This research was supported in part by the National Institutes of Health, Grant No. GM-12868 and partially by the Air Force Office of Scientific Research, Grant No. 68-1615.

ON THE ASYMPTOTIC PROPERTIES OF SOME ROBUST ESTIMATORS IN CERTAIN MULTIVARIATE STATIONARY AUTOREGRESSIVE PROCESSES

by

Kalyan Dutta
Department of Statistics and Biostatistics

University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 762

JULY 1971

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

It is a pleasure and privilege to acknowledge my deep feeling of indebtedness to my advisor, Professor P. K. Sen, for suggesting the problems in this dissertation. His guidance and assistance at every stage of the work have been illuminating and instructive.

I would also like to thank Professor Wassily Hoeffding, the chairman of my doctoral committee, for his invaluable suggestions and comments. Thanks are also due to Professor N. L. Johnson, Professor E. J. Wegman and Professor R. L. Davis, the other members of my doctoral committee, for going through the manuscript.

For financial support during my graduate training, I sincerely thank the Department of Statistics, the Department of Biostatistics, and the Air Force Office of Scientific Research and the National Institutes of Health. To these institutions I gratefully express my appreciation.

For her excellent typing of the final manuscript I extend my thanks to Mrs. Evelyn Nichols.

I wish to express my sincere gratitude to Professor R. C. Bose who was chiefly responsible for my decision to come over to Chapel Hill. I wish also to thank Professor J. E. Grizzle of the Department of Biostatistics for his encouragement during my graduate studies and Professor H. K. Nandi of the University of Calcutta for stimulating my interest in Mathematical Statistics.

Finally, these acknowledgements would be incomplete without mention of the active co-operation and enthusiasm of my grandfather, Mr. N. Dutta, and my parents without whose encouragement I would not have reached this stage of my academic career.

ABBREVIATIONS AND NOTATIONS

Abbreviations

iff	if and only if
r.v.	random variable
d.f.	distribution function
c.d.f.	cumulative distribution function
i.i.d.	independent and identically distributed
i.i.d.r.v.	independent and identically distributed random variables
a.s.	almost sure
p.d.	positive definite
C.L.T.	Central Limit Theorem
A.R.E.	Asymptotic Relative Efficiency

Notations

$N_q(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$	q variate normal with mean vector $\underset{\sim}{\mu}$ and dispersion matrix $\underset{\sim}{\Sigma}$
$\underset{p}{\rightarrow}$	convergence in probability
$\forall$	for every
identity matrix of order p	$\underset{\sim}{I}_p$ (unless otherwise mentioned)
$\underset{\sim}{J}_p$	square matrix of order p with each element unity

$$\begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix} = \underset{\sim}{A} = \underset{p\times q}{((a_{ij}))}$$

ABSTRACT

The following problems are considered in this dissertation.

In Chapters II and III, it is shown that Bahadur's almost sure asymptotic representation of a sample quantile for independent and identically distributed random variables holds under certain regularity conditions for a general class of stationary multivariate autoregressive processes. This yields the asymptotic (multi-) normality of the standardized forms of quantiles in autoregressive processes. Then the asymptotic (multi-) normality is generalized for a vector of linear compounds of the standardized forms of quantiles in autoregressive processes and the estimation of its dispersion matrix is also considered. Further, the theory of asymptotic (multi-) normality of a vector of linear compounds of the standardized forms of quantiles in autoregressive processes is extended for the random sample size.

In Chapter IV, for stationary multivariate autoregressive processes, the asymptotic distribution of Wilcoxon signed rank statistic is derived and subsequently the theory is generalized for a class of U-statistics.

Finally, in Chapter V, several rival estimators (median, 27% mid-range, median of the mid-ranges and mean) of location parameter for stationary autoregressive processes are considered and their relative performances are compared by their asymptotic relative efficiencies.

CHAPTER I

INTRODUCTION AND SUMMARY

1.1. Introduction and a Review of the Literature

The central limit theorem in its various forms, the characteristic functions and the laws of large numbers are the basic tools for the study of the asymptotic properties of linear estimators (i.e. linear functions of the unordered sample observations). However, in dealing with ordered observations or, in general, a linear function of ordered observations we require additional tools for the study of their asymptotic properties. For example, asymptotic normality of sample quantiles for i.i.d.r.v. can be deduced by the direct method given in Cramer (1946, pp. 367-369) or in Mosteller (1946). This approach becomes increasingly complicated as the number of quantiles increases. Also for the study of other large sample properties of quantiles (e.g., asymptotic normality in the multivariate case, almost sure convergence, law of iterative logarithm, etc.) this method becomes too complicated to provide the necessary results. A recent asymptotic representation of sample quantiles by Bahadur (1966) not only simplifies the proof of the asymptotic normality of quantiles in various univariate as well as multivariate situations but also provides the access to the study of more refined large sample properties of quantiles. The other notable non-linear estimators are the U-statistics, studied by Hoeffding (1948), and rank order statistics, studied by Chernoff and Savage (1958) and Hájek (1968) among others. For these statistics either a projection technique is used to decompose them into two parts where on the first part the classical C.L.T. and the laws of large numbers, etc., directly apply while the second part is asymptotically negligible in probability or they are related as functionals of empirical distributions and their large sample properties are studied with the aid of the large sample properties of the empirical distributions.

In this dissertation we are principally concerned with certain stochastic difference equations or stationary autoregressive processes where the successive observations are not independent but are subject to a chain of dependence. To be general we start with the following multivariate model:

$$\sum_{r=0}^{k} \underset{\sim}{A}_r \underset{\sim}{X}_{t-r} = \underset{\sim}{\varepsilon}_t, \quad t = 0, \pm 1, \ldots, \tag{1.1.1}$$

where $\underset{\sim}{X}_t = (X_{t,1}, \ldots, X_{t,q})'$ is a $q \times 1$ vector, $\underset{\sim}{A}_r = ((a_{js}^{(r)}))_{j,s=1,\ldots q}$, $r = 0, 1, \ldots, k(\geq 1)$ are $q \times q$ matrices of constants and $\underset{\sim}{\varepsilon}_t = (\varepsilon_{t,1}, \ldots, \varepsilon_{t,q})'$ (white noise) are i.i.d. $q \times 1$ random vectors with an absolutely continuous (q-variate) c.d.f., $G(\underset{\sim}{x})$, $\underset{\sim}{x} \in R^q$, the real q-space.

Meteorological and economic data provide perhaps the most obvious examples of such a model. Its adequacy has been investigated by Whittle (1953). He has successfully fitted the above model to data giving the total sunspot area for a series of 120 six-monthly periods, and for two belts of solar latitude (16°-21° N., and 16°-21° S.).

For autoregressive errors Hannan (1961) and Eicker (1965) studied large sample distribution properties of the classical least square estimators computed under the assumption of the independence of errors. Since these estimators are linear functions of the original observations, they are also linear in terms of the error components. These errors being independent, they were in a position to apply again the classical C.L.T. for their purpose. We briefly present their principal results below.

Consider a system of regressions

$$\tag{1.1.2} \underset{\sim}{X}_t = \underset{\sim}{\beta}^* \underset{\sim}{Y}_t + \underset{\sim}{\varepsilon}_t$$

where $\underset{\sim}{X}_t = (X_{t,1},\ldots, X_{t,q})'$, $\underset{\sim}{\varepsilon}_t = (\varepsilon_{t,1},\ldots,\varepsilon_{t,q})'$ are $q \times 1$ random vectors, $\underset{\sim}{Y}_t = (Y_{t,1},\ldots, Y_{t,p})'$ is a $p \times 1$ vector, $\underset{\sim}{\beta}^* = ((\beta^*_{ij}))_{\substack{i=1,\ldots,q \\ j=1,\ldots,p}}$ is a $(q \times p)$-matrix of regression coefficients and

$$\tag{1.1.3} \underset{\sim}{\varepsilon}_t = \sum_{j=-\infty}^{\infty} \underset{\sim}{C}_j \underset{\sim}{\eta}_{t-j}$$

where $\underset{\sim}{C}_j$ are a sequence of known $(q \times r)$-matrices and $\underset{\sim}{\eta}_t$ are $r \times 1$ i.i.d. random vectors with covariance matrix $\underset{\sim}{Q}$ and finite absolute moments of order $2 + \delta$, $\delta > 0$. The norm $||\underset{\sim}{C}||_\ell$ of a matrix $\underset{\sim}{C}$ is defined by

$$||\underset{\sim}{C}||_\ell = \sup_{||\underset{\sim}{x}||_\ell = 1} ||\underset{\sim}{C}\underset{\sim}{x}||_\ell$$

where for a vector $\underset{\sim}{x}$, $||\underset{\sim}{x}||_\ell = \{\sum_i |x_i|^\ell\}^{1/\ell}$. Then it is assumed that

$$\tag{1.1.4} \sum_{j=-\infty}^{\infty} ||\underset{\sim}{C}_j||_2 < \infty$$

and the components $Y_{t,j}$ of $\underset{\sim}{Y}_t$ are generated by a process such that, with probability one, as n increases

(1.1.5) $\lim_{n\to\infty} d^2_{n,j} = \infty, \quad d^2_{n,j} = \sum_{t=1}^{n} Y^2_{t,j}, \quad j = 1,\ldots,q.$

(1.1.6) $\lim_{n\to\infty} Y^2_{n,j} \Big/ d^2_{n,j} = 0, \quad j = 1,\ldots,q$

(1.1.7) $\lim_{n\to\infty} \left\{ \left(\sum_{t=1}^{n-h} Y_{t,j} Y_{t+h,k} \right) \Big/ d_{n,j} d_{n,k} \right\} = \rho_{jk}(h), \quad j,k = 1,\ldots,q$

If $\underset{\sim}{R}(h) = ((\rho_{jk}(h)))_{j,k=1,\ldots,q}$ then $\underset{\sim}{R}(h)$ can be written as

(1.1.8) $$\underset{\sim}{R}(h) = \int_{-\pi}^{\pi} e^{ih\lambda} d\underset{\sim}{M}(\lambda)$$

where $\underset{\sim}{M}(\lambda)$ is a matrix valued function whose increments are Hermitian non-negative. It is assumed that $\underset{\sim}{R}(0)$ is non-singular.

When the errors are independent, the least square estimate of $\underset{\sim}{B}^*$ is

(1.1.9) $$\hat{\underset{\sim}{B}}^* = ((b^*_{ij})) = \underset{\sim}{X}\underset{\sim}{Y}'(\underset{\sim}{Y}\underset{\sim}{Y}')^{-1}$$

where $\underset{\sim}{X}$ has $\underset{\sim}{X}_t$ and $\underset{\sim}{Y}$ has $\underset{\sim}{Y}_t$ as the t-th column, $t = 1,2,\ldots,n$. If (1.1.9) is expressed in the form of a column vector $\underset{\sim}{b}^*$ having b^*_{ij} in $\{q(i-1) + j\}$-th row, then

$$\underset{\sim}{b}^* = (\underset{\sim}{I}_q \otimes \underset{\sim}{Y}\underset{\sim}{Y}')^{-1}(\sum_t \underset{\sim}{X}_t \otimes \underset{\sim}{Y}_t)$$

where $\otimes$ denotes the Kronecker product. If $\underset{\sim}{\beta}^*$ is the vector obtained by arranging the rows of $\underset{\sim}{B}^*$ in the same way, then $E(\underset{\sim}{b}^*) = \underset{\sim}{\beta}^*$.

We denote by $\underset{\sim}{D}_n$, the diagonal matrix having $d_{n,j}$ in the j-th place. Then Rosenblatt (1956a) has shown that, as $n \to \infty$, process. These estimators being non-linear functions of the original observations are also non-linear in terms of the error components. Actually we are concerned here with a class of estimators different from the one considered by Hannan (1961) and Eicker (1965). Further, our proposed class of estimators being non-linear function of the errors, their technique does not seem to be readily applicable in this case.

Estimation and the asymptotic distribution of the estimates of the coefficients of autoregressive processes are available in the literature (see Grenander and Rosenblatt, 1957). Also, in the case of autoregressive schemes, the asymptotic distribution of sample mean is well known (see Hannan, 1961). However, for autoregressive processes, unlike the case of independent observations, we do not know much about the distribution theories of Hoeffding's U-statistics or sample quantiles and other related statistics. To bridge this gap we want to develop the distribution theory of sample quantiles, any linear combination of sample quantiles and a class of U-statistics for stationary autoregressive processes.

First, for a stationary autoregressive process, we consider the asymptotic distribution theory of a sample quantile. The chain of dependence by which the successive observations in the process are linked makes it difficult to apply the fundamental technique (such as in Cramér, 1946, pp.367-369) for deriving the asymptotic normality of a sample quantile. An alternative approach (see section 2.1) of deriving the asymptotic distribution of a sample quantile based on the properties of order statistics seems to be applicable for autoregressive processes, but this approach fails to give us more refined convergence results of sample quantiles. These difficulties have been avoided here by adopting the elegant Bahadur-representation of a quantile (cf. Bahadur, 1966) and extending it in the context of autoregressive processes. With this end in view, we first present the principal results of Bahadur (1966).

Let $\omega = (X_1, X_2, \ldots)$ be a sequence of i.i.d.r.v. with d.f. $F(x)$. If ξ be the population quantile of order p, then it is assumed that F has at least two derivatives in some neighbourhood of ξ, $F''(x)$ is bounded in the neighbourhood, and $F'(\xi) = f(\xi) > 0$. For each $n = 1, 2, \ldots$, let $Y_n(\omega)$ be the sample p-quantile when the sample is $(X_1, \ldots, X_n)$. Let $Z_n(\omega)$ be the number of observations X_i in the sample $(X_1, \ldots, X_n)$ such that $X_i > \xi$. Then

(1.1.11) $$Y_n(\omega) = \xi + [(Z_n(\omega) - nq)/nf(\xi)] + R_n(\omega)$$

where

$$q = 1 - p$$

and

(1.1.12) $$R_n(\omega) = O(n^{-3/4} \log n), \quad \text{as } n \to \infty,$$

with probability one.

One important feature of the Bahadur-representation of sample quantiles is its ability to yield the asymptotic normality of sample quantiles in various non-standard situations where either independence or homogeneity of the distribution functions or both may be vitiated. With this motivation Sen (1968 b) showed that Bahadur's asymptotic almost sure representation of the standardized form of a sample quantile is also valid for any m-dependent process. The simplified version of his results for stationary m-dependent processes are given below.

Let $\omega = (X_1, X_2, \ldots)$ be a sequence of random variables forming an m-dependent process, i.e. X_i and X_j are independent whenever $|i - j| > m$. The marginal c.d.f. of X_i is denoted by $F(x)$, and the joint c.d.f. of (X_i, X_{i+h}) by $F_h(x, y)$, for $h = 1, \ldots, m$. For any $p: 0 < p < 1$, let $Y_n(\omega)$ be the sample p-quantile when the sample is $(X_1, \ldots, X_n)$ and let the empirical c.d.f. $F_n(x)$ be defined by

$$(1.1.13) \qquad F_n(x) = n^{-1} \sum_{i=1}^{n} c(x - X_i), \quad -\infty < x < \infty ,$$

where

$$(1.1.14) \qquad c(u) = 1 \text{ or } 0 \text{ according as } u \geq \text{ or } < 0 .$$

If ξ be the population p-quantile of $F(x)$ then it is defined by

$$F(\xi) = p .$$

It is assumed that in the neighbourhood of ξ, $F(x)$ is absolutely continuous and that

a) $f(x) = (d/dx)F(x)$ is continuous in some neighbourhood of ξ, with $0 < f(\xi) < \infty$, and

b) $F''(x)$ is bounded in the same neighbourhood of ξ.

If $\nu^2_{n,m}$ is defined by

$$(1.1.15) \qquad \nu^2_{n,m} = p(1 - p) + 2 \sum_{h=1}^{m} (1 - h/n)\{F_h(\xi, \xi) - p^2\}$$

then it is shown that

$$n \operatorname{Var}\{F_n(\xi)\} = \nu^2_{n,m}$$

Finally, let

(1.1.16) $$I_n = \{x: \xi - a_n \leq x \leq \xi + a_n\}$$

where $a_n \sim n^{-1/2} \log n$, as $n \to \infty$.

With the above notations we have the following theorem.

Theorem 1.1.2. If the condition (a) is satisfied, then as $n \to \infty$,

(1.1.17) $$\sup\{|[F(x) - p] + [F_n(\xi) - F_n(x)]| : x \in I_n\} = O(n^{-3/4} \log n),$$

with probability one. If in addition, $\inf_n \nu^2_{n,m} > 0$,

(1.1.18) $$\mathcal{L}(n^{1/2} f(\xi)[Y_n(\omega) - \xi]/\nu_{n,m}) \to N(0, 1).$$

Finally, if both the conditions (a) and (b) are satisfied then

(1.1.19) $$[Y_n(\omega) - \xi]f(\xi) + [F_n(\xi) - p] = R_n(\omega)$$

where as $n \to \infty$, $R_n(\omega) = O(n^{-3/4} \log n)$, with probability one.

In the case of a stationary autoregressive process, it is shown here that for the asymptotic distribution theory of a sample quantile, we may as well (for large n only) replace the original process by an m_n-dependent process where $m_n \sim K \log n$, and K is some fixed positive number. Also, it is shown here that Sen's extension for stationary m-dependent processes of asymptotic almost sure representation of the standardized form of a sample quantile, as given above, can be further extended to the stationary m_n-dependent processes. This enables us to draw conclusions about the asymptotic normality of sample quantiles for a stationary autoregressive process.

Next, the distribution theory of any linear compound of several quantiles for stationary autoregressive processes is derived by extending our previous results for a particular quantile over the entire real line. Again, this extension is possible by extensive use of Sen and Ghosh's (1971) generalization of Bahadur's asymptotic almost sure representation of the standardized form of a sample quantile over the entire real line. Their generalization is briefly indicated here.

Let the distribution function F(x) be absolutely continuous with density function f(x) and $\sup_x f(x) = f_0 < \infty$. Let the empirical distribution function be $F_n(x)$ and

$$g_\kappa(n) = n^{-1/2}(\log n)^\kappa, \quad \kappa \geq 1.$$

Then Sen and Ghosh proved the following.

<u>Theorem</u> 1.1.3. <u>For every finite</u> s (>0) <u>there exists a positive constant</u> $C_s^{(1)}$ <u>and a sample size</u> n_s, <u>such that for</u> $n \geq n_s$,

$$(1.1.20) \qquad P\{\sup_x \sup_{|a|<g_\kappa(n)} n^{1/2}|F_n(x+a) - F_n(x) - F(x+a) + F(x)| > C_s^{(1)} n^{-1/4}(\log n)^\kappa\} \leq 4n^{-s}$$

Hence

$$(1.1.21)\qquad \sup_{x}\ \sup_{|a|<g_{\kappa}(n)} \{n^{1/2}|F_n(x+a) - F_n(x) - F(x+a) + F(x)|\} = O(n^{-1/4}(\log n)^{\kappa}),$$

with probability one as $n \to \infty$.

The above theorem is first extended to the stationary m_n-dependent process where $m_n \sim K \log n$, K being a positive number. This extension along with the asymptotic reduction of a stationary autoregressive process to a stationary m_n-dependent process enables us to prove the same theorem for a class of autoregressive processes. Then asymptotic normality of a linear compound of several quantiles for stationary autoregressive processes follows easily as a corollary of the theorem. An application of this result is made by finding the asymptotic distribution of the 27% mid-range estimate for a stationary autoregressive process and the result is also extended for the case of random sample size.

One of the mostly used statistics based on ranks is the so-called Wilcoxon score estimator which is expressed as the median of the mid-ranges. This estimator is derived from the Wilcoxon signed rank statistic (see Hodges and Lehmann, 1963) which can be expressed as a Hoeffding's (1948) U-statistic as well as a rank order statistic. The asymptotic distribution of Wilcoxon signed rank statistic for stationary autoregressive processes is derived by first asymptotically reducing it to a Wilcoxon signed rank statistic for an m_n-dependent stationary process and then using Rosenblatt's (1956b) C.L.T. for strongly mixing processes. Since Wilcoxon signed rank statistics can also be expressed as a U-statistic, attention is concentrated to generalize this theory to the study of the distribution theory of U-statistics for stationary autoregressive processes. This is done under certain regularity conditions. These regularity conditions are somewhat more restrictive than those in Hoeffding (1948) dealing with independent observations and appear to be necessary in view of infinite chain of dependence in the series of observations. It is intended to follow up the general case in near future.

In this generalization extensive use of Sen's (1963) extension of asymptotic normality of Hoeffding's (1948) U-statistics is made. First his extension is given below.

Let $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ be a vector valued sample of size n from an m-dependent stationary process having c.d.f. $F(\underset{\sim}{x})$ and $\phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r})$ be a statistic symmetric in the arguments $\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r}$; $\alpha_1 < \ldots < \alpha_r$. Let us define the parameter associated with the c.d.f. F as

$$(1.1.22) \qquad g(F|\ell, v_1, \ldots, v_\ell) = E\{\phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r})\}$$

where $\alpha_{i+1} - \alpha_i = v_i$, $0 < v_i \leq m$ for $i = i_1 \neq \ldots \neq i_\ell$, while for the remaining $r - \ell - 1$ values of i, $\alpha_{i+1} - \alpha_i > m$, for $\ell = 1, \ldots, r - 1$. For $\ell = 0$, the parameter (1.1.22) is reduced to the following:

$$(1.1.23) \quad g(F) = E\{\phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r})|\alpha_{i+1} - \alpha_i > m \text{ for } i = 1, \ldots, r - 1\}.$$

The statistic $\phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r})$ is termed a non-serial statistic, whenever $\alpha_{i+1} - \alpha_i > m$ for $i = 1, \ldots, r - 1$. The corresponding symmetric estimator $U_0(\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n)$ based on n observations is then defined by

(1.1.24) $$U_0(\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n) = \binom{n - rm + m}{r}^{-1} \sum_{S_0} \phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r})$$

where the summation S_0 extends over all possible $\binom{n - rm + m}{r}$ sets of $\alpha_1, \ldots, \alpha_r$, satisfying $\alpha_{i+1} - \alpha_i > m$, $i = 1, \ldots, r - 1$. Also let

(1.1.25) $$U(\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n) = \binom{n}{r}^{-1} \sum_{S} \phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r}),$$

where the summation S extends over all $1 \le \alpha_1 < \ldots < \alpha_r \le n$. Let us assume for all $(\alpha_1, \ldots, \alpha_r): 0 < \alpha_{i+1} - \alpha_i \le m + 1$, for $i = 1, \ldots, r - 1$,

(1.1.26) $$E\{|\phi(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_r})|^2\} < \infty.$$

Let us put

(1.1.27) $$\phi_a(\underset{\sim}{x}_{\alpha_1}, \ldots, \underset{\sim}{x}_{\alpha_a}) = E\{\phi(\underset{\sim}{x}_{\alpha_1}, \ldots, \underset{\sim}{x}_{\alpha_a}, \underset{\sim}{X}_{\alpha_{a+1}}, \ldots, \underset{\sim}{X}_{\alpha_r})\} - g(F)$$

for $a = 0, 1, \ldots, r$, where $\alpha_{i+1} - \alpha_i > m$ for all $i = 1, \ldots, r - 1$, and let

(1.1.28) $$\zeta_{a\cdot 0} = E\{\phi_a^2(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_a})\}$$

(1.1.29) $$\zeta_{a\cdot(h_1, \ldots, h_a)} = E\{\phi_a(\underset{\sim}{X}_{\alpha_1}, \ldots, \underset{\sim}{X}_{\alpha_a})\phi_a(\underset{\sim}{X}_{\beta_1}, \ldots, \underset{\sim}{X}_{\beta_a})\}$$

for $a = 0, 1, \ldots, r - 1$; where $|\alpha_i - \beta_i| = h_i$; $0 < h_i \le m$ for $i = 1, 2, \ldots, a$, but $|\alpha_i - \alpha_t| > m$, $|\beta_i - \beta_t| > m$ and $|\alpha_i - \beta_t| > m$ for all $i \ne t = 1, \ldots, m$. Finally, let

(1.1.30) $$\zeta_1 = \zeta_{1\cdot 0} + 2 \sum_{h=1}^{m} \zeta_{1\cdot h}$$

Then Sen (1963) proved the following.

Theorem 1.1.4. If $E\{|\phi_1(\underset{\sim}{X}_i)|^3\} < \infty$, and (1.1.26) holds, then both $n^{1/2}\{U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - g(F)\}$ and $n^{1/2}\{U_0(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - g(F)\}$ have asymptotically the same normal distribution with zero mean and variance $r^2\zeta_1$.

We may note that $\underset{\sim}{X}_t$ in (1.1.1) can be expressed as a linear compound of $\underset{\sim}{\varepsilon}_{t-r}$, $r = 0, 1, 2, \ldots, \infty$. Essentially Hannan (1961) and Eicker (1965) considered a truncation of this linear compound and wrote $\underset{\sim}{X}_t = \underset{\sim}{Y}_{n,t} + \underset{\sim}{R}_{n,t}$ where $\underset{\sim}{Y}_{n,t}$ is the truncated part and $\underset{\sim}{R}_{n,t}$ is the residual. Then they approximated the linear function of the $\underset{\sim}{X}_t$'s by the corresponding function of the $\underset{\sim}{Y}_{n,t}$'s. The latter function is again linear in the independent errors $\underset{\sim}{\varepsilon}_t$ over a certain number of terms depending on n, and thereby they were able to use the C.L.T. But the kernel $\phi(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_r)$ of a U-statistic is not, in general, a linear function of its arguments. Therefore the above decomposition is not adaptable. The non-linear nature of the U-statistic and the chain of dependence of the successive observations make it quite complicated to reduce the autoregressive process to an m_n-dependent process in which the above results of Sen (1963) can be incorporated and extended to prove our desired results. This, however, is done here through the following conditions:

i) For every $\varepsilon > 0$, we can find a positive number A such that if we define the qr-dimensional rectangle

$$A^* = \{\underset{\sim}{x}_{\alpha_1},\ldots,\underset{\sim}{x}_{\alpha_r} : |x_{\alpha_i,j}| < A, \text{ for } i = 1,\ldots,r;\ j = 1,\ldots,q\}$$

and I_{A^c} is the indicator function of $A^c = R^{qr} - A^*$, then

$$(1.1.31) \qquad \sup_n \sup_{1\le\alpha_1<\ldots<\alpha_r\le n} E[I_{A^c}|\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})|^3] < \varepsilon$$

and (ii)

$$(1.1.32) \qquad \sup_n \sup_{1\le\alpha_1<\ldots<\alpha_r\le n} E[|\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})|^3] < \infty.$$

Under the above conditions, Sen's extension of asymptotic normality of U-statistics for m-dependent stationary processes is further extended to the m_n-dependent stationary process, where $m_n \sim K \log n$ and K is a positive number. Again under the assumption that the kernel $\phi(\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_r)$ of a U-statistic is either continuous in its arguments or continuous everywhere except at a finite number of strips with finite discontinuities, the U-statistic for the original stationary autoregressive process is asymptotically reduced to a U-statistic for the above m_n-dependent process. This asymptotic reduction along with the extension of Sen's (1963) results enable us to draw conslusions about the asymptotic normality of a U-statistic for stationary autoregressive processes.

Finally, several rival estimators (whose asymptotic distributions are derived here) of location parameter for stationary autoregressive processes are considered and their relative performances are compared with that of process average.

1.2. A Summary of the Results in Chapters II-V

Let $(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n)$ be a sample of size n from the k-th order q-variate stationary autoregressive process denoted by (1.1.1). We assume that in addition to the conditions (cf. Whittle, 1953) necessary for the validity of the expression for $\underset{\sim}{X}_t$ as an infinite (but convergent) series in the $\underset{\sim}{\varepsilon}_{t-r}$, r = 0, 1,..., the following holds:

$$(1.2.1) \qquad E||\underset{\sim}{\varepsilon}_t||^{\delta} < \infty \quad \text{for some } \delta > 0,$$

where $||\underset{\sim}{\varepsilon}_t|| = [\underset{\sim}{\varepsilon}_t' \underset{\sim}{\varepsilon}_t]^{1/2}$. The joint c.d.f. of $\underset{\sim}{X}_t$ and the marginal c.d.f. of $X_{t,j}$ are denoted by F(x) and $F_{[j]}(x)$, and the j-th variate $p^{(j)}$-quantile is denoted by $\xi^{(j)}_{p(j)}$, where

(1.2.2) $\quad F_{[j]}\left\{\xi^{(j)}_{p(j)}\right\} = p^{(j)} : 0 < p^{(j)} < 1; \quad j = 1,\ldots, q.$

The j-th variate sample $p^{(j)}$-quantile is denoted by

(1.2.3) $\quad Z^{(j)}_n = X^{(j)}_{n,r}, \quad r = [np^{(j)}] + 1, \quad j = 1,\ldots, q$

where $X^{(j)}_{n,1} \le X^{(j)}_{n,2} \le \ldots \le X^{(j)}_{n,n}$ are the ordered values of $X_{1,j},\ldots, X_{n,j}$, for $j = 1,\ldots, q$. Let

(1.2.4) $\quad \underset{\sim}{Z}_n = \left(Z^{(1)}_n,\ldots, Z^{(q)}_n\right)', \quad \underset{\sim}{\xi}_{\underset{\sim}{p}} = \left(\xi^{(1)}_{p(1)},\ldots, \xi^{(q)}_{p(q)}\right)'$

For the j-th variate, define the empirical c.d.f.

