}E[(T_1 + \Delta_n T_3)sgn(T_2 + \Delta_n T_3)|T_3]. \end{aligned}$$

But

$$E[(T_1 + \Delta_n T_3)sgn(T_2 + \Delta_n T_3)|T_3 = t]$$

$$= E(T_1 + \Delta_n t)sgn(T_2 + \Delta_n t)$$

$$= E(T_1 + \Delta_n t)Esgn(T_2 + \Delta_n t) \quad \text{since } T_1 \text{ and } T_2 \text{ are independent.}$$

$$= \Delta_n t[P(T_2 + \Delta_n t > 0) - P(T_2 + \Delta_n t < 0)] \quad \text{since } ET_1 = 0.$$

$$= \Delta_n t[1 - 2P(T_2 + \Delta_n t < 0)]$$

$$= \Delta_n t[1 - 2P(T_2 < -\Delta_n t)]$$

$$= \Delta_n t[1 - 2F_2(-\Delta_n t)].$$

Thus,

$$(4.3.5) \qquad \theta_n^{(1)}(\underline{\beta}) = \Delta_n E^{T_3} T_3[1 - 2F_2(-\Delta_n T_3)]$$

$$= -2\Delta_n E^{T_3} T_3 F_2(-\Delta_n T_3) \quad \text{since } ET_3 = 0.$$

Let $\underline{s}$ be a fixed value of $\underline{S}_{1:n}$ say $(x_1, y_{1n}, y_{2n})'$ where $y_{1n}$ and $y_{2n}$ are of the form $y_{1n} = w_1 + \Delta_n w_3$ and $y_{2n} = w_2 + \Delta_n w_3$ for some $w_1, w_2$ and $w_3$. Then the limiting variance of $U_n^{(1)}(\underline{\beta})$ is calculated from the quantity $h_{1:n}^{(1)}(\underline{s}) = Eh^{(1)}[(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta})|\underline{S}_{1:n} = \underline{s}]$. Now

$$h_{1:n}^{(1)}(\underline{s}) = E[(y_{1n} - (W_{12} + \Delta_n W_{32}))sgn(y_{2n} - (W_{22} + \Delta_n W_{32}))]$$

$$= E^{W_{32}}[E[(y_{1n} - (W_{12} + \Delta_n W_{32}))sgn(y_{2n} - (W_{22} + \Delta_n W_{32}))|W_{32}].$$

But

$$E[(y_{1n} - (W_{12} + \Delta_n W_{32}))sgn(y_{2n} - (W_{22} + \Delta_n W_{32}))|W_{32} = w]$$

$$= E[(y_{1n} - (W_{12} + \Delta_n w))Esgn(y_{2n} - (W_{22} + \Delta_n w))]$$

$$= (y_{1n} - \Delta_n w - EW_{12})(2G_2(y_{2n} - \Delta_n w) - 1).$$

Thus

$$h_{1:n}^{(1)}(\underline{s}) = E^{W_{32}}[(y_{1n} - \Delta_n W_{32} - EW_{12})(2G_2(y_{2n} - \Delta_n W_{32}) - 1].$$

Then the asymptotic variance of $U_n^{(1)}(\underline{\beta})$ is given by

$$\sigma^{(1,1)} = \lim_{n\to\infty} 4Var\left[h_{1:n}^{(1)}(\underline{S}_{1:n})\right].$$

Noting that $Eh_{1:n}^{(1)}(\underline{S}_{1:n}) = \theta_n^{(1)}(\underline{\beta})$, we can write $\sigma^{(1,1)}$ as

$$\sigma^{(1,1)} = 4\lim_{n\to\infty}\left[Eh_{1:n}^{(1)^2}(\underline{S}_{1:n}) - \theta_n^{(1)^2}(\underline{\beta})\right]. \tag{4.3.6}$$

We shall obtain the above limit by appealing to the following result.

RESULT 4.3.7. *If $X_n \Rightarrow X$ and if $X_n$ is dominated by some $Y$ in $L^1$, then $EX_n \to EX$.*

The above result follows from the following 2 results.

(i) If $X_n \Rightarrow X$ and $X_n$ are uniformly integrable, then X is integrable and $EX_n \to EX$. (Billingsley, 1979, Theorem 25.12, p. 291)

(ii) If $X_n$ is dominated by some Y in $L^1$ then it is uniformly integrable. (Chung, 1974), Ex 7, p. 100)

First, from (4.3.5), since $E|T_3| < \infty$,

$$\lim_{n\to\infty} \theta_n^{(1)}(\underline{\beta}) = -2\lim_{n\to\infty} \Delta_n ET_3F_2(-\Delta_n T_3) = 0 \tag{4.3.8}$$

Next consider

$$h_{1:n}^{(1)}(\underline{S}_{1:n}) = E^{W_{32}}[(W_{11} - EW_{12} + \Delta_n(W_{31} - W_{32}))(2G_2(W_{21} + \Delta_n(W_{31} - W_{32})) - 1)].$$

Let $X_n = (W_{11} - EW_{12} + \Delta_n(W_{31} - W_{32}))(2G_2(W_{21} + \Delta_n(W_{31} - W_{32})) - 1)$. Then, since $G_2$ is continuous, we have $X_n \to X = (W_{11} - EW_{12})(2G_2(W_{21}) - 1)$. Also $|X_n| \le Y = |W_{11} - EW_{12}| + \Delta|W_{31} - W_{32}|$ and $E^{W_{32}}Y = |W_{11} - EW_{12}| + \Delta E^{W_{32}}|W_{31} - W_{32}| < \infty$. Hence by Result 4.3.7,

$$E^{W_{32}}X_n \to E^{W_{32}}X$$

$$\begin{aligned} i.e., \quad \lim_{n\to\infty} h^{(1)}_{1:n}(\underline{S}_{1:n}) &= E^{W_{32}}(W_{11} - EW_{12})(2G_2(W_{21}) - 1) \\ &= (W_{11} - EW_{12})(2G_2(W_{21}) - 1). \end{aligned}$$

From this it follows that

$$\lim_{n\to\infty} h^{(1)^2}_{1:n}(\underline{S}_{1:n}) = (W_{11} - EW_{12})^2(2G_2(W_{21}) - 1)^2.$$

Further,

$$|h^{(1)^2}_{1:n}(\underline{S}_{1:n})| \le [|W_{11} - EW_{12}| + \Delta E^{W_{32}}|W_{31} - W_{32}|]^2$$

and

$$\begin{aligned} &E[|W_{11} - EW_{12}| + \Delta E^{W_{32}}|W_{31} - W_{32}|]^2 \\ &= \sigma_1^2 + 2\Delta^2\sigma_3^2 + 2\Delta E|W_{11} - EW_{12}|E|W_{31} - W_{32}| \\ &< \infty. \quad \text{by assumption 4.3.1} \end{aligned}$$

Again by Result 4.3.7,

(4.3.9)

$$\begin{aligned} \lim_{n\to\infty} Eh^{(1)^2}_{1:n}(\underline{S}_{1:n}) &= E(W_{11} - EW_{12})^2(2G_2(W_{21}) - 1)^2 \\ &= 4E(W_{11} - EW_{12})^2 E(G_2(W_{21}) - \frac{1}{2})^2 \\ &= 4 \cdot \sigma_1^2 \frac{1}{12} \\ &= \frac{1}{3}\sigma_1^2. \end{aligned}$$

Hence by (4.3.6), (4.3.8) and (4.3.9), we have

$$\sigma^{(1,1)} = \frac{4}{3}\sigma_1^2 < \infty. \tag{4.3.10}$$

(iii)

$$\begin{aligned}
&E\left[|h_{1:n}^{(1)}(\underline{S}_{1:n}) - \theta_n^{(1)}(\underline{\beta})|^3\right] \\
&\leq 2^3 E\left[|h_{1:n}^{(1)}(\underline{S}_{1:n})|^3 + |\theta_n^{(1)}(\underline{\beta})|^3\right] \\
&\leq 2E\left[(|W_{11} - EW_{12}| + \Delta E|W_{31} - W_{32}|)^3 + (2\Delta E|T_3|)^3\right] \\
&= 2E\left[(|W_{11} - EW_{12}| + \Delta E|W_{31} - W_{32}|)^3 + 8\Delta^3(E|W_{31} - W_{32}|)^3\right] \\
&< \infty.
\end{aligned}$$

Thus $U_n^{(1)}(\underline{\beta})$ satisfies all three conditions of Theorem 5.3.10 of Randles and Wolfe (1979), and we have the asymptotic normality of $U_n^{(1)}(\underline{\beta})$ under the sequence of alternatives $\{\Delta_n\}$ given by the following theorem.

THEOREM 4.3.11. *Under conditions 4.3.1 - 4.3.3,*

$$n^{\frac{1}{2}}[U_n^{(1)}(\underline{\beta}) - \theta_n^{(1)}(\underline{\beta})] \Rightarrow N(0, \sigma^{(1,1)}) \quad \text{as } n \to \infty$$

*where $\theta_n^{(1)}(\underline{\beta})$ and $\sigma^{(1,1)}$ are given by (4.3.5) and (4.3.10) respectively.*

Now consider the U-statistic $U_n^{(2)}(\underline{\beta})$ with kernel given by $h^{(2)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta}) = |W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}|$. Then

(i)

$$\begin{aligned}
Eh^{(2)^2}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta}) &= E(T_1 + \Delta_n T_3)^2 \\
&\leq ET_1^2 + \Delta^2 ET_3^2 \\
&< \infty, \quad \text{by condition 4.3.1.}
\end{aligned}$$

(ii)

$$\theta_n^{(2)}(\underline{\beta}) = Eh^{(2)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\beta})$$

$$= E|T_1 + \Delta_n T_3|$$

$$= E(T_1 + \Delta_n T_3)sgn(T_1 + \Delta_n T_3)$$

$$= E^{T_3} E[(T_1 + \Delta_n T_3)sgn(T_1 + \Delta_n T_3)|T_3].$$

But

$$E[(T_1 + \Delta_n T_3)sgn(T_1 + \Delta_n T_3)|T_3 = t]$$

$$= E(T_1 + \Delta_n t)sgn(T_1 + \Delta_n t)$$

$$= ET_1 sgn(T_1 + \Delta_n t) + \Delta_n t E sgn(T_1 + \Delta_n t)$$

$$= ET_1 I_{(-\Delta_n t, \infty)} - ET_1 I_{(-\infty, -\Delta_n t)} + \Delta_n t[1 - 2F_1(-\Delta_n t)]$$

$$= -2ET_1 I_{(-\infty, -\Delta_n t)} + \Delta_n t[1 - 2F_1(-\Delta_n t)] \quad \text{since } ET_1 = 0.$$

Thus,

(4.3.12)

$$\theta_n^{(2)}(\underline{\beta}) = -2E^{T_3}[ET_1 I_{(-\infty, -\Delta_n T_3)}] + \Delta_n E^{T_3} T_3[1 - 2F_1(-\Delta_n T_3)]$$

$$= -2E^{T_3}[ET_1 I_{(-\infty, -\Delta_n T_3)}] - 2\Delta_n E^{T_3} T_3 F_1(-\Delta_n T_3)$$

Now consider

$$h_{1:n}^{(2)}(\underline{s}) = E[(y_{1n} - (W_{12} + \Delta_n W_{32}))sgn(y_{1n} - (W_{12} + \Delta_n W_{32}))]$$

$$= E^{W_{32}}[E[(y_{1n} - (W_{12} + \Delta_n W_{32}))sgn(y_{1n} - (W_{12} + \Delta_n W_{32}))|W_{32}].$$

But

$$E[(y_{1n} - (W_{12} + \Delta_n W_{32}))sgn(y_{1n} - (W_{12} + \Delta_n W_{32}))|W_{32} = w]$$

$$= -E[W_{12} sgn(y_{1n} - (W_{12} + \Delta_n w)]$$

$$+ (y_{1n} - \Delta_n w)E sgn(y_{1n} - (W_{12} + \Delta_n w))$$

$$= EW_{12} - 2EW_{12} I_{(-\infty, y_{1n} - \Delta_n w)}$$

$$+ (y_{1n} - \Delta_n w)[2G_1(y_{1n} - \Delta_n w) - 1].$$

Thus

$$h_{1:n}^{(2)}(\underline{s}) = EW_{12} - 2E^{W_{32}}[EW_{12}I_{(-\infty,y_{1n}-\Delta_n W_{32})}]$$

$$+ E^{W_{32}}[(y_{1n} - \Delta_n W_{32})(2G_1(y_{1n} - \Delta_n W_{32}) - 1].$$

Now consider the limiting variance $\sigma^{(2,2)}$ of $U_n^{(2)}(\underline{\beta})$ given by

(4.3.13) $$\sigma^{(2,2)} = 4 \lim_{n\to\infty} \left[ E{h_{1:n}^{(2)}}^2(\underline{S}_{1:n}) - {\theta_n^{(2)}}^2(\underline{\beta}) \right].$$

Since $E|T_1| < \infty$ and $E|T_3| < \infty$, from (4.3.12), we have

(4.3.14) $$\lim_{n\to\infty} \theta_n^{(2)}(\underline{\beta}) = -2ET_1 I_{(-\infty,0)} = E|T_1|.$$

Now consider

$$h_{1:n}^{(2)}(\underline{S}_{1:n}) = EW_{12} - 2E^{W_{32}}[EW_{12}I_{(-\infty,W_{11}+\Delta_n(W_{31}-W_{32}))}]$$

$$+ E^{W_{32}}[(W_{11} + \Delta_n(W_{31} - W_{32}))(2G_1(W_{11} + \Delta_n(W_{31} - W_{32})) - 1].$$

Note that the last term on the right hand side of the preceding equation is similar to $h_{1:n}^{(1)}(\underline{S}_{1:n})$, thus can be easily shown that it converges to $W_{11}(2G_1(W_{11}) - 1)$.

Since both $W_{12}$ and $E|W_{12}|$ are finite, an application Result 4.3.7 twice yields that the middle term on the right hand side of the preceding equation converges to $-2E^{W_{32}}[EW_{12}I_{(-\infty,W_{11})}]$ which is equal to $-2E^{W_{12}}W_{12}I_{(-\infty,W_{11})}$. Thus

$$\lim_{n\to\infty} h_{1:n}^{(2)}(\underline{S}_{1:n}) = EW_{12} - 2E^{W_{12}}W_{12}I_{(-\infty,W_{11})} + W_{11}(2G_1(W_{11}) - 1).$$

From this and condition 4.3.1 and another application of Result 4.3.7, it follows that

(4.3.15)

$$\lim_{n\to\infty} E{h_{1:n}^{(2)}}^2(\underline{S}_{1:n}) = E[EW_{12} - 2E^{W_{12}}W_{12}I_{(-\infty,W_{11})} + W_{11}(2G_1(W_{11}) - 1)]^2 < \infty.$$

Hence by (4.3.13),(4.3.14) and (4.3.15), we have

(4.3.16)

$$\sigma^{(2,2)} = 4\left[E[EW_{12} - 2E^{W_{12}}W_{12}I_{(-\infty,W_{11})} + W_{11}(2G_1(W_{11}) - 1)]^2 - (E|T_1|)^2\right]$$

$$< \infty.$$

Also since $h_{1:n}^{(2)}(\underline{S}_{1:n})$ is non degenerate, $\sigma^{(2,2)} > 0$.

(iii) Since the third moments of $W_1, W_2$ and $W_3$ are all finite by condition 4.3.1, we can easily show that

$$E\left[|h_{1:n}^{(2)}(\underline{S}_{1:n}) - \theta_n^{(2)}(\underline{\beta})|^3\right] < \infty.$$

Thus we have verified all three conditions of Theorem 5.3.10 of Randles and Wolfe (1979), for the U-statistic $U_n^{(2)}(\underline{\beta})$ under the sequence of alternatives $\{\Delta_n\}$ and thus we have the asymptotic normality of $U_n^{(2)}(\underline{\beta})$ given by the following theorem.

THEOREM 4.3.17. *Under conditions 4.3.1 - 4.3.3,*

$$n^{\frac{1}{2}}[U_n^{(2)}(\underline{\beta}) - \theta_n^{(2)}(\underline{\beta})] \Rightarrow N(0, \sigma^{(2,2)}) \quad \text{as } n \to \infty$$

*where $\theta_n^{(2)}(\underline{\beta})$ and $\sigma^{(2,2)}$ are given by (4.3.12) and (4.3.16) respectively.*

Now let $\hat{\underline{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$ be an estimate of $\underline{\beta} = (\beta_1, \beta_2)'$ obtained by the method discussed in Chapter III. We shall now apply Theorem 4.2.6 to obtain the asymptotic distribution of $U_n^{(i)}(\hat{\underline{\beta}}), i = 1, 2$ under a sequence of alternatives $\{\Delta_n\}$ approaching zero. For that we first need to verify conditions 4.2.3 and 4.2.4.

Condition 4.2.3 is discussed under condition 3.4.3 of the previous chapter.

Verification of condition 4.2.4 for $h^{(1)}$ is identical to that of condition 3.4.4 except that here $L_2(\cdot)(l_2(\cdot))$ denote the distribution(density) function of $\frac{T_2+\Delta_n T_3}{X_{31}-X_{32}}$.

Now we show that $\theta_n^{(1)}(\underline{\nu})$ is uniformly(in n) differentiable at $\underline{\nu} = \underline{\beta}$ and this differential is zero.

Denote by $B_i(\underline{\nu}) = (\nu_i - \beta_i)(X_{31} - X_{32})$ and $b_i(\underline{\nu}) = (\nu_i - \beta_i)(x_1 - x_2), i = 1, 2$, where $x_1$ and $x_2$ are some real numbers. Also recall $T_i$ represents the random variable $W_{i1} - W_{i2}, i = 1, 2, 3$. With these notations,

$$\theta_n^{(1)}(\underline{\nu}) = E_{\underline{\beta}}(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))sgn(T_2 + \Delta_n T_3 - B_2(\underline{\nu}))$$

$$= E^{T_3, X_{31}, X_{32}} E_{\underline{\beta}}[(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))sgn(T_2 + \Delta_n T_3 - B_2(\underline{\nu}))|T_3, X_{31}, X_{32}]$$

But

$$E_{\underline{\beta}}[(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))sgn(T_2 + \Delta_n T_3 - B_2(\underline{\nu}))|T_3 = t, X_{31} = x_1, X_{32} = x_2]$$

$$= E_{\underline{\beta}}[(T_1 + \Delta_n t - b_1(\underline{\nu}))sgn(T_2 + \Delta_n t - b_2(\underline{\nu}))]$$

$$= E_{\underline{\beta}}[(T_1 + \Delta_n t - b_1(\underline{\nu}))E_{\underline{\beta}} sgn(T_2 + \Delta_n t - b_2(\underline{\nu}))]$$

$$= (\Delta_n t - b_1(\underline{\nu}))[1 - 2F_2(-\Delta_n t + b_2(\underline{\nu})]$$

Thus

(4.3.18)
$$\theta_n^{(1)}(\underline{\nu}) = E^{T_3, X_{31}, X_{32}}[(\Delta_n T_3 - B_1(\underline{\nu}))(1 - 2F_2(-\Delta_n T_3 + B_2(\underline{\nu}))]$$

$$= E^{T_3, X_{31}, X_{32}}[\Delta_n T_3 - B_1(\underline{\nu})]$$

$$- 2E^{T_3, X_{31}, X_{32}}[(\Delta_n T_3 - B_1(\underline{\nu}))F_2(-\Delta_n T_3 + B_2(\underline{\nu})]$$

$$= -2E^{T_3, X_{31}, X_{32}}[(\Delta_n T_3 - B_1(\underline{\nu}))F_2(-\Delta_n T_3 + B_2(\underline{\nu})].$$

By definition (see for example, Serfling, 1980, p. 45), $\theta_n^{(1)}(\underline{\nu})$ is uniformly(in n) differentiable at $\underline{\nu} = \underline{\beta}$, if for any n, $\frac{\partial \theta_n^{(1)}(\underline{\nu})}{\partial \nu_i}$ exist at $\underline{\nu} = \underline{\beta}, i = 1, 2$, and if

$$\sum_{i=1}^{2} \left. \frac{\partial \theta_n^{(1)}(\underline{\nu})}{\partial \nu_i} \right|_{\underline{\nu}=\underline{\beta}} (\nu_i - \beta_i)$$

satisfies the property that for every $\epsilon > 0, \exists$ a neighborhood $N_\epsilon(\underline{\beta})$ of $\underline{\beta}$ and an integer $N_\epsilon$ such that for any $\underline{\nu} \in N_\epsilon(\underline{\beta})$ and for $n > N_\epsilon$,

$$\left| \theta_n^{(1)}(\underline{\nu}) - \theta_n^{(1)}(\underline{\beta}) - \sum_{i=1}^{2} \left. \frac{\partial \theta_n^{(1)}(\underline{\nu})}{\partial \nu_i} \right|_{\underline{\nu}=\underline{\beta}} (\nu_i - \beta_i) \right| \leq \epsilon \| \underline{\nu} - \underline{\beta} \|. \tag{4.3.19}$$

By conditions 4.3.2 and 4.3.3, we can pass the differentiation with respect to $\nu_i, i = 1, 2$, inside the expectation in (4.3.18) to obtain

$$\begin{aligned} \frac{\partial \theta_n^{(1)}(\underline{\nu})}{\partial \nu_1} &= -2E^{T_3, X_{31}, X_{32}} \frac{\partial [(\Delta_n T_3 - B_1(\underline{\nu}))F_2(-\Delta_n T_3 + B_2(\underline{\nu})]}{\partial \nu_1} \\ &= 2E^{T_3, X_{31}, X_{32}}[(X_{31} - X_{32})F_2(-\Delta_n T_3 + B_2(\underline{\nu}))] \end{aligned}$$

Thus

$$\begin{aligned} \left. \frac{\partial \theta_n^{(1)}(\underline{\nu})}{\partial \nu_1} \right|_{\underline{\nu}=\underline{\beta}} &= 2E^{T_3, X_{31}, X_{32}}[(X_{31} - X_{32})F_2(-\Delta_n T_3)] \\ &= 2E(X_{31} - X_{32})EF_2(-\Delta_n T_3) \quad \text{since } X_3 \text{ and } T_3 \text{ are independent} \\ &= 0 \quad \text{for every n.} \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{\partial \theta_n^{(1)}(\underline{\nu})}{\partial \nu_2} &= -2E^{T_3, X_{31}, X_{32}} \frac{\partial [(\Delta_n T_3 - B_1(\underline{\nu}))F_2(-\Delta_n T_3 + B_2(\underline{\nu})]}{\partial \nu_2} \\ &= -2E^{T_3, X_{31}, X_{32}}[(\Delta_n T_3 - B_1(\underline{\nu}))f_2(-\Delta_n T_3 + B_2(\underline{\nu}))(X_{31} - X_{32})] \end{aligned}$$

Thus

$$\left.\frac{\partial\theta_n^{(1)}(\underline{\nu})}{\partial\nu_2}\right|_{\underline{\nu}=\underline{\beta}} = -2E^{T_3,X_{31},X_{32}}[\Delta_n T_3 f_2(-\Delta_n T_3)(X_{31}-X_{32})]$$

$$= -2\Delta_n E[T_3 f_2(-\Delta_n T_3)]E(X_{31}-X_{32})$$

$$= 0 \quad \text{for every n.}$$

Now consider

$$\theta_n^{(1)}(\underline{\nu}) - \theta_n^{(1)}(\underline{\beta}) - \sum_{i=1}^{2}\left.\frac{\partial\theta_n^{(1)}(\underline{\nu})}{\partial\nu_i}\right|_{\underline{\nu}=\underline{\beta}}(\nu_i-\beta_i)$$

$$= -2E^{T_3,X_{31},X_{32}}[(\Delta_n T_3 - B_1(\underline{\nu}))F_2(-\Delta_n T_3 + B_2(\underline{\nu}))]$$

$$+ 2\Delta_n E^{T_3,X_{31},X_{32}}[T_3 F_2(-\Delta_n T_3)]$$

$$= -2\Delta_n E^{T_3,X_{31},X_{32}} T_3[F_2(-\Delta_n T_3 + B_2(\underline{\nu})) - F_2(-\Delta_n T_3)]$$

$$+ 2E^{T_3,X_{31},X_{32}}[B_1(\underline{\nu})F_2(-\Delta_n T_3 + B_2(\underline{\nu}))]$$

$$= -2\Delta_n E^{T_3,X_{31},X_{32}} T_3[F_2(-\Delta_n T_3 + B_2(\underline{\nu})) - F_2(-\Delta_n T_3)]$$

$$+ 2(\nu_1-\beta_1)E^{T_3,X_{31},X_{32}}[(X_{31}-X_{32})F_2(-\Delta_n T_3 + B_2(\underline{\nu}))]$$

Now

(4.3.20)

$$|\theta_n^{(1)}(\underline{\nu}) - \theta_n^{(1)}(\underline{\beta})|$$

$$\leq 2|\Delta_n|E^{T_3,X_{31},X_{32}}|T_3||F_2(-\Delta_n T_3 + B_2(\underline{\nu})) - F_2(-\Delta_n T_3)|$$

$$+ 2\|\underline{\nu}-\underline{\beta}\||E^{T_3,X_{31},X_{32}}[(X_{31}-X_{32})F_2(-\Delta_n T_3 + B_2(\underline{\nu}))]|$$

Consider the first term on the right hand side of (4.3.20)

$$2|\Delta_n|E^{T_3,X_{31},X_{32}}|T_3||F_2(-\Delta_n T_3 + B_2(\underline{\nu})) - F_2(-\Delta_n T_3)|$$

$$\leq 2|\Delta_n|E^{T_3,X_{31},X_{32}}|T_3||f_2(-\Delta_n T_3 + B_2(\underline{\nu}))B_2(\underline{\nu})|$$

$$\leq 2B_2|\Delta_n||\nu_2-\beta_2|E^{T_3,X_{31},X_{32}}|T_3||X_{31}-X_{32}|$$

$$\leq 8B_2|\Delta_n|\|\underline{\nu}-\underline{\beta}\|E|W_3|E|X_3|$$

where $B_2$ above represents the bound on $f_2$.

Let $\epsilon > 0$, then we can choose n sufficiently large ($\Delta_n$ sufficiently small), say, $n > N_1(\epsilon)$ such that

$$|\Delta_n| < \frac{\epsilon}{16B_2E|W_3|E|X_3|}$$

so that

$$(4.3.21)\quad 2|\Delta_n|E^{T_3,X_{31},X_{32}}|T_3||F_2(-\Delta_nT_3+B_2(\underline{\nu}))-F_2(-\Delta_nT_3)| \leq \frac{1}{2}\epsilon\|\underline{\nu}-\underline{\beta}\|.$$

Now consider the second term on the right hand side of (4.3.20). Note that

$$(X_{31}-X_{32})F_2(-\Delta_nT_3+B_2(\underline{\nu})) \to (X_{31}-X_{32})F_2(B_2(\underline{\nu}))$$

Further,

$$|(X_{31}-X_{32})F_2(-\Delta_nT_3+B_2(\underline{\nu}))| \leq |X_{31}-X_{32}|$$

and $E|X_{31}-X_{32}| < \infty$. Hence, by Dominated Convergence Theorem,

$$E^{T_3,X_{31},X_{32}}[(X_{31}-X_{32})F_2(-\Delta_nT_3+B_2(\underline{\nu})] \to E(X_{31}-X_{32})F_2(B_2(\underline{\nu})) \quad \text{as } n \to \infty.$$

Also note that, for $\underline{\nu}$ sufficiently close to $\underline{\beta}$, then $B_2(\underline{\nu})$ is close to zero and thus another application of Dominated Convergence Theorem gives,

$$\begin{aligned} E[(X_{31}-X_{32})F_2(B_2(\underline{\nu}))] &\to E(X_{31}-X_{32})F_2(0) \\ &= \frac{1}{2}E(X_{31}-X_{32}) \\ &= 0 \end{aligned}$$

Thus for the same $\epsilon > 0, \exists$ a neighborhood $N_\epsilon(\underline{\beta})$ of $\underline{\beta}$ and an integer $N_2(\epsilon)$ such that for $\underline{\nu} \in N_\epsilon(\underline{\beta})$ and $n > N_2(\epsilon)$.

$$|E^{T_3, X_{31}, X_{32}}(X_{31} - X_{32})F_2(-\Delta_n T_3 + B_2(\underline{\nu}))| \leq \frac{\epsilon}{4}$$

so that

$$(4.3.22)\quad 2\|\underline{\nu} - \underline{\beta}\||E^{T_3, X_{31}, X_{32}}[(X_{31} - X_{32})F_2(-\Delta_n T_3 + B_2(\underline{\nu}))]| \leq \frac{1}{2}\epsilon\|\underline{\nu} - \underline{\beta}\|.$$

Now for $\epsilon > 0$, let $N_\epsilon^{(1)} = max(N_1(\epsilon), N_2(\epsilon))$. Then from (4.3.21) and (4.3.22) $\exists$ a neighborhood $N_\epsilon^{(1)}(\underline{\beta})$ and $n > N_\epsilon^{(1)}$,

$$|\theta_n^{(1)}(\underline{\nu}) - \theta_n^{(1)}(\underline{\beta})| \leq \epsilon\|\underline{\nu} - \underline{\beta}\|$$

which establishes the uniform(in n) differentiability of $\theta_n(\underline{\nu})$ at $\underline{\beta}$ and that the differential is equal to zero.