(1.2.5) $\quad F_{nj}(x) = n^{-1} \sum_{i=1}^{n} c(x - X_{i,j}), \quad -\infty < x < \infty$

where c(u) is defined in (1.1.14). Let

(1.2.6) $\quad I^{(j)}_n = \{x : \xi^{(j)}_{p(j)} - a_n \le x \le \xi^{(j)}_{p(j)} + a_n\}, \quad j = 1,\ldots, q$

where $a_n \sim n^{-1/2} \log n$ as $n \to \infty$. Also define

(1.2.7) $\quad \underset{\sim}{\nu} = ((\nu_{jj'}))_{j,j'=1,\ldots,q};$

$$\nu_{jj'} = \lim_{n\to\infty}\left\{n \operatorname{cov}\left[F_{nj}\left(\xi^{(j)}_{p(j)}\right), F_{nj'}\left(\xi^{(j')}_{p(j')}\right)\right]\right\},$$

and

(1.2.8) $\quad \underset{\sim}{T} = ((\tau_{jj'}))_{j,j'=1,\ldots,q}; \quad \tau_{jj'} = \nu_{jj'}/\left[f_{[j]}\left(\xi^{(j)}_{p(j)}\right) f_{[j']}\left(\xi^{(j')}_{p(j')}\right)\right]$

where $f_{[j]}(x) = (d/dx)F_{[j]}(x); \quad j, j' = 1,\ldots, q.$

Then in connection with quantiles for stationary autoregressive processes, the following two main theorems are proved in chapter II.

Theorem 1.2.1. If in the neighbourhood of $\xi_{p(j)}^{(j)}$, $f_{[j]}(x)$ is finite continuous and positive, then as $n \to \infty$,

$$\text{(1.2.9)} \qquad \sup\left\{\left|\left[F_{nj}(x) - F_{nj}\left(\xi_{p(j)}^{(j)}\right)\right] - \left[F_{[j]}(x) - p^{(j)}\right]\right| : x \in I_n^{(j)}\right\} = O(n^{-3/4} \log n),$$

with probability one. Further, if $f'_{[j]}(x)$ is bounded in some neighbourhood of $\xi_{p(j)}^{(j)}$, then as $n \to \infty$,

$$\text{(1.2.10)} \qquad \left|n^{1/2}\left\{\left[Z_n^{(j)} - \xi_{p(j)}^{(j)}\right] f_{[j]}\left(\xi_{p(j)}^{(j)}\right) + \left[F_{nj}\left(\xi_{p(j)}^{(j)}\right) - p^{(j)}\right]\right\}\right| = O(n^{-1/4} \log n),$$

with probability one, for $j = 1, \ldots, q$.

Theorem 1.2.2. If $F(\underset{\sim}{x})$ is absolutely continuous at $\underset{\sim}{\xi}_{\underset{\sim}{p}}$ with finite, positive and continuous marginal densities in some neighbourhood of $\underset{\sim}{\xi}_{\underset{\sim}{p}}$, and if $\underset{\sim}{\nu}$ is positive definite, then

$$\text{(1.2.11)} \qquad \mathcal{L}(n^{1/2}[\underset{\sim}{Z}_n - \underset{\sim}{\xi}_{\underset{\sim}{p}}]) \to N_q(\underset{\sim}{0}, \underset{\sim}{T}), \text{ as } n \to \infty.$$

In the first part of chapter III, for stationary autoregressive processes Bahadur's asymptotic almost sure representation of the standardized form of a sample quantile is generalized over the entire real line. In connection with this, the following theorem is proved.

Let $g(n) = n^{-1/2} \log n$ and let us suppose that $F_{[j]}(x)$ be absolutely continuous and $\sup_x f_{[j]}(x) = f_{0j} < \infty$. Then we have

Theorem 1.2.3. As $n \to \infty$,

(1.2.12) $$\sup_x \sup_{|a|<g(n)} \{n^{1/2}|F_{nj}(x+a) - F_{nj}(x) - F_{[j]}(x+a) + F_{[j]}(x)|\} = O(n^{-1/4} \log n),$$

with probability one.

With the help of the above theorem and (1.2.10) the asymptotic multinormality, derived in chapter II of a sample quantile (vector) for a stationary autoregressive process, is generalized here for a vector of linear combinations of several quantiles. An application of this result is indicated and under certain conditions the result is also extended for the case of random sample size. Further, the chapter III contains a method of estimation of variances and covariances of sample quantiles and hence the variances and covariances of linear compounds of several quantiles. In connection with the estimation problem the following two theorems are proved.

Theorem 1.2.4. Let $\mu_n^{(j)} = np^{(j)} + o(n^{1/2} \log n)$, and let s_j be the $\mu_n^{(j)}$-th smallest observation among $(X_{1,j}, \ldots, X_{n,j})$. If

(1.2.13) $$\hat{\nu}_{jj'} = \{F_{njj'}(s_j, s_{j'}) - p^{(j)}p^{(j')}\} + \sum_{h=1}^{m_n} \{F_{njj'h}(s_j, s_{j'}) - p^{(j)}p^{(j')}\} + \sum_{h=1}^{m_n} \{F_{nj'jh}(s_j, s_{j'}) - p^{(j)}p^{(j')}\}$$

where

(1.2.14) $$F_{njj'}(u,v) = n^{-1} \sum_{i=1}^{n} c(u - X_{i,j}, v - X_{i,j'})$$

(1.2.15) $$F_{njj'h}(u,v) = n^{-1} \sum_{i=1}^{n} c(u - X_{i,j}, v - X_{i+h,j'})$$

$$(1.2.16) \qquad c(u,v) = \begin{cases} 1 & \text{for } u \geq 0,\ v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

then

$$(1.2.17) \qquad \hat{\nu}_{jj'} \xrightarrow{p} \nu_{jj'}, \quad j, j' = 1,\ldots,q.$$

Theorem 1.2.5. *Let* $\mu_{n,1}^{(j)} = np^{(j)} - c_{1n}(n^{1/2} \log n)$, $\mu_{n,2}^{(j)} = np^{(j)} + c_{2n}(n^{1/2} \log n)$ *where* c_{1n} *and* c_{2n} *are positive constants such that* $(\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)}) \to \infty$ *as* $n \to \infty$, *but* c_{1n} *and* c_{2n} *both converge to zero as* $n \to \infty$. *Also let* $s_{j,i}$ *be the* $\mu_{n,i}^{(j)}$*-th smallest observation among* $(X_{1,j},\ldots,X_{n,j})$. *Then for each* $j = 1,\ldots,q$, *as* $n \to \infty$,

$$(1.2.18) \qquad f_{[j]}\left(\xi_{p^{(j)}}^{(j)}\right) = (\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)})/n(s_{j,2} - s_{j,1}) \to f_{[j]}\left(\xi_{p^{(j)}}^{(j)}\right),$$

with probability one.

In chapter IV first, the asymptotic distribution of Wilcoxon signed rank statistic for stationary autoregressive processes is deduced and then the theory is generalized for a class of U-statistics. Let $\phi(\underset{\sim}{X}_{\alpha_1},\ldots,\underset{\sim}{X}_{\alpha_r})$ be a statistic symmetric in the arguments $\underset{\sim}{X}_{\alpha_1},\ldots,\underset{\sim}{X}_{\alpha_r}$; $\alpha_1 < \ldots < \alpha_r$, and

$$(1.2.19) \qquad U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) = \binom{n}{r}^{-1} \sum_S \phi(\underset{\sim}{X}_{\alpha_1},\ldots,\underset{\sim}{X}_{\alpha_r})$$

where the summation S extends over all $1 \leq \alpha_1 < \ldots < \alpha_r \leq n$

Let us define

(1.2.20) $g(\tilde{F}_n) = E\{\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r}) | \alpha_{i+1}-\alpha_i > m_n \text{ for } i=1,\ldots,r-1\}$

where $\underset{\sim}{Y}_{n,i}$ and m_n are defined in (2.3.5) and (2.3.6) respectively. Further, let $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_r$ be r.i.i.d.r.v. each having the d.f. F and $g(F) = E\{\phi(\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_r)\}$ then $\lim_{n\to\infty} g(\tilde{F}_n) = g(F)$. Let us put

(1.2.21) $$\phi_a^{(n)}(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_a}) = E\{\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_a},\underset{\sim}{Y}_{n,\alpha_{a+1}},\ldots,\underset{\sim}{Y}_{n,\alpha_r})\} - g(F)$$

for $a = 0,1,\ldots,r$, where $\alpha_{i+1} - \alpha_i > m_n$ for all $i=1,2,\ldots,r-1$, and let

(1.2.22) $$\zeta_{a\cdot 0}^{(n)} = E\{\phi_a^{(n)^2}(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_a})\}$$

(1.2.23) $$\zeta_{a\cdot(h_1,\ldots,h_a)}^{(n)} = E\{\phi_a^{(n)}(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_a})\phi_a^{(n)}(\underset{\sim}{Y}_{n,\beta_1},\ldots,\underset{\sim}{Y}_{n,\beta_a})$$

for $a = 0,1,\ldots,r$; where $|\alpha_1-\beta_i| = h_i$; $0 < \alpha_i \leq m_n$ for $i=1,\ldots,a$, but $|\alpha_i-\alpha_t| > m_n$, $|\beta_i-\beta_t| > m_n$ and $|\alpha_i-\beta_t| > m_n$ for all $i\neq t=1,\ldots,n$. Finally, let

(1.2.24) $$\zeta_1 = \lim_{n\to\infty} (\zeta_{1\cdot 0}^{(n)} + 2\sum_{h=1}^{\infty} \zeta_{1\cdot h}^{(n)})$$

Then in connection with U-statistics the following theorem is proved:

Theorem 1.2.6. *Under the assumptions (A) and (B), stated in section* 4.3, $n^{1/2}\{U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - g(F)\}$ *is asymptotically normally distributed with mean 0 and variance* $r^2\zeta_1$.

In the end of the chapter, the asymptotic distribution of the median of the mid-range estimate (vector) for a stationary autoregressive process is obtained as an application of the earlier results of the chapter.

In the final chapter several rival estimators (median, 27% mid-range, median of the mid-ranges and mean) of location parameter for stationary autoregressive processes are considered and their relative performances are compared by their asymptotic relative efficiencies. Some A.R.E. values are tabulated for univariate Gaussian autoregressive processes.

CHAPTER II

BAHADUR-REPRESENTATION OF SAMPLE QUANTILES IN SOME STATIONARY MULTIVARIATE AUTOREGRESSIVE PROCESSES.

2.1. Introduction

In this chapter, we are concerned with the limiting properties of sample quantiles in a general class of stationary multivariate autoregressive processes where there is dependence among the successive observations. The standard technique of deriving the asymptotic distribution of a sample quantile for i.i.d.r.v. [cf. Cramér (1946, pp. 367-369)] usually encounters considerable difficulties in the multivariate case or in the case of dependent random variables. An alternative simple approach is as follows. Let $X_{n,r}$ be the r-th order statistic of a sample $X_1, \ldots, X_n$ of size n from a distribution F(x). We let $r = np + o(\sqrt{n})$ $(0 < p < 1)$ and denote by ξ_p the p-quantile of F. Suppose $F'(\xi_p) = f(\xi_p)$ exists and is positive. Then, with c(u) defined in (1.1.14), we have for every fixed u,

$$(2.1.1) \quad P[X_{n,r} \leq \xi_p + n^{-1/2}u]$$

$$= P[\text{at least } r \text{ of } X_1, \ldots, X_n \leq \xi_p + n^{-1/2}u]$$

$$= P\left[n^{-1/2} \sum_{i=1}^{n} \{c(\xi_p + n^{-1/2}u - X_i) - F(\xi_p + n^{-1/2}u)\} \geq \sqrt{n}\, \{\frac{r}{n} - F(\xi_p + n^{-1/2}u)\}\right]$$

$$= P\left[n^{-1/2} \sum_{i=1}^{n} \frac{\{c(\xi_p + n^{-1/2}u - X_i) - F(\xi_p + n^{-1/2}u)\}}{\{F(\xi_p + n^{-1/2}u)(1 - F(\xi_p + n^{-1/2}u))\}^{1/2}} \geq - \frac{uf(\xi_p)}{\{p(1 - p)\}^{1/2}} + o(1)\right]$$

and hence, the asymptotic normality can be obtained by using an appropriate form of a central limit theorem applicable to the double sequence of random variables $\{c(\xi_p + n^{-1/2}u - X_i),\ i = 1,\dots,n\}$, $n \geq 1$. This approach appears to be applicable for multivariate as well as dependent random variables. However, in refined statistical analysis, we are not merely satisfied with this weak convergence of sample quantiles. For i.i.d.r.v.'s, Bahadur (1966) has considered an elegant asymptotic almost sure representation of a sample quantile, which is further extended to the case of m-dependent processes by Sen (1968b). An important by product of this a.s. representation is an alternative proof of the asymptotic normality of the standardized form of a sample quantile; the cases of several quantiles or of a linear function of quantiles also follow more easily from this representation.

We derive an analogous representation for an m_n-dependent process, where we let $m_n \sim K \log n$, K being a positive number. Also, we show that insofar as the asymptotic behaviour of a sample quantile is concerned, a stationary autoregressive process may as well be replaced by a suitable m_n-dependent process, with $m_n \sim K \log n$. Combining the above, our conclusions about the limiting behaviour of a sample quantile, including its asymptotic normality, follow readily.

2.2. Preliminary Notations and Assumptions

Consider a sequence $\{\underset{\sim}{X}_t,\ t = 0, \pm 1, \pm 2, \ldots\}$ of stochastic $q(\geq 1)$-vectors from the multivariate autoregressive process defined in (1.1.1). Using the forward operator E as

$$E^r X_{t,j} = X_{t+r,j}, \tag{2.2.1}$$

we may rewrite the set of equations in (1.1.1) as

$$\sum_{\ell=1}^{q} \sum_{r=0}^{k} a_{j\ell}^{(r)} E^{k-r} X_{t-k,\ell} = \varepsilon_{t,j},\ j = 1, \ldots, q;\ \text{for all } t. \tag{2.2.2}$$

Again, if we write (when $a_{j\ell}^{(0)} \neq 0$, otherwise take a lower degree polynomial)

$$v_{j\ell} = a_{j\ell}^{(0)}\left(E - e_{1\ell}^{(j)}\right)\cdots\left(E - e_{k\ell}^{(j)}\right),\quad j,\ \ell = 1, \ldots, q, \tag{2.2.3}$$

where $e_{1\ell}^{(j)}, \ldots, e_{k\ell}^{(j)}$ are the roots of

$$\sum_{r=0}^{k} a_{j\ell}^{(r)} E^{k-r} = 0,\quad j,\ \ell = 1, \ldots, q, \tag{2.2.4}$$

then (1.1.1) may be rewritten as

$$\underset{\sim}{V}\underset{\sim}{X}_{t-k} = \underset{\sim}{\varepsilon}_t \text{ where } \underset{\sim}{V} = ((v_{j\ell}))_{j,\ell= ,\ldots,q}. \tag{2.2.5}$$

Our first assumption is that all the $|e_{r\ell}^{(j)}|$, $r = 1, \ldots, k$; $j, \ell = 1, \ldots, q$ lie in the semi-closed interval $[0,1)$, i.e.,

$$0 \leq \max_{r=1,\ldots,k}\ \max_{j,\ell=1,\ldots,q} |e_{r\ell}^{(j)}| = e^* < 1. \tag{2.2.6}$$

Under (2.2.6) and proceeding as in Whittle (1953), we have

$$\underset{\sim}{X}_t = \sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\varepsilon}_{t-r},\ \text{for all } t, \tag{2.2.7}$$

where $\underset{\sim}{B}_r = \left(\left(b_{j\ell}^{(r)}\right)\right)_{j,\ell=1,\ldots,q}$, r = 0, 1,... are all q × q matrices of constant coefficients, where $b_{r\ell}^{(j)}$ is a polynomial of degree r in the roots $e_{s\ell}^{(j)}$, s = 1,..., k; j, ℓ = 1,..., q. If the errors $\underset{\sim}{\varepsilon}_t$ have finite second moments, it can be shown (along the lines of Doob (1953, p. 503)) that

$$(2.2.8) \qquad |b_{j\ell}^{(r)}| \le Cr^g(e^*)^r \text{ , for all } j, \ell, r,$$

where C and g are two positive constants. We shall see later on that we do not require the existence of the second moments of $\underset{\sim}{\varepsilon}_t$. Hence, we make (2.2.8) a part of our basic assumptions. For clarification of ideas, we touch briefly the univariate case where q = 1. Here g = k - 1. Thus, when k = 2 and the two roots e_1 and e_2 are equal (to e), $b^{(r)} = (r+1)e^r$, while if $e_1 \ne e_2$, $|b^{(r)}| \le (r+1)(e^*)^r$. In general, $|b^{(r)}| \le \binom{k+r-1}{r}(e^*)^r \le Cr^{k-1}(e^*)^r$. Note that (2.2.8) insures that for every δ > 0,

$$(2.2.9) \qquad \sum_{r=0}^{\infty} |b_{j\ell}^{(r)}|^{\delta} < \infty \text{ for all } j, \ell = 1,\ldots, q.$$

Our third assumption is that

$$(2.2.10) \qquad E||\underset{\sim}{\varepsilon}_t||^{\delta} < \infty, \text{ for some } \delta > 0 \text{ (need not be } \ge 1).$$

In fact, when (2.2.10) holds for δ = 2, and we denote by

$$(2.2.11) \qquad \underset{\sim}{\Gamma}_h = E[\underset{\sim}{X}_t \underset{\sim}{X}_{t-h}'], \; h = 0, 1, \ldots,$$

then the matrices $\underset{\sim}{A}_0$, $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_k$ in (1.1.1) satisfy the generalized Yule-Walker equations:

$$(2.2.12) \qquad \sum_{r=0}^{k} \underset{\sim}{A}_r \underset{\sim}{\Gamma}_{r+h} = \underset{\sim}{0} \quad \text{for } h = 0, 1, \ldots, k,$$

and the $\underset{\sim}{B}_r$ can then be expressed in terms of the elements $\underset{\sim}{A}_0, \ldots, \underset{\sim}{A}_k$.

The joint c.d.f. of $\underset{\sim}{X}_t$ is denoted by $F(\underset{\sim}{x})$, $\underset{\sim}{x} \in R^q$, the q-dimensional real space and the marginal c.d.f. of $X_{t,j}$, the j-th component of $\underset{\sim}{X}_t$, is denoted by $F_{[j]}(x)$, $j = 1, \ldots, q$. By the assumed (absolute) continuity of G, F and all the $F_{[j]}$ are also (absolutely) continuous. Thus, there exists a unique vector $\underset{\sim}{\xi}_{\underset{\sim}{p}} = \left(\xi^{(1)}_{p(1)}, \ldots, \xi^{(q)}_{p(q)}\right)'$ (where $0 < p^{(j)} < 1$ for all $j = 1, \ldots, q$), such that for every $\varepsilon > 0$,

$$(2.2.13) \qquad F_{[j]}\left(\xi^{(j)}_{p(j)} - \varepsilon\right) < p^{(j)} = F_{[j]}\left(\xi^{(j)}_{p(j)}\right) < F_{[j]}\left(\xi^{(j)}_{p(j)} + \varepsilon\right), \quad j = 1, \ldots, q.$$

$\underset{\sim}{\xi}_{\underset{\sim}{p}}$ is termed the vector of population (co-ordinate wise) $p^{(j)}$-quantiles.

For a finite time interval $T_n = \{t: 1 \le t \le n\}$, let $\{\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n\}$ correspond to the chance variables associated with the sample of size n from the process in (1.1.1). The ordered random variables on the j-th variate are denoted by $X^{(j)}_{n,1} < \ldots < X^{(j)}_{n,n}$; by the continuity of F (and hence of $F_{[j]}$), ties among the observations can be neglected, with probability one, for all $j = 1, \ldots, q$. The sample $\underset{\sim}{p}$-quantile $\underset{\sim}{Z}_n = \left(Z^{(1)}_n, \ldots, Z^{(q)}_n\right)'$ is then defined by

$$(2.2.14) \qquad Z^{(j)}_n = X^{(j)}_{n,\mu_j}, \quad j = 1, \ldots, q, \text{ where } \mu_j = [np^{(j)}] + 1,$$

and [s] is the largest integer contained in s. We are primarily interested in the asymptotic behaviour of $n^{1/2}(\underset{\sim}{Z}_n - \underset{\sim}{\xi}_{\underset{\sim}{p}})$. For this, we introduce the following notations.

For the j-th variate, the empirical c.d.f. is defined by

$$\text{(2.2.15)} \qquad F_{nj}(x) = n^{-1} \sum_{i=1}^{n} c(x - X_{i,j}), \ -\infty < x < \infty, \ j = 1,\ldots, q.$$

We denote the dispersion matrix of $n^{1/2}\left[F_{n1}\left(\xi^{(1)}_{p(1)}\right) - p^{(1)}, \ldots, F_{nq}\left(\xi^{(q)}_{p(q)}\right) - p^{(q)}\right]$ by $\underset{\sim}{\nu}_n = ((\nu_{njj'}))_{j,j'=1,\ldots,q}$. Later, we shall show that $\lim_{n\to\infty} \underset{\sim}{\nu}_n = \underset{\sim}{\nu}$ exists and is finite. Also, we denote by

$$\text{(2.2.16)} \qquad f_{[j]}(x) = (d/dx)F_{[j]}(x), \quad f'_{[j]}(x) = (d/dx)f_{[j]}(x),$$

and assume that both $f_{[j]}(x)$ and $f'_{[j]}(x)$ exist for every $j = 1,\ldots, q$. Other notations will be introduced as and when necessary.

2.3. Almost Sure (Bahadur) representation of $\underset{\sim}{Z}_n$

As in Bahadur (1966), we let

$$\text{(2.3.1)} \quad I_n^{(j)} = \left\{x : \xi^{(j)}_{p(j)} - a_n \le x \le \xi^{(j)}_{p(j)} + a_n\right\}, \quad j = 1,\ldots, q,$$

where

$$\text{(2.3.2)} \qquad a_n \sim n^{-1/2} \log n \quad \text{as } n \to \infty.$$

Then, the main theorem of this section is the following.

Theorem 2.3.1. If in the neighbourhood of $\xi^{(j)}_{p(j)}$, $f_{[j]}(x)$ is finite continuous and positive, then under (2.2.6), (2.2.8) and (2.2.10), as $n \to \infty$,

$$\text{(2.3.3)} \quad \sup\left\{\left|\left[F_{nj}(x) - F_{nj}\left(\xi^{(j)}_{p(j)}\right)\right] - \left[F_{[j]}(x) - p^{(j)}\right]\right| : x \in I_n^{(j)}\right\}$$
$$= O(n^{-3/4} \log n),$$

with probability one. Further, if $f'_{[j]}(x)$ *is bounded in some neighbourhood of* $\xi^{(j)}_{p(j)}$, *then as* $n \to \infty$,

$$(2.3.4)\quad \left| n^{1/2}\left\{\left[Z^{(j)}_n - \xi^{(j)}_{p(j)}\right] f_{[j]}\left(\xi^{(j)}_{p(j)}\right) + \left[F_{nj}\left(\xi^{(j)}_{p(j)}\right) - p^{(j)}\right]\right\}\right| = O(n^{-1/4}\log n),$$

with probability one, for $j = 1,\ldots, q$.

For proving the theorem, we require certain basic results, which are considered first. Let us define

$$(2.3.5)\quad \underset{\sim}{Y}_{n,i} = \sum_{r=0}^{m_n} \underset{\sim}{B}_r \underset{\sim}{\varepsilon}_{i-r},\ \underset{\sim}{R}_{n,i} = \sum_{r=m_n+1}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\varepsilon}_{i-r},\quad i = 1, 2,\ldots, n,$$

where the $\underset{\sim}{B}_r$ are defined by (2.2.7), and where m_n is a positive integer such that

$$(2.3.6)\qquad m_n \sim K \log n \text{ as } n \to \infty;\ \delta\rho^* K = c > 8,$$

K is a positive constant, δ is defined by (2.2.10) and we write

$$(2.3.7)\qquad e^* = e^{-\rho^*},$$

where by (2.2.6), $\rho^* > 0$. Let then

$$(2.3.8)\quad F^*_{nj}(x) = n^{-1}\sum_{i=1}^{n} c\left(x - Y^{(j)}_{n,i}\right),\ -\infty < x < \infty,\quad j = 1,\ldots, q,$$

be the empirical c.d.f.'s of the q components of the $\underset{\sim}{Y}_{n,i}$. The true c.d.f. of $Y^{(j)}_{n,i}$ is denoted by

$$(2.3.9)\quad \tilde{F}_{n[j]}(x) = P\{Y^{(j)}_{n,i} \le x\},\ -\infty < x < \infty,\quad j = 1,\ldots, q.$$

The joint c.d.f. of $(X_{i,j}, X_{i+h,j'})$ is denoted by

(2.3.10) $$F_{[j,j']h}(x, y) = P\{X_{i,j} \le x, X_{i+h,j'} \le y\},$$

and the joint c.d.f. of $\left(Y_{n,i}^{(j)}, Y_{n,i+h}^{(j')}\right)$ is denoted by

(2.3.11) $$\tilde{F}_{n[j,j']h}(x, y) = P\{Y_{n,i}^{(j)} \le x, Y_{n,i+h}^{(j')} \le y\},$$

$j, j' = 1,\ldots, q$; $h = 0, 1,\ldots$. We assume that the joint density

(2.3.12) $$f_{[j,j']h}(x, y) = (\delta^2/\delta x\delta y)F_{[j,j']h}(x, y)$$

also exists for every $j, j' = 1,\ldots, q$; $h = 0, 1,\ldots$ and is finite.

Lemma 2.3.1. Under (2.2.6), (2.2.8) and (2.2.10), for c_1 $(0 < c_1 < c/_2)$, there exist two positive numbers c_2 and c_3 such that

(2.3.13) $$P\{|R_{n,i}^{(j)}| \ge c_2 n^{-c_1/\delta}\} \le c_3 n^{-(c-c_1)}(\log n)^{g\delta}, \quad j = 1,\ldots, q,$$

where $R_{n,i}^{(j)}$ is the j-th component of $\underset{\sim}{R}_{n,i}$ and c is defined in (2.3.6). [Throughout this chapter, we shall exclusively deal with $\delta: 0 < \delta \le 1$, with a remark that for $\delta > 1$, the proofs become comparatively simpler and we can still work with some $\delta' \le 1$.]

Proof. By (2.2.8), (2.2.9) and (2.3.5),

(2.3.14) $$E|R_{n,i}^{(j)}|^{\delta} \le \sum_{r>m_n} \sum_{s=1}^{q} |b_{js}^{(r)}|^{\delta} E|\varepsilon_{i-r,s}|^{\delta}$$

$$= \sum_{s=1}^{q} E|\varepsilon_{i,s}|^{\delta}\{\sum_{r>m_n} |b_{js}^{(r)}|^{\delta}\}$$

$$\le C^{\delta} \sum_{s=1}^{q} E|\varepsilon_{i,s}|^{\delta}\{\sum_{r>m_n} r^{g\delta}(e^*)^{r\delta}\}$$

$$= O\left([m_n]^{\delta g}(e^*)^{\delta m_n}\right)$$

$$= O\left((\log n)^{\delta g}\, e^{-c \log n}\right),$$

as $m_n \sim K \log n$, $e^* = e^{-\rho^*}$ and $c = \delta\rho^* K$. Hence, by the Markov inequality,

$$P\{|R_{n,i}^{(j)}| \geq c_2 n^{-c_1/\delta}\} = P\{|R_{n,i}^{(j)}|^\delta \geq c_2^\delta\, n^{-c_1}\}$$

$$\leq c_2^{-\delta}\, n^{c_1}\, E|R_{n,i}^{(j)}|^\delta$$

$$= O\left(n^{-(c-c_1)}(\log n)^{\delta g}\right). \qquad \text{Q.E.D.}$$

Lemma 2.3.2. Under (2.2.6), (2.2.8) and (2.2.10), for every $c = K\rho^*\delta\,(>8)$,

(2.3.15) $$\sup\{|F_{[j]}(x) - \tilde{F}_{n[j]}(x)| : x \in I_n^{(j)}\} = O(n^{-d}),\ j = 1,\ldots,q,$$

(2.3.16) $$\sup_{h=1,\ldots,m_n}\{\sup[|F_{[j,j']h}(x,y) - \tilde{\tilde{F}}_{n[j,j']h}(x,y)| : x \in I_n^{(j)}, y \in I_n^{(j')}]\}$$

$$= O(n^{-d}) \text{ for all } j,j' = 1,\ldots,q,$$

where $0 < d < c/2$ and d can be made greater than 1 by proper choice of K.