We shall now verify condition 4.2.4 for $h^{(2)}$. For that we consider,

(4.3.23)

$$|h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}') - h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|$$

$$= |(T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32}))sgn(T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32}))$$

$$- (T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32}))sgn(T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32}))|$$

$$= |(T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32}))[sgn(T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32}))$$

$$- sgn(T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32}))]$$

$$- ({\nu_1}' - \nu_1)(X_{31} - X_{32})sgn(T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32}))|$$

$$\leq |T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32})||sgn(T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32}))$$

$$- sgn(T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32}))|$$

$$+ |{\nu_1}' - \nu_1||X_{31} - X_{32}|$$

$$\leq 2|T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32})|I_A + |{\nu_1}' - \nu_1||X_{31} - X_{32}|$$

where A is the set of $(X_{31} - X_{32}, T_1 + \Delta_n T_3)$ between the lines $T_1 + \Delta_n T_3 = (min(\nu_1, {\nu_1}') - \beta_1)(X_{31} - X_{32})$ and $T_1 + \Delta_n T_3 = (max(\nu_1, {\nu_1}') - \beta_1)(X_{31} - X_{32})$.

Now if B denotes the set of $(X_{31} - X_{32}, T_1 + \Delta_n T_3)$ between the lines $T_1 + \Delta_n T_3 = (\nu_1 - d - \beta_1)(X_{31} - X_{32})$ and $T_1 + \Delta_n T_3 = (\nu_1 + d - \beta_1)(X_{31} - X_{32})$, then

$$\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}') - h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|$$

$$\leq 2I_B \sup_{\underline{\nu}' \in D(\underline{\nu},d)} |T_1 + \Delta_n T_3 - ({\nu_1}' - \beta_1)(X_{31} - X_{32})|$$

$$+ d|X_{31} - X_{32}|$$

$$\leq 2|X_{31} - X_{32}|I_B \sup_{\underline{\nu}' \in D(\underline{\nu},d)} \left|\frac{T_1 + \Delta_n T_3}{X_{31} - X_{32}} - ({\nu_1}' - \beta_1)\right|$$

$$+ d|X_{31} - X_{32}|.$$

Note that

$$\left|\frac{T_1+\Delta_n T_3}{X_{31}-X_{32}}-({\nu_1}'-\beta_1)\right| = \left|\frac{T_1+\Delta_n T_3}{X_{31}-X_{32}}-(\nu_1-\beta_1)-({\nu_1}'-\nu_1)\right|$$
$$\leq \left|\frac{T_1+\Delta_n T_3}{X_{31}-X_{32}}-(\nu_1-\beta_1)\right| + |{\nu_1}'-\nu_1|$$

and on the set B,

$$\left|\frac{T_1+\Delta_n T_3}{X_{31}-X_{32}}-(\nu_1-\beta_1)\right| \leq d.$$

Consequently,

$$I_B \sup_{\underline{\nu}'\in D(\underline{\nu},d)} \left|\frac{T_1+\Delta_n T_3}{X_{31}-X_{32}}-({\nu_1}'-\beta_1)\right| \leq d+d=2d.$$

Thus,

$$\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}')-h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|$$
$$\leq 2|X_{31}-X_{32}|(2d)+d|X_{31}-X_{32}|$$
$$\leq 5d|X_{31}-X_{32}|.$$

From the above, it follows that for every n,

$$E\left[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}')-h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|\right] \leq K_2 d$$

where $K_2 = 5E|X_{31}-X_{32}| < \infty$.

Further, from (4.3.23),

$$|h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}')-h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|^2$$
$$\leq 4|T_1+\Delta_n T_3-({\nu_1}'-\beta_1)(X_{31}-X_{32})|^2 I_A$$
$$+|{\nu_1}'-\nu_1|^2|X_{31}-X_{32}|^2$$
$$+4|{\nu_1}'-\nu_1||X_{31}-X_{32}||T_1+\Delta_n T_3-({\nu_1}'-\beta_1)(X_{31}-X_{32})|I_A$$

and thus

$$\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}') - h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|^2$$

$$\leq 4|X_{31} - X_{32})|^2(2d)^2 + d^2|X_{31} - X_{32}|^2$$

$$+ 4d|X_{31} - X_{32}||X_{31} - X_{32}|(2d)$$

$$= 25d^2|X_{31} - X_{32}|^2.$$

Hence

$$\lim_{d\to 0} E\left[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}') - h^{(2)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|^2\right] = 0$$

uniformly in n. This verifies condition 4.2.4 for $h^{(2)}$.

Now we shall show that $\theta_n^{(2)}(\underline{\nu})$ is uniformly (in n) differentiable at $\underline{\nu} = \underline{\beta}$ and the differential is zero.

$$\theta_n^{(2)}(\underline{\nu}) = E_{\underline{\beta}}(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))sgn(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))$$

$$= E^{T_3,X_{31},X_{32}} E_{\underline{\beta}}[(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))sgn(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))|T_3, X_{31}, X_{32}]$$

But

$$E_{\underline{\beta}}[(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))sgn(T_1 + \Delta_n T_3 - B_1(\underline{\nu}))|T_3 = t, X_{31} = x_1, X_{32} = x_2]$$

$$= E_{\underline{\beta}}[(T_1 + \Delta_n t - b_1(\underline{\nu}))sgn(T_1 + \Delta_n t - b_1(\underline{\nu}))]$$

$$= E_{\underline{\beta}}[T_1 sgn(T_1 + \Delta_n t - b_1(\underline{\nu}))$$

$$+ (\Delta_n t - b_1(\underline{\nu}))E_{\underline{\beta}} sgn(T_1 + \Delta_n t - b_1(\underline{\nu}))$$

$$= -2ET_1 I_{(-\infty, -\Delta_n t + b_1(\underline{\nu}))}$$

$$+ (\Delta_n t - b_1(\underline{\nu}))[1 - 2F_1(-\Delta_n t + b_1(\underline{\nu}))]$$

Thus

(4.3.25)

$$\theta_n^{(2)}(\underline{\nu}) = E^{T_3,X_{31},X_{32}}[-2ET_1 I_{(-\infty,-\Delta_n T_3+B_1(\underline{\nu}))}$$

$$+ (\Delta_n T_3 - B_1(\underline{\nu}))[1 - 2F_1(-\Delta_n T_3 + B_1(\underline{\nu}))]]$$

$$= -2E^{T_3,X_{31},X_{32}}[ET_1 I_{(-\infty,-\Delta_n T_3+B_1(\underline{\nu}))}$$

$$+ (\Delta_n T_3 - B_1(\underline{\nu}))F_1(-\Delta_n T_3 + B_1(\underline{\nu}))]$$

Again by conditions 4.3.2 and 4.3.3, we can pass the differentiation with respect to $\nu_i, i = 1,2$, inside the expectation to get,

$$\frac{\partial \theta_n^{(2)}(\underline{\nu})}{\partial \nu_1} = -2E^{T_3,X_{31},X_{32}}[(-\Delta_n T_3 + B_1(\underline{\nu}))f_1(-\Delta_n T_3 + B_1(\underline{\nu}))(X_{31} - X_{32})$$

$$+(\Delta_n T_3 - B_1(\underline{\nu}))f_1(-\Delta_n T_3 + B_1(\underline{\nu}))(X_{31} - X_{32})$$

$$-F_1(-\Delta_n T_3 + B_1(\underline{\nu}))(X_{31} - X_{32})]$$

$$= 2E^{T_3,X_{31},X_{32}}F_1(-\Delta_n T_3 + B_1(\underline{\nu}))(X_{31} - X_{32}).$$

Thus,

$$\left.\frac{\partial \theta_n^{(2)}(\underline{\nu})}{\partial \nu_1}\right|_{\underline{\nu}=\underline{\beta}} = 2E^{T_3,X_{31},X_{32}}F_1(-\Delta_n T_3)(X_{31} - X_{32})$$

$$= 2EF_1(-\Delta_n T_3)E(X_{31} - X_{32})$$

$$= 0.$$

Also, since $\theta_n^{(2)}(\underline{\nu})$ is free from $\nu_2$, $\frac{\partial \theta_n^{(2)}(\underline{\nu})}{\partial \nu_2} = 0$.

Thus, $\theta_n^{(2)}(\underline{\nu})$ is differentiable at $\underline{\nu} = \underline{\beta}$ with the differential being zero.

In order to show that $\theta_n^{(2)}(\underline{\nu})$ is uniformly differentiable at $\underline{\beta}$, consider

$$\begin{aligned}
&\theta_n^{(2)}(\underline{\nu}) - \theta_n^{(2)}(\underline{\beta}) \\
&= -2E^{T_3,X_{31},X_{32}}\left[ET_1 I_{(-\infty,-\Delta_n T_3+B_1(\underline{\nu}))} - ET_1 I_{(-\infty,-\Delta_n T_3)}\right] \\
&\quad - 2\Delta_n E^{T_3,X_{31},X_{32}} T_3[F_1(-\Delta_n T_3 + B_1(\underline{\nu})) - F_1(-\Delta_n T_3)] \\
&\quad + 2E^{T_3,X_{31},X_{32}} B_1(\underline{\nu}) F_1(-\Delta_n T_3 + B_1(\underline{\nu})) \\
&= -2E^{T_3,X_{31},X_{32}} ET_1 I_{(-\Delta_n T_3,-\Delta_n T_3+B_1(\underline{\nu}))} \\
&\quad - 2\Delta_n E^{T_3,X_{31},X_{32}} T_3[F_1(-\Delta_n T_3 + B_1(\underline{\nu})) - F_1(-\Delta_n T_3)] \\
&\quad + 2(\nu_1 - \beta_1)E^{T_3,X_{31},X_{32}}(X_{31} - X_{32})F_1(-\Delta_n T_3 + B_1(\underline{\nu}))
\end{aligned}$$

Hence

(4.3.26)

$$\begin{aligned}
&|\theta_n^{(2)}(\underline{\nu}) - \theta_n^{(2)}(\underline{\beta})| \\
&\le 2E^{T_3,X_{31},X_{32}}\left|ET_1 I_{(-\Delta_n T_3,-\Delta_n T_3+B_1(\underline{\nu}))}\right| \\
&\quad + 2|\Delta_n| E^{T_3,X_{31},X_{32}}|T_3||F_1(-\Delta_n T_3 + B_1(\underline{\nu})) - F_1(-\Delta_n T_3)| \\
&\quad + 2\|\underline{\nu} - \underline{\beta}\| \left|E^{T_3,X_{31},X_{32}}(X_{31} - X_{32})F_1(-\Delta_n T_3 + B_1(\underline{\nu}))\right|
\end{aligned}$$

Note that since $E|T_1| < \infty$, by Dominated convergence theorem

$$E^{T_3,X_{31},X_{32}}\left|ET_1 I_{(-\Delta_n T_3,-\Delta_n T_3+B_1(\underline{\nu}))}\right| \to E^{X_{31},X_{32}}\left|ET_1 I_{(0,B_1(\underline{\nu}))}\right|.$$

But,

$$
\begin{aligned}
&E^{X_{31},X_{32}}\left|ET_1 I_{(0,B_1(\underline{\nu}))}\right| \\
&= E^{X_{31},X_{32}}\left|\int_0^{|B_1(\underline{\nu})|} x f_1(x)dx\right| \text{since } f_1(x) \text{ symmetric and hence } x f_1(x) \text{ is odd} \\
&\leq E^{X_{31},X_{32}}|B_1(\underline{\nu})||F_1(|B_1(\underline{\nu})|) - F_1(0)| \\
&= |\nu_1 - \beta_1| E^{X_{31},X_{32}}|X_{31} - X_{32}|\left|F_1(|B_1(\underline{\nu})|) - \frac{1}{2}\right| \\
&\leq \frac{1}{2}\|\underline{\nu} - \underline{\beta}\| E^{X_{31},X_{32}}|X_{31} - X_{32}||1 - 2F_1(|B_1(\underline{\nu})|)|
\end{aligned}
$$

For $\underline{\nu}$ close to $\underline{\beta}$, we can bound $E^{X_{31},X_{32}}|X_{31} - X_{32}||1 - 2F_1(|B_1(\underline{\nu})|)|$ by $\frac{\epsilon}{3}$, so that

$$E^{X_{31},X_{32}}\left|ET_1 I_{(0,B_1(\underline{\nu}))}\right| \leq \frac{\epsilon}{6}\|\underline{\nu} - \underline{\beta}\|.$$

So we can find an integer $N_1(\epsilon)$ such that for $n > N_1(\epsilon)$,

$$2E^{T_3,X_{31},X_{32}}\left|ET_1 I_{(-\Delta_n T_3, -\Delta_n T_3 + B_1(\underline{\nu}))}\right| \leq \frac{\epsilon}{3}\|\underline{\nu} - \underline{\beta}\|. \tag{4.3.27}$$

The second and third terms on the right hand side of (4.3.26) are same as the first and second terms on the right hand side of (4.3.20) respectively with $F_1(B_1)$ are replaced by $F_2(B_2)$. Thus there are integers $N_2(\epsilon)$ and $N_3(\epsilon)$ such that

$$2|\Delta_n| E^{T_3,X_{31},X_{32}}|T_3||F_1(-\Delta_n T_3 + B_1(\underline{\nu})) - F_1(-\Delta_n T_3)| \leq \frac{\epsilon}{3}\|\underline{\nu} - \underline{\beta}\| \tag{4.3.28}$$

for all $n > N_2(\epsilon)$ and

$$2\|\underline{\nu} - \underline{\beta}\|\left|E^{T_3,X_{31},X_{32}}(X_{31} - X_{32})F_1(-\Delta_n T_3 + B_1(\underline{\nu}))\right| \leq \frac{\epsilon}{3}\|\underline{\nu} - \underline{\beta}\| \tag{4.3.29}$$

for all $n > N_3(\epsilon)$ and $\underline{\nu}$ in some neighborhood $N_\epsilon(\underline{\beta})$ of $\underline{\beta}$.

Now for $\epsilon > 0$, let $N_\epsilon^{(2)} = max(N_1(\epsilon), N_2(\epsilon), N_3(\epsilon))$. Then from (4.3.27)-(4.3.29), $\exists$ a neighborhood $N_\epsilon^{(2)}(\underline{\beta})$ such that if $\underline{\nu} \in N_\epsilon^{(2)}(\underline{\beta})$ and $n > N_\epsilon^{(2)}$,

$$|\theta_n^{(2)}(\underline{\nu}) - \theta_n^{(2)}(\underline{\beta})| \leq \epsilon \|\underline{\nu} - \underline{\beta}\|.$$

This establishes the uniform differentiability of $\theta_n^{(2)}(\underline{\nu})$ in at $\underline{\nu} = \underline{\beta}$.

Thus all the conditions of Theorem 4.2.6 have been verified for both the U-statistics $U_n^{(1)}(\hat{\underline{\beta}})$ and $U_n^{(2)}(\hat{\underline{\beta}})$ and we have the following theorem.

THEOREM 4.3.30. *Under assumptions 4.3.1 - 4.3.3,*

$$n^{\frac{1}{2}}[U_n^{(i)}(\hat{\underline{\beta}}) - \theta_n^{(i)}(\underline{\beta})] \Rightarrow N(0, \sigma^{(i,i)}), \quad \text{as } n \to \infty, i = 1, 2$$

*where* $\theta_n^{(1)}(\underline{\beta}), \sigma^{(1,1)}, \theta_n^{(2)}(\underline{\beta})$ *and* $\sigma^{(2,2)}$ *are given by (4.3.5),(4.3.10),(4.3.12) and (4.3.16) respectively.*

Now we shall establish the asymptotic normality of $T_n(\hat{\underline{\beta}})$ under a sequence of alternatives converging to null.

THEOREM 4.3.31. *Under conditions 4.3.1 - 4.3.3,*

$$n^{\frac{1}{2}}[T_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] \Rightarrow N(0, \tau^2), \quad \text{as } n \to \infty,$$

*under a sequence of alternatives* $\{\Delta_n\}$ *converging to zero, where*

$$\theta_n(\underline{\beta}) = \frac{\theta_n^{(1)}(\underline{\beta})}{\theta_n^{(2)}(\underline{\beta})} = \tau^*,$$

*and*

$$\tau^2 = \frac{4}{3} \frac{\sigma_1^2}{(E|W_{11} - W_{12}|)^2}.$$

PROOF: Note that

(4.3.32)

$$
\begin{aligned}
&n^{\frac{1}{2}}[T_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] \\
&= n^{\frac{1}{2}}\left[\frac{U_n^{(1)}(\hat{\underline{\beta}})}{U_n^{(2)}(\hat{\underline{\beta}})} - \frac{\theta_n^{(1)}(\underline{\beta})}{\theta_n^{(2)}(\underline{\beta})}\right] \\
&= \frac{n^{\frac{1}{2}}[U_n^{(1)}(\hat{\underline{\beta}})\theta_n^{(2)}(\underline{\beta}) - U_n^{(2)}(\hat{\underline{\beta}})\theta_n^{(1)}(\underline{\beta})]}{U_n^{(2)}(\hat{\underline{\beta}})\theta_n^{(2)}(\underline{\beta})} \\
&= \frac{n^{\frac{1}{2}}[U_n^{(1)}(\hat{\underline{\beta}}) - \theta_n^{(1)}(\underline{\beta})]}{U_n^{(2)}(\hat{\underline{\beta}})} - \frac{\theta_n^{(1)}(\underline{\beta})}{U_n^{(2)}(\hat{\underline{\beta}})\theta_n^{(2)}(\underline{\beta})} \cdot n^{\frac{1}{2}}[U_n^{(2)}(\hat{\underline{\beta}}) - \theta_n^{(2)}(\underline{\beta})]
\end{aligned}
$$

We have by Theorem 4.3.30,

$$n^{\frac{1}{2}}[U_n^{(i)}(\hat{\underline{\beta}}) - \theta_n^{(i)}(\underline{\beta})] \Rightarrow N(0, \sigma^{(i,i)}), \quad \text{as } n \to \infty, \text{ under } \{\Delta_n\}, i = 1, 2.$$

From this it follows that

$$U_n^{(2)}(\hat{\underline{\beta}}) - \theta_n^{(2)}(\underline{\beta}) \to 0 \quad \text{in probability.}$$

Further, from (4.3.8) and (4.3.14), we have

$$\theta_n^{(1)}(\underline{\beta}) \to 0 \quad \text{and} \quad \theta_n^{(2)}(\underline{\beta}) \to E|T_1| \quad \text{as } n \to \infty.$$

This implies that $U_n^{(2)}(\hat{\underline{\beta}}) \to E|T_1|$ in probability.

Then by Slutsky's theorem and (4.3.32), we have

$$n^{\frac{1}{2}}[T_n(\hat{\underline{\beta}}) - \theta_n(\underline{\beta})] \Rightarrow \frac{1}{E|T_1|}N(0, \sigma^{(1,1)}), \quad \text{as } n \to \infty.$$

But the distribution on the right hand side above is nothing but $N(0, \tau^2)$ where

$$
\begin{aligned}
\tau^2 &= \frac{1}{(E|T_1|)^2}\sigma^{(1,1)} \\
&= \frac{4}{3}\frac{\sigma_1^2}{(E|T_1|)^2} \\
&= \frac{4}{3}\frac{\sigma_1^2}{(E|W_{11} - W_{12}|)^2}
\end{aligned}
$$

and hence the theorem.

## 4.4 Asymptotic Distribution of $R_n$

In this Section, we shall obtain the asymptotic distribution of $R_n$, the Pearson's Partial Correlation Coefficient between $X_1$ and $X_2$ holding $X_3$ fixed. Recall that under the model (2.1.1), $R_n$ can be written as

$$R_n(\underline{\hat{\beta}}) = \frac{\sum_{i=1}^{n}(Z_{1i}-\overline{Z}_1)(Z_{2i}-\overline{Z}_2)}{[\sum_{i=1}^{n}(Z_{1i}-\overline{Z}_1)^2\sum_{i=1}^{n}(Z_{2i}-\overline{Z}_2)^2]^{\frac{1}{2}}}$$

where $Z_{1i} = X_{1i} - \hat{\beta}_1 X_{3i}$ and $Z_{2i} = X_{2i} - \hat{\beta}_2 X_{3i}, i = 1,2,\ldots,n$ are the 'least squares' residuals from the model (2.1.1).

Noting that

$$\sum_{i=1}^{n}(Z_{1i}-\overline{Z}_1)(Z_{2i}-\overline{Z}_2) = \frac{1}{n}\sum_{i<j}(Z_{1i}-Z_{1j})(Z_{2i}-Z_{2j}),$$

$R_n(\underline{\hat{\beta}})$ can be written in a more convenient form as follows:

$$\begin{aligned} R_n(\underline{\hat{\beta}}) &= \frac{\sum_{i<j}(Z_{1i}-Z_{1j})(Z_{2i}-Z_{2j})}{[\sum_{i<j}(Z_{1i}-Z_{1j})^2\sum_{i<j}(Z_{2i}-Z_{2j})^2]^{\frac{1}{2}}} \\ &= \frac{R_n^{(12)}(\underline{\hat{\beta}})}{[R_n^{(11)}(\underline{\hat{\beta}})R_n^{(22)}(\underline{\hat{\beta}})]^{\frac{1}{2}}} \end{aligned}$$

where, $R_n^{(12)}, R_n^{(11)}$ and $R_n^{(22)}$ are U-statistics given by

$$\begin{aligned} &R_n^{(12)}(\underline{\nu}) \\ &= \frac{1}{\binom{n}{2}}\sum_{i<j}(X_{1i}-X_{1j}-\nu_1(X_{3i}-X_{3j}))(X_{2i}-X_{2j}-\nu_2(X_{3i}-X_{3j})) \\ &= \frac{1}{\binom{n}{2}}\sum_{i<j}(e_{1i}-e_{1j}-(\nu_1-\beta_1)(X_{3i}-X_{3j}))(e_{2i}-e_{2j}-(\nu_2-\beta_2)(X_{3i}-X_{3j})), \end{aligned}$$

$$R_n^{(11)}(\underline{\nu}) = \frac{1}{\binom{n}{2}}\sum_{i<j}(e_{1i}-e_{1j}-(\nu_1-\beta_1)(X_{3i}-X_{3j}))^2$$

and

$$R_n^{(22)}(\underline{\nu}) = \frac{1}{\binom{n}{2}} \sum_{i<j} (e_{2i} - e_{2j} - (\nu_2 - \beta_2)(X_{3i} - X_{3j}))^2.$$

We shall obtain the asymptotic normality of $R_n(\hat{\underline{\beta}})$ under the model proposed earlier, i.e., $e_1$ and $e_2$ are related by

$$e_{1i} = W_{1i} + \Delta W_{3i}$$

$$e_{2i} = W_{2i} + \Delta W_{3i}, \quad i = 1, 2, \dots, n.$$

Note that the hypothesis of independence of $e_1$ and $e_2$ is equivalent to the hypothesis $H_0 : \Delta = 0$. For efficiency considerations we assume that $\Delta$ is a function of n which tends to zero as n tends to infinity.

Also note that, $R_n$ is a function of the 3 U-statistics $R_n^{(12)}, R_n^{(11)}$, and $R_n^{(22)}$ were it not for the fact that they contain an estimated parameter of $\underline{\beta}$. We first show the asymptotic normality of each one of the U-statistics mentioned above with estimated parameters by applying Theorem 4.2.6. Then, we deduce the asymptotic normality of $R_n(\hat{\underline{\beta}})$ by Slutsky's theorem.

To establish the results of this section, we shall assume the following conditions.

CONDITION 4.4.1. *$W_{1i}, W_{2i}, W_{3i}, i = 1, 2, \dots, n$, are three independent random samples having the same distribution as the continuous random variables $W_1, W_2$ and $W_3$ with distribution functions $G_1(\cdot), G_2(\cdot)$, and $G_3(\cdot)$, respectively.*

CONDITION 4.4.2. *The variables $W_1, W_2, W_3$ have finite variance.*

CONDITION 4.4.3. *$X_3$ is a continuous random variable with finite fourth moment.*

Let $\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}, \cdots, \underset{\sim}{S}_{n:n}$ denote a random sample from a trivariate distribution with distribution function $H_n(\cdot,\cdot,\cdot)$ depending on n, where

$$\underset{\sim}{S}_{i:n} = \begin{pmatrix} X_{3i} \\ W_{1i} + \Delta_n W_{3i} \\ W_{2i} + \Delta_n W_{3i} \end{pmatrix}.$$

The symmetric kernels of degree 2 for the 3 U-statistics $R_n^{(12)}$, $R_n^{(11)}$ and $R_n^{(22)}$ defined earlier are respectively,

$$\begin{aligned} h^{(12)}(\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}; \underline{\nu}) =& (W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) - (\nu_1 - \beta_1)(X_{31} - X_{32}))\cdot \\ & (W_{21} - W_{22} + \Delta_n(W_{31} - W_{32}) - (\nu_2 - \beta_2)(X_{31} - X_{32})), \\ h^{(11)}(\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}; \underline{\nu}) =& (W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) - (\nu_1 - \beta_1)(X_{31} - X_{32}))^2, \\ h^{(22)}(\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}; \underline{\nu}) =& (W_{21} - W_{22} + \Delta_n(W_{31} - W_{32}) - (\nu_2 - \beta_2)(X_{31} - X_{32}))^2. \end{aligned}$$

Let

$$\begin{aligned} \xi_n^{(12)}(\underline{\nu}) &= E_{\underline{\beta}} h^{(12)}(\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}; \underline{\nu}) \\ \xi_n^{(11)}(\underline{\nu}) &= E_{\underline{\beta}} h^{(11)}(\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}; \underline{\nu}) \\ \xi_n^{(22)}(\underline{\nu}) &= E_{\underline{\beta}} h^{(22)}(\underset{\sim}{S}_{1:n}, \underset{\sim}{S}_{2:n}; \underline{\nu}). \end{aligned}$$

For $i, j = 1, 2, 3; i \leq j$, denote by

$$r^{(ij)} = \frac{1}{\binom{n}{2}} \sum_{k<l} (W_{ik} - W_{il})(W_{jk} - W_{jl})$$

and

$$\rho^{(ij)} = E(W_{i1} - W_{i2})(W_{j1} - W_{j2}).$$

Note that

(i)

$$R_n^{(ij)}(\underline{\beta}) = r^{(ij)} + \Delta_n r^{(i3)} + \Delta_n r^{(j3)} + \Delta_n^2 r^{(33)}$$

(ii)

$$\xi_n^{(ij)}(\underline{\beta}) = \rho^{(ij)} + \Delta_n \rho^{(i3)} + \Delta_n \rho^{(j3)} + \Delta_n^2 \rho_n^{(33)}, i,j = 1,2; i \leq j.$$

So, for $i,j = 1,2; i \leq j$,

(4.4.4)

$$\sqrt{n}(R_n^{(ij)}(\underline{\beta}) - \xi_n^{(ij)}(\underline{\beta})) = \sqrt{n}(r^{(ij)} - \rho^{(ij)}) + \Delta_n \sqrt{n}(r^{(i3)} - \rho^{(i3)})$$
$$+ \Delta_n \sqrt{n}(r^{(j3)} - \rho^{(j3)}) + \Delta_n^2 \sqrt{n}(r^{(33)} - \rho^{(33)})$$

Further, notice that, $r^{(ij)}, i,j = 1,2,3; i \leq j$, is a U-statistic with symmetric kernel of degree 2 and with mean $\rho^{(ij)}$. So by Theorem 3.6.9 of Randles and Wolfe (1979),

$$\sqrt{n}(r^{(ij)} - \rho^{(ij)}) \Rightarrow N(0, \sigma^{(i,j)}) \quad \text{as } n \to \infty. \tag{4.4.5}$$

where

$$\begin{aligned} \sigma^{(i,j)} &= 4Var[E((W_{i1} - W_{i2})(W_{j1} - W_{j2}) | W_{i1}, W_{j1})] \\ &= 4Var[(W_{i1} - EW_{i2})(W_{j1} - EW_{j2})] \\ &= 4Var[(W_{i1} - EW_{i1})(W_{j1} - EW_{j1})]. \end{aligned} \tag{4.4.6}$$

By condition 4.4.1 and 4.4.2, this variance is positive and finite. Since $\Delta_n$ converges to zero, by (4.4.4), $\sqrt{n}(R_n^{(ij)}(\underline{\beta}) - \xi_n^{(ij)}(\underline{\beta}))$ has the same distribution as $\sqrt{n}(r^{(ij)} - \rho^{(ij)})$. Thus we have the following theorem.

THEOREM 4.4.7. *Under conditions 4.4.1 and 4.4.2, for $i,j = 1,2; i \leq j$,*

$$\sqrt{n}(R_n^{(ij)}(\underline{\beta}) - \xi_n^{(ij)}(\underline{\beta})) \Rightarrow N(0, \sigma^{(i,j)}) \quad \text{as } n \to \infty$$

*where $\sigma^{(i,j)}$ is given by (4.4.6).*

PROOF: Immediate form (4.4.4) and (4.4.5).

We shall now apply Theorem 4.2.6 to obtain the asymptotic normality of $\sqrt{n}(R_n^{(ij)}(\hat{\underline{\beta}}) - \xi_n^{(ij)}(\underline{\beta})), i,j = 1,2; i \leq j$, under a sequence of alternatives approaching the null. To that effect, we first need to verify the conditions of Theorem 4.2.6.

I. Conditions 4.4.2 and 4.4.3 ensure that the OLS estimators of $\beta_1$ and $\beta_2$ satisfy condition 4.2.3 of Theorem 4.2.6.

II. Now we shall verify condition 4.2.4 for $h^{(ij)}$ and show that $\xi_n^{(ij)}(\underline{\nu})$ is uniformly (in n) differentiable at $\underline{\nu} = \underline{\beta}$ and the differential is zero, assuming conditions 4.4.1-4.4.3.