Proof. By definition in (2.2.7) and (2.3.5), for all $x \in I_n^{(j)}$

(2.3.17) $$\tilde{F}_{n[j]}(x) = P\{Y_{n,i}^{(j)} \leq x\} = P\{X_{i,j} \leq x + R_{n,i}^{(j)}\}$$

$$\leq P\{X_{i,j} \leq x + R_{n,i}^{(j)},\ R_{n,i}^{(j)} < c_2 n^{-c_1/\delta}\}$$

$$+ P\{R_{n,i}^{(j)} \geq c_2 n^{-c_1/\delta}\}$$

$$\leq P\{X_{i,j} \leq x + c_2 n^{-c_1/\delta}\} + P\{R_{n,i}^{(j)} \geq c_2 n^{-c_1/\delta}\}$$

$$= F_{[j]}(x + c_2 n^{-c_1/\delta}) + 0(n^{-(c-c_1)}(\log n)^{g\delta}), \text{ by Lemma 2.3.1.}$$

$$= F_{[j]}(x) + 0(n^{-c_1/\delta}) + 0(n^{-c_1}) \text{ as } c_1 < c/2,$$

and as $f_{[j]}(x)$ is bounded in $I_n^{(j)}$. By assumption $\delta \leq 1$, and hence,

$$\text{(2.3.18)} \qquad \tilde{F}_{n[j]},x. \leq F_{[j]}(x) + 0(n^{-c_1}), \; \forall\, x\varepsilon I_n^{(j)} .$$

In a similar manner, it follows that as $n \to \infty$,

$$\text{(2.3.19)} \quad \tilde{F}_{n[j]}(x) \geq F_{[j]}(x) - 0(n^{-c_1}), \; \forall\, x \in I_n^{(j)}$$

Thus, (2.3.15) readily follows from (2.3.18) and (2.3.19), by letting $d \leq c_1$.

Similar to (2.3.17), for all $x \in I_n^{(j)}$, $y \in I_n^{(j')}$, we have for large n,

$$\begin{aligned}
\tilde{F}_{n[j,j']h}(x,y) &= P\{Y_{n,i}^{(j)} \leq x,\; Y_{n,i}^{(j')} \leq y\} \\
&= P\{X_{i,j} \leq x + R_{n,i}^{(j)},\; X_{i+h,j'} \leq y + R_{n,i+h}^{(j')}\} \\
&\leq P\{X_{i,j} \leq x + R_{n,i}^{(j)},\; X_{i+h,j} \leq y + R_{n,i+h}^{(j')},\; R_{n,i}^{(j)} \\
&\qquad < c_2 n^{-c_1/\delta},\; R_{n,i+h}^{(j')} < c_2 n^{-c_1/\delta}\} \\
&\quad + P\{R_{n,i}^{(j)} \geq c_2 n^{-c_1/\delta},\; R_{n,i+h}^{(j')} \geq c_2 n^{-c_1/\delta}\} \\
&\leq P\{X_{i,j} \leq x + c_2 n^{-c_1/\delta}, X_{i+h,j'} \leq y + c_2 n^{-c_1/\delta}\} \\
&\quad + P\{R_{n,i}^{(j)} \geq c_2 n^{-c_1/\delta}\} + P\{R_{n,i+h}^{(j')} \geq c_2 n^{-c_1/\delta}\} \\
&= F_{[j,j']h}(x + c_2 n^{-c_1/\delta}, y + c_2 n^{-c_1/\delta}) \\
&\qquad + 0(n^{-(c-c_1)}(\log n)^{g\delta}),
\end{aligned}$$

by lemma (2.3.1)

$$= F_{[j,j']h}(x,y) + O(n^{-c_1/\delta}) + O(n^{-c_1}), \text{ as } c_1 < c/2$$

and as $f_{[j,j']h}(x,y)$ is bounded for $x \in I_n^{(j)}$, $y \in I_n^{(j')}$. Again $\delta \leq 1$ implies

(2.3.20) $\tilde{F}_{n[j,j']h}(x,y) \leq F_{[j,j']h}(x,y) + O(n^{-c_1}), \forall\, x \in I_n^{(j)}, y \in I_n^{(j')}$.

Similarly, it also follows that for large n,

$$\tilde{F}_{n[j,j']h}(x,y) \geq F_{[j,j']h}(x,y) - O(n^{-c_1}), \forall\, x \in I_n^{(j)}, y \in I_n^{(j')} .$$

Hence the above result and (2.3.20) imply (2.3.16) where $d \leq c_1$.

Q.E.D.

Let us now define

(2.3.21) $$\underset{\sim}{\nu}_n^*(\underset{\sim}{\xi}_{\underset{\sim}{p}}) = ((\nu^*_{njj'}(\underset{\sim}{\xi}_{\underset{\sim}{p}})))_{j,j'=1,\ldots,q} ,$$

where

(2.3.22) $$\nu^*_{njj'}(\underset{\sim}{\xi}_{\underset{\sim}{p}}) = n \operatorname{covar}[F^*_{nj}(\xi^{(j)}_{p(j)}), F^*_{nj'}(\xi^{(j')}_{p(j')})], \; j,j'=1,\ldots,q.$$

Further, if $\underset{\sim}{\nu}_n$ is defined as in section 2.2, then since F_{nj} involves bounded values random variables and the $X_{i,j}$ satisfy the strong mixing or ϕ-mixing condition (viz, pp. 166-168 of Billingsley, 1968) it is easy to show that the conditions of his lemma 3 on page 172 hold, which ensure the existence of $\lim_{n\to\infty} \underset{\sim}{\nu}_n$. Then let us define $\lim_{n\to\infty} \underset{\sim}{\nu}_n = \underset{\sim}{\nu}$, where

(2.3.23) $$\underset{\sim}{\nu} = ((\nu_{jj'})); \; \nu_{jj'} = \left\{F_{[j,j']}\left(\xi^{(j)}_{p(j)}, \xi^{(j')}_{p(j')}\right) - p^{(j)}p^{(j')}\right\}$$

$$+ \sum_{h=1}^{\infty} \left\{ F_{[j,j']h}\left(\xi^{(j)}_{p(j)}, \xi^{(j')}_{p(j')}\right)\right.$$

$$+ F_{[j',j]h}\left(\xi^{(j)}_{p(j)}, \xi^{(j')}_{p(j')}\right) - 2p^{(j)}p^{(j')}\Big\},$$

$$j,j'=1,\ldots,q,$$

and for $j = j'$, $h = 0$, $F_{[j,j]0}\left(\xi^{(j)}_{p(j)}, \xi^{(j)}_{p(j)}\right)$ is equal to $p^{(j)}$. Then, we have the following.

Lemma 2.3.3. Under (2.2.6), (2.2.8) and (2.2.10), $\underset{\sim}{\nu}$ is finite and

$$\lim_{n \to \infty} \underset{\sim}{\nu}_n = \lim_{n \to \infty} \underset{\sim}{\nu}^*_n(\underset{\sim}{\xi}_{\underset{\sim}{p}}) = \underset{\sim}{\nu} \tag{2.3.24}$$

where $\underset{\sim}{\nu}$ and $\nu_{jj'}$ are given by (2.3.23).

Proof. $\underset{\sim}{\nu}$ being a matrix, to prove its finiteness, we have to show that $\nu_{jj'}$ is finite for every $j,j' = 1,\ldots,q$. We first show that ν_{jj} is finite for every $j = 1,\ldots,q$. Then, this will imply finiteness of $\nu_{jj'}$ for every $j,j' = 1,\ldots,q$.

We note that

$$\nu_{jj} = p^{(j)}(1 - p^{(j)}) + 2\sum_{h=1}^{\infty}\left\{F_{[j,j]h}\left(\xi^{(j)}_{p(j)}, \xi^{(j)}_{p(j)}\right) - p^{(j)^2}\right\}$$

and its first term is bounded by 1/4. To find a bound for the second term, we write

$$F_{[j,j]h}\left(\xi^{(j)}_{p(j)}, \xi^{(j)}_{p(j)}\right) = P\{X_{i,j} \le \xi^{(j)}_{p(j)},\ X_{i+h,j} \le \xi^{(j)}_{p(j)}\} \tag{2.3.25}$$

$$= P\Big\{\sum_{r=0}^{\infty}\sum_{s=1}^{q} b^{(r)}_{js}\varepsilon_{i-r,s} \le \xi^{(j)}_{p(j)},\ \sum_{r=0}^{h-1}\sum_{s=1}^{q} b^{(r)}_{js}\varepsilon_{i+h-r,s}$$

$$+ \sum_{r=h}^{\infty}\sum_{s=1}^{q} b^{(r)}_{js}\varepsilon_{i+h-r,s} \le \xi^{(j)}_{p(j)}\Big\}$$

$$\le P\{\sum_{r=0}^{\infty}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i-r,s} \le \xi_{p(j)}^{(j)},$$

$$\sum_{r=0}^{h-1}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i+h-r,s} \le \xi_{p(j)}^{(j)} + C\sum_{r=h}^{\infty}\sum_{s=1}^{q} r^{g}(e^*)^{r}|\varepsilon_{i+h-r,s}|\},$$

by (2.2.8) .

Let us write

$$U = \sum_{r=0}^{\infty}\sum_{s=1}^{q} (h+r)^{g}(e^*)^{r}|\varepsilon_{i-r,s}| \tag{2.3.26}$$

and proceeding as in (2.3.14), we have

$$E|U|^{\delta} \le \sum_{r=0}^{\infty}\sum_{s=1}^{q} E|\varepsilon_{i-r,s}|^{\delta}(h+r)^{g\delta}(e^*)^{r\delta} < \infty ,$$

so that by the Markov inequality,

$$P\{U > (e^*)^{-h/2}\} \le E|U|^{\delta}(e^*)^{\delta h/2} = O(e^*)^{\delta h/2} \tag{2.3.27}$$

By (2.3.25), (2.3.26), and (2.3.27), we get

$$F_{[j,j]h}\left(\xi_{p(j)}^{(j)},\xi_{p(j)}^{(j)}\right) \tag{2.3.28}$$

$$\le P\left\{\sum_{r=0}^{\infty}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i-r,s} \le \xi_{p(j)}^{(j)}, \sum_{r=0}^{h-1}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i+h-r,s} \le \xi_{p(j)}^{(j)} + C(e^*)^{h}\, U\right\}$$

$$\le P\left\{\sum_{r=0}^{\infty}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i-r,s} \le \xi_{p(j)}^{(j)}, \sum_{r=0}^{h-1}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i+h-r,s} \le \xi_{p(j)}^{(j)} + C(e^*)^{h}U,\right.$$

$$\left. U \le (e^*)^{-h/2}\right\} + P\left\{U > (e^*)^{-h/2}\right\}$$

$$\leq P\{\sum_{r=0}^{\infty}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i-r,s} \leq \xi_{p(j)}^{(j)}\}P\{\sum_{r=0}^{h-1}\sum_{s=1}^{q} b_{js}^{(r)}\varepsilon_{i+h-r,s} \leq \xi_{p(j)}^{(j)} + C(e^*)^{h/2}\}$$

$$+ P\{U > (e^*)^{-h/2}\}$$

$$\leq P\{X_{i,j} \leq \xi_{p(j)}^{(j)}\}P\{X_{i+h,j} \leq \xi_{p(j)}^{(j)} + 2\,C(e^*)^{h/2}\} + 2P\{U > (e^*)^{-h/2}\}$$

$$= F_{[j]}\left(\xi_{p(j)}^{(j)}\right) F_{[j]}\left(\xi_{p(j)}^{(j)} + O(e^*)^{h/2}\right) + O(e^*)^{\delta h/2}$$

and using the Taylor theorem, we have from (2.3.28),

$$(2.3.29) \qquad F_{[j,j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j)}^{(j)}\right) \leq F_{[j]}^2\left(\xi_{p(j)}^{(j)}\right) + O(e^*)^{\delta h/2}$$

Proceeding in the similar way we can easily show

$$(2.3.30) \qquad F_{[j,j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j)}^{(j)}\right) \geq F_{[j]}^2\left(\xi_{p(j)}^{(j)}\right) - O(e^*)^{\delta h/2}$$

Since $F_{[j]}\left(\xi_{p(j)}^{(j)}\right) = p^{(j)}$, from (2.3.29) and (2.3.30), we get

$$(2.3.31) \qquad \left|F_{[j,j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j)}^{(j)}\right) - p^{(j)2}\right| = O(e^*)^{\delta h/2}$$

and by (2.3.31), the second term of ν_{jj} is also bounded.

By use of the Schwarz inequality and the finiteness of ν_{jj}, we get $|\nu_{jj'}| < \infty, \forall\, j,j' = 1,\ldots,q$. This completes the proof of finiteness of $\underset{\sim}{\nu}$.

Since $\lim_{n\to\infty} \underset{\sim}{\nu}_n = \lim_{n\to\infty} \underset{\sim}{\nu}_n^*(\underset{\sim}{\xi}_p)$, for proof of (2.3.24), only we have to show that $\lim_{n\to\infty} \nu_{njj'}^*(\underset{\sim}{\xi}_p) = \nu_{jj'}$, for every $j,j' = 1,\ldots,q$. First $\nu_{jj'}$ being finite, we see that, we must have

$$(2.3.32) \qquad \sum_{h=1}^{\infty}\left|F_{[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - p^{(j)}p^{(j')}\right| < \infty$$

so that

$$\text{(2.3.33)} \quad \sum_{m_n+1}^{\infty} \left| F_{[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - p^{(j)} p^{(j')} \right| \to 0 \text{ as } n \to \infty.$$

Next using the definition (2.3.8), we obtain for every $j, j' = 1, \ldots, q$,

$$\text{(2.3.34)} \qquad E\left[F_{nj}^{*}\left(\xi_{p(j)}^{(j)}\right)\right] = \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)$$

and

$$\text{(2.3.35)} \quad \nu_{njj'}^{*}(\underset{\sim}{\xi}_{p}) = n \operatorname{covar}\left[F_{nj}^{*}\left(\xi_{p(j)}^{(j)}\right), F_{nj'}^{*}\left(\xi_{p(j')}^{(j')}\right)\right]$$

$$= n^{-1} E\left[\left\{\sum_{i=1}^{n}\left(c\left(\xi_{p(j)}^{(j)} - Y_{n,i}^{(j)}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\right)\right\}\left\{\sum_{i=1}^{n}\left(c\left(\xi_{p(j')}^{(j')} - Y_{n,i}^{(j')}\right) - \tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right)\right\}\right]$$

$$= n^{-1}\left[\sum_{i=1}^{n} E\left\{c\left(\xi_{p(j)}^{(j)} - Y_{n,i}^{(j)}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\right\}\left\{c\left(\xi_{p(j')}^{(j')} - Y_{n,i}^{(j')}\right) - \tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\}\right.$$

$$+ \sum_{i \neq \ell = 1}^{n} E\left\{c\left(\xi_{p(j)}^{(j)} - Y_{n,i}^{(j)}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\right\}$$

$$\left.\left\{c\left(\xi_{p(j')}^{(j')} - Y_{n,\ell}^{(j')}\right) - \tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\}\right]$$

$$= \tilde{F}_{n[j,j']}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)$$

$$+ n^{-1}\sum_{h=1}^{m_n} \sum_{i=1}^{n-h}\left\{\tilde{F}_{n[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\}$$

$$+ n^{-1} \sum_{h=1}^{m_n} \sum_{i=1}^{n-h} \{\tilde{F}_{n[j',j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\}$$

$$= \{\tilde{F}_{n[j,j']}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\}$$

$$+ \sum_{h=1}^{m_n} (1 - \frac{h}{n})\{\tilde{F}_{n[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\}$$

$$+ \sum_{h=1}^{m_n} (1 - \frac{h}{n})\{\tilde{F}_{n[j',j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\}$$

so that we can write

$$(2.3.36) \quad \lim_{n\to\infty} \nu^*_{njj'}(\underset{\sim}{\xi}_p) = \lim_{n\to\infty} J_1 + \lim_{n\to\infty} J_2 - \lim_{n\to\infty} J_3 + \lim_{n\to\infty} J_4 - \lim_{n\to\infty} J_5$$

where

$$J_1 = \tilde{F}_{n[j,j']}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right),$$

$$J_2 = \sum_{h=1}^{m_n}\left\{\tilde{F}_{n[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\},$$

$$J_3 = \sum_{h=1}^{m_n} \frac{h}{n}\left\{\tilde{F}_{n[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\},$$

$$J_4 = \sum_{h=1}^{m_n}\left\{\tilde{F}_{n[j',j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\},$$

$$J_5 = \sum_{h=1}^{m_n} \frac{h}{n}\left\{\tilde{F}_{n[j',j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right\}$$

Since $\left|\tilde{F}_{n[j,j']h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - \tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right)\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right)\right| \le 1$,

$$(2.3.37)\quad |J_3| \le \sum_{h=1}^{m_n} h/n \sim (K\log n)(K\log n+1)/2n \to 0 \text{ as } n \to \infty$$

and similarly $|J_5| \to 0$ as $n \to \infty$. By (2.3.15) and (2.3.16), we have

$$(2.3.38)\quad \lim_{n\to\infty} J_1 = F_{[j,j']}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right) - p^{(j)}p^{(j')} .$$

Also by use of (2.3.33), we get

$$(2.3.39)\quad \left|J_2 - \sum_{h=1}^{\infty}\left\{F_{[j,j']h}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right) - p^{(j)}p^{(j')}\right\}\right|$$

$$\le \sum_{h=1}^{m_n}\left|\tilde F_{n[j,j']h}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right) - F_{[j,j']h}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right)\right|$$

$$+ \sum_{h=1}^{m_n}\left|\tilde F_{n[j]}\left(\xi^{(j)}_{p(j)}\right)\tilde F_{n[j']}\left(\xi^{(j')}_{p(j')}\right) - p^{(j)}p^{(j')}\right|$$

$$+ \sum_{h=m_n+1}^{\infty}\left|F_{[j,j']h}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right) - p^{(j)}p^{(j')}\right|$$

$$\to 0 \text{ as } n \to \infty,$$

since by (2.3.15) and (2.3.16),

$$(2.3.40)\quad \sum_{n=1}^{m_n}\left|\tilde F_{n[j,j']h}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right) - F_{[j,j']h}\left(\xi^{(j)}_{p(j)},\xi^{(j')}_{p(j')}\right)\right|$$

$$= O(n^{-d}\log n) \to 0 \text{ as } n \to \infty,$$

and

$$(2.3.41)\quad \sum_{h=1}^{m_n}\left|\tilde F_{n[j]}\left(\xi^{(j)}_{p(j)}\right)\tilde F_{n[j']}\left(\xi^{(j')}_{p(j')}\right) - p^{(j)}p^{(j')}\right|$$

$$= \sum_{h=1}^{m_n}\left|\tilde F_{n[j']}\left(\xi^{(j')}_{p(j')}\right)\left\{\tilde F_{n[j]}\left(\xi^{(j)}_{p(j)}\right) - p^{(j)}\right\}\right.$$

$$\left. + p^{(j)}\{\tilde F_{n[j']}\left(\xi^{(j')}_{p(j')}\right) - p^{(j')}\}\right|$$

$$\leq \sum_{h=1}^{m_n} \{ |\tilde{F}_{n[j]}\left(\xi_{p(j)}^{(j)}\right) - p^{(j)}| + |\tilde{F}_{n[j']}\left(\xi_{p(j')}^{(j')}\right) - p^{(j')}| \}$$

$$= O(n^{-d} \log n) \to 0 \text{ as } n \to \infty.$$

Similarly

(2.3.42) $$|J_4 - \sum_{h=1}^{\infty} \{F_{[j',j]h}\left(\xi_{p(j)}^{(j)}, \xi_{p(j')}^{(j')}\right) - p^{(j)}p^{(j')}\}|$$

$$\to 0 \text{ as } n \to \infty.$$

The proof of (2.3.24) follows readily from (2.3.36) - (2.3.42).

Lemma 2.3.4. Let $\{Z_i\}$ be a sequence of m_n-dependent binomial random variables, where $EZ_i = p$, $i \geq 1$, and $m_n \sim K \log n$ as $n \to \infty$. Let then $n_j^* = [(n + m_n + 1 - j)/(m_n + 1)]$, $j = 1,\ldots,m_n$, $n^* = \min_{j=1,\ldots,m_n} (n_j^*)$, and

(2.3.43) $$\gamma_n = c_1^*(n^*)^{-1/2}\{(\log n^*)p(1 - p)\}^{1/2},\ c_1^{*2} > 2.$$

Then

(2.3.44) $$\sum_{i=1}^{\infty} P\{|n^{-1} \sum_{i=1}^{n} Z_i - p| > \gamma_n\} < \infty.$$

Proof. Consider the partial sums

$$S_{j,n} = Z_j + Z_{j+(m_n+1)} + \ldots + Z_{j+(n_j^*-1)(m_n+1)},\ j = 1,2,\ldots,m_n+1.$$

Then the lemma follows in the similar way by putting p for $p_{j,n}^*$ in Lemma 2.1 of Sen (1968b). The only difference is that instead of (2.22) of Sen (1968b), we get in this case

(2.3.45) $$n_j^* \gamma_n^2/2p(1 - p) > (1 + \delta)\log n^*; \delta > 0,\ (m_n + 1)n^* \sim n.$$

As $m_n \sim K \log n$ for large n, $n^* \sim n/K \log n$ so that for adequately large values of n, we have

(2.3.46) $$P\{n^{-1} \sum_{i=1}^{n} Z_i - p \geq \gamma_n\} \leq c_4(\log n)^{1+\delta}/n^{1+\delta}$$

where c_4 is a finite positive quentity. Similarly, for adequately large n,

$$(2.3.47)\qquad P\{n^{-1}\sum_{i=1}^{n} Z_i - p \leq -\gamma_n\} \leq c_4(\log n)^{1+\delta}/n^{1+\delta}$$

For $0 < \delta_0 < \delta$, we can write

$$\begin{aligned} c_4(\log n)^{1+\delta}/n^{1+\delta} &= c_4[(\log n)^{1+\delta}/n^{\delta_0}]/n^{1+(\delta-\delta_0)} \\ &= c_4 u_n/n^{1+\gamma}, \end{aligned}$$

where $\gamma = \delta - \delta_0 > 0$ and $u_n = (\log n)^{1+\delta}/n^{\delta_0}$.

Now $u_n \to 0$ as $n \to \infty$ so that for every fixed $\varepsilon > 0$, there exists a n_0 such that $u_n \leq \varepsilon$ for $n \geq n_0$. Hence for adequately large n,

$$(2.3.48)\qquad c_4(\log n)^{1+\delta}/n^{1+\delta} \leq c_4\varepsilon/n^{1+\gamma}$$

(2.3.46), (2.3.47) and (2.3.48) along with the fact $\sum_{n\geq 1} n^{-(1+\gamma)} < \infty$, imply (2.3.44).

<u>Lemma</u> 2.3.5. <u>Let</u> $\mu_{nj} = np^{(j)} + o(n^{1/2}\log n)$, $0 < p^{(j)} < 1$. <u>Then</u>, <u>if</u> $f_{[j]}\left(\xi_{p(j)}^{(j)}\right) > 0$, $X_{n,\mu_{nj}}^{(j)} \in I_n^{(j)}$, $j = 1,\ldots,q$, <u>with probability one</u>, <u>as</u> $n \to \infty$.

Using Lemma 2.3.4, the proof follows along the same line as in Lemma 2.4 of Sen (1968b), and hence, is omitted.

<u>Proof of theorem</u> 2.3.1. In Lemma 2.3.2, by proper choice of K, we take $d \geq 1$. Thus, to prove (2.3.3), it suffices to prove that as $n \to \infty$, for every $j = 1,\ldots,q$,

$$(2.3.49)\qquad \sup\{|[F_{nj}^*(x) - F_{nj}^*(\xi_{p(j)}^{(j)})] - [\tilde{F}_{n[j]}(x) - \tilde{F}_{n[j]}(\xi_{p(j)}^{(j)})]| : x\in I_n^{(j)}\}$$
$$= o(n^{-3/4}\log n),$$

with probability one, and

$$\text{(2.3.50)} \qquad \sup\{|F_{nj}(x) - F^*_{nj}(x)| : x \in I_n^{(j)}\} = O(n^{-3/4} \log n),$$

with probability one.

To prove (2.3.49), we consider a sequence $\{b_n\}$ of positive integers such that $b_n \sim n^{1/4}$ as $n \to \infty$. Then the proof follows by the same technique as in Lemma 1 of Bahdur (1966), provided his (13) also holds for the m_n-dependent case where $m_n \sim K \log n$ for large n. By his definition, $G_n(\eta_r,\omega)$ (see (5) and (13) of Bahadur) is equal to

$$\Big\{[F^*_{nj}\big(\xi^{(j)}_{p(j)} + rn^{-3/4} \log n\big) - F^*_{nj}\big(\xi^{(j)}_{p(j)}\big)]$$

$$- [\tilde{F}_{n[j]}\big(\xi^{(j)}_{p(j)} + rn^{-3/4} \log n\big) - \tilde{F}_{n[j]}\big(\xi^{(j)}_{p(j)}\big)]\Big\}$$

where r ($\leq b_n$) is a positive integer. The above can also be expressed as $n^{-1} \sum_{i=1}^{n} (Z_i^{(j)} - p^{(j)})$, where $Z_i^{(j)}$'s are m_n-dependent binomial variables and $p^{(j)} = \tilde{F}_{n[j]}\big(\xi^{(j)}_{p(j)} + rn^{-3/4} \log n\big) - \tilde{F}_{n[j]}\big(\xi^{(j)}_{p(j)}\big)$. It therefore follows from $0 < f_{[j]}\big(\xi^{(j)}_{p(j)}\big) < \infty$ that for n sufficiently large, $0 < p^{(j)} \leq c_2^* n^{-1/2} \log n$ for all $r \leq b_n$. Hence by (2.3.44) we can select $\gamma_n = c_1^* n^{-3/4} \log n$, with $c_1^{*2}/2c_2^*$ sufficiently large, and this along with (2.3.45) and the Bonferroni inequality gives us

$$\text{(2.3.51)} \qquad \max_{|r| \leq b_n} P\{|G_n(\eta_r,\omega)| \geq \gamma_n\} \leq 4e^{-v_n}$$

where

$$\text{(2.3.52)} \qquad v_n = -\tfrac{1}{4} \log n + c_1^{*2}\, n^{1/2} (\log n)^2 / 2[c_2^* n^{1/2} \log n + c_1^* n^{1/4} \log n]$$

$$= \log n^{-1/4 + c_1^{*2}/2c_2^*[1 + o(1)]}$$

If $-1/4 + c_1^{*2}/2c_2^*$ is chosen greater than 1, then use of (9) of Bahadur (1966) completes the proof of (2.3.49).

To prove (2.3.50), we consider the same sequence $\{b_n\}$ as defined above. Let us then write $\eta_{r,n}^{(j)} = \xi_{p(j)}^{(j)} + ra_n/b_n$ (where a_n is defined by (2.3.2)), for $r = 0, \pm 1, \ldots, \pm b_n$, and let $I_{r,n}^{(j)}$ be the interval $(\eta_{r,n}^{(j)}, \eta_{r+1,n}^{(j)})$, $r = -b_n, \ldots, b_n - 1$. Since both F_{nj} and F_{nj}^* are non-decreasing, we have for all $x \varepsilon I_{r,n}^{(j)}$,

$$F_{nj}(\eta_{r,n}^{(j)}) - F_{nj}^*(\eta_{r+1,n}^{(j)}) \leq F_{nj}(x) - F_{nj}^*(x) \leq F_{nj}(\eta_{r+1,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})$$

and consequently, we get

$$(2.3.53) \quad |F_{nj}(x) - F_{nj}^*(x)| \leq \max_{s=r,r+1} |F_{nj}(\eta_{s,n}^{(j)}) - F_{nj}^*(\eta_{s,n}^{(j)})| + |F_{nj}^*(\eta_{r+1,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})|$$

This implies

$$(2.3.54) \quad \sup_{x\varepsilon I_n^{(j)}} |F_{nj}(x) - F_{nj}^*(x)| \leq \max_{|r|\leq b_n} |F_{nj}(\eta_{r,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})| + \max_{-b_n \leq r \leq b_n - 1} |F_{nj}^*(\eta_{r+1,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})|$$

Let us now denote by $G_{n[j]}(x) = P\{|R_{n,i}^{(j)}| \leq x\}$, $0 \leq x < \infty$. Then,

$$(2.3.55) \quad E|F_{nj}(\eta_{r,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})|$$
$$\leq n^{-1} \sum_{i=1}^{n} E|c(\eta_{r,n}^{(j)} - Y_{n,i}^{(j)} - R_{n,i}^{(j)}) - c(\eta_{r,n}^{(j)} - Y_{n,i}^{(j)})|$$
$$\leq \int_0^{c_2 n^{-c_1/\delta}} P\{\eta_{r,n}^{(j)} - x \leq Y_{n,i}^{(j)} \leq \eta_{r,n}^{(j)}$$
$$\text{or } \eta_{r,n}^{(j)} \leq Y_{n,i}^{(j)} \leq \eta_{r,n}^{(j)} + x\} dG_{n[j]}(x)$$

$$+ P\{|R_{n,i}^{(j)}| \geq c_2 n^{-c_1/\delta}\}$$

$$= O(n^{-c_1/\delta}) + O(n^{-c_1}), \text{ by Lemma 2.3.1,}$$

where $0 < \delta \leq 1$. We now choose a δ': $1/3 < \delta' < 3/4$ and K such that $c = \delta\rho^* K > 8$, so that in Lemma 2.3.1 we may take $c_1 > 4$. This implies $c_1\delta' > 4/3 \Rightarrow \sum_{n\geq 1} n^{-c_1\delta' + 1/4} < \infty$. Also, $c_1(1 - \delta') > 1$. Thus, by using the Chebychev inequality and (2.3.55), we obtain that as $n \to \infty$,

$$\text{(2.3.56)} \qquad P\{|F_{nj}(\eta_{r,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})| \geq (K_2 n^{-c_1})^{1-\delta'}\} \leq K_2^{\delta'} n^{-c_1\delta'},$$

for all $r = 0, \pm 1, \ldots, \pm b_n$. Hence, by the Bonferroni inequality, as $n \to \infty$,

$$\text{(2.3.57)} \qquad P\{\max_{|r|\leq b_n} |F_{nj}(\eta_{r,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})| \geq K_2^{1-\delta'} n^{-1}\} \leq K_2^{\delta'} n^{-1-\varepsilon}, \varepsilon > 0.$$

The last equation implies that as $n \to \infty$,

$$\text{(2.3.58)} \qquad \max_{|r|\leq b_n} |F_{nj}(\eta_{r,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})| = O(n^{-1}),$$

with probability one.