(i) For the kernel $h^{(12)}$ of $R_n^{(12)}$

$$
\begin{aligned}
&|h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})| \\
&= |(W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) - ({\nu_1}' - \beta_1)(X_{31} - X_{32}))\cdot \\
&\qquad (W_{21} - W_{22} + \Delta_n(W_{31} - W_{32}) - ({\nu_2}' - \beta_2)(X_{31} - X_{32})) \\
&- (W_{11} - W_{12} + \Delta_n(W_{31} - W_{32}) - (\nu_1 - \beta_1)(X_{31} - X_{32}))\cdot \\
&\qquad (W_{21} - W_{22} + \Delta_n(W_{31} - W_{32}) - (\nu_2 - \beta_2)(X_{31} - X_{32}))|
\end{aligned}
$$

With the usual notations $T_i = W_{i1} - W_{i2}, i = 1,2,3$, we have

$$
\begin{aligned}
&|h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})| \\
&= |(T_1 + \Delta_n T_3) - (T_2 + \Delta_n T_3 - ({\nu_2}' - \beta_2)(X_{31} - X_{32})) \\
&\quad - ({\nu_1}' - \beta_1)(X_{31} - X_{32})(T_2 + \Delta_n T_3 - ({\nu_2}' - \beta_2)(X_{31} - X_{32})) \\
&\quad - (T_1 + \Delta_n T_3) - (T_2 + \Delta_n T_3 - (\nu_2 - \beta_2)(X_{31} - X_{32}))\cdot \\
&\quad + (\nu_1 - \beta_1)(X_{31} - X_{32})(T_2 + \Delta_n T_3 - (\nu_2 - \beta_2)(X_{31} - X_{32}))|
\end{aligned}
$$

$$= | - (T_1 + \Delta_n T_3)(X_{31} - X_{32})(\nu_2{}' - \nu_2) - (T_2 + \Delta_n T_3)(X_{31} - X_{32})(\nu_1{}' - \nu_1)$$

$$+ (X_{31} - X_{32})^2[(\nu_1{}' - \beta_1)(\nu_2{}' - \beta_2) + (\nu_1 - \beta_1)(\nu_2 - \beta_1)]|$$

$$= | - (T_1 + \Delta_n T_3)(X_{31} - X_{32})(\nu_2{}' - \nu_2) - (T_2 + \Delta_n T_3)(X_{31} - X_{32})(\nu_1{}' - \nu_1)$$

$$+ (X_{31} - X_{32})^2[(\nu_1{}' - \beta_1)(\nu_2{}' - \nu_2) + (\nu_2 - \beta_2)(\nu_1{}' - \nu_1)]|$$

$$\leq |T_1 + \Delta_n T_3||X_{31} - X_{32}||\nu_2{}' - \nu_2| + |T_2 + \Delta_n T_3||X_{31} - X_{32}||\nu_1{}' - \nu_1|$$

$$+ |X_{31} - X_{32}|^2[|\nu_1{}' - \beta_1||\nu_2{}' - \nu_2| + |\nu_2 - \beta_2||\nu_1{}' - \nu_1|]|$$

Now for any arbitrary neighborhood of $\underline{\beta}$, say $K(\underline{\beta})$ with diameter of $K(\underline{\beta})$ say m, if $\nu \in K(\underline{\beta})$ and $D(\underline{\nu}, d)$ is a sphere centered at $\underline{\nu}$ with radius d satisfying $D(\underline{\nu}, d) \subset K(\underline{\beta})$, then

$$E\left[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|\right]$$
$$\leq E(|T_1 + \Delta_n T_3||X_{31} - X_{32}|) \cdot d + E(|T_2 + \Delta_n T_3||X_{31} - X_{32}|) \cdot d$$
$$+ E(X_{31} - X_{32})^2(m \cdot d + m \cdot d)$$

As we are interested in the behavior of the statistic for $\Delta_n$ in the neighborhood of 0, we can assume without loss of generality that $|\Delta_n| \leq \Delta$ for some $\Delta > 0$.

Then

$$E\left[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(12)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|\right] \leq K_1 d$$

for every n, where $K_1 = (E|T_1| + \Delta E|T_3|)E|X_{31} - X_{32}| + (E|T_2| + \Delta E|T_3|)E|X_{31} - X_{32}| + 2mE(X_{31} - X_{32})^2$.

Further,

$$\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(12)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}') - h^{(12)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|^2$$

$$\leq [(|T_1 + \Delta_n T_3|)|X_{31} - X_{32}| \cdot d + (|T_2 + \Delta_n T_3|)|X_{31} - X_{32}| \cdot d$$

$$+ |X_{31} - X_{32}|^2 \cdot 2md]^2$$

$$\leq [(|T_1| + \Delta|T_3|)|X_{31} - X_{32}| + (|T_2| + \Delta|T_3|)|X_{31} - X_{32}|$$

$$+ 2m|X_{31} - X_{32}|^2]^2 \cdot d^2$$

Hence

$$\lim_{d\to 0} E\left[\sup_{\underline{\nu}'\in D(\underline{\nu},d)} |h^{(12)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu}') - h^{(12)}(\underline{S}_{1:n},\underline{S}_{2:n};\underline{\nu})|^2\right] = 0,$$

uniformly in n.

This verifies condition 4.2.4 for $h^{(12)}$.

Now to show that $\xi^{(12)}(\underline{\nu})$ is uniformly differentiable at $\underline{\nu} = \underline{\beta}$,

$\xi_n^{(12)}(\underline{\nu})$

$$= E_{\underline{\beta}}[(T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32}))(T_2 + \Delta_n T_3 - (\nu_2 - \beta_2)(X_{31} - X_{32}))]$$

$$= E[(T_1 + \Delta_n T_3)(T_2 + \Delta_n T_3)] + (\nu_1 - \beta_1)(\nu_2 - \beta_2)E(X_{31} - X_{32})^2$$

$$= \Delta_n^2 E T_3^2 + (\nu_1 - \beta_1)(\nu_2 - \beta_2)E(X_{31} - X_{32})^2$$

Clearly,

$$\frac{\partial \xi_n^{(12)}(\underline{\nu})}{\partial \nu_1} = (\nu_2 - \beta_2)E(X_{31} - X_{32})^2$$

$$\frac{\partial \xi_n^{(12)}(\underline{\nu})}{\partial \nu_2} = (\nu_1 - \beta_1)E(X_{31} - X_{32})^2$$

so that

$$\left.\frac{\partial \xi_n^{(12)}(\underline{\nu})}{\partial \nu_i}\right|_{\underline{\nu}=\underline{\beta}} = 0 \quad \text{for } i = 1, 2.$$

Thus $\xi_n^{(12)}(\underline{\nu})$ is differentiable at $\underline{\nu} = \underline{\beta}$ with the differential being zero.

In order to show that $\xi_n^{(12)}(\underline{\nu})$ is uniformly (in n) differentiable at $\underline{\beta}$, consider

$$|\xi_n^{(12)}(\underline{\nu}) - \xi_n^{(12)}(\underline{\beta})| = |\Delta_n^2 ET_3^2 + (\nu_1 - \beta_1)(\nu_2 - \beta_2)E(X_{31} - X_{32})^2 - \Delta_n^2 ET_3^2|$$

$$= |(\nu_1 - \beta_1)(\nu_2 - \beta_2)E(X_{31} - X_{32})^2|$$

$$\leq \|\underline{\nu} - \underline{\beta}\|\|\underline{\nu} - \underline{\beta}\|E(X_{31} - X_{32})^2.$$

Now for every $\epsilon > 0, \exists \quad N_\epsilon(\underline{\beta}) = D(\underline{\beta}, \frac{\epsilon}{E(X_{31}-X_{32})^2})$ such that for any $\underline{\nu} \in N_\epsilon(\underline{\beta})$ and for all n,

$$|\xi_n^{(12)}(\underline{\nu}) - \xi_n^{(12)}(\underline{\beta})| \leq \epsilon\|\underline{\nu} - \underline{\beta}\|.$$

This verifies the uniform differentiability of $\xi_n^{(12)}(\underline{\nu})$ at $\underline{\nu} = \underline{\beta}$.

(ii) For the kernel $h^{(11)}$ of $R_n^{(11)}$,

$$|h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|$$

$$\leq 2|T_1 + \Delta_n T_3||X_{31} - X_{32}|\nu_1{}' - \nu_1|$$

$$+ |X_{31} - X_{32}|^2|\nu_1{}' - \nu|(|\nu_1{}' - \beta_1| + |\nu_1 - \beta_1|)$$

Now for any arbitrary neighborhood of $\underline{\beta}$ say $K(\underline{\beta})$ with diameter m, if $\underline{\nu} \in K(\underline{\beta})$ and $D(\underline{\nu}, d)$ is a sphere centered at $\underline{\nu}$ with radius d satisfying $D(\underline{\nu}, d) \subset K(\underline{\beta})$, then

$$\sup_{\underline{\nu}' \in D(\underline{\nu}, d)} |h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|$$
$$\leq 2|T_1 + \Delta_n T_3||X_{31} - X_{32}| \cdot d + 2m|X_{31} - X_{32}|^2 \cdot d$$

and assuming $|\Delta_n| \leq \Delta$, for some $\Delta$,

$$E\left[\sup_{\underline{\nu}' \in D(\underline{\nu}, d)} |h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|\right]$$
$$\leq [2E(|T_1| + \Delta E|T_3|)E|X_{31} - X_{32}| + 2mE|X_{31} - X_{32}|^2] \cdot d.$$

This verifies the first part of 4.2.4 with $K_1 = 2E(|T_1| + \Delta|T_3|)E|X_{31} - X_{32}| + 2mE|X_{31} - X_{32}|^2$. Further more,

$$\begin{aligned}&\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|^2 \\ &\leq [2|T_1 + \Delta_n T_3||X_{31} - X_{32}| + 2m|X_{31} - X_{32}|^2]^2 \cdot d^2\end{aligned}$$

and thus

$$\lim_{d \to 0} E\left[\sup_{\underline{\nu}' \in D(\underline{\nu},d)} |h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu}') - h^{(11)}(\underline{S}_{1:n}, \underline{S}_{2:n}; \underline{\nu})|^2\right] = 0,$$

uniformly in n. This completes the verification of the second part of 4.2.4 for $h^{(11)}$.

We now show that $\xi_n^{(11)}(\underline{\nu})$ is uniformly (in n) differentiable at $\underline{\nu} = \underline{\beta}$.

$$\begin{aligned}\xi_n^{(11)}(\underline{\nu}) &= E_{\underline{\beta}}[T_1 + \Delta_n T_3 - (\nu_1 - \beta_1)(X_{31} - X_{32})]^2 \\ &= E(T_1 + \Delta_n T_3)^2 + (\nu_1 - \beta_1)^2 E(X_{31} - X_{32})^2 \\ &= ET_1^2 + \Delta_n^2 ET_3^2 + (\nu_1 - \beta_1)^2 E(X_{31} - X_{32})^2\end{aligned}$$

Clearly,

$$\frac{\partial \xi_n^{(11)}(\underline{\nu})}{\partial \nu_1} = 2(\nu_1 - \beta_1)E(X_{31} - X_{32})^2$$

$$\frac{\partial \xi_n^{(11)}(\underline{\nu})}{\partial \nu_2} = 0$$

so that

$$\left.\frac{\partial \xi_n^{(11)}(\underline{\nu})}{\partial \nu_i}\right|_{\underline{\nu}=\underline{\beta}} = 0 \quad \text{for } i = 1, 2.$$

Thus $\xi_n^{(11)}(\underline{\nu})$ is differentiable at $\underline{\nu} = \underline{\beta}$ with the differential being zero.

In order to show that $\xi_n^{(11)}$ is uniformly (in n) differentiable at $\underline{\beta}$, consider

$$\begin{aligned}|\xi_n^{(11)}(\underline{\nu}) - \xi_n^{(11)}(\underline{\beta})| &= (\nu_1 - \beta_1)^2 E(X_{31} - X_{32})^2 \\ &\leq \|\underline{\nu} - \underline{\beta}\|^2 E(X_{31} - X_{32})^2.\end{aligned}$$

Thus for every $\epsilon > 0$, and $\underline{\nu} \in D(\underline{\beta}, \frac{\epsilon}{E(X_{31}-X_{32})^2})$,

$$|\xi_n^{(11)}(\underline{\nu}) - \xi_n^{(11)}(\underline{\beta})| \leq \epsilon \|\underline{\nu} - \underline{\beta}\|,$$

for all n, and thus verifying the uniform differentiability of $\xi_n^{(11)}$ at $\underline{\beta}$.

(iii) Verification of conditions for $h^{(22)}$ and $\xi_n^{(22)}$ are exactly same as in (ii).

From Theorem 4.4.7 and I and II above we have the following theorem.

THEOREM 4.4.8. *Under conditions 4.4.1 - 4.4.3, for $i, j = 1, 2; i \leq j$,*

$$\sqrt{n}(R_n^{(ij)}(\hat{\underline{\beta}}) - \xi_n^{(ij)}(\underline{\beta})) \Rightarrow N(0, \sigma^{(i,j)}) \quad \text{as } n \to \infty$$

*where $\sigma^{(i,j)}$ is given by (4.4.6).*

REMARK 4.4.9. *From Theorem 4.4.8, it follows that for $i, j = 1, 2; i \leq j$,*

$$R_n^{(ij)}(\hat{\underline{\beta}}) - \xi_n^{(ij)}(\underline{\beta}) \to 0 \quad \textit{in probability, as } n \to \infty.$$

But

$$\xi_n^{(12)}(\underline{\beta}) = \Delta_n^2 ET_3^2 = 2\Delta_n^2 \sigma_3^2 \to 0,$$

$$\xi_n^{(11)}(\underline{\beta}) = ET_1^2 + \Delta_n^2 ET_3^2 = 2\sigma_1^2 + 2\Delta_n^2 \sigma_3^2 \to 2\sigma_1^2$$

$$\xi_n^{(22)}(\underline{\beta}) = ET_2^2 + \Delta_n^2 ET_3^2 = 2\sigma_2^2 + 2\Delta_n^2 \sigma_3^2 \to 2\sigma_2^2$$

Thus we have, as n tends to infinity,

$$R_n^{(12)}(\hat{\underline{\beta}}) \to 0 \quad \text{in probability,}$$

$$R_n^{(11)}(\hat{\underline{\beta}}) \to 2\sigma_1^2 > 0 \quad \text{in probability,}$$

and

$$R_n^{(22)}(\hat{\underline{\beta}}) \to 2\sigma_2^2 > 0 \quad \text{in probability.}$$

Let

$$\rho_n(\underline{\beta}) = \frac{\xi_n^{(12)}(\underline{\beta})}{\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}} = \frac{\Delta_n^2\sigma_3^2}{\sqrt{(\sigma_1^2+\Delta_n^2\sigma_3^2)(\sigma_2^2+\Delta_n^2\sigma_3^2)}}. \tag{4.4.10}$$

Now let's prove the main theorem of this section which gives the asymptotic distribution of $R_n(\hat{\underline{\beta}})$ under a sequence of alternatives.

THEOREM 4.4.11. *Under conditions 4.4.1 - 4.4.3,*

$$\sqrt{n}(R_n(\hat{\underline{\beta}}) - \rho_n(\underline{\beta})) \Rightarrow N(0,1), \qquad \text{as } n \to \infty,$$

*where $\rho_n(\underline{\beta})$ is given by (4.4.10).*

PROOF: Write

(4.4.12)

$$\begin{aligned}
&\sqrt{n}(R_n(\hat{\underline{\beta}}) - \rho_n(\underline{\beta}))\\
&= \sqrt{n}\left[\frac{R_n^{(12)}(\hat{\underline{\beta}})}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}} - \frac{\xi_n^{(12)}(\underline{\beta})}{\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}}\right]\\
&= \frac{\sqrt{n}\left[R_n^{(12)}(\hat{\underline{\beta}})\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})} - \xi_n^{(12)}(\underline{\beta})\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}\right]}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}}\\
&= \frac{\sqrt{n}\left[R_n^{(12)}(\hat{\underline{\beta}})\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})} - \xi_n^{(12)}(\underline{\beta})\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}\right]}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}}\\
&+ \frac{\sqrt{n}\left[\xi_n^{(12)}(\underline{\beta})\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})} - \xi_n^{(12)}(\underline{\beta})\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}\right]}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}}
\end{aligned}$$

$$= \frac{\sqrt{n}(R_n^{(12)}(\hat{\underline{\beta}}) - \xi_n^{(12)}(\underline{\beta}))}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}}$$
$$+ \frac{\sqrt{n}\xi_n^{(12)}(\underline{\beta})(\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta}) - R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}}))}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}\left(\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})} + \sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}\right)}$$
$$= \frac{\sqrt{n}(R_n^{(12)}(\hat{\underline{\beta}}) - \xi_n^{(12)}(\underline{\beta}))}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}}$$
$$- \frac{\xi_n^{(12)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})\sqrt{n}(R_n^{(11)}(\hat{\underline{\beta}}) - \xi_n^{(11)}(\underline{\beta}))}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}\left(\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})} + \sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}\right)}$$
$$- \frac{\xi_n^{(12)}(\underline{\beta})R_n^{(11)}(\hat{\underline{\beta}})\sqrt{n}(R_n^{(22)}(\hat{\underline{\beta}}) - \xi_n^{(22)}(\underline{\beta}))}{\sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})}\left(\sqrt{\xi_n^{(11)}(\underline{\beta})\xi_n^{(22)}(\underline{\beta})} + \sqrt{R_n^{(11)}(\hat{\underline{\beta}})R_n^{(22)}(\hat{\underline{\beta}})}\right)}$$

From (4.4.2), Theorem 4.4.8 and Slutsky's theorem,

$$\text{(4.4.13)} \qquad \sqrt{n}(R_n(\hat{\underline{\beta}}) - \rho_n(\underline{\beta})) \Rightarrow \frac{1}{2\sigma_1\sigma_2}N(0, \sigma^{(1,2)}), \quad \text{as } n \to \infty.$$

From (4.4.6),

$$\begin{aligned} \sigma^{(1,2)} &= 4Var[(W_{11} - EW_{11})(W_{21} - EW_{21})] \\ &= 4E(W_{11} - EW_{11})^2 E(W_{21} - EW_{21})^2 \\ &= 4\sigma_1^2\sigma_2^2. \end{aligned}$$

Thus the random variable on the right hand side of (4.4.13) has a N(0,1) distribution and this completes the proof of the theorem.

REMARK 4.4.14. *If $\Delta_n = 0$ for all $n$, then*

$$\sqrt{n}(R_n(\hat{\underline{\beta}}) - \rho_n(\underline{\beta})) \Rightarrow N(0,1), \quad \text{as } n \to \infty,$$

*where $\rho_n(\underline{\beta}) = 0$, which gives the asymptotic normality of $R_n(\hat{\underline{\beta}})$ under the null hypothesis.*

## 4.5 Asymptotic Relative Efficiency of $T_n$ with respect to $R_n$

In this section we shall obtain the asymptotic relative efficiency of the test of the hypothesis $H_0$ based on $T_n$ with respect to the test procedure based on $R_n$. In this regard, we apply an extension of Pitman's theorem by Noether (1955) which is stated below.

THEOREM 4.5.1 (NOETHER). *Assume that we want to test the null hypothesis $H_0 : \theta = \theta_0$ against alternatives $H : \theta > \theta_0$. Assume also that a particular alternative $\theta = \theta_n$ changes with the sample size n in such a way that*

$$\lim_{n\to\infty} \theta_n = \theta_0.$$

*Let the test be based on the statistic $T_n = T(x_1, x_2, \ldots, x_n)$ and let $\psi_n(\theta)$ and $\sigma_n(\theta)$ be functions of $\theta$ satisfying,*

*A.*

$$\psi_n{}^{(1)}(\theta_0) = \cdots = \psi_n{}^{(m-1)}(\theta_0) = 0, \quad \psi_n{}^{(m)}(\theta_0) > 0,$$

*B.*

$$\lim_{n\to\infty} \frac{n^{-m\delta}\psi_n^{(m)}(\theta_0)}{\sigma_n(\theta_0)} = c > 0,$$

*for some $\delta > 0$.*

*The indicated derivatives are assumed to exist. Consider $H_1 : \theta_n = \theta_0 + \frac{k}{n^\delta}$ where k is an arbitrary positive constant.*

*In addition to A and B, assume*

*C.*

$$\lim_{n\to\infty} \frac{\psi_n^{(m)}(\theta_n)}{\psi_n^{(m)}(\theta_0)} = 1, \quad \lim_{n\to\infty} \frac{\sigma_n^{(m)}(\theta_n)}{\sigma_n^{(m)}(\theta_0)} = 1,$$

*D. The distribution of* $\frac{T_n - \psi_n(\theta)}{\sigma_n(\theta)}$ *tends to the normal distribution with mean 0 and variance 1, both under* $H_1$ *and under the null hypothesis* $\theta_n = \theta_0$.

*Then the quantity* $R_n^{\frac{1}{m\delta}}(\theta_0)$, *where* $R_n(\theta) = \frac{\psi_n^{(m)}(\theta)}{\sigma_n(\theta)}$, *is called the efficacy of* $T_n$ *in testing the hypothesis* $H_0 : \theta = \theta_0$.

*Further if* $T_{1n}$ *and* $T_{2n}$ *are two statistics for testing* $H_0$ *satisfying A, B, C, and D, with* $\delta_1 = \delta_2 = \delta$ *and* $m_1 = m_2 = m$, *then the asymptotic relative efficiency of the two tests denoted by* $ARE(T_{1n}, T_{2n})$ *is given by,*

$$ARE(T_{1n}, T_{2n}) = \lim_{n\to\infty} \frac{R_{1n}^{\frac{1}{m\delta}}(\theta_0)}{R_{2n}^{\frac{1}{m\delta}}(\theta_0)}. \tag{4.5.2}$$

We want to test the hypothesis, $H_0 : \Delta = 0$ versus $H : \Delta \neq 0$, where $\Delta$ is such that

$$e_{1i} = W_{1i} + \Delta W_{3i}$$

$$e_{2i} = W_{2i} + \Delta W_{3i}, \quad i = 1, 2, \ldots, n.$$

In addition to the assumptions 4.3.1 - 4.3.3 of section 4.3 and assumptions 4.4.1 - 4.4.3 of section 4.4, we need the following assumption.

CONDITION 4.5.3. *The density function* $f_k(\cdot)$ *of* $T_k = W_{k1} - W_{k2}, k = 1, 2, 3$ *have continuous, bounded derivatives and* $f_k(0) > 0$.

Next, we shall verify the conditions of Theorem 4.5.1 for each of the

statistics $T_n$ and $R_n$. Here we shall let $\{\Delta_n\}$ denote a sequence of parameters converging to the null. i.e., $\lim_{n\to\infty} \Delta_n = 0$.

Verification of conditions of Theorem 4.5.1 for the statistic $T_n$ :

Let $\theta = \Delta, \theta_0 = 0, T_n = T_n(\hat{\underline{\beta}})$ and

$$\psi_n(\Delta) = \frac{\theta^{(1)}(\Delta)}{\theta^{(2)}(\Delta)}$$

where $\theta^{(1)}$ and $\theta^{(2)}$ are given by (4.3.5) and (4.3.12) with $\Delta_n$ replaced by $\Delta$, and

$$\sigma_n{}^2(\Delta) = \frac{4}{3n}\frac{\sigma_1{}^2}{(E|W_{11} - W_{12}|)^2}.$$

A. Finding derivatives of the function $\psi_n$ involves taking derivatives of the integrals of the form$\int k(x,\Delta)dF_3(x)$. We shall once again use Result 3.4.18 in order to interchange the order of differentiation and integration.

(a) Consider $E^{T_3}T_3F_2(-\Delta T_3) = \int xF_2(-\Delta x)dF_3(x)$ and let $k(x,\Delta) = xF_2(-\Delta x)$ for $\Delta \in (-1,1)$. It is easy to see that conditions (i) and (iii) of Result 3.4.18 are satisfied by the function $k(x,\Delta)$ and

$$\begin{aligned}\left|\frac{\partial k(x,\Delta)}{\partial\Delta}\right| &= |-x^2 f_2(-\Delta x)| \\ &\leq B_2 x^2,\end{aligned}$$

where $B_2$ is the upper bound on $f_2$.

Further,

$$B_2\int x^2 dF_3(x) = B_2 E T_3{}^2 < \infty.$$

$$\int x^2 f_2(-\Delta x)dF_3(x).$$

Hence, by Result 3.4.18,

$$\frac{d}{d\Delta}\int xF_2(-\Delta x)dF_3(x) = -\int x^2 f_2(-\Delta x)dF_3(x).$$

$$\int xF_2(-\Delta x)dF_3(x)$$

Thus

$$(4.5.4) \qquad \frac{d}{d\Delta}\theta^{(1)}(\Delta) = 2\Delta \int x^2 f_2(-\Delta x)dF_3(x) - 2\int xF_2(-\Delta x)dF_3(x).$$

(b) Consider now $E^{T_3}(E^{T_1}T_1 I_{(-\infty,\Delta T_3)}) = \int(\int_{-\infty}^{-\Delta x} y\, dF_1(y))dF_3(x)$ and let $k(x,\Delta) = \int_{-\infty}^{-\Delta x} y\, dF_1(y)$ for $\Delta \in (-1,1)$. Again it is easy to see that conditions (i) and (iii) of Result 3.4.18 are readily satisfied by $k(x,\Delta)$ and

$$\begin{aligned} \left|\frac{\partial k(x,\Delta)}{\partial \Delta}\right| &= |\Delta x^2 f_1(-\Delta x)| \\ &\leq B_1 x^2, \end{aligned}$$

where $B_1$ is the upper bound on $f_1$. Further,

$$B_1 \int x^2 dF_3(x) = B_1 E{T_3}^2 < \infty.$$

Hence by Result 3.4.18,

$$\frac{d}{d\Delta}\int \left(\int_{-\infty}^{-\Delta x} ydF_1(y)\right) dF_3(x) = \Delta \int x^2 f_1(-\Delta x)dF_3(x).$$

Thus,

(4.5.5)

$$\begin{aligned} \frac{d}{d\Delta}\theta^{(2)}(\Delta) &= -2\Delta \int x^2 f_1(-\Delta x)dF_3(x) \\ &\qquad + 2\Delta \int x^2 f_1(-\Delta x)dF_3(x) - 2\int xF_1(-\Delta x)dF_3(x) \\ &= -2\int xF_1(-\Delta x)dF_3(x) \end{aligned}$$

Now,

$$(4.5.6) \qquad \psi_n'(\Delta) = \frac{\theta^{(2)}(\Delta)\frac{d}{d\Delta}\theta^{(1)}(\Delta) - \theta^{(1)}(\Delta)\frac{d}{d\Delta}\theta^{(2)}(\Delta)}{(\theta^{(2)}(\Delta))^2}$$

and we note that $\theta^{(1)}(0) = 0, \theta^{(2)}(0) = E|T_1|; \frac{d}{d\Delta}\theta^{(1)}(\Delta)|_{\Delta=0} = -2F_2(0)ET_3 = 0$ and $\frac{d}{d\Delta}\theta^{(2)}(\Delta)|_{\Delta=0} = -2F_1(0)ET_3 = 0$.

Substituting in (4.5.6), we have,

$$\psi_n'(0) = 0. \tag{4.5.7}$$

Now finding the second derivative of $\psi_n$ again involves differentiating an integral of the form $\int k(x,\Delta)dF_3(x)$.

(c) Let $k(x,\Delta) = x^2 f_2(-\Delta x)$. Clearly $k(x,\Delta)$ is absolutely continuous function of $\Delta$ for each x with

$$\frac{\partial k(x,\Delta)}{\partial \Delta} = -x^3 f_2'(-\Delta x).$$

and

$$\left|\frac{\partial k(x,\Delta)}{\partial \Delta}\right| \leq C_2|x|^3$$

where $C_2$ is the upper bound on $f_2'$. Further,

$$C_2 \int |x|^3 dF_3(x) = C_2 E|T_3|^3 < \infty.$$

Hence, by Result 3.4.18,

$$\frac{d}{d\Delta}\int x^2 f_2(-\Delta x)dF_3(x) = -\int x^3 f_2'(-\Delta x)dF_3(x).$$

Thus from (4.5.4),

(4.5.8)

$$\begin{aligned}\frac{d^2}{d\Delta^2}\theta^{(1)}(\Delta) &= -2\Delta \int x^3 f_2'(-\Delta x)dF_3(x) \\ &\quad + 2\int x^2 f_2(-\Delta x)dF_3(x) + 2\int x^2 f_2(-\Delta x)dF_3(x) \\ &= -2\Delta \int x^3 f_2'(-\Delta x)dF_3(x) + 4\int x^2 f_2(-\Delta x)dF_3(x)\end{aligned}$$

Similarly, since $f_1'$ is bounded, we can pass the differentiation inside the integral to obtain

$$\frac{d}{d\Delta}\int x^2 f_1(-\Delta x)dF_3(x) = -\int x^3 f_1'(-\Delta x)dF_3(x)$$

so that from (4.5.5),

(4.5.9) $$\frac{d^2}{d\Delta^2}\theta^{(2)}(\Delta) = 2\int x^2 f_1(-\Delta x)dF_3(x)$$

Now,

(4.5.10)

$$\psi_n''(\Delta)$$
$$= \frac{1}{[\theta^{(2)}(\Delta)]^4}\Big[\Big\{\theta^{(2)}(\Delta)\Big\}^2\Big\{(\theta^{(2)}(\Delta)\frac{d^2}{d\Delta^2}\theta^{(1)}(\Delta) - \theta^{(1)}(\Delta)\frac{d^2}{d\Delta^2}\theta^{(2)}(\Delta)\Big\}$$
$$-\Big\{\theta^{(2)}(\Delta)\frac{d}{d\Delta}\theta^{(1)}(\Delta) - \theta^{(1)}(\Delta)\frac{d}{d\Delta}\theta^{(2)}(\Delta)\Big\}\cdot 2\theta^{(2)}(\Delta)\frac{d}{d\Delta}\theta^{(2)}(\Delta)\Big]$$

Recall that $\theta^{(1)}(0) = 0, \theta^{(1)}(0) = E|T_1|, \frac{d}{d\Delta}\theta^{(1)}(\Delta)|_{\Delta=0} = 0, \frac{d}{d\Delta}\theta^{(2)}(\Delta)|_{\Delta=0} = 0$ and from (4.5.8) and (4.5.9), we have

$$\frac{d^2}{d\Delta^2}\theta^{(1)}(\Delta)\Big|_{\Delta=0} = 4f_2(0)ET_3^2 = 8f_2(0)\sigma_3^2.$$

and

$$\frac{d^2}{d\Delta^2}\theta^{(2)}(\Delta)\Big|_{\Delta=0} = 2f_1(0)ET_3^2 = 4f_1(0)\sigma_3^2.$$

Substituting for these quantities in (4.5.10), we get

(4.5.11) $$\psi_n''(0) = \frac{8f_2(0)\sigma_3^2}{E|T_1|} > 0$$

From (4.5.7) and (4.5.11), condition A is satisfied with m=2.