Now, for every j ($= 1,\ldots,q$), $\{Y_{n,i}^{(j)}, i=1,2,\ldots\}$ forms an m_n-dependent process. Hence by (2.3.49), it follows that as $n \to \infty$,

$$\text{(2.3.59)} \qquad \max_{-b_n\leq r\leq b_n - 1} |F_{nj}^*(\eta_{r+1,n}^{(j)}) - F_{nj}^*(\eta_{r,n}^{(j)})| = O(n^{-3/4} \log n),$$

with probability one. Then (2.3.50) follows from (2.3.54), (2.3.58) and (2.3.59). This completes the proof of (2.3.3).

Finally, if $f'_{[j]}(x)$ is bounded in some neighbourhood of $\xi_{p(j)}^{(j)}$, (2.3.4) follows directly from (2.3.3), as, under this condition, Lemma 3 of Bahadur (1966) extends directly to our case. This completes the proof when $Z_n^{(j)}$ is defined by a single order statistic. If $Z_n^{(j)}$ is defined as an average of two successive order statistics, say, the μ_n-th and (μ_n+1)-th ones, (where $\mu_n \le np^{(j)} \le (\mu_{n+1})$), upon noting that the difference of empirical c.d.f's at these two points is equal to 1/n, it follows from (2.3.3) and Lemma 2.3.5 that the difference between the values of $F_{[j]}(x)$ at these two points is also $o(n^{-3/4} \log n)$, with probability one, as $n \to \infty$. Consequently (2.3.4) holds for the general case when $Z_n^{(j)}$ is any inner point of two successive order statistics. Hence the proof of Theorem 2.3.1 is completed.

2.4. Asymptotic Joint Normality of $\underset{\sim}{Z}_n$

If we let

$$(2.4.1) \qquad \underset{\sim}{T} = ((\tau_{jj'})); \; \tau_{jj'} = \nu_{jj'} \Big/ \left[f_{[j]}\left(\xi_{p(j)}^{(j)}\right) f_{[j']}\left(\xi_{p(j')}^{(j')}\right)\right], \qquad j,j' = 1,\ldots,q,$$

where $\nu_{jj'}$ and $f_{[j]}\left(\xi_{p(j)}^{(j)}\right)$ are defined in section 2.2, then we have the following.

<u>Theorem</u> 2.4.1. <u>If</u> $F(\underset{\sim}{x})$ <u>is absolutely continuous at</u> $\underset{\sim}{\xi}_{\underset{\sim}{p}}$ <u>with finite, positive and continuous marginal densities in some neighbourhood of</u> $\underset{\sim}{\xi}_{\underset{\sim}{p}}$<u>, and if</u> $\underset{\sim}{\nu}$ <u>is positive definite, then under (2.2.6)- (2.2.10),</u>

$$(2.4.2) \qquad \mathcal{L}(n^{1/2}[\underset{\sim}{Z}_n - \underset{\sim}{\xi}_{\underset{\sim}{p}}]) \to N_q(\underset{\sim}{0},\underset{\sim}{T}), \quad \text{as } n \to \infty.$$

<u>Proof</u>. By virtue of (2.3.3), Lemma 2.3.5, (2.3.49) and (2.3.50), as $n \to \infty$,

$$(2.4.3) \quad \sqrt{n}\left|\{F_{[j]}(Z_n^{(j)}) - p^{(j)}\} + \{F_{nj}^{*}\left(\xi_{p(j)}^{(j)}\right) - p^{(j)}\}\right| \xrightarrow{p} 0, \quad j=1,\ldots,q.$$

This implies that $[n^{1/2}\{Z_n^{(j)} - \xi_{p(j)}^{(j)}\}f_{[j]}\left(\xi_{p(j)}^{(j)}\right), j = 1,\ldots,q]$ and $[n^{1/2}\{F_{nj}^*\left(\xi_{p(j)}^{(j)}\right) - p^{(j)}\}, j = 1,\ldots,q]$ have the same limiting distribution, if they have one at all. Now, by (2.3.8), F_{nj}^* involves an average over zero-one valued m_n-dependent random variables, on which a direct multivariate extension of the Hoeffding-Robbins (1948) C.L.T. for m-dependent processes (with a straightforward extension for an m_n-dependent process with $m_n \sim K \log n$) or a multivariate extension of the C.L.T. for strongly mixing processes by Rosenblatt (1956 b) yields that $[n^{1/2}\{F_{nj}^*\left(\xi_{p(j)}^{(j)}\right) - p^{(j)}\}, j = 1,\ldots,q]$ has asymptotically, a multinormal distribution with mean $\underset{\sim}{0}$ and dispersion matrix $\underset{\sim}{\nu}_n^*(\underset{\sim}{\xi}_{\underset{\sim}{p}})$, defined by (2.3.21) and (2.3.22). The rest of the proof follows trivally by a use of Lemma 2.3.3 along with (2.4.3) and (2.4.1).

Q.E.D.

Note: In the above theorem if the condition that $\underset{\sim}{\nu}$ is positive definite is not satisfied and $\underset{\sim}{\nu}$ is positive semidefinite with rank $f(< q)$ then we choose f linearly independent variables and prove asymptotic normality for this subset. Since the other variables will be linearly dependent on them, in this case, the asymptotic distribution for the whole set of variables will be a singular normal distribution with rank f.

CHAPTER III

ASYMPTOTIC DISTRIBUTION OF LINEAR COMPOUND OF SOME FIXED NUMBER OF QUANTILES FOR STATIONARY MULTIVARIATE AUTOREGRESSIVE PROCESSES

3.1 Introduction

In the first part of the chapter, for stationary autoregressive processes Bahadur's asymptotic a.s. representation of the standardized form of a sample quantile is extended over the entire real line. With this extension, the asymptotic multinormality, derived in Chapter II of a sample quantile (vector) for a stationary autoregressive process, is generalized in section 3.3 for a vector of linear combinations of several quantiles and an application of this result is considered in the next section. The section 3.5 contains a method of estimation of variances and covariances of sample quantiles and hence the variances and covariances of the linear compounds of several quantiles. The chapter is concluded with an extension of the above asymptotic multinormality of the vector of linear compounds of several quantiles for random sample sizes.

3.2. Extension of Bahadur-representation of Sample Quantiles over the Entire Real Line.

For a fixed j, if $p^{(j)}$ is varied in (2.2.13) over the open interval (0,1) then we get a quantile process (see Kiefer 1967, 1970).

In this section the result (2.3.3) (extension of Bahadur-representation of sample quantiles to the autoregressive case) is first extended for a quantile process. Then asymptotic normality for linear function of several quantiles follows immediately as a corollary of this result.

Let us suppose that $F_{[j]}(x)$ is absolutely continuous and $\sup_x f_{[j]}(x) = f_{oj} < \infty$.

Similar to (2.3.15), we have

$$\tilde{F}_{n[j]}(x) \leq F_{[j]}(x) + O(n^{-d})$$

and

$$\tilde{F}_{n[j]}(x) \geq F_{[j]}(x) - O(n^{-d})$$

uniformly in x, where we take $d \geq 1$. Hence it follows that

$$\sup_x |F_{[j]}(x) - \tilde{F}_{n[j]}(x)| = O(n^{-d}), \quad d \geq 1. \tag{3.2.1}$$

Then absolute continuity of $\tilde{F}_{n[j]}(x)$ is obtained by (3.2.1) and the absolute continuity of $F_{[j]}(x)$. Hence the random variables $Z_{n,i}^{(j)} = \tilde{F}_{n[j]}(Y_{n,i}^{(j)})$, $i = 1,\dots,n$ are distributed uniformly over (0,1). Define the empirical process $[V_{nj}^*(t): 0 \leq t \leq 1]$ by

$$V_{nj}^*(t) = n^{1/2}[G_{nj}^*(t) - t]; \quad G_{nj}^*(t) = n^{-1} \sum_{i=1}^{n} c(t - Z_{n,i}^{(j)}), \quad 0 \leq t \leq 1 . \tag{3.2.2}$$

Let $g(n) = n^{-1/2} \log n$. Define

$$K_{nj}(t) = \sup_a \{|V_{nj}^*(t) - V_{nj}^*(t + a)|: |a| \leq g(n)\}, \; 0 < t < 1; \tag{3.2.3}$$

$$K_{nj}^* = \sup_{0<t<1} \{K_{nj}(t)\} . \tag{3.2.4}$$

Then we have the following theorem.

Theorem 3.2.1. For every $j = 1,\dots,q$, as $n \to \infty$,

$$\text{(3.2.5)} \qquad \sup_{x} \sup_{|a|<g(n)} n^{1/2} |F_{nj}(x+a) - F_{nj}(x) - F_{[j]}(x+a) + F_{[j]}(x)| = O(n^{-1/4} \log n),$$

with probability one, provided (2.2.6), (2.2.8) and (2.2.10) hold.

The proof of the above theorem depends on the following lemmas.

Lemma 3.2.1. For every finite $s (> 0)$ and j $(= 1,\dots,q)$, there exist a positive constant $C_s^{(1)}$ and a sample size $n_{s,j}$, such that for $n \geq n_{s,j}$,

$$\text{(3.2.6)} \qquad P\{K_{nj}^* \geq C_s^{(1)} n^{-1/4} \log n\} \leq 4n^{-s}$$

Hence as $n \to \infty$, $K_{nj}^* = O(n^{-1/4} \log n)$, with probability one.

Proof. Let $\xi_{\ell,n} = \ell[n^{1/2}]^{-1}$ $(0 \leq \ell \leq [n^{1/2}])$, $[x]$ being the largest integer less than or equal to x. For uniform distribution the density function being equal to 1 $(0 \leq t \leq 1)$, by the same technique as in (2.3.51) and using (9) of Bahadur (1966) we get for large n,

$$\text{(3.2.7)} \qquad P\{K_{nj}(\xi_{\ell,n}) > c_1^* n^{-1/4} \log n\} \leq 4e^{-v_n}, \ell = 1,2,\dots[n^{1/2}]$$

where $c_1^* > 0$

and $\quad v_n = -1/4 \log n + c_1^{*2}/2 \log n (1 + c_1^* n^{-1/4})^{-1}$.

The rest of the proof of (3.2.6) is exactly the same as in Theorem 4.1 of Sen and Ghosh (1971).

The second part of the lemma follows by taking $s > 1$ in (3.2.6).

Note: Actually the above lemma is the extension of Theorem 4.1 (extension of Bahadur's (1966) result to the entire real line) of Sen and Ghosh (1971) to the m_n-dependent process where $m_n \sim K \log n$ for large n.

Lemma 3.2.2. Let $U_{i,j} = F(X_{i,j})$, $i = 1,2,\ldots,n$, $G_{nj}(t)=n^{-1}\sum_{i=1}^{n} c(t - U_{i,j})$. Then for every $j = 1,\ldots,q$, as $n \to \infty$,

$$\sup_{0<t<1} n^{1/2}|G_{nj}(t) - G^*_{nj}(t)| = O(n^{-1/4} \log n), \tag{3.2.8}$$

with probability one.

Proof. As in Lemma 3.2.1, let us take the grid points $\xi_{\ell,n} = \ell[n^{1/2}]^{-1}$, $0 \le \ell \le [n^{1/2}]$. Now $G_{nj}(t)$ and $G^*_{nj}(t)$ being non-decreasing in t, for $t \in [\xi_{\ell,n},\xi_{\ell+1,n}]$, we have

$$G_{nj}(\xi_{\ell,n}) \le G_{nj}(t) \le G_{nj}(\xi_{\ell+1,n})$$

$$G^*_{nj}(\xi_{\ell,n}) \le G^*_{nj}(t) \le G^*_{nj}(\xi_{\ell+1,n})$$

so that

$$\begin{aligned} n^{1/2}[G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell+1,n})] &\le n^{1/2}[G_{nj}(t) - G^*_{nj}(t)] \\ &\le n^{1/2}[G_{nj}(\xi_{\ell+1,n}) - G^*_{nj}(\xi_{\ell,n})] \end{aligned}$$

This implies

$$\begin{aligned} \sup_{0<t<1} n^{1/2}|G_{nj}(t) - G^*_{nj}(t)| &\le \max_{\ell=0,1,\ldots,[n^{1/2}]} n^{1/2}|G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell,n})| \\ &+ \max_{\ell=0,1,\ldots[n^{1/2}]} n^{1/2}|G^*_{nj}(\xi_{\ell+1,n}) - G^*_{nj}(\xi_{\ell,n}) \end{aligned} \tag{3.2.9}$$

By Lemma 3.2.1, as $n \to \infty$, the second term is of order $n^{-1/4} \log n$, with probability one. Hence for (3.2.9) it is sufficient to prove that as $n \to \infty$,

$$(3.2.10)\qquad \max_{\ell=0,1,\ldots,[n^{1/2}]} n^{1/2}|G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell,n})| = O(n^{-1/4}),$$

with probability one.

Similar to (2.3.55), here we get

$$(3.2.11)\qquad E[n^{\frac{1}{2}}|G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell,n})|] = O(n^{\frac{1}{2} - c_1}).$$

We now choose a $\delta_1 : 3/7 < \delta_1 < 13/14$ and as before K such that $c_1 > 4$. This implies $(c_1 - \frac{1}{2})\delta_1 > 3/2 \Rightarrow \sum_{n\geq 1} n^{-\{(c_1 - \frac{1}{2})\delta_1 - \frac{1}{2}\}} < \infty$ and also

$(c_1 - \frac{1}{2})(1 - \delta_1) > \frac{1}{4}$. Thus, by using the Chebychev inequality and

(3.2.11), we get as $n \to \infty$,

$$(3.2.12)\qquad P\left\{n^{\frac{1}{2}}|G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell,n})| \geq \left(K_3 n^{\frac{1}{2} - c_1}\right)^{1-\delta_1}\right\} \leq \left(K_3 n^{\frac{1}{2} - c_1}\right)^{\delta_1}$$

for all $\ell = 0,1,\ldots,[n^{1/2}]$. Hence, by the Bonferroni inequality, as $n \to \infty$,

$$(3.2.13)\qquad P\left\{\max_{\ell= 0,1,\ldots,[n^{1/2}]} n^{1/2}|G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell,n})| \geq K_3^{1-\delta_1} n^{-1/4}\right\} \leq K_3^{\delta_1} n^{-1-\varepsilon}, \quad \varepsilon > 0.$$

(3.2.10) is implied by the Borel-Cantelli Lemma and (3.2.13). Q.E.D.

<u>Lemma</u> 3.2.3. <u>With the notations used above, for every $j = 1,\ldots,q$, as</u> $n \to \infty$,

$$(3.2.14)\qquad \sup_t \sup_{|a|<g(n)} n^{1/2}|G_{nj}(t + a) - G^*_{nj}(t + a)| = O(n^{-1/4} \log n),$$

<u>with probability one</u>.

Proof. We have

$$(3.2.15) \quad \sup_{0<t<1} \sup_{|a|<g(n)} n^{1/2}|G_{nj}(t+a) - G^*_{nj}(t+a)|$$

$$\leq \sup_{0<t<1} \sup_{|a|<g(n)} n^{1/2}|G_{nj}(t+a) - G_{nj}(t)|$$

$$+ \sup_{0<t<1} n^{1/2}|G_{nj}(t) - G^*_{nj}(t)|$$

$$+ \sup_{0<t<1} \sup_{|a|<g(n)} n^{1/2}|G^*_{nj}(t+a) - G^*_{nj}(t)|$$

By Lemma 3.2.2 and Lemma 3.2.1 it follows that the second term and the third term of (3.2.15) are of order $n^{-1/4} \log n$, with probability one, as $n \to \infty$. Hence for proof of Lemma 3.2.3, it is sufficient to show that, as $n \to \infty$,

$$(3.2.16) \quad \sup_{0<t<1} \sup_{|a|<g(n)} n^{1/2}|G_{nj}(t+a) - G_{nj}(t)| = 0(n^{-1/4} \log n),$$

with probability one.

If we denote by

$$H_{nj}(t) = \sup_{|a|<g(n)} n^{1/2}|G_{nj}(t+a) - G_{nj}(t)|$$

and by

$$H^*_{nj} = \sup_{0<t<1} H_{nj}(t)$$

and consider the grid points $\xi_{\ell,n} = \ell[n^{1/2}]^{-1}$, $\ell = 1, \ldots [n^{1/2}]$, then as in (4.8) of Sen and Ghosh (1971), we have

$$(3.2.17) \qquad H^*_{nj} \leq 3[\max_{1 \leq \ell \leq [n^{1/2}]} H_{nj}(\xi_{\ell,n})]$$

Hence for (3.2.16), it is enough to show that, as $n \to \infty$, (3.2.18) $$\max_{1\le \ell \le [n^{1/2}]} H_{nj}(\xi_{\ell,n}) = O(n^{-1/4} \log n),$$

with probability one.

Again

(3.2.19) $$\max_{1\le \ell \le [n^{1/2}]} H_{nj}(\xi_{\ell,n})$$

$$= \max_{1\le \ell \le [n^{1/2}]} \sup_{|a|<g(n)} n^{1/2}|G_{nj}(\xi_{\ell,n} + a) - G_{nj}(\xi_{\ell,n})|$$

$$\le \max_{1\le \ell \le [n^{1/2}]} \sup_{|a|<g(n)} n^{1/2}|G_{nj}(\xi_{\ell,n} + a) - G^*_{nj}(\xi_{\ell,n} + a)|$$

$$+ \max_{1\le \ell \le [n^{1/2}]} \sup_{|a|<g(n)} n^{1/2}|G^*_{nj}(\xi_{\ell,n} + a) - G^*_{nj}(\xi_{\ell,n})|$$

$$+ \max_{1\le \ell \le [n^{1/2}]} n^{1/2}|G_{nj}(\xi_{\ell,n}) - G^*_{nj}(\xi_{\ell,n})|$$

By Lemma 3.2.1 and Lemma 3.2.2, the second term and the third term in (3.2.19) are of order $n^{-1/4} \log n$, with probability one, as $n \to \infty$ and by the same technique as in (2.3.54) and (2.3.59), we obtain, as $n \to \infty$,

(3.2.20) $$\max_{1\le \ell \le [n^{1/2}]} \sup_{|a|<g(n)} n^{1/2}|G_{nj}(\xi_{\ell,n} + a) - G^*_{nj}(\xi_{\ell,n} + a)|$$

$$\le \max_{1\le \ell \le [n^{1/2}]} \max_{|r|\le b_n} n^{1/2}|G_{nj}(\xi_{\ell,n} + ra_n/b_n) - G^*_{nj}(\xi_{\ell,n} + ra_n/b_n)|$$
$$+ O(n^{-1/4} \log n),$$

with probability one.

Therefore for proof of (3.2.18), it is only required to show that, as $n \to \infty$,

$$(3.2.21)\quad \max_{1\le \ell \le [n^{1/2}]} \max_{|r|\le b_n} n^{1/2}|G_{nj}(\xi_{\ell,n}+ ra_n/b_n) - G^*_{nj}(\xi_{\ell,n}+ ra_n/b_n)| = O(n^{-1/4}),$$

with probability one.

We choose a δ_2: $1/2 < \delta_2 < 13/14$ and as before K such that $c_1 > 4$. This gives $(c_1 - 1/2)\delta_2 > 7/4 \Rightarrow \sum_{n\ge 1} n^{-\{(c_1 - 1/2)\delta_2 - 3/4\}} < \infty$ and also

$(c_1 - 1/2)(1 - \delta_2) > 1/4.$

Hence, similar to (3.2.12), we have by the Chebychev inequality

$$(3.2.22)\quad P\left\{n^{1/2}|G_{nj}(\xi_{\ell,n}+ ra_n/b_n) - G^*_{nj}(\xi_{\ell,n}+ ra_n/b_n)| \ge (K_4 n^{1/2-c_1})^{1-\delta_2}\right\} \le (K_4 n^{1/2 - c_1})^{\delta_2}$$

for all $\ell = 1,\dots,[n^{1/2}]$; $r = 0,\pm 1,\dots,\pm b_n$. Hence, by the Bonferroni inequality, as $n \to \infty$,

$$(3.2.23)\quad P\left\{\max_{1\le \ell \le [n^{1/2}]} \max_{|r|\le b_n} n^{1/2}|G_{nj}(\xi_{\ell,n}+ ra_n/b_n)-G^*_{nj}(\xi_{\ell,n}+ ra_n/b_n)| \ge K_4^{1-\delta_2} n^{-1/4}\right\} \le 2K_4^{\delta_2} n^{-\{(c_1 -1/2)\delta_2 -3/4\}}$$

and by use of the Borel-Cantelli lemma, (3.2.23) implies (3.2.21).

Q.E.D.

<u>Proof of theorem</u> 3.2.1.

As $\sup_x f_{[j]}(x) = f_{oj} < \infty$, we have $|F_{[j]}(x+a) - F_{[j]}(x)| \le f_{oj}|a|$, $-\infty < x < \infty$. Now since

$$(3.2.24)\quad \sup_x \sup_{|a|<g(n)} n^{1/2}|F_{nj}(x+a) - F_{nj}(x) - F_{[j]}(x+a) + F_{[j]}(x)|$$

$$\leq \sup_x \sup_{|a|<g(n)} n^{1/2} |[F^*_{nj}(x+a) - \tilde{F}_{n[j]}(x+a)] - [F^*_{nj}(x) - \tilde{F}_{n[j]}(x)]|$$

$$+ \sup_x n^{1/2} |[F_{nj}(x) - F_{[j]}(x)] - [F^*_{nj}(x) - \tilde{F}_{n[j]}(x)]|$$

$$+ \sup_x \sup_{|a|<g(n)} n^{1/2} |[F_{nj}(x+a) - F_{[j]}(x+a)] - [F^*_{nj}(x+a) - \tilde{F}_{n[j]}(x+a)]|,$$

by use of Lemma 3.2.1, Lemma 3.2.2 and Lemma 3.2.3, the theorem 3.2.1 easily follows.

3.3 Asymptotic Normality of a Linear Compound of Several Quantiles

Let us suppose that $Z^{(j)}_{n,p_i^{(j)}}$ denotes the sample quantile of order $p_i^{(j)}$ for the j-th variate and $\xi^{(j)}_{p_i^{(j)}}$ be the corresponding population quantile, $j = 1,\ldots,q$. We want to determine the asymptotic distribution of the vector of linear compounds

$$\left(n^{1/2} \sum_{i=1}^{t} a^{(1)}_{p_i^{(1)}} \left(Z^{(1)}_{n,p_i^{(1)}} - \xi^{(1)}_{p_i^{(1)}}\right), \ldots, n^{1/2} \sum_{i=1}^{t} a^{(q)}_{p_i^{(q)}} \left(Z^{(q)}_{n,p_i^{(q)}} - \xi^{(q)}_{p_i^{(q)}}\right)\right)'$$

where the $a^{(j)}_{p_i^{(j)}}$'s ,$i=1,\ldots,t$; $j=1,\ldots q$, are known constants

Let us assume that $f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right) > 0$, $j=1,\ldots,q$, and let us denote by

$$I^{(j)}_{n,p_i^{(j)}} = \{x: \xi^{(j)}_{p_i^{(j)}} - a_n \leq x \leq \xi^{(j)}_{p_i^{(j)}} + a_n\}, \quad j = 1,\ldots,q, \tag{3.3.1}$$

where as before $a_n \sim n^{-1/2} \log n$ as $n \to \infty$. By use of Lemma 2.3.5, we get $Z^{(j)}_{n,p_i^{(j)}} \in I^{(j)}_{n,p_i^{(j)}}$ for $i=1,\ldots,t$; $j=1,\ldots,q$, with probability one, as $n \to \infty$ and under the condition that $\sup_{1\leq j\leq q} \sup_x f'_{[j]}(x)$ is finite, Lemma 3 of Bahadur (1966) extends directly to our case.

Similar to (2.3.4), in this case, by (3.2.5) and Lemma 3 of Bahadur (1966) we get simultaneously for $i = 1,\ldots,t$, $j = 1,\ldots q$, as $n \to \infty$,

$$(3.3.2) \quad n^{1/2}\left\{[Z^{(j)}_{n,p_i^{(j)}} - \xi^{(j)}_{p_i^{(j)}}]f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right) + [F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}]\right\} = O(n^{-1/4}\log n),$$

with probability one. This implies that for every $\varepsilon > 0$ there exists an $\eta > 0$ such that, as $n \to \infty$,

$$(3.3.3) \quad P\{n^{1/2}|[Z^{(j)}_{n,p_i^{(j)}} - \xi^{(j)}_{p_i^{(j)}}]f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right) + [F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}]| \leq \varepsilon \text{ for } i=1,\ldots t,\ j=1,\ldots,q\} \geq 1-\eta$$

From Lemma 3.2.2, it follows that

$$(3.3.4) \quad \sup_{1\leq j\leq q} \sup_{x} n^{1/2}|F_{nj}(x) - F^*_{nj}(x)| \xrightarrow{p} 0 \text{ as } n \to \infty,$$

and this implies

$$(3.3.5) \quad \mathcal{L}(n^{1/2}[F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}],\ i = 1,\ldots,t,\ j = 1,\ldots,q)$$
$$\sim \mathcal{L}(n^{1/2}[F^*_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}],\ i=1,\ldots,t,\ j=1,\ldots,q).$$

Now, by (2.3.8), F^*_{nj} involves an average over zero-one valued m_n-dependent random variables, on which a direct multivariate extension of the Hoeffding-Robbins (1948) central limit theorem for m-dependent processes (with a straightforward extension for an m_n-dependent process with $m_n \sim K \log n$) or a multivariate extension of the central limit theorem for strongly mixing processes by Rosenblatt (1956b) yields that $[n^{1/2}\{F^*_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}\},\ i=1,\ldots,t\ j=1,\ldots,q]$ has asymptotically a multinormal distribution with mean $\underset{\sim}{0}$ and dispersion matrix

matrix

$$(3.3.6)\qquad \left(\left(\nu^{*(jj')}_{nii'}\right)\right)_{\substack{i,i'=1,\ldots,t\\ j,j'=1,\ldots,q}}$$

where

$$(3.3.7)\qquad \nu^{*(jj')}_{nii'} = n\ \mathrm{covar}\left[F^{*}_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right),\ F^{*}_{nj'}\left(\xi^{(j')}_{p_{i'}^{(j')}}\right)\right]$$

Hence by (3.3.5), Lemma 2.3.3 and above, it follows that $[n^{1/2}\{F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}\}\ ,\ i=1,\ldots,t; j=1,\ldots,q]$ has asymptotically a multinormal distribution with mean $\underset{\sim}{0}$ and dispersion matrix

$$(3.3.8)\qquad \left(\left(\nu^{(jj')}_{ii'}\right)\right)_{\substack{i,i'=1,\ldots,t\\ j,j'=1,\ldots,q}}$$

where

$$(3.3.9)\qquad \nu^{(jj')}_{ii'} = \lim_{n\to\infty} \nu^{*(jj')}_{nii'},\ \forall\ i,i'=1,\ldots,t; j,j'=1,\ldots,q$$

and its exact expression is given by (2.3.23) with $p^{(j)}$ and $p^{(j')}$ replaced by $p_i^{(j)}$ and $p_{i'}^{(j')}$ respectively. We assume the matrix $\left(\left(\nu^{(jj')}_{ii'}\right)\right)$ to be positive definite.

Finally, the asymptotic distribution of $[n^{1/2}\{F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}\}$, $i=1,\ldots,t$, $j=1,\ldots,q]$ being multivariate normal, $[-n^{1/2}\{F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}\}$, $i=1,\ldots,t$, $j=1,\ldots,q]$ and $[n^{1/2}\{F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}\}, i=1,\ldots,t, j=1,\ldots,q]$ have the same asymptotic distribution.

Therefore, by (3.3.3) and the above arguments, it follows that as $n \to \infty$,

$$(3.3.10)\qquad \mathcal{L}\left(n^{1/2}[Z^{(j)}_{n,p_i^{(j)}} - \xi^{(j)}_{p_i^{(j)}}],\ i=1,\ldots,t; j=1,\ldots,q\right) \to N_{t,q}(\underset{\sim}{0},\underset{\sim}{H})$$

where

$$(3.3.11)\qquad \underset{\sim}{H} = \left(\left(\eta^{(jj')}_{ii'}\right)\right)_{\substack{i,i'=1,\ldots,t\\ j,j'=1,\ldots,q}}$$

and

$$\eta_{ii'}^{(jj')} = \nu_{ii'}^{(jj')} \Big/ f_{[j]}\left(\xi_{p_i}^{(j)}{}_{(j)}\right) f_{[j']}\left(\xi_{p_{i'}}^{(j')}{}_{(j')}\right), \quad i,i'=1,\ldots,t \quad j,j'=1,\ldots,q. \tag{3.3.12}$$

Hence for any vector $\underset{\sim}{d} = (d_1,\ldots,d_q)'$, the asymptotic distribution of the linear combination

$$n^{1/2} \sum_{j=1}^{q} \sum_{i=1}^{t} d_j a_{p_i}^{(j)}{}_{(j)} \left(Z_{n,p_i}^{(j)}{}_{(j)} - \xi_{p_i}^{(j)}{}_{(j)} \right)$$

is univariate normal with mean zero and variance

$$\sum_{j=1}^{q} \sum_{j'=1}^{q} \sum_{i=1}^{t} \sum_{i'=1}^{t} d_j d_{j'} a_{p_i}^{(j)}{}_{(j)} a_{p_{i'}}^{(j')}{}_{(j')} \eta_{ii'}^{(jj')} \tag{3.3.13}$$

and this implies that the asymptotic distribution of the vector

$$\left(n^{1/2} \sum_{i=1}^{t} a_{p_i}^{(1)}{}_{(1)} \left(Z_{n,p_i}^{(1)}{}_{(1)} - \xi_{p_i}^{(1)}{}_{(1)} \right), \ldots, n^{1/2} \sum_{i=1}^{t} a_{p_i}^{(q)}{}_{(q)} \left(Z_{n,p_i}^{(q)}{}_{(q)} - \xi_{p_i}^{(q)}{}_{(q)} \right) \right)'$$

is multivariate normal with mean vector $\underset{\sim}{0}$ and dispersion matrix

$$\underset{\sim}{T}^{*} = \left(\left(\tau_{jj'}^{*} \right)\right)_{j,j'=1,\ldots,q} \tag{3.3.14}$$

where

$$\tau_{jj'}^{*} = \sum_{i=1}^{t} \sum_{i'=1}^{t} a_{p_i}^{(j)}{}_{(j)} a_{p_{i'}}^{(j')}{}_{(j')} \eta_{ii'}^{(jj')}, \quad j,j'=1,\ldots,q. \tag{3.3.15}$$

3.4. A Useful Application of the Results of Section 3.3.

Mosteller (1946) considered the case of more than one quantiles with the idea of studying the properties of estimates based on a subset of quantiles. He termed such estimates as inefficient statistics. In the book by Sarhan and Greenberg (1962) for small sample sizes and various known parent distributions optimal or best linear ordered estimators of location and scale parameters are considered. These estimators demand the complete knowledge of the covariance matrix of the sample order statistics up to a unknown multiplicative constant. In large samples it becomes extremely laborious to compute the above covariance matrix. There are two different approaches to the large sample problem. (i) To consider a subset of fixed number of quantiles and to base our estimate as a linear function of the corresponding sample quantiles. These include the so-called mid-ranges which are often more efficient than the sample median. (ii) To consider a subset of quantiles of size depending on the sample size and to base our estimate as a linear function of the corresponding sample quantiles. These include the so-called Trimmed mean and Winsorized mean which sometimes provide useful estimates.

Often, for symmetrical density functions, we employ the sample median as an estimate of the location parameter. It is known that for the class of regular density functions, we can always find more efficient estimates from the class of mid-ranges, provided that the parametric form of the distribution is known. If the variables are independently normally distributed, it is known that the 27% mid-range is optimum (See Sen, 1961). In our case, for each $j = 1,\ldots,q$, the 27% mid-range estimator for the location parameter is $U_n^{(j)} = \left(Z^{(j)}_{n,p_1^{(j)}} + Z^{(j)}_{n,p_2^{(j)}}\right)/2$. Here $\underset{\sim}{U}_n = (U_n^{(1)},\ldots,U_n^{(q)})'$ being a vector of linear combination of two quantiles, $\sqrt{n}\underset{\sim}{U}_n$ is asymptotically multinormally distributed with mean $\underset{\sim}{\xi}_{1/2}$ (population median vector) and dispersion matrix $\underset{\sim}{T}^*$, where $\underset{\sim}{T}^*$ is given by (3.3.14) and (3.3.15) with $t = 2$, $p_1{}^{(j)} = .27$, $p_2{}^{(j)} = .73$ and $a^{(j)}_{p_1{}^{(j)}} = a^{(j)}_{p_2{}^{(j)}} = 1/2$ for $j,j' = 1,\ldots,q$.

Note: By the theory, developed in 3.3 for linear combination of fixed number of sample quantiles, the asymptotic distribution of Trimmed mean and Winsorized mean cannot be obtained for stationary autoregressive processes. For such statistics our strong convergence results do not hold. However, the distribution of such statistics for stationary autoregressive processes may be derived by some weak convergence results. Such studies are proposed for future work.

3.5. Consistent Estimate of the Dispersion Matrix of the Vector of Linear Compounds of Several Quantiles.

The following theorems enable us to estimate the variances and covariances of sample quantiles and hence the dispersion matrix of the vector of their linear compounds.

Theorem 3.5.1. *Let* $\mu_n^{(j)} = np^{(j)} + o(n^{1/2} \log n)$, *and let* s_j *be the* $\mu_n^{(j)}$*-th smallest observation among* $(X_{1,j},\ldots,X_{n,j})$. *If*

$$(3.5.1)\qquad \hat{\nu}_{jj'} = \{F_{njj'}(s_j,s_{j'})-p^{(j)}p^{(j')}\} + \sum_{h=1}^{m_n}\{F_{njj'h}(s_j,s_{j'})-p^{(j)}p^{(j')}\} + \sum_{h=1}^{m_n}\{F_{nj'jh}(s_j,s_{j'}) -p^{(j)}p^{(j')}\}$$

where

$$(3.5.2)\qquad F_{njj'}(u,v) = n^{-1}\sum_{i=1}^{n} c(u - X_{i,j},\ v - X_{i,j'})$$

$$(3.5.3)\qquad F_{njj'h}(u,v) = n^{-1}\sum_{i=1}^{n} c(u - X_{i,j},\ v - X_{i+h,j'})$$

$$(3.5.4)\qquad c(u,v) = \begin{cases} 1 & \text{for } u \geq 0,\ v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

then

$$(3.5.5)\qquad \hat{\nu}_{jj'} \xrightarrow[p]{} \nu_{jj'}\ .$$

Remark: In practice, of course, we take $\mu_n^{(j)} = [np^{(j)}] + 1$ or $[(n + 1)p^{(j)}]$.

Proof..By lemma 2.3.5, as $n \to \infty$, $s_j \in I_n^{(j)}$ and $s_{j'} \in I^{(j')}$ with probability one. Clearly,

$$
\begin{aligned}
(3.5.6)\quad \hat{\nu}_{jj'} - \nu_{jj'} = {} & F_{njj'}(s_j,s_{j'}) - F_{[j,j']}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right) \\
& + \sum_{h=1}^{m_n}\{F_{njj'h}(s_j,s_{j'}) - F_{[j,j']}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right)\} \\
& - \sum_{h=m_n+1}^{\infty}\left\{F_{[j,j']h}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right) - p^{(j)}p^{(j')}\right\} \\
& + \sum_{h=1}^{m_n}\left\{F_{nj'jh}(s_j s_{j'}) - F_{[j',j]}(s_j,s_{j'})\right\} \\
& - \sum_{h=m_n+1}^{\infty}\left\{F_{[j',j]h}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right) - p^{(j)}p^{(j')}\right\}
\end{aligned}
$$

By (2.3.33), the third term and the fifth term in (3.5.6) tend to zero as $n \to \infty$. Also the second term and the fourth term in (3.5.6) are of similar nature. Hence to prove (3.5.5), it is sufficient to show that

$$
(3.5.7)\quad \sup_{\substack{s_j \in I_n^{(j)} \\ s_{j'} \in I_n^{(j')}}} \left|\sum_{h=0}^{m_n}\left\{F_{njj'}(s_j,s_{j'}) - F_{[j,j']h}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right)\right\}\right| \underset{p}{\to} 0\,,
$$

as $n \to \infty$. Actually the left hand side of (3.5.7) is the combination of the first and the second term in (3.5.6).

Again to prove (3.5.7) we see that

$$
\begin{aligned}
(3.5.8)\quad & \sup_{\substack{s_j \in I_n^{(j)} \\ s_{j'} \in I_n^{(j')}}} \left|\sum_{h=0}^{m_n}\{F_{njj'h}(s_j,s_{j'}) - F_{[j,j']h}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right)\}\right| \\
& \le \sup_{\substack{s_j \in I_n^{(j)} \\ s_{j'} \in I_n^{(j')}}} \sum_{h=0}^{m_n} \left|F_{njj'h}(s_j,s_{j'}) - F_{[j,j']h}\left(\xi_{p(j)}^{(j)},\xi_{p(j')}^{(j')}\right)\right|
\end{aligned}
$$

$$\leq \sup_{\substack{s_j \varepsilon I_n^{(j)} \\ s_{j'} \varepsilon I_n^{(j')}}} \sum_{h=0}^{m_n} \Big\{ \left|F_{njj'h}(s_j s_{j'}) - F_{[j,j']h}(s_j, s_{j'})\right| + \left|F_{[j,j']h}(s_j, s_{j'}) - F_{[j,j']h}\left(\xi^{(j)}_{p^{(j)}}, \xi^{(j')}_{p^{(j')}}\right)\right| \Big\}$$

Since with probability one, $s_j \in I_n^{(j)}$ and $s_{j'} \in I_n^{(j')}$, we can write $s_j = \xi^{(j)}_{p^{(j)}} + \theta_{1j} n^{-1/2} \log n$ and $s_{j'} = \xi^{(j')}_{p^{(j')}} + \theta_{2j'} n^{-1/2} \log n$, where $|\theta_{1j}| \leq 1$ and $|\theta_{2j'}| \leq 1$. Then for all s_j, $s_{j'}$,

$$(3.5.9) \quad \left|F_{[j,j']h}(s_j, s_{j'}) - F_{[j,j']h}\left(\xi^{(j)}_{p^{(j)}}, \xi^{(j')}_{p^{(j')}}\right)\right| \leq \left|F_{[j]}(s_j) - F_{[j]}\left(\xi^{(j)}_{p^{(j)}}\right)\right| + \left|F_{[j']}(s_{j'}) - F_{[j']}\left(\xi^{(j')}_{p^{(j')}}\right)\right| = O(n^{-1/2} \log n),$$

with probability one as $n \to \infty$ and use of (3.5.9) in (3.5.8) yields, as $n \to \infty$,

$$(3.5.10) \quad \sup_{\substack{s_j \varepsilon I_n^{(j)} \\ s_{j'} \varepsilon I_n^{(j')}}} \left| \sum_{h=0}^{m_n} \left\{F_{njj'h}(s_j, s_{j'}) - F_{[j,j']h}\left(\xi^{(j)}_{p^{(j)}}, \xi^{(j')}_{p^{(j')}}\right)\right\}\right| \leq \sup_{\substack{s_j \varepsilon I_n^{(j)} \\ s_{j'} \varepsilon I_n^{(j')}}} \sum_{h=0}^{m_n} \left|F_{njj'h}(s_j, s_{j'}) - F_{[j,j']h}(s_j, s_{j'})\right| + O\{n^{-1/2} (\log n)^2\},$$

with probability one. Hence, it is sufficient to show that,

$$(3.5.11) \quad \sup_{\substack{s_j \varepsilon I_n^{(j)} \\ s_{j'} \varepsilon I_n^{(j')}}} \sum_{h=0}^{m_n} \left|F_{njj'h}(s_j, s_{j'}) - F_{[j,j']h}(s_j, s_{j'})\right| \underset{p}{\to} 0,$$

as $n \to \infty$.

To prove (3.5.11) consider a particular n and as in (2.3.50) take $\eta^{(j)}_{r,n} = \xi^{(j)}_{p^{(j)}} + r\, a_n/b_n$, $r = 0, \pm 1, \ldots, \pm b_n$. Since $F_{njj'h}(s_j, s_{j'})$ and $F_{[j,j']h}(s_j, s_{j'})$ are non-decreasing in s_j and $s_{j'}$, for $s_j \in I^{(j)}_{r,n}$ and

$s_{j'} \varepsilon I^{(j')}_{r,n}$,

$$F_{njj'h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) \leq F_{njj'h}(s_j,s_{j'}) \leq F_{njj'h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n})$$

and

$$F_{[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) \leq F_{[j,j']h}(s_j,s_{j'}) \leq F_{[j,j']h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n})$$

Therefore,

$$(3.5.12) \quad F_{njj'h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) - F_{[j,j']h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n})$$
$$\leq F_{njj'}(s_j,s_{j'}) - F_{[j,j']h}(s_j,s_{j'})$$
$$\leq F_{njj'h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n}) - F_{[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})$$

which gives us

$$(3.5.13) \quad |F_{njj'h}(s_j,s_{j'}) - F_{[j,j']h}(s_j,s_{j'})|$$
$$\leq \max_{s=r,r+1} |F_{njj'h}(\eta^{(j)}_{s,n},\eta^{(j')}_{s,n}) - F_{[j,j']h}(\eta^{(j)}_{s,n},\eta^{(j')}_{s,n})|$$
$$+ |F_{[j,j']h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n}) - F_{[j,j']}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})|$$

This implies

$$(3.5.14) \quad \sup_{\substack{s_j \varepsilon I^{(j)}_n \\ s_{j'} \varepsilon I^{(j')}_n}} |F_{njj'h}(s_j,s_{j'}) - F_{[j,j']h}(s_j,s_{j'})|$$
$$\leq \max_{|r| \leq b_n} |F_{njj'h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) - F_{[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})|$$
$$+ \max_{-b_n \leq r \leq b_n - 1} |F_{[j,j']h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n}) - F_{[j,j']}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})|$$

Since for each $r = 0, \pm 1, \ldots, \pm b_n$,

$$(3.5.15) \quad |F_{[j,j']h}(\eta^{(j)}_{r+1,n},\eta^{(j')}_{r+1,n}) - F_{[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})|$$
$$\leq |F_{[j]}(\eta^{(j)}_{r+1,n}) - F_{[j]}(\eta^{(j)}_{r,n})| + |F_{[j']}(\eta^{(j')}_{r+1,n}) - F_{[j']}(\eta^{(j')}_{r,n})|$$
$$= O(a_n/b_n) = O(n^{-3/4} \log n),$$

by the Bonferroni inequality, the second term in (3.5.14) is $O(n^{-1/2}\log n)$ and consequently we get,

$$(3.5.16) \quad \sum_{h=0}^{m_n} \sup_{\substack{s_j \varepsilon I_n^{(j)} \\ s_{j'} \varepsilon I_n^{(j')}}} |F_{njj'h}(s_j, s_{j'}) - F_{[j,j']h}(s_j, s_{j'})|$$

$$\leq \sum_{h=0}^{m_n} \max_{|r| \leq b_n} |F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F_{[j,j']h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})|$$

$$+ O\{n^{-1/2}(\log n)^2\}$$

Hence for (3.5.11), it is only required to show that

$$(3.5.17) \quad \sum_{h=0}^{m_n} \max_{|r| \leq b_n} |F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F_{[j,j']h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})| \underset{p}{\to} 0,$$

as $n \to \infty$.

Let us now write

$$F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) = n^{-1} \sum_{i=1}^{n} c(\eta_{r,n}^{(j)} - Y_{n,i}^{(j)}, \eta_{r,n}^{(j')} - Y_{n,i}^{(j')})$$

where $c(u,v)$ is given by (3.5.4) and let us denote by

$$G_{n[j,j']h}(x,y) = P\left\{|R_{n,i}^{(j)}| \leq x,\ |R_{n,i+h}^{(j')}| \leq y\right\}, \quad \begin{matrix} 0 \leq x < \infty \\ 0 \leq y < \infty \end{matrix}$$

Then,

$$(3.5.18) \quad E|F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})|$$

$$\leq n^{-1} \sum_{i=1}^{n} E|c(\eta_{r,n}^{(j)} - Y_{n,i}^{(j)} - R_{n,i}^{(j)}, \eta_{r,n}^{(j')} - Y_{n,i}^{(j')} - R_{n,i}^{(j')})$$

$$- c(\eta_{r,n}^{(j)} - Y_{n,i}^{(j)}, \eta_{r,n}^{(j')} - Y_{n,i}^{(j')})|$$

$$\leq \int_0^{c_2 n^{-c_1/\delta}} \int_0^{c_2 n^{-c_1/\delta}} P\left\{ \begin{matrix} \eta_{r,n}^{(j)} - x \leq Y_{n,i}^{(j)} \leq \eta_{r,n}^{(j)},\ \eta_{r,n}^{(j')} - y \leq Y_{n,i+h}^{(j')} \leq \eta_{r,n}^{(j')} \\ \text{or} \\ \eta_{r,n}^{(j)} \leq Y_{n,i}^{(j)} \leq \eta_{r,n}^{(j)} + x,\ \eta_{r,n}^{(j')} \leq Y_{n,i+h}^{(j')} \leq \eta_{r,n}^{(j')} + y \end{matrix} \right\} dG_{n[j,j']h}(x,y)$$

$$+P\{|R_{n,i}^{(j)}| \geq c_2 n^{-c_1/\delta}\} + P\{|R_{n,i+h}^{(j')}| \geq c_2 n^{-c_1/\delta}\}$$
$$= O(n^{-2c_1/\delta}) + O(n^{-c_1}), \text{ by Lemma 2.3.1,}$$

where $0 < \delta \leq 1$. We now choose a δ_3: $5/16 < \delta_3 < 3/4$ and as before K such that $c_1 > 4$. This implies $c_1\delta_3 > 5/4 \Rightarrow \sum_{n \geq 1} n^{-c_1\delta_3 + \frac{1}{4}} < \infty$ and also $c_1(1 - \delta_3) > 1$. Thus, by using the Chebychev inequality and (3.5.18), we have, as $n \to \infty$,

$$(3.5.19) \quad P\{|F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})| \geq (K_5 n^{-c_1})^{1-\delta_3}\}$$
$$\leq (K_5 n^{-c_1})^{\delta_3},$$

for all $r = 0, \pm 1, \ldots, \pm b_n$. Hence, by the Bonferroni inequality, as $n \to \infty$,

$$(3.5.20) \quad P\{\max_{|r| \leq b_n} |F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})| \geq K_5^{1-\delta_3} n^{-1}\}$$
$$\leq K_5^{\delta_3} n^{-1-\varepsilon}, \quad \varepsilon > 0.$$

The last equation implies that as $n \to \infty$,

$$(3.5.21) \quad \max_{|r| \leq b_n} |F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})| = O(n^{-1}),$$

with probability one. Also by Lemma 2.3.2, for all $r = 0, \pm 1, \ldots, \pm b_n$,

$$(3.5.22) \quad |\tilde{F}_{n[j,j']h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F_{[j,j']h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})| = O(n^{-d})$$

where we take $d \geq 1/4$, by proper choice of K. Therefore (3.5.21), (3.5.22) and the Bonferroni inequality imply that, as $n \to \infty$,

$$(3.5.23) \quad \sum_{h=0}^{m_n} \max_{|r| \leq b_n} \{|F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F_{[j,j']h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})|$$
$$\leq \sum_{h=0}^{m_n} \max_{|r| \leq b_n} \{|F_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})$$
$$+ |F^*_{njj'h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')}) - \tilde{F}_{n[j,j']h}(\eta_{r,n}^{(j)}, \eta_{r,n}^{(j')})|$$

$$+ \left|\tilde{F}_{n[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) - F_{[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})\right|\Big\}$$

$$\leq O(n^{-1}\log n) + O(n^{-d+\frac{1}{4}}\log n)$$

$$+ \sum_{h=0}^{m_n} \max_{|r|\leq b_n} |F^*_{njj'h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) - \tilde{F}_{n[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})|$$

with probability one. Hence for (3.5.17), it is only required to show that

$$(3.5.24) \quad \sum_{h=0}^{m_n} \max_{|r|\leq b_n} |F^*_{njj'h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n}) - \tilde{F}_{n[j,j']h}(\eta^{(j)}_{r,n},\eta^{(j')}_{r,n})| \underset{p}{\to} 0, \text{ as } n \to \infty.$$

For proof of (3.5.24), we see that for all u,v, we have

$$(3.5.25) \quad E\{F^*_{njj'h}(u,v)\} = n^{-1}\sum_{i=1}^{n} E\, c(u - Y^{(j)}_{n,i},\ v - Y^{(j')}_{n,i+h}) = \tilde{F}_{n[j,j']h}(u,v)$$

and

$$(3.5.26) \quad \mathrm{Var}\{F^*_{njj'h}(u,v)\}$$

$$= n^{-2} E\Big[\sum_{i=1}^{n}\{c(u-Y^{(j)}_{n,i},\ v-Y^{(j')}_{n,i+h}) - \tilde{F}_{n[j,j']h}(u,v)\}\Big]^2$$

$$= n^{-2}\sum_{i=1}^{n}\Big[\{1-\tilde{F}_{n[j,j']h}(u,v)\}^2\tilde{F}_{n[j,j']h} + \tilde{F}^2_{n[j,j']h}(u,v)\{1-\tilde{F}_{n[j,j']h}(u,v)\}\Big]$$

$$+ 2n^{-2}\sum_{h=1}^{m_n}\sum_{i=1}^{n-h} E\{c(u-Y^{(j)}_{n,i},v-Y^{(j')}_{n,i+h})c(u-Y^{(j)}_{n,i+h},v-Y^{(j')}_{n,i+2h}) - \tilde{F}^2_{n[j,j']h}(u,v)\}$$

Since $|c(u-Y^{(j)}_{n,i},v-Y^{(j')}_{n,i+h})c(u-Y^{(j)}_{n,i+h},v-Y^{(j')}_{n,i+2h})-\tilde{F}^2_{n[j,j']h}(u,v)| \leq 1$ a.s., we have

$$E|c(u-Y_{n,i}^{(j)},v-Y_{n,i+h}^{(j')})c(u-Y_{n,i+h}^{(j)},v-Y_{n,i+h}^{(j')}) - \tilde{F}_{n[j,j']h}^{2}(u,v)| \leq 1$$

and this implies

$$\text{(3.5.27)}\quad \text{Var}\{F_{njj'h}^{*}(u,v)\} \leq n^{-1}[\tilde{F}_{n[j,j']h}(u,v)\{1 - \tilde{F}_{n[j,j']h}(u,v)\}] + 2n^{-1}\sum_{h=1}^{m_n}(1 - h/n)$$

$$\leq 1/4n \;+ 2n^{-1} K \log n\{1 - (K \log n + 1)/2n\}$$

$$= O(n^{-1} \log n)$$

We see that (3.5.27) holds for $(\eta_{r,n}^{(j)},\eta_{r,n}^{(j')})$, for all $r = 1,\ldots,b_n$. Hence (3.5.24) also holds, since by the Chebychev inequality, for every $\varepsilon > 0$,

$$P\left\{\sum_{h=0}^{m_n} \max_{|r|\leq b_n} |F_{njj'h}^{*}(\eta_{r,n}^{(j)},\eta_{r,n}^{(j')}) - \tilde{F}_{n[j,j']h}(\eta_{r,n}^{(j)},\eta_{r,n}^{(j')})| \geq \varepsilon\right\}$$

$$\leq \sum_{h=0}^{m_n}\sum_{r=-b_n}^{b_n} P\left\{|F_{njj'h}^{*}(\eta_{r,n}^{(j)},\eta_{r,n}^{(j')}) - \tilde{F}_{n[j,j']h}(\eta_{r,n}^{(j)},\eta_{r,n}^{(j')})| \geq \varepsilon/m_n\right\}$$

$$\leq 2b_n(m_n+1)m_n^2\, V\{F_{njj'h}^{*}(\eta_{r,n}^{(j)},\eta_{r,n}^{(j')})\}/\varepsilon^2$$

$$= O\{n^{-3/4}(\log n)^4\}$$

$$\to 0, \text{ as } n \to \infty. \qquad \text{Q.E.D.}$$

<u>Theorem</u> 3.5.2. <u>Let</u> $\mu_{n,1}^{(j)} = np^{(j)} - c_{1n}(n^{1/2} \log n)$, $\mu_{n,2}^{(j)} = np^{(j)} + c_{2n}(n^{1/2} \log n)$ <u>where</u> c_{1n} <u>and</u> c_{2n} <u>are positive constants such that</u> $(\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)}) \to \infty$ <u>as</u> $n \to \infty$, <u>but</u> c_{1n} <u>and</u> c_{2n} <u>both converge to zero as</u> $n \to \infty$. <u>Also let</u> $s_{j,i}$ <u>be the</u> $\mu_{n,i}^{(j)}$-th <u>smallest observation among</u> $(X_{1,j},\ldots,X_{n,j})$. <u>Then for each</u> $j = 1,\ldots,q$, <u>as</u> $n \to \infty$,

$$\hat{f}_{[j]}(\xi_{p(j)}^{(j)}) = (\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)})/n(s_{j,2} - s_{j,1}) \to f_{[j]}(\xi_{p(j)}^{(j)}) ,$$

with probability one.

Proof . Since $\mu_{n,2}^{(j)}/n = F_{nj}(s_{j,2})$ and $\mu_{n,1}^{(j)}/n = F_{nj}(s_{j,1})$, we can write

$$\begin{aligned} n^{-1}(\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)}) \\ &= F_{nj}(s_{j,2}) - F_{nj}(s_{j,1}) \\ &= F_{[j]}(s_{j,2}) - F_{[j]}(s_{j,1}) \\ &\quad + [\{F_{[j]}(s_{j,1})-p^{(j)}\} + \{F_{nj}(\xi_{p(j)}^{(j)}) - F_{nj}(s_{j,1})\}] \\ &\quad - [\{F_{[j]}(s_{j,2})-p^{(j)}\} + \{F_{nj}(\xi_{p(j)}^{(j)}) - F_{nj}(s_{j,2})\}] \end{aligned}$$

and by Lemma 2.3.5, both $s_{j,2}$ and $s_{j,1}$ belong to $I_n^{(j)}$. Hence by (2.3.3), as $n \to \infty$,

$$n^{-1}(\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)}) = F_{[j]}(s_{j,2}) - F_{[j]}(s_{j,1}) + O(n^{-3/4} \log n),$$

with probability one. This implies, as $n \to \infty$,

$$\begin{aligned} \hat{f}_{[j]}(\xi_{p(j)}^{(j)}) &= (\mu_{n,2}^{(j)} - \mu_{n,1}^{(j)})/n(s_{j,2} - s_{j,1}) \\ &= \lim_{n \to \infty} [\{F_{[j]}(s_{j,2})- F_{[j]}(s_{j,1})\}/(s_{j,2} - s_{j,1})] \\ &= f_{[j]}(\xi_{p(j)}^{(j)}) , \end{aligned}$$

with probability one. Q.E.D.

Now if we write $\hat{\tau}_{jj'} = \hat{\nu}_{jj'}/[\hat{f}_{[j]}(\xi_{p(j)}^{(j)})\hat{f}_{[j']}(\xi_{p(j')}^{(j')})]$, $j,j'=1,\ldots,q$, then due to Theorem 3.5.1 and Theorem 3.5.2, as $n \to \infty$, $\hat{\tau}_{jj'} \xrightarrow{p} \tau_{jj'}$ for each $j,j' = 1,\ldots,q$ and hence a consistent estimate of $\underset{\sim}{T}$, the variance covariance matrix of $\sqrt{n}\,\underset{\sim}{Z}_n$, is $\hat{\underset{\sim}{T}}$, where $\hat{\underset{\sim}{T}} = ((\hat{\tau}_{jj'}))_{j,j'=1,\ldots,q}$.