B. We have

$$\lim_{n\to\infty} n^{-\frac{1}{2}} \frac{\psi_n{}''(0)}{\sigma_n(0)} = 4\sqrt{3} f_2(0) \frac{\sigma_3^2}{\sigma_1} > 0.$$

This satisfied condition B with $\delta = \frac{1}{4}$.

Let $\Delta_n = \frac{k}{n^{1/4}}$ for some constant k.

C. Because of condition 4.5.3, $\theta^{(1)}, \theta^{(2)}, \frac{d}{d\Delta}\theta^{(1)}, \frac{d}{d\Delta}\theta^{(2)}, \frac{d^2}{d\Delta^2}\theta^{(1)}$and $\frac{d^2}{d\Delta^2}\theta^{(2)}$ are all continuous functions of $\Delta$ at 0 and so is $\psi_n''(\Delta)$, which implies

$$\lim_{n\to\infty} \frac{\psi_n{}''(\Delta_n)}{\psi_n{}''(0)} = 1.$$

Since $\sigma_n$ is free from $\Delta, \lim_{n\to\infty} \frac{\sigma_n(\Delta_n)}{\sigma_n(0)} = 1$ is immediate. Thus, condition C is verified.

D. In sections 3.4 and 4.3 respectively, we have shown that$\frac{T_n - \psi_n(\Delta)}{\sigma_n(\Delta)}$ converges in distribution to N(0,1) as $n \to \infty$, under the null hypothesis and under a sequence of alternatives converging to null, thereby satisfying condition D.

Thus the efficacy of the test of $H_0$ based on $T_n$ is given by,

(4.5.12)
$$\begin{aligned} R_n{}^{\frac{1}{n^\delta}}(0) &= \left[\frac{\psi_n{}''(0)}{\sigma_n(0)}\right]^2 \\ &= \left[4\sqrt{3} f_2(0) \frac{\sigma_3^2}{\sigma_1}\right]^2 \\ &= 48 n f_2^2(0) \frac{\sigma_3{}^4}{\sigma_1{}^2}. \end{aligned}$$

Verification of conditions of Theorem 4.5.1 for the statistics $R_n$ :

Let $\theta = \Delta, \theta_0 = 0, T_n = R_n(\hat{\underline{\beta}}), \psi_n(\Delta) = \rho_n(\underline{\beta})$ and $\sigma_n^2(\Delta) = \frac{1}{n}$

A. Note that

$$\psi_n(\Delta) = \frac{\Delta^2 {\sigma_3}^2}{[({\sigma_1}^2 + \Delta^2 {\sigma_3}^2)({\sigma_2}^2 + \Delta^2 \sigma_3^2)]^{\frac{1}{2}}}$$
$$= c_{12} \frac{\Delta^2}{[(1 + c_1\Delta^2)(1 + c_2\Delta^2)]^{\frac{1}{2}}}$$

where $c_1 = \frac{{\sigma_3}^2}{{\sigma_1}^2}, c_2 = \frac{{\sigma_3}^2}{{\sigma_2}^2}$ and $c_{12} = \frac{{\sigma_3}^2}{\sigma_1 \sigma_2}$.

Note that $\psi_n(0) = 0$ and $\psi_n(\Delta)$ is differentiable with respect to $\Delta$ with

$$\begin{aligned}
&\psi_n'(\Delta) \\
&= c_{12} \left\{ -\frac{1}{2}\Delta^2[(1 + c_1\Delta^2)(1 + c_2\Delta^2)]^{-\frac{3}{2}}[(1 + c_1\Delta^2)(2c_2\Delta) \right. \\
&\qquad \left. +(1 + c_2\Delta^2)(2c_1\Delta)] + 2\Delta[(1 + c_1\Delta^2)(1 + c_2\Delta^2)]^{-\frac{1}{2}} \right\} \\
&= c_{12} \left\{ -\Delta^3[(1 + c_1\Delta^2)(1 + c_2\Delta^2)]^{-\frac{3}{2}}[c_2(1 + c_1\Delta^2) + c_1(1 + c_2\Delta^2)] \right. \\
&\qquad \left. +2\Delta[(1 + c_1\Delta^2)(1 + c_2\Delta^2)]^{-\frac{1}{2}} \right\} \\
&= -c_{12}c_2\Delta^3(1 + c_1\Delta^2)^{-\frac{1}{2}}(1 + c_2\Delta^2)^{-\frac{3}{2}} \\
&\qquad - c_{12}c_1\Delta^3(1 + c_1\Delta^2)^{-\frac{3}{2}}(1 + c_2\Delta^2)^{-\frac{1}{2}} \\
&\qquad + 2c_{12}\Delta(1 + c_1\Delta^2)^{-\frac{1}{2}}(1 + c_2\Delta^2)]^{-\frac{1}{2}}
\end{aligned}$$

so that

$${\psi_n}'(0) = 0.$$

Further, $\psi_n{}'(\Delta)$ is differentiable with respect to $\Delta$ giving,

$$
\begin{aligned}
&\psi_n{}''(\Delta)\\
&= -c_{12}c_2\Delta^3(1+c_1\Delta^2)^{-\frac{1}{2}}(-\frac{3}{2})(1+c_2\Delta^2)]^{-\frac{5}{2}}(2c_2\Delta)\\
&\quad - c_{12}c_2(1+c_2\Delta^2)^{-\frac{3}{2}}[\Delta^3(-\frac{1}{2})(1+c_1\Delta^2)^{-\frac{3}{2}}(2c_1\Delta)+(1+c_1\Delta^2)^{-\frac{1}{2}}3\Delta^2]\\
&\quad - c_{12}c_1\Delta^3(1+c_2\Delta^2)^{-\frac{1}{2}}(-\frac{3}{2})(1+c_1\Delta^2)]^{-\frac{5}{2}}(2c_1\Delta)\\
&\quad - c_{12}c_1(1+c_1\Delta^2)^{-\frac{3}{2}}[\Delta^3(-\frac{1}{2})(1+c_2\Delta^2)^{-\frac{3}{2}}(2c_2\Delta)+(1+c_2\Delta^2)^{-\frac{1}{2}}3\Delta^2]\\
&\quad + 2c_{12}\Delta(1+c_1\Delta^2)^{-\frac{1}{2}}(-\frac{1}{2})(1+c_2\Delta^2)]^{-\frac{3}{2}}(2c_2\Delta)\\
&\quad + 2c_{12}(1+c_2\Delta^2)^{-\frac{1}{2}}[\Delta(-\frac{1}{2})(1+c_1\Delta^2)^{-\frac{3}{2}}(2c_1\Delta)+(1+c_1\Delta^2)^{-\frac{1}{2}}]
\end{aligned}
$$

Substituting for $\Delta = 0$ in the above equation,

$$\psi_n{}''(0) = 2c_{12} = 2\frac{{\sigma_3}^2}{\sigma_1\sigma_2} > 0.$$

This satisfies condition A with m=2.

B.

$$
\begin{aligned}
\lim_{n\to\infty} n^{-\frac{1}{2}}\frac{\psi_n{}''(0)}{\sigma_n(0)} &= \lim_{n\to\infty} n^{-\frac{1}{2}} 2\frac{{\sigma_3}^2}{\sigma_1\sigma_2} n^{\frac{1}{2}}\\
&= 2\frac{{\sigma_3}^2}{\sigma_1\sigma_2} = c > 0.
\end{aligned}
$$

So condition B is satisfied with $\delta = \frac{1}{4}$.

C. Condition C is obvious as both $\psi_n{}''(\Delta)$ and $\sigma_n(\Delta)$ both are free from $\Delta$.

D. In section 4.4, we have shown that $\frac{T_n - \psi_n(\Delta)}{\sigma_n(\Delta)}$ converges in distribution to $N(0,1)$ as $n \to \infty$, both under the null hypothesis and under a sequence of alternatives converging to null, thereby satisfying condition D.

Thus the efficacy of the test based $R_n$ is,

$$R_n^{\frac{1}{m\delta}}(0) = \left[\frac{\psi_n''(0)}{\sigma_n(0)}\right]^2 = 4n\frac{{\sigma_3}^4}{{\sigma_1}^2{\sigma_2}^2}. \tag{4.5.13}$$

THEOREM 4.5.14. *Under assumptions 4.3.1 - 4.3.3, 4.4.1 - 4.4.3, and assumption 4.5.3, the asymptotic relative efficiency of $T_n$ with respect to $R_n$ for testing $H_0$ is,*

$$ARE(T_n, R_n) = 12{\sigma_2}^2 f_2^2(0). \tag{4.5.15}$$

PROOF: Immediate from (4.5.2), (4.5.12) and (4.5.13).

Note: In (4.5.15), $f_2$ represents the density of the difference of two independent copies of $W_2$, which is assumed to have a density $g_2$.

By convolution, the density $f_2$ of $W_{21} - W_{22}$ is,

$$f_2(x) = \int g_2(y) g_2(y - x)\, dy$$

and hence

$$f_2(0) = \int {g_2}^2(y)\, dy.$$

Thus the efficiency formula in (4.5.15) can be written in terms of the variables $W_i, i = 1, 2, 3$, as,

$$ARE(T_n, R_n) = 12{\sigma_2}^2 \left(\int {g_2}^2(y)\, dy\right)^2. \tag{4.5.16}$$

Using the relation (4.5.16), $ARE(T_n, R_n)$ is calculated in the case $W_1, W_2$, and $W_3$ have the same distribution for several well known distributions. The results are given in the following tables.

Table 4.5.17

| Common Distribution of $W_1$, $W_2$ and $W_3$ | $\sigma_2^2$ | $\int g_2^2(y)dy$ | $ARE(T_n, R_n) = 12\sigma_2^2(\int g_2^2(y)dy)^2$ |
|---|---|---|---|
| Standard normal | 1 | $\frac{1}{2\sqrt{\pi}}$ | 0.955 |
| Logistic | $\frac{\pi^2}{3}$ | $\frac{1}{6}$ | 1.100 |
| Double Exponential | 2 | $\frac{1}{4}$ | 1.500 |
| Triangular | $\frac{1}{6}$ | $\frac{2}{3}$ | 0.890 |
| Uniform | $\frac{1}{12}$ | 1 | 1.000 |
| Exponential | 1 | $\frac{1}{2}$ | 3.000 |
| Lognormal | $e(e-1)$ | $\frac{\sqrt[4]{e}}{2\sqrt{\pi}}$ | 7.350 |
| Pareto, $\alpha = 3,\ \beta = 1$ | $\frac{11}{4}$ | $\frac{9}{7}$ | 54.55 |
| Contaminated Normal $(1-\varepsilon)\Phi(x) + \varepsilon\Phi(\frac{x}{\sigma})$ | $1 - \varepsilon + \varepsilon\sigma^2$ | $\frac{(1-\varepsilon)^2}{2\sqrt{\pi}} + \frac{2\varepsilon(1-\varepsilon)}{\sqrt{2\pi}\sqrt{1+\sigma^2}} + \frac{\varepsilon^2}{2\sqrt{\pi}\sigma}$ | $12[1-\varepsilon+\varepsilon\sigma^2] \times \left[\frac{(1-\varepsilon^2}{2\sqrt{\pi}} + \frac{2\varepsilon(1\varepsilon)}{\sqrt{2\pi}\sqrt{1+\sigma^2}} + \frac{\varepsilon^2}{2\sqrt{\pi}\sigma}\right]^2$ |

Table 4.5.18 $ARE(T_n, R_n)$ for Contaminated normal distribution $(1-\epsilon)\Phi(x) + \epsilon\Phi(\frac{x}{\sigma})$

| $\sigma$ | $\epsilon$ | $ARE(T_n, R_n)$ |
|---|---|---|
| | 0.01 | 0.9574468 |
| | 0.03 | 0.9621882 |
| $\sqrt{2}$ | 0.05 | 0.9665998 |
| | 0.08 | 0.9726232 |
| | 0.10 | 0.9762558 |
| | 0.15 | 0.9840713 |
| | 0.01 | 0.9626593 |
| | 0.03 | 0.9771515 |
| $\sqrt{3}$ | 0.05 | 0.9904497 |
| | 0.08 | 1.0082670 |
| | 0.10 | 1.0187960 |
| | 0.15 | 1.0407310 |
| | 0.01 | 0.9692364 |
| | 0.03 | 0.9959236 |
| 2 | 0.05 | 1.0201930 |
| | 0.08 | 1.0523180 |
| | 0.10 | 1.0710923 |
| | 0.15 | 1.0922600 |
| | 0.01 | 1.0087570 |
| | 0.03 | 1.1077960 |
| 3 | 0.05 | 1.1959810 |
| | 0.08 | 1.3093680 |
| | 0.10 | 1.3732970 |
| | 0.15 | 1.4966790 |
| | 0.01 | 1.0696430 |
| | 0.03 | 1.2790110 |
| 4 | 0.05 | 1.4631900 |
| | 0.08 | 1.6960120 |
| | 0.10 | 1.8246570 |
| | 0.15 | 2.0644620 |

# CHAPTER V

# A GENERAL CORRELATION PROBLEM

## 5.1 Introduction

In this chapter we study the association between the two variables X and Y based on a random sample $(X_1, Y_1)', (X_2, Y_2)', \ldots, (X_n, Y_n)'$ from some bivariate continuous distribution H with marginal cdfs F and G. This problem is commonly known as the test for independence since the available testing procedures based on statistics such as Hoeffding's D, Pearson's R, Spearman's rho and Kendall's tau, all test the null hypothesis of independence,

$$H_0 : X \quad \text{and} \quad Y \quad \text{are independent.} \tag{5.1.1}$$

We shall propose yet another test procedure based on the statistics $T_n$ calculated from $(X_1, Y_1)', (X_2, Y_2)', \ldots, (X_n, Y_n)'$, for the above hypothesis and compare it with some of the above mentioned tests.

In section 5.2, we shall review the class of bivariate distribution function defined by Farlie (1960) and study how Pearson's product moment correlation coefficient$(R_n)$ and Kendall's alternative rank coefficient$(t_n)$ compare in efficiency with each other and with a maximum likelihood estimator of the parameter of association. In section 5.3, we shall obtain an estimate of the parameter of association for the above bivariate distribution based on $T_n$ and obtain the

asymptotic relative efficiency of $T_n$ relative to $R_n$, $t_n$ and a maximum likelihood estimator.