Again for each $i,i' = 1,\ldots,t$; $j,j' = 1,\ldots,q$, if we write

$$\hat{\nu}_{ii'}^{(jj')} = F_{njj'}(s_{j,i}, s_{j',i'}) - p_i^{(j)} p_{i'}^{(j')} + \sum_{h=1}^{m_n} \{F_{njj'h}(s_{j,i}, s_{j',i'}) - p_i^{(j)} p_{i'}^{(j')}\} + \sum_{h=1}^{m_n} \{F_{nj'jh}(s_{j,i}, s_{j',i'}) - p_i^{(j)} p_{i'}^{(j')}\}$$

where $\mu_{n,i}^{*(j)} = np_i^{(j)} + o(n^{1/2}\log n)$ and $s_{j,i}$ is the $\mu_{n,i}^{*(j)}$-th smallest observation among $(X_{1,j},\ldots,X_{n,j})$, then by the same arguments as in Theorem 3.5.1, it follows for every $i,i' = 1,\ldots,t; j,j' = 1,\ldots,q$, that

$$\hat{\nu}_{ii'}^{(jj')} \underset{p}{\rightarrow} \nu_{ii'}^{(jj')}, \text{ as } n \rightarrow \infty$$

and consequently we have

$$\hat{\eta}_{ii'}^{(jj')} \underset{p}{\rightarrow} \eta_{ii'}^{(jj')}, \text{ as } n \rightarrow \infty,$$

where

$$\hat{\eta}_{ii'}^{(jj')} = \hat{\nu}_{ii'}^{(jj')} / \hat{f}_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right) \hat{f}_{[j']}\left(\xi^{(j')}_{p_{i'}^{(j')}}\right)$$

and $\nu_{ii'}^{(jj')}$ and $\eta_{ii'}^{(jj')}$ are given by (3.3.9) and (3.3.12) respectively. Hence a consistent estimate of the dispersion matrix (3.3.14) is given by $\hat{\underset{\sim}{T}}^* = ((\hat{\tau}^*_{jj'}))_{j,j'=1,\ldots,q}$, where

$$\hat{\tau}^*_{jj'} = \sum_{i=1}^{t} \sum_{i'=1}^{t} a^{(j)}_{p_i^{(j)}} a^{(j')}_{p_{i'}^{(j')}} \hat{\eta}_{ii'}^{(jj')}, \quad j,j'=1,\ldots,q.$$

3.6. Asymptotic Normality of a Linear Compound of Several Quantiles for Random Sample Size.

Let $\{n_r\}$ be an increasing sequence of positive integers tending to infinity, and let $\{N_r\}$ be a sequence of proper random variables taking positive integer values such that $N_r/n_r \to 1$, in probability, as $r \to \infty$. In conformity with the notations used earlier, for sample of size N_r, $X^{(j)}_{N_r,1} < \dots < X^{(j)}_{N_r,N_r}$ are regarded as the ordered random variables for the j-th variate, $j = 1,\dots,q$.

Let us suppose that $Z^{(j)}_{N_r,p_i^{(j)}}$ denotes the sample quantile of order $p_i^{(j)}$ for the j-th variate and $\xi^{(j)}_{p_i^{(j)}}$ be the corresponding population quantile and $\underset{\sim}{Z}_{N_r,i} = \left(Z^{(1)}_{N_r,p_i^{(1)}},\dots,Z^{(q)}_{N_r,p_i^{(q)}}\right)'$, $\underset{\sim}{\xi}_{p_i} = \left(\xi^{(1)}_{p_i^{(1)}},\dots,\xi^{(q)}_{p_i^{(q)}}\right)'$.

We want to determine the distribution of the vector of linear compounds

$$\left(N_r^{1/2}\sum_{i=1}^{t} a^{(1)}_{p_i^{(1)}}\left(Z^{(1)}_{N_r,p_i^{(1)}} - \xi^{(1)}_{p_i^{(1)}}\right),\dots,N_r^{1/2}\sum_{i=1}^{t} a^{(q)}_{p_i^{(q)}}\left(Z^{(q)}_{N_r,p_i^{(q)}} - \xi^{(q)}_{p_i^{(q)}}\right)\right)'$$

as $r \to \infty$ where the $a^{(j)}_{p_i^{(j)}}$'s, $i=1,\dots,t;j=1,\dots,q$, are constants.

Let us assume that $f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right)>0$, $i=1,\dots,t;j=1,\dots,q$, and let us denote by

$$I^{(j)}_{N_r,p_i^{(j)}} = \{x:\xi^{(j)}_{p_i^{(j)}} - a_{N_r} \le x \le \xi^{(j)}_{p_i^{(j)}} + a_{N_r}\}\ ,\quad \begin{matrix} i=1,\dots,t \\ j=1,\dots,q \end{matrix}$$

where $a_{N_r} \sim N_r^{-1/2}\log N_r$ as $r \to \infty$. By Lemma 2.3.5, $Z^{(j)}_{N_r,p_i^{(j)}} \varepsilon I^{(j)}_{N_r,p_i^{(j)}}$ for $i = 1,\dots,t;j = 1,\dots,q$, with probability one, as $r \to \infty$ and under the condition that $\sup_{1\le j\le q}\ \sup_{x} f'_{[j]}(x)$ is finite, Lemma 3 of Bahadur (1966) extends directly to our case.

By (3.2.5) and Lemma 3 of Bahadur (1966) we get that for every $\varepsilon > 0$ there exists an $\eta > 0$ such that as $r \to \infty$,

$$(3.6.1)\quad P\left\{ N_r^{1/2} \left| [Z^{(j)}_{N_r, p_i^{(j)}} - \xi^{(j)}_{p_i^{(j)}}] f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right) + [F_{nj}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}] \right| > \varepsilon \ \text{for at least one } i=1,\ldots,t, j=1,\ldots,q \right\} < \eta$$

This implies that as $r \to \infty$,

$$(3.6.2)\qquad N_r^{1/2}(\underset{\sim}{Z}_{N_{r,i}} - \underset{\sim}{\xi}_{p_i}) \xrightarrow{p} \underset{\sim}{U}_{N_{r,i}} = (U^{(1)}_{N_{r,i}}, \ldots, U^{(q)}_{N_{r,i}})'$$

where

$$(3.6.3)\qquad U^{(j)}_{N_{r,i}} = \frac{\sqrt{N_r}}{f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right)} [F_{N_r j}\left(\xi^{(j)}_{p_i^{(j)}}\right) - p_i^{(j)}]$$

$$= \sum_{u=1}^{N_r} v^{(j)}_{u,i}\left(\xi^{(j)}_{p_i^{(j)}}\right) \Big/ \sqrt{N_r}\, f_{[j]}\left(\xi^{(j)}_{p_i^{(j)}}\right);$$

$$v^{(j)}_{u,i}\left(\xi^{(j)}_{p_i^{(j)}}\right) = c\left(\xi^{(j)}_{p_i^{(j)}} - X_{u,j}\right) - p_i^{(j)}$$

By (2.2.7) and (2.2.8) the sequence $\{\underset{\sim}{X}_u, u=0,\pm 1,\pm 2,\ldots\}$ satisfies the condition of a ϕ-mixing process with $\sum_n \phi_n^{1/2} < \infty$. Again for any $j=1,\ldots,q$, $v^{(j)}_{u,i}\left(\xi^{(j)}_{p_i^{(j)}}\right)$ being a Borel-measurable function of $X_{u,j}$ is also a ϕ-mixing process. Moreover since $X_{u,j}$ decreases exponentially we shall still have $\sum_n \phi_n^{1/2} < \infty$. Also

$$E\, v^{(j)}_{u,i}\left(\xi^{(j)}_{p_i^{(j)}}\right) = 0 \text{ and } \operatorname{var}\left\{v^{(j)}_{u,i}\left(\xi^{(j)}_{p_i^{(j)}}\right)\right\} = p_i^{(j)}(1 - p_i^{(j)}).$$

For any vector $\underset{\sim}{d} = (d_1,\ldots,d_q)'$, we have by (3.6.2) and (3.6.3)

$$(3.6.4)\qquad N_r^{1/2} \sum_{j=1}^{q} \sum_{i=1}^{t} d_j a^{(j)}_{p_i^{(j)}} \left(Z^{(j)}_{N_r, p_i^{(j)}} - \xi^{(j)}_{p_i^{(j)}}\right)$$

$$\underset{p}{\rightarrow} N_r^{-1/2} \sum_{j=1}^{q} \sum_{u=1}^{N_r} \sum_{i=1}^{t} \left\{ d_j a_{p_i}^{(j)}{}_{(j)} / f_{[j]}\left(\xi_{p_i}^{(j)}{}_{(j)}\right)\right\} v_{u,i}^{(j)}\left(\xi_{p_i}^{(j)}{}_{(j)}\right)$$

$$= L, \text{ say,}$$

and similar to (3.3.7), (3.3.9) and (3.3.13), here we also have

$$(3.6.5) \quad \sigma_L^2 = \mathrm{var}(L) = \sum_{j=1}^{q} \sum_{j'=1}^{q} \sum_{i=1}^{t} \sum_{i'=1}^{t} d_j d_{j'} a_{p_i}^{(j)}{}_{(j)} a_{p_{i'}}^{(j')}{}_{(j')} \eta_{ii'}^{(jj')}$$

where $\eta_{ii'}^{(jj')}$ is defined in (3.3.12). Let us denote by

$$(3.6.6) \quad Y_{N_r}(\underset{\sim}{\xi}_p) = N_r^{-1/2} \sigma_L^{-1} \sum_{j=1}^{q} \sum_{u=1}^{N_r} \sum_{i=1}^{t} \left\{ d_j a_{p_i}^{(j)}{}_{(j)} / f_{[j]}\left(\xi_{p_i}^{(j)}{}_{(j)}\right)\right\} v_{u,i}^{(j)}\left(\xi_{p_i}^{(j)}{}_{(j)}\right)$$

Then by Theorem 20.3 of Billingsley (1968),

$$(3.6.7) \quad \mathcal{L}\left\{Y_{N_r}(\underset{\sim}{\xi}_p)\right\} \rightarrow N(0,1), \quad \text{as } r \rightarrow \infty .$$

Hence by (3.6.4), as $r \rightarrow \infty$, the distribution of the vector

$$\left(N_r^{1/2} \sum_{i=1}^{t} a_{p_i}^{(1)}{}_{(1)}\left(Z_{N_r,p_i}^{(1)}{}_{(1)} - \xi_{p_i}^{(1)}{}_{(1)}\right), \ldots, N_r^{1/2} \sum_{i=1}^{t} a_{p_i}^{(q)}{}_{(q)}\left(Z_{N_r,p_i}^{(q)}{}_{(q)} - \xi_{p_i}^{(q)}{}_{(q)}\right)\right)'$$

is multivariate normal with mean vector $\underset{\sim}{0}$ and same dispersion matrix as given in (3.3.14) and (3.3.15).

CHAPTER IV

ASYMPTOTIC PROPERTIES OF THE WILCOXON SIGNED RANK STATISTIC AND RELATED ESTIMATORS FOR MULTIVARIATE STATIONARY AUTOREGRESSIVE PROCESSES

4.1. Introduction

In this chapter, first, the asymptotic multinormality of the Wilcoxon signed rank statistic (vector) for multivariate stationary autoregressive processes is derived. Since the Wilcoxon signed rank statistic can be expressed as a particular case of Hoeffding's (1948) U-statistics, in the next section we extend the above theory for a class of U-statistics. As an application of the earlier results of the chapter, in the last section, the asymptotic distribution of the median of the mid-range estimator (vector) is deduced from the asymptotic multinormality of the Wilcoxon signed rank statistic (vector) and a straight forward multivariate extension of the Hodges and Lehmann's (1963) techniques.

4.2. Asymptotic Properties of the Wilcoxon Signed Rank Statistic

Let $\{\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n\}$ be the chance variables associated with the sample of size n from the process defined in (1.1.1). Let $R_{1,j},\ldots,R_{n,j}$ be the ranks of $|X_{1,j}|,\ldots,|X_{n,j}|$ where $X_{1,j},\ldots,X_{n,j}$ are the sample observations for the j-th variate and let $S_{ij} = 1$ or 0 according as $X_{i,j} >$ or ≤ 0.

Then let us define the vector of Wilcoxon signed rank statistic as $\underset{\sim}{T}_n = (T_{n1},\ldots,T_{nq})'$ where

$$(4.2.1) \qquad T_{nj} = \{n(n+1)\}^{-1} \sum_{i=1}^{n} R_{i,j} S_{ij}, \quad j = 1,\ldots,q.$$

If the empirical c.d.f. for the absolute value of the j-th variate is defined by

$$(4.2.2) \qquad H_{nj}(x) = n^{-1} \sum_{i=1}^{n} c(x - |X_{i,j}|), \; j = 1,\ldots,q,$$

then the Wilcoxon signed rank statistic for the j-th variate can be written as

$$(4.2.3.) \qquad T_{nj} = n(n+1)^{-1} \int_0^\infty H_{nj}(x) dF_{nj}(x)$$

For the m_n-dependent stationary process, where $m_n \sim K \log n$ and K is a positive number, if the empirical c.d.f. for the absolute value of the j-th variate is defined by

$$(4.2.4) \qquad H^*_{nj}(x) = n^{-1} \sum_{i=1}^{n} c(x - |Y^{(j)}_{n,i}|), \quad j = 1,\ldots,q,$$

then the corresponding Wilcoxon signed rank statistic for the j-th variate can be written as

$$(4.2.5) \qquad T^*_{nj} = n(n+1)^{-1} \int_0^\infty H^*_{nj}(x) dF^*_{nj}(x), \quad j = 1,\ldots,q.$$

Let the vector $\underset{\sim}{T}^*_n = (T^*_{n1},\ldots,T^*_{nq})'$.

The true c.d.f. of $|X_{i,j}|$ is defined by

$$(4.2.6) \qquad H_{[j]}(x) = P\{|X_{i,j}| \le x\}, \quad 0 < x < \infty, \quad j = 1,\ldots,q,$$

and for the m_n-dependent process the true c.d.f. of $|Y^{(j)}_{n,i}|$ is denoted by

$$(4.2.7) \qquad \tilde{H}_{n[j]}(x) = P\{|Y^{(j)}_{n,i}| \le x\}, \quad 0 < x < \infty, \quad j = 1,\ldots,q.$$

For the derivation of the asymptotic normality of the Wilcoxon signed rank statistic, let us first prove some lemmas.

Lemma 4.2.1. For every $j = 1,\ldots,q$, as $n \to \infty$,

(4.2.8) $$\sup_x \; n^{1/2}|H_{nj}(x) - H^*_{nj}(x)| = O(n^{-1/4}\log n),$$

with probability one.

Proof. Upon noting that $H_{nj}(x) = F_{nj}(x) - F_{nj}(-x)$ and $H^*_{nj}(x) = F^*_{nj}(x) - F^*_{nj}(-x)$, the proof follows from Lemma 3.2.2.

Lemma 4.2.2. For any $j = 1,\ldots,q$, as $n \to \infty$,

(4.2.9) $$n^{1/2}|T_{nj} - T^*_{nj}| = O(n^{-1/4}\log n),$$

with probability one.

Proof. For any $j = 1,\ldots,q$, we can write by (4.2.5) and integration by parts,

$$\begin{aligned} T_{nj} &= n(n+1)^{-1}\Big[\int_0^\infty (H_{nj}(x) - H^*_{nj}(x))dF_{nj}(x) \\ &\quad + \int_0^\infty H^*_{nj}(x)d(F_{nj}(x)-F^*_{nj}(x)) + \int_0^\infty H^*_{nj}(x)dF^*_{nj}(x)\Big] \\ &= T^*_{nj} + n(n+1)^{-1}\int_0^\infty (H_{nj}(x)-H^*_{nj}(x))dF_{nj}(x) \\ &\quad + n(n+1)^{-1}[H^*_{nj}(x)(F_{nj}(x)-F^*_{nj}(x))]_0^\infty \\ &\quad - n(n+1)^{-1}\int_0^\infty (F_{nj}(x) - F^*_{nj}(x))dH^*_{nj}(x) \\ &= T^*_{nj} + n(n+1)^{-1}\int_0^\infty (H_{nj}(x) - H^*_{nj}(x))dF_{nj}(x) \\ &\quad - n(n+1)^{-1}\int_0^\infty (F_{nj}(x) - F^*_{nj}(x))dH^*_{nj}(x) \end{aligned}$$

so that

$$(4.2.10)\quad n^{1/2}|T_{nj} - T^*_{nj}| \leq n^{3/2}(n+1)^{-1}\int_0^\infty |H_{nj}(x) - H^*_{nj}(x)|dF_{nj}(x)$$
$$+ n^{3/2}(n+1)^{-1}\int_0^\infty |F_{nj}(x)-F^*_{nj}(x)|dH^*_{nj}(x)$$

By Lemma 4.2.1, as $n \to \infty$, the first term in (4.2.12) is of order $(n^{-1/4} \log n)$, with probability one, and by Lemma 3.2.2, as $n \to \infty$, the second term is of order $(n^{-1/4} \log n)$, with probability one. Q.E.D.

Lemma 4.2.3. For large n,

$$(4.2.11)\qquad \mathscr{L}(n^{1/2}[\underset{\sim}{T}^*_n - \underset{\sim}{\tilde{\mu}}_n]) \sim N_q(\underset{\sim}{0},\underset{\sim}{\tilde{\Gamma}}),$$

where

$$(4.2.12)\quad \underset{\sim}{\tilde{\mu}}_n = (\tilde{\mu}_{n1},\ldots,\tilde{\mu}_{nq})', \quad \underset{\sim}{\tilde{\Gamma}} = ((\tilde{\gamma}_{jj'}))_{j,j'=1,\ldots,q},$$

$$(4.2.13)\quad \tilde{\mu}_{nj} = \int_0^\infty \tilde{H}_{n[j]}(x)d\tilde{F}_{n[j]}(x), \quad j = 1,\ldots,q,$$

$$(4.2.14)\quad \tilde{\gamma}_{jj'} = \tilde{\delta}_{jj'} + \sum_{h=1}^{m_n}(1-\frac{h}{n})\tilde{\delta}_{jj'h} + \sum_{h=1}^{m_n}(1-\frac{h}{n})\tilde{\delta}_{j'jh}, \quad j,j'=1,\ldots,q,$$

$$(4.2.15)\quad \tilde{\delta}_{jj'h} = \tilde{\delta}^{(11)}_{jj'h} - \tilde{\delta}^{(12)}_{jj'h} - \tilde{\delta}^{(21)}_{jj'h} + \tilde{\delta}^{(22)}_{jj'h}, \quad h = 0,1,2,\ldots,$$

$$(4.2.16)\quad \tilde{\delta}^{(11)}_{jj'h} = \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(x,y)-\tilde{F}_{n[j]}(x)\tilde{F}_{n[j']}(y)\}d\tilde{F}_{n[j]}(-x)d\tilde{F}_{n[j']}(-y),$$

$$(4.2.17)\quad \tilde{\delta}^{(12)}_{jj'h} = \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(x,-y)-\tilde{F}_{n[j]}(x)\tilde{F}_{n[j']}(-y)\}d\tilde{F}_{n[j]}(-x)d\tilde{F}_{n[j']}(y),$$

$$(4.2.18)\quad \tilde{\delta}^{(21)}_{jj'h} = \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(-x,y) - \tilde{F}_{n[j]}(-x)\tilde{F}_{n[j']}(y)\}d\tilde{F}_{n[j]}(x)d\tilde{F}_{n[j']}(-y),$$

$$(4.2.19)\quad \tilde{\delta}^{(22)}_{jj'h} = \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(-x,-y) - \tilde{F}_{n[j]}(-x)\tilde{F}_{n[j']}(-y)\}d\tilde{F}_{n[j]}(x)d\tilde{F}_{n[j']}(y).$$

Proof. We can write for $j = 1,\ldots,q$,

$$T^*_{nj} = \int_0^\infty H^*_{nj}(x)dF^*_{nj}(x) + O(n^{-1})$$

$$= \int_0^\infty (H^*_{nj}(x) - \tilde{H}_{n[j]}(x))dF^*_{nj}(x) + \int_0^\infty \tilde{H}_{n[j]}(x)d(F^*_{nj}(x)-\tilde{F}_{n[j]}(x)$$

$$+ \int_0^\infty \tilde{H}_{n[j]}(x)d\tilde{F}_{n[j]}(x) + O(n^{-1}).$$

For every $j = 1,\ldots,q$, if we write

$$R_{nj} = \int_0^\infty (H^*_{nj}(x) - \tilde{H}_{n[j]}(x))d(F^*_{nj}(x) - \tilde{F}_{n[j]}(x)), \tag{4.2.20}$$

then by (4.2.13), integration by parts and the relations $\tilde{H}_{n[j]}(x) = \tilde{F}_{n[j]}(x) - \tilde{F}_{n[j]}(-x)$ and $H^*_{nj}(x) = F^*_{nj}(x) - F^*_{nj}(-x)$, we have

$$\begin{aligned} T^*_{nj} &= \tilde{\mu}_{nj} + [\tilde{H}_{n[j]}(x)(F^*_{nj}(x) - \tilde{F}_{n[j]}(x))]_0^\infty \\ &\quad - \int_0^\infty (F^*_{nj}(x) - \tilde{F}_{n[j]}(x))d\tilde{H}_{n[j]}(x) \\ &\quad + \int_0^\infty (H^*_{nj}(x) - \tilde{H}_{n[j]}(x))d\tilde{F}_{n[j]}(x) + R_{nj} + O(n^{-1}) \\ &= \tilde{\mu}_{nj} + \int_0^\infty (F^*_{nj}(x) - \tilde{F}_{n[j]}(x))d\tilde{F}_{n[j]}(-x) \\ &\quad - \int_0^\infty (F^*_{nj}(-x) - \tilde{F}_{n[j]}(-x))d\tilde{F}_{n[j]}(x) \\ &\quad + R_{nj} + O(n^{-1}) \end{aligned} \tag{4.2.21}$$

so that

$$n^{1/2}(T^*_{nj} - \tilde{\mu}_{nj}) = n^{-1/2} \sum_{i=1}^{n} B(Y^{(j)}_{n,i}) + n^{1/2}R_{nj} + O(n^{-1/2}) \tag{4.2.22}$$

where

$$B(Y^{(j)}_{n,i}) = B_1(Y^{(j)}_{n,i}) - B_2(Y^{(j)}_{n,i}), \tag{4.2.23}$$

$$B_1(Y^{(j)}_{n,i}) = \int_0^\infty \{c(x-Y^{(j)}_{n,i}) - \tilde{F}_{n[j]}(x)\}d\tilde{F}_{n[j]}(-x) , \tag{4.2.24}$$

and

$$B_2(Y^{(j)}_{n,i}) = \int_0^\infty \{c(-x-Y^{(j)}_{n,i}) - \tilde{F}_{n[j]}(-x)\}d\tilde{F}_{n[j]}(x) . \tag{4.2.25}$$

To show that the remainder term $n^{1/2}R_{nj}$ converges in probability to zero, we write by (4.2.20),

$$n^{1/2}R_{nj} = n^{-3/2} \sum_{i=1}^{n} \sum_{\ell=1}^{n} \int_0^\infty \{c(x-|Y_{n,i}^{(j)}|) - \tilde{H}_{n[j]}(x)\} d\{c(x-Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\} \tag{4.2.26}$$

which implies by Fubini's theorem and the fact that $Y_{n,i}^{(j)}$'s are m_n-dependent,

$$\begin{aligned}
E(n^{1/2}R_{nj})^2 = {} & n^{-3} \sum_{i=1}^{n} \sum_{\ell=1}^{n} \int_0^\infty E[\{c(x-|Y_{n,i}^{(j)}|) - \tilde{H}_{n[j]}(x)\}^2 d\{c(x-Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\}] \\
& + n^{-3} \sum_{h=1}^{m_n} \sum_{i=1}^{n-h} \sum_{\ell=1}^{n} \int_0^\infty E[\{c(x-|Y_{n,i}^{(j)}|) - \tilde{H}_{n[j]}(x)\} \\
& \qquad \{c(x-|Y_{n,i+h}^{(j)}|) - \tilde{H}_{n[j]}(x)\} \\
& \qquad d\{c(x-Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\}] \\
& + n^{-3} \sum_{i=1}^{n} \sum_{h'=1}^{m_n} \sum_{\ell=1}^{n-h'} \int_0^\infty \int_0^\infty E[\{c(x-|Y_{n,i}^{(j)}|) - \tilde{H}_{n[j]}(x)\} \\
& \qquad \{c(y-|Y_{n,i}^{(j)}|) - \tilde{H}_{n[j]}(y)\} \\
& \qquad d\{c(x-Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\} d\{c(y-Y_{n,\ell+h'}^{(j)}) - \tilde{F}_{n[j]}(y)\}] \\
& + n^{-3} \sum_{h=1}^{m_n} \sum_{i=1}^{n-h} \sum_{h'=1}^{m_n} \sum_{\ell=1}^{n-h'} \int_0^\infty \int_0^\infty E[\{c(x-|Y_{n,i}^{(j)}|) \\
& \qquad - \tilde{H}_{n[j]}(x)\}\{c(y-|Y_{n,i+h}^{(j)}|) - \tilde{H}_{n[j]}(y)\} \\
& \qquad d\{c(x-Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\} d\{c(y-Y_{n,\ell+h'}^{(j)}) - \tilde{F}_{n[j]}(y)\}]
\end{aligned} \tag{4.2.27}$$

where we note that for any $i,i',\ell,\ell' = 1,2,\ldots$ and $h,h' > m_n$,

$$\int_0^\infty\int_0^\infty E[\{c(x-|Y_{n,i}^{(j)}|-\tilde{H}_{n[j]}(x)\}\{c(y-|Y_{n,i+h}^{(j)}|) - \tilde{H}_{n[j]}(y)\}$$

$$d\{c(x-Y_{n,\ell}^{(j)})-\tilde{F}_{n[j]}(x)\}d\{c(y-Y_{n,\ell+h'}^{(j)})-\tilde{F}_{n[j]}(y)\}] = 0.$$

Using the fact that the absolute value of each of the integrands in (4.2.27) is bounded by 1 and the inequalities

$$(4.2.28)\qquad E|d\{c(x - Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\}| \leq 2d\tilde{F}_{n[j]}(x),\ \ell = 1,\ldots,n,$$

and

$$(4.2.29)\qquad E|d\{c(x - Y_{n,\ell}^{(j)}) - \tilde{F}_{n[j]}(x)\}d\{c(y-Y_{n,\ell+h'}^{(j)}) - \tilde{F}_{n[j]}(y)\}|$$

$$\leq d\tilde{F}_{n[j,j]h'}(x,y)+3d\tilde{F}_{n[j]}(x)d\tilde{F}_{n[j]}(y),\quad \begin{array}{l}\ell=1,\ldots,n-h',\\ h'=1,\ldots,m_n,\end{array}$$

we have,

$$(4.2.30)\qquad E(n^{1/2} R_{nj})^2 = O\{n^{-1}(\log n)^2\}$$

Hence (4.2.30) and the Chebyshev inequality imply that, as $n\to\infty$,

$$(4.2.31)\qquad n^{1/2} R_{nj} \underset{p}{\to} 0 .$$

Now using the definitions (2.3.9), (2.3.11) and the Fubini theorem we see that for every $i = 1,\ldots,n$; $h = 0,1,\ldots,m_n$, and $j,j' = 1,\ldots,q$,

$$E\{B_s(Y_{n,i}^{(j)})\} = 0,\quad s = 1,2$$

which imply

$$(4.2.32)\qquad E\ \{B(Y_{n,i}^{(j)})\} = 0$$

and

$$E\{B_s(Y_{n,i}^{(j)})B_{s'}(Y_{n,i}^{(j)})\}, \qquad s,s' = 1,2$$

$$= E\int_0^\infty\int_0^\infty \{c((-1)^{s-1}x - Y_{n,i}^{(j)}) - \tilde{F}_{n[j]}((-1)^{s-1}x)\}$$

$$\{c((-1)^{s'-1}y - Y_{n,i}^{(j')}) - \tilde{F}_{n[j']}((-1)^{s-1}y)\}$$

$$d\tilde{F}_{n[j]}((-1)^s x)\, d\tilde{F}_{n[j']}((-1)^{s'}y)$$

$$= \int_0^\infty \int_0^\infty \{\tilde{F}_{n[j,j']h}((-1)^{s-1}x, (-1)^{s'-1}y)$$

$$-\tilde{F}_{n[j]}((-1)^{s-1}x)\tilde{F}_{n[j']}((-1)^{s'-1}y)\}$$

$$d\tilde{F}_{n[j]}((-1)^s x)\, d\tilde{F}_{n[j]}((-1)^{s'}y)$$

$$= \tilde{\delta}_{jj'h}^{(ss')}, \text{ by (4.2.16) - (4.2.19)},$$

so that

$$(4.2.33) \qquad E\{B(Y_{n,i}^{(j)})B(Y_{n,i+h}^{(j')})\} = \tilde{\delta}_{jj'h}^{(11)} - \tilde{\delta}_{jj'h}^{(12)} - \tilde{\delta}_{jj'h}^{(21)} + \tilde{\delta}_{jj'h}^{(22)}$$

$$= \tilde{\delta}_{jj'h}, \text{ by (4.2.15)}.$$

Here note that if $h = 0$ and $j = j'$ then,

$$(4.2.34) \qquad \tilde{\delta}_{jj} = E\{B^2(Y_{n,i}^{(j)})\} = \tilde{\delta}_{jj}^{(11)} + \tilde{\delta}_{jj}^{(22)} - 2\tilde{\delta}_{jj}^{(21)}.$$

Since $B(Y_{n,i}^{(j)})$'s are m_n-dependent, the dispersion matrix of $\{n^{-1/2}\sum_{i=1}^{n} B(Y_{n,i}^{(1)}),\ldots,n^{-1/2}\sum_{i=1}^{n} B(Y_{n,i}^{(q)})\}$ is given by

$$\underset{\sim}{\tilde{\Gamma}} = ((\tilde{\gamma}_{jj'}))_{j,j'=1,\ldots,q},$$

where

$$\tilde{\gamma}_{jj'} = n^{-1}[\sum_{i=1}^{n} \text{cov}\{B(Y_{n,i}^{(j)})\ B(Y_{n,i}^{(j')})\} + \sum_{h=1}^{m_n}\sum_{i=1}^{n-h} \text{cov}\{B(Y_{n,i}^{(j)})B(Y_{n,i+h}^{(j')})\}$$

$$+ \sum_{h=1}^{m_n} \sum_{i=1}^{n-h} \operatorname{cov}\{B(Y_{n,i}^{(j')})B(Y_{n,i+h}^{(j)})\}]$$

$$= \tilde{\delta}_{jj'} + \sum_{h=1}^{m_n} (1 - \frac{h}{n})\tilde{\delta}_{jj'h} + \sum_{h=1}^{m_n} (1 - \frac{h}{n})\tilde{\delta}_{j'jh}$$

Now for any $j = 1,\ldots,q$, the $B(Y_{n,i}^{(j)})$'s are bounded valued random variables and they satisfy the condition of a ϕ-mixing process with $\phi_n = 0$ for $n/\log n > K$. Hence by a multivariate extension of the central limit theorem for strongly mixing processes by Rosenblatt (1956 b), the lemma follows. Q.E.D.