## 5.2 A General Bivariate Distribution

The form of the bivariate distribution considered in this section is the class given by Farlie (1960).

Let F(x) be the marginal continuous distribution function of X, G(y) be the marginal continuous distribution function of Y, A(F(x)) be a function of x that tends to zero as F(x)→1, B(G(y)) be a function of y that tends to zero as G(y)→1.

Then

$$H(x,y) = F(x)G(y)\{1 + \alpha A(F)B(G)\} \tag{5.2.1}$$

is a suitably general class of bivariate distribution functions for which the marginal distribution functions are F(x) and G(y). Note that for the above model, the hypothesis of independence in (5.1.1) is equivalent to the hypothesis $\alpha = 0$. Farlie showed that the functions A(F) and B(G) are not completely random but it will suffice if we choose for A(F) and B(G), functions that are bounded and have bounded first differential coefficients with respect to their arguments. Since $\alpha$ is a parameter measuring association, we can assume without loss generality that the least upper bounds on $|A(F)|$ and $|B(G)|$ are both 1. The expression for H(x,y) is then unique.

Let $(X_1, Y_1)', (X_2, Y_2)', \ldots, (X_n, Y_n)'$ be a random sample from a bivariate distribution with distribution function H given by (5.2.1). Farlie considered Kendall's tau $(t_n)$ and Pearson's product moment correlation coefficient $(R_n)$ among others for comparison with each other and with the maximum likelihood estimator $\hat{\alpha}$ of the parameter of association $\alpha$. He showed that the mean values of $t_n, R_n$ and $\hat{\alpha}$ when $\alpha$ is small are:

$$\begin{aligned}
E(t_n) &= 8\alpha \int_0^1 FAdF \int_0^1 GBdG; \\
E(R_n) &= \frac{\alpha}{\sigma_x \sigma_y} \int_0^1 X\frac{d(FA)}{dF}dF \int_0^1 Y\frac{d(GB)}{dG}dG; \\
E(\hat{\alpha}) &= \alpha,
\end{aligned}$$

and the variances of $t_n, R_n$ and $\hat{\alpha}$ under the hypothesis $\alpha = 0$ are:

$$\begin{aligned}
Var_0(t_n) &= \frac{4}{9n}; \\
Var_0(R_n) &= \frac{1}{n}; \\
Var_0(\hat{\alpha}) &= \frac{1}{n\int_0^1 \left(\frac{d(FA)}{dF}\right)^2 dF \int_0^1 \left(\frac{d(GB)}{dG}\right)^2 dG}.
\end{aligned}$$

The asymptotic relative efficiencies relative to the maximum likelihood estimate are, for $t_n$,

$$\frac{144\left(\int_0^1 FAdF \int_0^1 GBdG\right)^2}{\int_0^1 \left(\frac{d(FA)}{dF}\right)^2 dF \int_0^1 \left(\frac{d(GB)}{dG}\right)^2 dG} \tag{5.2.2}$$

and for $R_n$,

$$\frac{\left(\int_0^1 X\frac{d(FA)}{dF}dF \int_0^1 Y\frac{d(GB)}{dG}dG\right)^2}{\sigma_x^2\sigma_y^2 \int_0^1 \left(\frac{d(FA)}{dF}\right)^2 dF \int_0^1 \left(\frac{d(GB)}{dG}\right)^2 dG}. \tag{5.2.3}$$

Farlie also showed that along the line $A = 1-F$, Kendall's tau is better than the product moment correlation coefficient; along $A = \frac{\int_0^x x dF}{\int_0^x dF}$, the product moment correlation coefficient is better than Kendall's tau. At the intersection, each is equally efficient with the maximum likelihood estimator.

## 5.3 The Performance of $T_n$ as a measure of correlation

In what follows, we shall derive the asymptotic relative efficiency of our proposed new estimate of the parameter of association, relative to the maximum likelihood estimate as well as the product moment correlation coefficient and Kendall's tau. But, first we obtain an expression for the population characteristic $\tau^*$ under the bivariate distribution given by (5.2.1). Note that,

$$\tau^* = \frac{E(X_1 - X_2)sgn(Y_1 - Y_2)}{E|X_1 - X_2|}. \tag{5.3.1}$$

As pointed out by Yitzhaki and Olkin (1987), an alternative form for $E|X_1 - X_2|$ is $4Cov(X, F(X))$. As we expect from this relation, $E(X_1 - X_2)sgn(Y_1 - Y_2)$

can be shown to be $4Cov(X, G(Y))$ as follows:

$$
\begin{aligned}
&E(X_1 - X_2)sgn(Y_1 - Y_2)\\
&= \int_{x_1}\int_{y_1}\int_{x_2}\int_{y_2}(x_1 - x_2)sgn(y_1 - y_2)h(x_1, y_1)h(x_2, y_2)dx_1dy_1dx_2dy_2\\
&= 2\int_{x_1}\int_{y_1}\int_{x_2}\int_{y_2} x_1 sgn(y_1 - y_2)h(x_1, y_1)h(x_2, y_2)dx_1dy_1dx_2dy_2\\
&= 2\int_{x_1}\int_{y_1} x\left[\int_{y_2<y_1} 1\cdot g(y_2)dy_2 + \int_{y_2>y_1}(-1)\cdot g(y_2)dy_2\right]h(x_1, y_1)dx_1dy_1\\
&= 2\int_{x_1}\int_{y_1} x_1[G(y_1) - (1 - G(y_1))]h(x_1, y_1)dx_1dy_1\\
&= 2\int_{x_1}\int_{y_1} x_1[2G(y_1) - 1]h(x_1, y_1)dx_1dy_1\\
&= 4\int_{x_1}\int_{y_1} x_1\left[G(y_1) - \frac{1}{2}\right]h(x_1, y_1)dx_1dy_1\\
&= 4Cov(X, G(Y)).
\end{aligned}
$$

Thus $\tau^*$ in (5.3.1) can be written as

$$\tau^* = \frac{Cov(X, G(Y))}{Cov(X, F(X))}. \tag{5.3.2}$$

Under model (5.2.1),

$$
\begin{aligned}
EXG(Y) &= \int_0^1 XG\left\{1 + \alpha\frac{d(FA)}{dF}\cdot\frac{d(GB)}{dG}\right\}dFdG\\
&= \int_0^1 XdF\int_0^1 GdG + \alpha\int_0^1 X\frac{d(FA)}{dF}dF\int_0^1 G\frac{d(GB)}{dG}dG\\
&= EXEG(Y) + \alpha\int_0^1 X\frac{d(FA)}{dF}dF\int_0^1 G\frac{d(GB)}{dG}dG\\
&= EXEG(Y) - \alpha\int_0^1 X\frac{d(FA)}{dF}dF\int_0^1 GBdG.
\end{aligned}
$$

The last equation follows from the fact that $\int G\frac{d(GB)}{dG}dG = G^2B - \int GBdG$ and $G^2B$ is zero at both ends of the interval [0,1]. From this, it it follows that,

$$Cov(X, G(Y)) = -\alpha\int_0^1 X\frac{d(FA)}{dF}dF\int_0^1 GBdG.$$

Thus from (5.3.2),

$$\tau^* = -\frac{\alpha \int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG}{\int_0^1 X \left[F - \frac{1}{2}\right] dF}. \tag{5.3.3}$$

Note that $\alpha = 0$ implies $\tau^* = 0$.

For large n, $ET_n$ can be approximated by $\tau^*$ and thus an estimator of $\alpha$ based on $T_n$ is given by

$$\hat{\alpha}_T = -\frac{\int_0^1 X \left[F - \frac{1}{2}\right] dF}{\int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG} \cdot T_n \tag{5.3.4}$$

and in the null case this estimate has variance

$$\begin{aligned}
Var_0(\hat{\alpha}_T) &= \frac{\left(\int_0^1 X \left[F - \frac{1}{2}\right] dF\right)^2}{\left(\int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG\right)^2} \cdot Var_0(T_n) \\
&= \frac{\left(\int_0^1 X \left[F - \frac{1}{2}\right] dF\right)^2}{\left(\int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG\right)^2} \cdot \frac{1}{(E|X_1 - X_2|)^2} \frac{4}{3n} \sigma_x^2 \\
&= \frac{\left(\int_0^1 X \left[F - \frac{1}{2}\right] dF\right)^2}{\left(\int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG\right)^2} \cdot \frac{1}{\left(4 \int_0^1 X \left[F - \frac{1}{2}\right] dF\right)^2} \frac{4}{3n} \sigma_x^2 \\
&= \frac{1}{12n} \frac{\sigma_x^2}{\left(\int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG\right)^2}.
\end{aligned}$$

Thus the asymptotic relative efficiency of $T_n$ relative to the maximum likelihood estimate is

(5.3.5)

$$\begin{aligned}
&ARE(T_n, MLE) \\
&= \frac{\textit{Asymptotic null variance of MLE of } \alpha}{\textit{Asymptotic null variance of an estimate of } \alpha \textit{ based on } T_n} \\
&= \frac{12 \left(\int_0^1 X \frac{d(FA)}{dF} dF \int_0^1 GBdG\right)^2}{\sigma_x^2 \left(\int_0^1 \left(\frac{d(FA)}{dF}\right)^2 dF \int_0^1 \left(\frac{d(GB)}{dG}\right)^2 dG\right)^2}.
\end{aligned}$$

On lines similar to those of Farlie, the functions F(x) and A(F(x)), G(y) and B(G(y)) that maximize the relative efficiency of $T_n$ when compared with the maximum likelihood estimator are:

$$A(x) = \frac{\int_{-\infty}^{x} x dF}{\int_{-\infty}^{x} dF}$$

and

$$B(y) = 1 - G(y).$$

The asymptotic relative efficiency of $T_n$ relative to the product moment correlation coefficient ($R_n$) and Kendall's tau ($t_n$) are respectively,

(5.3.6) $$ARE(T_n, R_n) = 12\sigma_y^2 \frac{\left(\int_0^1 GBdG\right)^2}{\left(\int_0^1 Y \frac{d(GB)}{dG} dG\right)^2},$$

and

(5.3.7) $$ARE(T_n, t_n) = \frac{1}{12\sigma_x^2} \frac{\left(X \frac{d(FA)}{dF} dF\right)^2}{\left(\int_0^1 FAdF\right)^2}.$$

Note that $ARE(T_n, R_n)$ is a function of the marginal distribution of Y only and $ARE(T_n, t_n)$ is a function of the marginal distribution of X only. Also a similar argument to that of Farlie, shows that $T_n$ is better than $R_n$ along the line B(y)=1- G(y) but not as good as $t_n$ and $T_n$ is better than $t_n$ along the line $A(x) = \frac{\int_0^x x dF}{\int_0^x dF}$, but not as good as $R_n$. At the intersection, each of the three is equally efficient with the maximum likelihood estimate.

BIBLIOGRAPHY


Anderson, T.W., *An Introduction to Multivariate Analysis.* New York: John Wiley 1984

Billingsley, P., *Probability and Measure.* New York: John Wiley 1979

Bhuchongkul, S., A class of nonparametric tests for independence in bivariate populations. *AMS 35* 1956, 138-149

Chung, K.L., *A Course in Probability Theory.* New York: Academic Press 1967

Farlie, D.J.G., The performance of some correlation coefficients for a general bivariate distribution. *Biom 47* 1960, 307-323

Gokhale, D.V., Some problems in independence and dependence. *Phd Thesis* University of California: Berkeley 1966

Gumbel, E.J., *Comptes Rendus 246* 1958, 2717-2719

Hájek, J., and Šidák, Z., *Theory of Rank Tests.* New York: Academic Press 1967

Hampel, F.R., The influence curve and its role in robust estimation *JASA 69* 1974, 383-393

Hampel, F.R., Rousseeuw, E.M., Ronchetti, E.M., and Stahel, W.A., *Robust Statistics: The Approach based on Influence Functions.* New York: John Wiley 1985

Hettmansperger, T.P., *Statistical Inference Based on Ranks.* New York: John Wiley 1984

Hettmansperger, T.P., and McKean, J.W., A robust alternative based on ranks to least squares in analyzing linear models. *TC 19* 1975, 275-284

Jaeckel, L., Estimating regression coefficients by minimizing the dispersion function . *AMS 43* 1972, 1449-1458

Jurečkova', J., Nonparametric estimate of regression coefficients. *AMS 42* 1971,

1328-1338

Kendall, M.G., *Rank Correlation Methods.* London: Griffin 1970

Konijn, H.S., On the power of certain tests for independence in bivariate populations. *AMS 27* 1956, 300-323

Lees, M., *Several Variable Calculus.* New York: Worth Publishers, Inc. 1971

McKean, J.W., and Hettmansperger, T.P., Tests of hypothesis based on ranks in the general linear model. *Comm. St. A* 1976, 693-709

Morgenstern, D., *Mitt. Math. Statist. 8* 1956, 234-235

Muirhead, R.J., *Aspects of Multivariate Statistical Analysis.* New York: John Wiley 1982

Noether, G.E., On a theorem of Pitman. *AMS 26* 1955, 64-68

Pitman, E.J.G., *Unpublished Lecture Notes.* University of North Carolina, 1948

Quade, D., *Measurements in the Social Sciences: Theories and Strategies, edited by H.M. Blalock, Jr.* Chicago: Aldine Publishing Company 1974

Randles, R.H., On tests applied to Residuals *JASA 79* 1984, 349-354

Randles, R.H., On the asymptotic normality of statistics with estimated parameters *AS 10* 1982, 462-474

Randles, R.H., and Wolfe, D.A., *Introduction to the Theory of Nonparametric Statistics.* New York: John Wiley 1979

Rao, C.R., *Linear Statistical Inference and its Applications.* New York: John Wiley 1973

Raveh, A., On measures of monotone association. *The American Statistician 40* 1986, 117-123

Robinson, E.A., *Least Squares Regression Analysis in terms of Linear Algebra.* Houston: Goose Pond Press 1981

Samara, B., Tests for correlation and partial correlation based on Kendall's tau. *Phd Thesis* University of Florida 1985

Schollenberger, J., Agresti, A., and Wackerly, D.D., Measuring association and modelling relationships between interval and ordinal variables.*ASA Proceedings of Social Statistics* 1979, 624-626

Serfling, R.J., *Approximation Theorems of Mathematical Statistics.* New York: John Wiley 1980

Shirahata, S., Tests for partial correlation in a linear model. *Biom 64* 1977, 162-164

Sievers, G.L., A weighted dispersion function for estimation in linear models. *Comm. St. A* 1983, 1161-1179

Sievers, G.L., Weighted Rank Statistics for simple linear regression. *JASA 73* 1978, 628-631

Sievers, G.L., A Robust Multiple Correlation Coefficient for the Rank Analysis of Linear Models. *Unpublished manuscript* 1983

Somers, R.H., A new Asymmetric measure of association for ordinal variables. *American Sociological Review* 1959, 799-811

Steel, R.G.D., and Torrie, J.H., *Principles and Procedures of Statistics.* New York: McGraw-Hill 1960

Yitzhaki, S., and Olkin, I., Gini Regression Analysis. *Unpublished manuscript* 1987