Theorem 4.2.1. Under (2.2.8) and (2.2.10) as $n \to \infty$,

(4.2.35) $$\mathcal{L}(n^{1/2}[\underset{\sim}{T}_n - \underset{\sim}{\mu}]) \to N_q(\underset{\sim}{0},\underset{\sim}{\Gamma})$$

where

(4.2.36) $$\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_q)', \quad \underset{\sim}{\Gamma} = ((\gamma_{jj'}))_{j,j'=1,\ldots,q},$$

and μ_j and $\gamma_{jj'}$ are defined in the same way as $\tilde{\mu}_{nj}$ and $\tilde{\gamma}_{jj'}$ in (4.2.13) and (4.2.14) - (4.2.19) with $\tilde{H}_{n[j]}(x)$, $\tilde{F}_{n[j]}(x)$ and $\tilde{F}_{n[j,j']h}(x,y)$ replaced by $H_{[j]}(x)$, $F_{[j]}(x)$ and $F_{[j,j']h}(x,y)$ respectively.

Proof. Upon noting that $H_{[j]}(x) = F_{[j]}(x) - F_{[j]}(-x)$ and $\tilde{H}_{n[j]}(x) = \tilde{F}_{n[j]}(x) - \tilde{F}_{n[j]}(-x)$, under (2.2.8) and (2.2.10), we have by (3.2.1), as $n \to \infty$,

(4.2.37) $$\sup_x |H_{[j]}(x) - \tilde{H}_{n[j]}(x)| = O(n^{-d}), \quad d \geq 1, \; j = 1,\ldots,q.$$

Also by a straight forward extension of (2.3.16), similar to the extension of (2.3.15) to (3.2.1), we have under (2.2.8) and (2.2.10), as $n \to \infty$,

$$(4.2.38)\quad \sup_{h=0,1,\ldots,m_n} \{\sup_{x,y} |F_{[j,j']h}(x,y)-F_{n[j,j']h}(x,y)|\} = O(n^{-d}),$$

$$d \geq 1,$$

$$j,j' = 1,\ldots,q.$$

For every $j = 1,\ldots,q$, we can write

$$(4.2.39)\quad \tilde{\mu}_{nj} = \int_0^\infty \tilde{H}_{n[j]}(x)dF_{[j]}(x) + \int_0^\infty \tilde{H}_{n[j]}(x)d(\tilde{F}_{n[j]}(x) - F_{[j]}(x))$$

First integrating by parts and then using (3.2.1), it is seen that the second term in (4.2.39) is of order (n^{-d}), $d \geq 1$, as $n \to \infty$. By (4.2.37) and the dominated convergence theorem, as $n \to \infty$,

$$\int_0^\infty \tilde{H}_{n[j]}(x)dF_{[j]}(x) \to \int_0^\infty H_{[j]}(x)dF_{[j]}(x) = \mu_j,$$

for every $j = 1,\ldots,q$. Hence by (4.2.39) and the above arguments, as $n \to \infty$,

$$(4.2.40)\quad \tilde{\underline{\mu}}_n \to \underline{\mu}$$

where for each $j = 1,\ldots,q$, μ_j is defined similar to $\tilde{\mu}_{nj}$ with $\tilde{H}_{n[j]}(x)$ and $\tilde{F}_{n[j]}(x)$ replaced by $H_{[j]}(x)$ and $F_{[j]}(x)$ respectively.

Again for every $j,j' = 1,\ldots,q$ we can write

$$(4.2.41)\quad \tilde{\delta}^{(11)}_{jj'h} = \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(x,y)-\tilde{F}_{n[j]}(x)\tilde{F}_{n[j']}(y)\}\, dF_{[j]}(-x)dF_{[j]}(-y)$$

$$+ \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(x,y)-\tilde{F}_{n[j]}(x)\tilde{F}_{n[j']}(y)\}\, d(\tilde{F}_{n[j]}(-x)-F_{[j]}(-x))dF_{[j]}(-y)$$

$$+ \int_0^\infty\int_0^\infty \{\tilde{F}_{n[j,j']h}(x,y) - \tilde{F}_{n[j]}(x)\tilde{F}_{n[j']}(y)\}$$

$$d\tilde{F}_{n[j]}(-x)d(\tilde{F}_{n[j]}(-y) - F_{[j]}(-y))$$

Integrating by parts and then using (3.2.1), it is seen that each of the second and third terms in (4.2.41) is of order (n^{-d}), $d \geq 1$, as $n \to \infty$. Also by (3.2.1), (4.2.38) and the dominated convergence theorem, as $n \to \infty$, the first term in (4.2.41) tends to $\delta^{(11)}_{jj'h}$ where $\delta^{(11)}_{jj'h}$ is defined as $\tilde{\delta}^{(11)}_{jj'h}$ with $\tilde{F}_{n[j]}(x)$ and $\tilde{F}_{n[j,j']h}(x,y)$ replaced by $F_{[j]}(x)$ and $F_{[j,j']h}(x,y)$ respectively. By similar arguments, as $n \to \infty$, $\tilde{\delta}^{(12)}_{jj'h} \to \delta^{(12)}_{jj'h}$, $\tilde{\delta}^{(21)}_{jj'h} \to \delta^{(21)}_{jj'h}$ and $\tilde{\delta}^{(22)}_{jj'h} \to \delta^{(22)}_{jj'h}$, where $\delta^{(12)}_{jj'h}$, $\delta^{(21)}_{jj'h}$ and $\delta^{(22)}_{jj'h}$ are defined similarly. Therefore, for every $j,j' = 1,\ldots,q$, $h = 0,1,\ldots$, as $n \to \infty$,

$$(4.2.42) \qquad \tilde{\delta}_{jj'h} \to \delta_{jj'h}$$

where $\delta_{jj'h}$ is defined similar to $\tilde{\delta}_{jj'h}$ with $\tilde{F}_{n[j]}(x)$ and $\tilde{F}_{n[j,j']h}(x,y)$ replaced by $F_{[j]}(x)$ and $F_{[j,j']h}(x,y)$ respectively. For $h = 0$ and $j = j'$, also similarly, as $n \to \infty$,

$$(4.2.43) \qquad \tilde{\delta}_{jj} \to \delta_{jj}$$

where δ_{jj} is defined similar to $\tilde{\delta}_{jj}$ with the above replacement.

Further, the fact that $|\tilde{F}_{n[j,j']h}(x,y)-\tilde{F}_{n[j]}(x)\tilde{F}_{n[j']}(y)| \leq 1$, for all x,y, gives us $|\tilde{\delta}_{jj'h}| \leq 4$, for $j,j' = 1,\ldots,q$, $h = 0,1,\ldots,m_n$, which in turn implies

$$(4.2.44) \qquad |\sum_{h=1}^{m_n} (h/n)\tilde{\delta}_{jj'h}| \leq 2n^{-1}m_n(m_n+1) \to 0, \quad \text{as } n \to \infty,$$

for every $j,j' = 1,\ldots,q$.

Hence (4.2.42) - (4.2.44) imply that

$$(4.2.45) \qquad \tilde{\Sigma} \to \Sigma, \text{ as } n \to \infty,$$

where $\underset{\sim}{\Gamma}$ is defined similar to $\tilde{\underset{\sim}{\Gamma}}$ with the above replacement.

Now the theorem follows readily from (4.2.40),(4.2.45), Lemma 4.2.2 and Lemma 4.2.3. Q.E.D.

<u>Remark</u>: In Lemma 4.2.3 and Theorem 4.2.1 we have proved the asymptotic normality for fixed F, the marginal distribution of the vector $\underset{\sim}{X}_i$. If we consider a double sequence $\underset{\sim}{U}_{n,1}, \underset{\sim}{U}_{n,2},\ldots,\underset{\sim}{U}_{n,n}$, defined for each n, where $\underset{\sim}{U}_{n,i} = \underset{\sim}{U}_n(\underset{\sim}{X}_i)$, $i = 1,\ldots,n$ and $\underset{\sim}{U}_{n,i}$ has distribution depending on n then Lemma 4.2.3 and Theorem 4.2.1 also hold provided the dispersion matrix $\underset{\sim}{\Gamma}_n$ defined on $\underset{\sim}{U}_{n,i}$ is positive definite in the sense that

(4.2.46) $\liminf_n \{\text{minimum eigenvalue of } \underset{\sim}{\Gamma}_n\} \geq c > 0,$

since then for a vector $\underset{\sim}{\ell} \neq \underset{\sim}{0}$,

$$E|\underset{\sim}{\ell}'\underset{\sim}{U}_{n,i}|^2 = \underset{\sim}{\ell}'E(\underset{\sim}{U}_{n,i}\underset{\sim}{U}_{n,i}')\underset{\sim}{\ell}$$

$$= \underset{\sim}{\ell}'\underset{\sim}{\Gamma}_n\underset{\sim}{\ell} > c||\underset{\sim}{\ell}|| > 0 ,$$

so that if

$$E|\underset{\sim}{\ell}'\underset{\sim}{U}_{n,i}|^3 \leq \sum_{j=1}^{q} |\ell_j{}^3|E|U_{n,i}^{(j)}| < \infty ,$$

then Liapounov's condition (see Loève, 1963, p. 275) is automatically satisfied. In our case $B(Y_{n,i}^{(j)})$'s being bounded valued random variables, Liapounov's condition (see Loève, 1963, p. 275) is automatically satisfied whenever the dispersion matrix is p.d. in the above sense and this implies the condition (2) of Theorem 3 of

Gnedenko and Kolmogorov (1967, pp. 101-103). Hence the lemma and the theorem hold whenever the dispersion matrix is p.d. in the above sense.

4.3. Asymptotic Properties of a Class of U-statistics.

It is well known that the Wilcoxon signed rank statistic (vector) can be viewed on as a Hoeffding's (1948) U-statistic. Under certain regularity conditions, the asymptotic normality of the Wilcoxon signed rank statistic obtained in section 4.2 is extended here to a class of U-statistics for stationary autoregressive processes. These regularity conditions are somewhat more restrictive than those in Hoeffding (1948) dealing with independent observations, and appear to be necessary in view of the infinite chain of dependence in the series of observations. It is intended to follow up the general case in the near future. However, for the sake of completeness, we will summarize the results on U-statistics obtained in this chapter.

We may note that in (2.2.7) $\underset{\sim}{X}_t$ is expressed as a linear compound of $\underset{\sim}{\varepsilon}_{t-r}$, $r = 0,1,\ldots,\infty$. Essentially Hannan (1961) and Eicker (1965) considered a similar truncation of this linear compound as is done in (2.3.5) and approximated the linear function of the $\underset{\sim}{X}_t$'s by the corresponding function of $\underset{\sim}{Y}_{n,t}$'s. The latter function is again linear in the independent errors $\underset{\sim}{\varepsilon}_t$ over a certain number of terms depending on n, and thereby they were able to use the C.L.T. But the kernel $\phi(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_r)$ of a U-statistic is not, in general, a linear function of its arguments. Therefore the decomposition (2.3.5) is not adaptable. The non-linear nature of the U-statistic and the chain of dependence of the successive observations makes it quite complicated to reduce the autoregressive process to an m_n-dependent process in which the results of Sen (1963) can be incorporated and extended to prove our desired results. This, however, is done here through the following conditions.

i) For every $\varepsilon > 0$, we can find a positive number A such that if we define the qr-dimensional rectangle

$$A^* = \{\underset{\sim}{x}_{\alpha_1},\ldots,\underset{\sim}{x}_{\alpha_r} : |x_{\alpha_{i,j}}| < A, \text{ for } i=1,\ldots,r;\ j=1,\ldots,q\}$$

and $A^c = R^{qr} - A^*$, i.e., the complement of A^* and I_{A^c} is the indicator function of A^c, then

$$(4.3.1) \qquad \sup_n \sup_{1\le\alpha_1<\ldots<\alpha_r\le n} E[I_{A^c}|\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})|^3] < \varepsilon$$

and

ii)

$$(4.3.2) \qquad \sup_n \sup_{1\le\alpha_1< \quad <\alpha_r\le n} E[|\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})|^3] < \infty.$$

For a finite time interval $T_n = \{t: 1 \le t \le n\}$, let $\{\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n\}$ be the ramdom variables associated with the sample of size n from the process defined in (1.1.1). In addition to satisfying the above assumptions, let the statistic $\phi(\underset{\sim}{X}_{\alpha_1},\ldots,\underset{\sim}{X}_{\alpha_r})$ be symmetric in the arguments $\underset{\sim}{X}_{\alpha_1},\ldots,\underset{\sim}{X}_{\alpha_r}; \alpha_1<\ldots<\alpha_r$ and let us define the U-statistic based on the sample of size n as

$$(4.3.3) \qquad U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) = \binom{n}{r}^{-1} \sum_S \phi(\underset{\sim}{X}_{\alpha_1},\ldots,\underset{\sim}{X}_{\alpha_r})$$

where the summation S extends over all $1 \le \alpha_1 < \ldots < \alpha_r \le n$. The U-statistic for the m_n-dependent process, where $m_n \sim K \log n$ and K is a positive number, is defined as

$$(4.3.4) \qquad U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}) = \binom{n}{r}^{-1} \sum_{S} \phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r}).$$

Let us define

$$(4.3.5) \qquad g(\tilde{F}_n) = E\{\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r}) | \alpha_{i+1}-\alpha_i > m_n, \text{ for } i = 1,\ldots,r-1\}$$

and

$$(4.3.6) \qquad g(\tilde{F}_n | \ell, v_1,\ldots,v_\ell) = E\{\phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})\}$$

where $\alpha_{i+1}-\alpha_i = v_i$, $0 < v_i \leq m_n$ for $i = i_1 \neq \ldots \neq i_\ell$, while for the remaining $r - \ell - 1$ values of i, $\alpha_{i+1} - \alpha_i > m_n$ for $i = 1,2,\ldots,r-1$. Obviously $g(\tilde{F}_n) = g(\tilde{F}_n | 0)$.

The symmetric estimator $U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})$ based on $\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}$ is defined by

$$(4.3.7) \qquad U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}) = \binom{n-rm_n + m_n}{r}^{-1} \sum_{S_0} \phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})$$

where the summation S_0 extends over all possible $\binom{n-rm_n + m_n}{r}^{-1}$ sets of $\alpha_1,\ldots,\alpha_r$ satisfying $\alpha_{i+1}-\alpha_i > m_n$ for $i = 1,\ldots,r-1$.

The following two assumptions are made concerning the kernel $\phi(\underset{\sim}{x}_{\alpha_1},\ldots,\underset{\sim}{x}_{\alpha_r})$, defined on R^{qr}.

Assumption (A): It is assumed that the kernel $\phi(\underset{\sim}{x}_{\alpha_1},\ldots,\underset{\sim}{x}_{\alpha_r})$ is continuous everywhere.

Assumption (B): Further assume the following Lipschitz type condition, namely, there exists a $\delta_0 > 0$, a function $g(\cdot)$ on R^{qr} and a constant M such that if $\underset{\sim}{x}_{\underset{\sim}{\alpha}} = (\underset{\sim}{x}'_{\alpha_1},\ldots,\underset{\sim}{x}'_{\alpha_r})'$ and $\underset{\sim}{x}^*_{\underset{\sim}{\alpha}} = (\underset{\sim}{x}^{*\prime}_{\alpha_1},\ldots,\underset{\sim}{x}^{*\prime}_{\alpha_r})'$ be any two points on R^{qr} with $||\underset{\sim}{x}_{\underset{\sim}{\alpha}} - \underset{\sim}{x}^*_{\underset{\sim}{\alpha}}|| < \delta_0$ then

$$(4.3.8) \qquad |\phi(\underset{\sim}{x}_{\underset{\sim}{\alpha}}) - \phi(\underset{\sim}{x}^*_{\underset{\sim}{\alpha}})| \leq ||\underset{\sim}{x}_{\underset{\sim}{\alpha}} - \underset{\sim}{x}^*_{\underset{\sim}{\alpha}}||g(\underset{\sim}{x}_{\underset{\sim}{\alpha}}),$$

and

$$Eg(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) \leq M, \text{for all } \underset{\sim}{\alpha}, \text{ all } n$$

Lemma 4.3.1. Under (2.2.6), (2.2.8), (2.2.10) and the assumptions (A) and (B), as $n \to \infty$,

$$(4.3.9) \qquad n^{1/2}\{U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})\} \underset{p}{\to} 0.$$

Proof. Under the notations defined above,

$$\phi(\underset{\sim}{x}_{\alpha_1},\ldots,\underset{\sim}{x}_{\alpha_r}) \text{ and } \phi(\underset{\sim}{y}_{n,\alpha_1},\ldots,\underset{\sim}{y}_{n,\alpha_r})$$

can be viewed on as functions of the $rq \times 1$ vectors $\underset{\sim}{x}_{\underset{\sim}{\alpha}}$ and $\underset{\sim}{y}_{n,\underset{\sim}{\alpha}}$ respectively. Hence by (4.3.3), (4.3.4), (4.3.8) and (2.3.5), for every $\varepsilon > 0$,

$$(4.3.10) \qquad P\{n^{1/2}|U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - U(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})| > \varepsilon\}$$

$$\leq P\{n^{1/2}|U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - U(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})| > \varepsilon, \sup_{1\leq\alpha_1<\ldots<\alpha_r\leq n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| < \delta_0\}$$

$$+ P\{\sup_{1\leq\alpha_1<\ldots<\alpha_r\leq n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| > \delta_0\}$$

$$\leq P\{n^{1/2}\binom{n}{r}^{-1} \sum_S |\phi(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) - \phi(\underset{\sim}{Y}_{n,\underset{\sim}{\alpha}})| > \varepsilon, \sup_{1\leq\alpha_1<\ldots<\alpha_r\leq n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| < \delta_0\}$$

$$+ P\{\sup_{1\leq\alpha_1<\ldots<\alpha_r\leq n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| > \delta_0\}$$

$$\leq P\{n^{1/2}\binom{n}{r}^{-1} \sum_S ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}||g(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) > \varepsilon\}$$

$$+ P\{\sup_{1\le\alpha_1<\cdots<\alpha_r\le n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| > \delta_0\}$$

$$\le P\{n^{1/2} \sup_{1\le\alpha_1<\cdots<\alpha_r\le n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| \binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) > \varepsilon\}$$

$$+ P\{\sup_{1\le\alpha_1<\cdots<\alpha_r\le n} ||\underset{\sim}{X}_{\underset{\sim}{\alpha}} - \underset{\sim}{Y}_{n,\underset{\sim}{\alpha}}|| > \delta_0\}$$

$$\le P\{n^{1/2} \sup_{1\le\alpha_1<\cdots<\alpha_r\le n} \left(\sum_{i=1}^{r}\sum_{j=1}^{q} R_{n,\alpha_i}^{(j)2}\right)^{1/2} \binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) > \varepsilon\}$$

$$P\{\sup_{1\le\alpha_1<\cdots<\alpha_r\le n} \left(\sum_{i=1}^{r}\sum_{j=1}^{q} R_{n,\alpha_i}^{(j)^2}\right)^{1/2} > \delta_0\}$$

$$\le P\{(nrq)^{1/2} \sup_{i=1,\ldots,n} \max_{j=1,\ldots,q} |R_{n,i}^{(j)}| \binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) > \varepsilon\}$$

$$+ P\{(rq)^{1/2} \sup_{i=1,\ldots,n} \max_{j=1,\ldots,q} |R_{n,i}^{(j)}| > \delta_0\}$$

The first thing to be noted is that, since $E[\binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}})] \le M < \infty$, there exist a small $\eta(>0)$ and a large K_η such that

(4.3.11) $P\{\binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) > K\} \le \frac{M}{K} < \eta$ for $K > K_\eta$,

which implies that $\binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}})$ is bounded in probability. Secondly, by Lemma 2.3.1, a constant c_1 (>0) can be so chosen that

(4.3.12) $\quad P\{|R_{n,\alpha_i}^{(j)}| \ge c_2 n^{-c_1/\delta}\} \le c_3 n^{-c_1}, \quad c_3 < \infty,$

where $0 < \delta \le 1$ and c_2, c_3 are also positive constants and this implies that

(4.3.13) $\quad P\{\sup_{i=1,\ldots,n} \max_{j=1,\ldots,q} |R_{n,i}^{(j)}| \ge c_2 n^{-c_1/\delta}\}$

$$\leq P\left\{\bigcup_{\substack{i=1,\ldots,n\\ j=1,\ldots,q}} \left(|R_{n,i}^{(j)}| \geq c_2 n^{-c_1/\delta}\right)\right\}$$

$$\leq \sum_{i=1}^{n} \sum_{j=1}^{q} P\left\{|R_{n,i}^{(j)}| \geq c_2 n^{-c_1/\delta}\right\}$$

$$\leq qc_3 n^{-(c_1-1)} .$$

In (4.3.13) let us choose $c_1 > 1$ and, since $0 < \delta \leq 1$, this implies $c_1/\delta > 1$. Then by (4.3.13), $(nrq)^{1/2} \sup_{i=1,\ldots,n} \max_{j=1,\ldots,q} |R_{n,i}^{(j)}|$ is also bounded in probability. This and (4.3.11) imply that the first term in (4.3.10), i.e.,

$$P\{(nrq)^{1/2} \sup_{i=1,\ldots,n} \max_{j=1,\ldots,q} |R_{n,i}^{(j)}| \binom{n}{r}^{-1} \sum_S g(\underset{\sim}{X}_{\underset{\sim}{\alpha}}) > \varepsilon\}$$

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

Again, if δ_0 is kept fixed then also by (4.3.13) the second term in (4.3.10), i.e.,

$$P\{(rq)^{1/2} \sup_{i=1,\ldots,n} \max_{j=1,\ldots,q} |R_{n,i}^{(j)}| > \delta_0\} \to 0, \text{ as } n \to \infty.$$

This completes the proof of the lemma.

Lemma 4.3.2. Under (2.2.6) and (2.2.8), as $n \to \infty$,

(i) $\quad E\{U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})\} = g(\tilde{F}_n) + O(n^{-1} \log n)$

and

(ii) $\quad n^{1/2}\{U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}) - U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})\} \underset{p}{\to} 0.$

Proof. Similar to (2.4) in Sen (1963), here we have

(4.3.14) $\quad E\{U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}) - g(\tilde{F}_n)\}$

$$= \binom{n}{r}^{-1}\left\{\sum_{\ell=1}^{r-1}\binom{r-1}{\ell}\sum_{v_1=1}^{m_n}\cdots\sum_{v_\ell=1}^{m_n}\binom{n-v_1-\cdots-v_\ell-(r-\ell-1)m_n}{r-\ell}[F(\tilde{F}_n|\ell,v_1,\ldots,v_\ell)\right.$$

$$\left.-g(\tilde{F}_n)]\right.$$

By the Liapounov's inequality of moments and (4.3.1), $g(\tilde{F}_n|\ell,v_1,\ldots,v_\ell)-g(\tilde{F}_n)|$ is bounded for all $(\alpha_1,\ldots,\alpha_r)$ with $\alpha_{i+1}-\alpha_i = v_i \le m_n+1$, for all $i = 1,\ldots,\ell-1$. Let

$$g_s^{(n)}(\tilde{F}_n) = \sup_{\alpha_1,\ldots,\alpha_r} |g(\tilde{F}_n|\ell,v_1,\ldots,v_\ell) - g(\tilde{F}_n)|$$

Then similar to (2.6) in Sen (1963), we have by (4.3.2),

$$(4.3.15)\quad |E\{U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})-g(\tilde{F}_n)| \le \binom{n}{r}^{-1}\left[\binom{n}{r}-\binom{n-(r-1)m_n}{r}\right]g_s^{(n)}(\tilde{F}_n)$$

$$\le n^{-1}r(r-1)m_n g_s^{(n)}(\tilde{F}_n)$$

$$= O(n^{-1}\log n),$$

which proves (i).

Also

$$(4.3.16)\quad n^{1/2}\{U(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})-\binom{n}{r}^{-1}\binom{n-(r-1)m_n}{r}U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})\}$$

$$= rn^{-1/2}\binom{n-1}{r-1}^{-1}\sum_{S-S_0}\phi(Y_{n,\alpha_1},\ldots,Y_{n,\alpha_r})$$

where the summation extends over all possible $\left\{\binom{n}{r}-\binom{n-(r-1)m_n}{r}\right\}$ terms with $\alpha_{i+1}-\alpha_i \le m_n$, for at least one $i = 1,\ldots,r-1$. Since from Lemma 4.7 of Sen (1963), it follows that

$$\binom{n-1}{r-1}^{-1}\left[\binom{n}{r}-\binom{n-(r-1)m_n}{r}\right] \le m_n(r-1),$$

similar to (2.8) of Sen (1963), here we get

$$(4.3.17) \qquad E\{[rn^{-1/2}\binom{n-1}{r-1}^{-1} \sum_{S-S_0} \phi(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_r})]^2\} = O\{n^{-2}(\log n)^2\}$$

which by the Chebyshev inequality and a well known convergence theorem due to Cramér (1946, pp. 253-254) implies (ii). Q.E.D.

Let us now put

$$(4.3.18) \qquad \phi_a^{(n)}(\underset{\sim}{y}_{n,\alpha_1},\ldots,\underset{\sim}{y}_{n,\alpha_a}) = E\{\phi(\underset{\sim}{y}_{n,\alpha_1},\ldots,\underset{\sim}{y}_{n,\alpha_a},\underset{\sim}{Y}_{n,\alpha_{a+1}}, \ldots,\underset{\sim}{Y}_{n,\alpha_r}\} - g(\tilde{F}_n)$$

for $a = 0,\ldots,r$; where $\alpha_{i+1} - \alpha_i > m_n$ for all $i = 1,2,\ldots,r-1$ and let

$$(4.3.19) \qquad \zeta_{a\cdot 0}^{(n)} = E\{\phi_a^{(n)2}(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_a})\}$$

$$(4.3.20) \qquad \zeta_{a\cdot(h_1,\ldots,h_a)}^{(n)} = E\{\phi_a^{(n)}(\underset{\sim}{Y}_{n,\alpha_1},\ldots,\underset{\sim}{Y}_{n,\alpha_a})\phi_a^{(n)}(\underset{\sim}{Y}_{n,\beta_1},\ldots,\underset{\sim}{Y}_{n,\beta_a})\}$$

for $a = 0,1,\ldots,r$; where $|\alpha_i - \beta_i| = h_i$; $0 < h_i \leq m_n$ for $i = 1,\ldots,a$, but $|\alpha_i - \alpha_t| > m_n$, $|\beta_i - \beta_t| > m_n$ and $|\alpha_i - \beta_t| > m_n$ for all $i \neq t = 1,\ldots,n$. Finally, let

$$(4.3.21) \qquad \zeta_1^{(n)} = \zeta_{1\cdot 0}^{(n)} + 2\sum_{h=1}^{m_n} \zeta_{1\cdot h}^{(n)} .$$

Then we have the following.

<u>Lemma</u> 4.3.3. <u>Under (2.2.6) and (2.2.8)</u>, $n^{1/2}\{U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})\} - g(\tilde{F}_n)$ <u>is asymptotically normally distributed with zero mean and variance</u> $r^2\zeta_1^{(n)}$.

<u>Proof</u>. By Lemma 4.3.2, it is sufficient to show that $n^{1/2}\{U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}) - g(\tilde{F}_n)\}$ has asymptotically the normal distribution with zero mean and variance $r^2\zeta_1^{(n)}$. By the same techniques as used in Sen

(1963) to obtain the variance of a U-statistic for an m_n-dependent process, here we have

$$\mathrm{Var}\{U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n})\} = n^{-1}r^2\zeta_1^{(n)} + O(n^{-2}\log n) . \tag{4.3.22}$$

Again, if we define

$$Y_n = rn^{-1/2}\sum_{i=1}^{n}\phi_1^{(n)}(\underset{\sim}{Y}_{n,i})$$

where

$$\mathrm{Var}\{Y_n\} = r^2\zeta_1^{(n)} + O(n^{-1}\log n),$$

then by exactly the same procedure as in (2.22) of Sen (1963), we get

$$Y_n - n^{1/2}\{U_0(\underset{\sim}{Y}_{n,1},\ldots,\underset{\sim}{Y}_{n,n}) - g(\tilde{F}_n)\} \underset{p}{\to} 0 . \tag{4.3.23}$$

Now $\{\phi_1^{(n)}(\underset{\sim}{Y}_{n,i})\}$ form an m_n-dependent stationary stochastic process, whose third (absolute) moment is finite by the assumption (4.3.2). Hence by the C.L.T. for strongly mixing processes by Rosenblatt (1956b), it follows that Y_n has asymptotically the normal distribution with zero mean and variance $r^2\zeta_1^{(n)}$. The above arguments along with (4.3.38) complete the proof of the lemma. Q.E.D.

<u>Theorem</u> 4.3.1. <u>Under the assumptions (A), (B), (2.2.6), (2.2.8), and (2.2.10),</u> $n^{1/2}\{U(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n) - g(F)\}$ <u>is asymptotically normally distributed with mean zero and variance</u> $r^2\zeta_1$, <u>where</u>

$$g(F) = \int\ldots\int_{R^{qr}} \phi(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_r)dF(\underset{\sim}{x}_1)\ldots dF(\underset{\sim}{x}_r), \tag{4.3.24}$$

$$\zeta_1 = \zeta_{1\cdot 0} + 2\sum_{h=0}^{\infty}\zeta_{1\cdot h} , \tag{4.3.25}$$

$$(4.3.26)\quad \zeta_{1.0} = \int_{R^{q(2r-1)}}\cdots\int \{\phi(\underset{\sim}{x}_1,\underset{\sim}{y}_2,\ldots,\underset{\sim}{y}_r)-g(F)\}\{\phi(\underset{\sim}{x}_1,\underset{\sim}{y}_2^*,\ldots,\underset{\sim}{y}_r^*)-g(F)\}$$

$$dF(\underset{\sim}{y}_2)\ldots dF(\underset{\sim}{y}_r)dF(\underset{\sim}{y}_2^*)\ldots dF(\underset{\sim}{y}_r^*)dF(\underset{\sim}{x}_1),$$

$$(4.3.27)\quad \zeta_{1.h} = \int_{R^{2qr}}\cdots\int \{\phi(\underset{\sim}{x}_1,\underset{\sim}{y}_2,\ldots,\underset{\sim}{y}_r)-g(F)\}\{\phi(\underset{\sim}{x}_2,\underset{\sim}{y}_2,\ldots,\underset{\sim}{y}_r)-g(F)\}$$

$$dF(\underset{\sim}{y}_2)\ldots dF(\underset{\sim}{y}_r)dF(\underset{\sim}{y}_2^*)\ldots dF(\underset{\sim}{y}_r^*)dF_h(\underset{\sim}{x}_1,\underset{\sim}{x}_2)$$

<u>and</u> $F_h(\underset{\sim}{x}_1,\underset{\sim}{x}_2)$ is the joint c.d.f. of $(\underset{\sim}{X}_i,\underset{\sim}{X}_{i+h})$ $h = 1,2,\ldots$.

<u>Proof.</u> If we denote by $\tilde{F}_n(\underset{\sim}{x}_1)$ and $\tilde{F}_{nh}(\underset{\sim}{x}_1,\underset{\sim}{x}_2)$, the joint c.d.f. of $\underset{\sim}{Y}_{n,i}$ and $(\underset{\sim}{Y}_{n,i},\underset{\sim}{Y}_{n,i+h})$ respectively, then by straight forward generalizations of (3.2.1), it is easily seen that $\tilde{F}_n$ and $\tilde{F}_{nh}$ converges weakly to F and F_h respectively.

Since $\Phi(\cdot)$ is continuous,

$$g(\tilde{F}_n) = \int_{R^{qr}}\cdots\int \phi(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_r)d\tilde{F}_n(x_1)\ldots d\tilde{F}_n(\underset{\sim}{x}_r)$$

and (4.3.1) holds, using a result in Cramér (1946, p. 74), we get,

$$(4.3.28)\qquad \lim_{n\to\infty} g(\tilde{F}_n) = \int_{R^{qr}}\cdots\int \phi(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_r)dF(\underset{\sim}{x}_1)\ldots dF(\underset{\sim}{x}_r)$$

$$= g(F),\quad \text{by } (4.3.24).$$

Again since,

$$\zeta_{1.h}^{(n)} = \int_{r^{2q}}\cdots\int \phi_1^{(n)}(\underset{\sim}{x}_1)\phi_1^{(n)}(\underset{\sim}{x}_2)d\tilde{F}_{nh}(\underset{\sim}{x}_1,\underset{\sim}{x}_2)$$

$$= \int_{R^{2qr}}\cdots\int \phi(\underset{\sim}{x}_1,\underset{\sim}{y}_2,\ldots,\underset{\sim}{y}_r)\phi(\underset{\sim}{x}_2,\underset{\sim}{y}_2^*,\ldots,\underset{\sim}{y}_r^*)$$

$$d\tilde{F}_n(\underset{\sim}{y}_2)\ldots d\tilde{F}_n(\underset{\sim}{y}_r)d\tilde{F}_n(\underset{\sim}{y}_2^*)\ldots d\tilde{F}_n(\underset{\sim}{y}_r^*)dF_{nh}(\underset{\sim}{x}_1,\underset{\sim}{x}_2)-g^2(\tilde{F}_n)$$

by the same arguments as given above,

$$(4.3.29) \qquad \lim_{n \to \infty} \zeta_{1\cdot h}^{(n)} = \int_{R^{2qr}} \cdots \int \{\phi(\underset{\sim}{x}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_r) - g(F)\}\{\phi(\underset{\sim}{x}_2, \underset{\sim}{y}_2^*, \ldots, \underset{\sim}{y}_r^*) - g(F)\}$$

$$dF(\underset{\sim}{y}_2) \ldots dF(\underset{\sim}{y}_r) dF(\underset{\sim}{y}_2^*) \ldots dF(\underset{\sim}{y}_r^*) dF_h(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$$

$$= \zeta_{1\cdot h}, \text{ by } (4.3.27), \quad h = 1, 2, \ldots .$$

Similarly,

$$(4.3.30) \qquad \lim_{n \to \infty} \zeta_{1\cdot 0}^{(n)} = \zeta_{1\cdot 0}$$

where $\zeta_{1\cdot 0}$ is defined in (4.3.26).

Now under (2.2.8) and (4.3.2),

$$\sup_n \{ \sum_{h=0}^{\infty} |\zeta_{1\cdot h}^{(n)}| \} \text{ is finite}$$

and

$$|\zeta_{1\cdot h}^{(n)}| < a_h, \quad \text{for all } n,$$

where

$$\sum_{h=0}^{\infty} a_h < \infty .$$

Hence by above and bounded convergence theorem, we have

$$(4.3.31) \qquad \lim_{n \to \infty} \zeta_1^{(n)} = \zeta_{1\cdot 0} + 2 \sum_{h=1}^{\infty} \zeta_{1\cdot h} = \zeta_1, \text{ by } (4.3.25).$$

Therefore the proof of the theorem follows readily from (4.3.28), (4.3.31), Lemma 4.3.1, Lemma 4.3.2 and Lemma 4.3.3.

4.4. Asymptotic Distribution of the Median of the Mid-range Estimator.

Let $\{\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n\}$ be a sample of size n from the stationary autoregressive process defined in (1.1.1). We assume that the marginal c.d.f. of $\underset{\sim}{X}_1$ is diagonally symmetric about its location parameter.

Then the median of the mid-range estimator (vector) for the location parameter is $\underset{\sim}{\theta}_n = (\theta_{n,1},\ldots,\theta_{n,q})'$ where $\theta_{n,j} = \underset{1\le t\le t'\le n}{\text{med}} \{\frac{1}{2}(X_{t,j}+X_{t',j})\}$, for $j = 1,\ldots,q$. From the asymptotic normality of the Wilcoxon signed rank statistic (vector), the asymptotic distribution of the vector $\sqrt{n}\underset{\sim}{\theta}_n$ is derived here by a straight forward multivariate extension of the Hodges and Lehmann's (1963) techniques.

Using the relation $H_{[j]}(x) = F_{[j]}(x) - F_{[j]}(-x)$, $j = 1,\ldots,q$, we see that if the c.d.f. F is diagonally symmetrical about $\underset{\sim}{0}$, which implies $F_{[j]}(x) + F_{[j]}(-x) = 1$ for every x and for every $j = 1,\ldots,q$,

(4.4.1) $$\mu_j = \int_0^\infty H_{[j]}(x)dF_{[j]}(x) = 1/4, \quad j = 1,\ldots,q,$$

and making use of the substitutions -y to y, -x to x and -x to x and -y to y in the second, third and fourth integrals respectively,

(4.4.2) $$\delta_{jj'h} = \delta^{(11)}_{jj'h} - \delta^{(12)}_{jj'h} - \delta^{(21)}_{jj'h} + \delta^{(22)}_{jj'h}$$

$$= \int_0^\infty\int_0^\infty \sum_{s,s'=0,1} \{F_{[j,j']h}((-1)^s x,(-1)^{s'} y) - F_{[j]}((-1)^s x)F_{[j']}((-1)^{s'} y)\} dF_{[j]}((-1)^{s+1} x)\, dF_{[j']}((-1)^{s'+1} y)$$

$$= \int_{-\infty}^\infty \int_{-\infty}^\infty \{F_{[j,j']h}(x,y) - F_{[j]}(x)F_{[j']}(y)\, dF_{[j]}(x)\}\, dF_{[j']}(y)$$

and for $h = 0$, $j = j' = 1,\ldots,q$, it is equal to

$$(4.4.3) \qquad \delta_{jj} = \int_0^\infty F_{[j]}(x)(1-F_{[j]}(x))dF_{[j]}(-x) + \int_0^\infty F_{[j]}(-x)(1-F_{[j]}(-x))\,dF_{[j]}(x)$$

$$-2\int_0^\infty F_{[j]}(-x)(1-F_{[j]}(x))dF_{[j]}(x)$$

$$= 1/12,$$

where $\delta_{jj'h}^{(s,s')}$, $s,s' = 1,2$, are defined after (4.2.36) in the same manner as in (4.2.15) - (4.2.19).

Hence for symmetrical population, $\gamma_{jj'}$ is as usual given by (4.2.36), where for every $j,j' = 1,\ldots,q$, $h = 0,1,2,\ldots$, $\delta_{jj'h}$ and δ_{jj} are given by (4.4.2) and (4.4.3) respectively.

For each $j = 1,\ldots,q$, let us now consider the variables $X'_{i,j} = X_{i,j} - a_j$, $i = 1,\ldots,n$ where a_j is a number chosen according to our convenience. Let us take $a_j = n^{-1/2}u_j$, $j = 1,\ldots,q$, where u_j's are some fixed numbers. Then the Wilcoxon signed rank statistic (vector) based on the variables $X'_{i,j}$, $i = 1,\ldots,n$, $j = 1,\ldots,q$ is given by

$$(4.4.4) \qquad \underset{\sim}{T}_n(\underset{\sim}{a}) = (T_{n1}(a_1),\ldots,T_{nq}(a_q))',$$

where

$$(4.4.5) \qquad T_{nj}(a_j) = n(n+1)^{-1}\int_0^\infty H_{nj}(x-n^{-1/2}u_j)dF_{nj}(x),\ j = 1,\ldots,q.$$

Similar to (4.2.39), here we get

$$ET_{nj}(a_j) = n(n+1)^{-1}\int_0^\infty H_{[j]}(x-n^{-1/2}u_j)dF_{[j]}(x)$$

$$= \int_0^\infty H_{[j]}(x)dF_{[j]}(x) - n^{-1/2}u_j\int_0^\infty H'_{[j]}(x)dF_{[j]}(x) + o(n^{-1/2})$$

$$= \frac{1}{4} - n^{-1/2}\,u_j\int_{-\infty}^\infty f^2_{[j]}(x)dx + o(n^{-1/2})$$

so that

$$(4.4.6)\qquad \sqrt{n}\, E\{T_{nj}(a_j) - \tfrac{1}{4}\} = - u_j \int_{-\infty}^{\infty} f^2_{[j]}(x) + o(1)$$

$$= \mu_j(a_j),\ \text{say},\ j = 1,\ldots,q.$$

If we denote the corresponding $\delta_{jj'h}$ by $\delta'_{jj'h}$ then by (4.4.2) and the Taylor series expansion, we can write

$$(4.4.7)\quad \delta'_{jj'h} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{F_{[j,j']h}(x-n^{-1/2}u_j, y-n^{-1/2}u_j) - F_{[j]}(x-n^{-1/2}u_j)$$

$$F_{[j']}(y-n^{-1/2}u_{j'})\}dF_{[j]}(x)dF_{[j]}(y)$$

$$= \delta_{jj'h} + O(n^{-1/2})$$

Hence

$$(4.4.8)\qquad \text{Var}\{\sqrt{n}\, \underset{\sim}{T}_n(\underset{\sim}{a})\} = \underset{\sim}{\Gamma}_n = ((\gamma_{njj'}))_{j,j'=1,\ldots,q}$$

where

$$(4.4.9)\qquad \gamma_{njj'} = \gamma_{jj'} + O(n^{-1/2})$$

(4.4.9) implies, as $n \to \infty$,

$$(4.4.10)\qquad \underset{\sim}{\Gamma}_n \to \underset{\sim}{\Gamma}.$$

Without any loss of generality we take the location parameter to be $\underset{\sim}{0}$ and suppose $P_{\underset{\sim}{0}}$ denotes the probability when $\underset{\sim}{0}$ is the true value of the parameter. Then by a straight forward multivariate extension of Theorem 4 of Hodges and Lehmann (1963), we have

$$(4.4.11)\qquad \lim_{n\to\infty} P_{\underset{\sim}{0}}\{\sqrt{n}\theta_{n,j} \le u_j,\ \ j = 1,\ldots,q\}$$

$$= \lim_{n\to\infty} P_{\underset{\sim}{0}}\{\sqrt{n}[T_{nj}(a_j) - \tfrac{1}{4}] \le 0,\ j = 1,\ldots,q\}$$

$$= \lim_{n\to\infty} P_{\underset{\sim}{0}}\{\sqrt{n}[T_{nj}(a_j) - \tfrac{1}{4}] - \mu_j(a_j) \le -\mu_j(a_j),\ j=1,\ldots,q\}$$

$$= \lim_{n \to \infty} P_{\underset{\sim}{0}}\left\{\frac{\sqrt{n}[T_{nj}(a_j) - \frac{1}{4}] - \mu_j(a_j)}{\int_{-\infty}^{\infty} f_{[j]}^2(x)\,dx} \leq u_j + o(1),\ j = 1,\ldots,q\right\}$$

By (4.4.10) and the remark in the end of Theorem 4.2.1, as $n \to \infty$, the right hand side of (4.4.11) converges to the c.d.f. of a multivariate normal distribution with mean vector $\underset{\sim}{0}$ and dispersion matrix

$$(4.4.12) \qquad \underset{\sim}{T}^{**} = ((\tau_{jj'}^{**}))_{j,j'=1,\ldots,q}$$

where

$$(4.4.13) \qquad \tau_{jj'}^{**} = \gamma_{jj'}/\left(\int_{-\infty}^{\infty} f_{[j]}^2(x)\,dx\right)\left(\int_{-\infty}^{\infty} f_{[j']}^2(x)\,dx\right)$$

Hence by (4.4.11) and the above arguments, the vector $\sqrt{n}\underset{\sim}{\theta}_n$ is asymptotically multinormally distributed with mean vector $\underset{\sim}{0}$ and dispersion matrix $\underset{\sim}{T}^{**}$, where $\underset{\sim}{T}^{**}$ is defined in (4.4.12) and (4.4.13).

CHAPTER V

COMPARISON OF THE PERFORMANCES OF SEVERAL ESTIMATORS OF LOCATION FOR STATIONARY AUTOREGRESSIVE PROCESSES

5.1. Introduction

In this chapter we consider several rival estimators of location parameter for stationary autoregressive processes and compare their relative performances by their asymptotic relative efficiencies. In section 5.2, A.R.E. of the median (vector) with respect to the mean (vector) for the general stationary multivariate autoregressive processes are considered. The details of this A.R.E. value for the univariate case are studied in section 5.3. In section 5.3, the A.R.E. of the 27% mid-range estimator (vector) with respect to the mean (vector) and the A.R.E. of the median of the mid-range estimator (vector) with respect to the mean (vector) for the general stationary multivariate autoregressive processes are considered in a nut shell. However, emphasis is given on the study of the A.R.E. values for the univariate case. In the end of the section the A.R.E. values for univariate Gaussian autoregressive processes for the above estimators are tabulated. We conclude this chapter with a description of some related problems not considered in this dissertation and is expected to be studied later.

5.2. Asymptotic Relative Efficiency of the Median with Respect to the Mean.

In this section first, for samples from general stationary multivariate autoregressive processes the A.R.E. of the sample median (vector) with respect to the sample mean (vector) is obtained and then some bounds for the A.R.E. are worked out for some particular cases. The A.R.E. values for samples from univariate (normal) autoregressive processes are studied in the next section and some values of it are also tabulated there.

Let $\underset{\sim}{X}_t$, $t = 1,2,\ldots,n$ be the sample observations ($q \times 1$ vectors) where $\underset{\sim}{X}_t$ has the representation (2.2.7). Here we assume (2.2.10) to be true with some $\delta \geq 2$ and let us suppose that $\text{var}(\underset{\sim}{\varepsilon}_t) = \underset{\sim}{\Sigma}_\varepsilon = ((\sigma_{\varepsilon ss'}))_{s,s'=1,\ldots,q}$ for all t and $0 < |\underset{\sim}{\Sigma}_\varepsilon| < \infty$. Define $\underset{\sim}{\overline{X}}_n = (\overline{X}_{n,1},\ldots,\overline{X}_{n,q})'$, where $\overline{X}_{n,j} = n^{-1}\sum_{t=1}^{n} X_{t,j}$, $j = 1,\ldots,q$ and denote by $\underset{\sim}{t}_n$, the sample median vector, and by $\underset{\sim}{\xi}_{1/2}$, the population median vector.

By (2.2.7), we have

$$(5.2.1) \qquad E(\underset{\sim}{X}_t) = \sum_{r=0}^{\infty} \underset{\sim}{B}_r E(\underset{\sim}{\varepsilon}_{t-r}) = \underset{\sim}{0},$$

$$(5.2.2) \qquad E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h}) = E\left(\sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\varepsilon}_{t-r}\right)\left(\sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\varepsilon}_{t-r+h}\right)'$$

$$= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \underset{\sim}{B}_r E(\underset{\sim}{\varepsilon}_{t-r}\underset{\sim}{\varepsilon}'_{t-s+h})\underset{\sim}{B}_s$$

$$= \sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\Sigma}_\varepsilon \underset{\sim}{B}_{r+h}, \quad h = 0,1,\ldots,$$

and (5.2.1) and (5.2.2) imply

$$(5.2.3)\qquad E(\bar{\underset{\sim}{X}}_n) = \underset{\sim}{0}$$

and

$$(5.2.4)\qquad E(\bar{\underset{\sim}{X}}_n\bar{\underset{\sim}{X}}_n') = n^{-2}\left[\sum_{t=1}^{n} E(\underset{\sim}{X}_t\underset{\sim}{X}_t') + 2\sum_{h=1}^{n}\sum_{t=1}^{n-h} E(\underset{\sim}{X}_t\underset{\sim}{X}_{t+h}')\right]$$

$$= n^{-1}\left[\sum_{r=0}^{\infty} \underset{\sim}{B}_r\underset{\sim}{\Sigma}_\varepsilon\underset{\sim}{B}_r + 2\sum_{h=1}^{n}(1-\frac{h}{n})\sum_{r=0}^{\infty}\underset{\sim}{B}_r\underset{\sim}{\Sigma}_\varepsilon\underset{\sim}{B}_{r+h}\right].$$

If we write $\underset{\sim}{\Sigma}^{(h)} = \sum_{r=0}^{\infty} \underset{\sim}{B}_r\underset{\sim}{\Sigma}_\varepsilon\underset{\sim}{B}_{r+h}$, then since $|\underset{\sim}{\Sigma}^{(h)}| \leq \prod_{j=1}^{q}\sigma_{jj}^{(h)}$, for finiteness of $\left|\sum_{h=1}^{\infty}\sum_{r=0}^{\infty}\underset{\sim}{B}_r\underset{\sim}{\Sigma}_\varepsilon\underset{\sim}{B}_{r+h}\right|$, it is sufficient to show that $\sum_{h=1}^{\infty}\max_{j=1,\ldots,q}\sigma_{jj}^{(h)} < \infty$. Let us write $\sigma_0 = \max_{s=1,\ldots,q}\sigma_{\varepsilon ss}$. Then by Schwarz's inequality and (2.2.8), we have for any $j = 1,\ldots,q$,

$$(5.2.5)\qquad \sigma_{jj}^{(h)} = \sum_{r=0}^{\infty}\sum_{s=1}^{q}\sum_{s'=1}^{q} b_{js}^{(r)}b_{js'}^{(r+h)}\sigma_{\varepsilon ss'}$$

$$\leq q\sum_{r=0}^{\infty}\left[\sum_{s=1}^{q}\left\{b_{js}^{(r)}\right\}^2\sigma_{\varepsilon ss}\right]^{1/2}\left[\sum_{s'=1}^{q}\left\{b_{js'}^{(r+h)}\right\}^2\sigma_{\varepsilon s's'}\right]^{1/2}$$

$$\leq C^2q^3\sigma_0\sum_{r=0}^{\infty} r^g(r+h)^g(e^*)^{2r+h}$$

so that

$$(5.2.6)\qquad \sum_{h=1}^{\infty}\max_{j=1,\ldots,q}\sigma_{jj}^{(h)} \leq C^2q^4\sigma_0\sum_{h=1}^{\infty}\sum_{r=0}^{\infty} r^g(r+h)^g(e^*)^{2r+h} < \infty,$$

which proves finiteness of $\left|\sum_{h=1}^{\infty}\sum_{r=0}^{\infty}\underset{\sim}{B}_r\underset{\sim}{\Sigma}_\varepsilon\underset{\sim}{B}_{r+h}\right|$. Again by (5.2.5),

$$(5.2.7)\qquad \left|\sum_{h=1}^{n}\frac{h}{n}\sum_{r=0}^{\infty}\underset{\sim}{B}_r\underset{\sim}{\Sigma}_\varepsilon\underset{\sim}{B}_{r+h}\right|$$

$$\leq \sum_{h=1}^{n}\frac{h}{n}\prod_{j=1}^{q}\sigma_{jj}^{(h)}$$

$$\leq (C^2q^4/n)\sum_{h=1}^{n}\sum_{r=0}^{\infty} hr^g(r+h)^g(e^*)^{2r+h} \to 0 \text{ as } n\to\infty,$$

since the double series

$$\sum_{h=1}^{\infty} \sum_{r=0}^{\infty} hr^{g}(r + h)^{g}(e^{*})^{2r+h}$$

converges for $0 < e^{*} < 1$. Therefore as $n \to \infty$,

(5.2.8) $$nE(\underset{\sim}{\overline{X}}_n \underset{\sim}{\overline{X}}_n') \to \left[\sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\Sigma}_\varepsilon \underset{\sim}{B}_r + 2 \sum_{h=1}^{\infty} \sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\Sigma}_\varepsilon \underset{\sim}{B}_{r+h}\right] = \underset{\sim}{\Sigma} \text{ , say.}$$

If $\underset{\sim}{\Sigma} = ((\sigma_{jj'}))_{j,j'=1,\ldots,q}$ then we assume that

(5.2.9) $$\underset{\sim}{\Sigma} \text{ is non-null.}$$

By (2.2.8) it follows that

(5.2.10) $$\sum_{r=0}^{\infty} ||\underset{\sim}{B}_r||_2 < \infty$$

where

$$||\underset{\sim}{B}_r||_2 = \sup_{||\underset{\sim}{x}||_2=1} ||\underset{\sim}{B}_r \underset{\sim}{x}||,$$

and $\underset{\sim}{x}$ is a $q \times 1$ vector with $||\underset{\sim}{x}||_2 = \{\sum_{i=1}^{q} x_i^2\}^{1/2}$ and satisfying the condition $||\underset{\sim}{x}||_2 = 1$.

For our case let us take $\underset{\sim}{B}^* = \underset{\sim}{0}$, a $q \times 1$ null vector, and $y_t = 1$ in the model (1.1.2), considered by Hannan (1961). Then Hannan's (c3), (c4) and (c5) conditions are also satisfied. Hence by Hannan's (1961) C.L.T., $\sqrt{n}\, \underset{\sim}{\overline{X}}_n$ has asymptotically a multivariate normal distribution with mean vector $\underset{\sim}{0}$ and dispersion matrix $\underset{\sim}{\Sigma}$.

Also by (2.4.2),

$$\mathcal{L}(n^{1/2}[\underset{\sim}{t}_n - \underset{\sim}{\xi}_{1/2}]) \to N_q(\underset{\sim}{0}, \underset{\sim}{T}), \quad \text{as } n \to \infty,$$

where $\underset{\sim}{T}$ is given by (2.4.1) with $p^{(j)} = 1/2$ and $F_{[j]}\left(\xi^{(j)}_{p^{(j)}}\right) = 1/2$ for each $j = 1,\ldots,q$.

In the above model we have $E(\bar{\underset{\sim}{X}}_n) = \underset{\sim}{0}$. Instead of $\underset{\sim}{0}$, if $E(\underset{\sim}{\varepsilon}_t) = \underset{\sim}{\mu}_\varepsilon$ then $E(\bar{\underset{\sim}{X}}_n) = \sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\mu}_\varepsilon = \underset{\sim}{\mu}$, say. If however, $\underset{\sim}{\mu} = \underset{\sim}{\xi}_{1/2}$ then the sample mean (vector) and the sample median (vector) estimates the same parameter. In particular if the c.d.f. G is symmetric about $\underset{\sim}{0}$, then $\underset{\sim}{\mu} = \underset{\sim}{\xi}_{1/2}$. In that case the A.R.E. of the sample median (vector) with respect to the sample mean (vector) is the ratio of the reciprocal of their generalized variances as considered by Wilks. For a detailed definition of A.R.E. see chapter 6 of Puri and Sen (1971). Hence the A.R.E. of the sample median (vector) with respect to the sample mean (vector) is given by

$$\{|\underset{\sim}{\Sigma}|/|\underset{\sim}{T}|\}^{1/q} \tag{5.2.11}$$

Now we use the following well-known result, known as Courant's theorem.

If $\underset{\sim}{A}$ and $\underset{\sim}{B}$ be two square matrices of same order and $\underset{\sim}{B}$ is non-singular then

$$\sup_{\underset{\sim}{a} \neq \underset{\sim}{0}} \left(\frac{\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{a}}{\underset{\sim}{a}'\underset{\sim}{B}\underset{\sim}{a}}\right) = \text{maximum eigenvalue of } (\underset{\sim}{A}\underset{\sim}{B}^{-1})$$

and

$$\inf_{\underset{\sim}{a} \neq \underset{\sim}{0}} \left(\frac{\underset{\sim}{a}'\underset{\sim}{A}\underset{\sim}{a}}{\underset{\sim}{a}'\underset{\sim}{B}\underset{\sim}{a}}\right) = \text{minimum eigenvalue of } (\underset{\sim}{A}\underset{\sim}{B}^{-1}).$$

By the above theorem, the expression (5.2.11) is bounded below and above by the q-th root of minimum and maximum eigenvalues of $\underset{\sim}{\Sigma}\underset{\sim}{T}^{-1}$.

We consider, now, some particular cases.

a) We assume $\varepsilon_{t,j}$'s are independent for $j = 1,\ldots,q$ and $\underset{\sim}{B}_r$'s are diagonal matrices for $r = 0,1,2,\ldots$. Then $X_{t,j}$'s are also independent. This gives us $\sigma_{jj'} = 0$ for $j \neq j'$ and also $\tau_{jj'} = 0$ for $j \neq j'$. In this case the efficiency (5.2.11) is given by

$$\{|\underset{\sim}{\Sigma}|/|\underset{\sim}{T}|\}^{1/q} = (\prod_{j=1}^{q} \sigma_{jj}/\tau_{jj})^{1/q} . \tag{5.2.12}$$

Actually σ_{jj}/τ_{jj} correspond to the A.R.E. of sign test with respect to t-test for the univariate case. Hence by an application of Hodges and Lehmann's (1956) lower bound for the A.R.E. of sign test with respect to t-test, the A.R.E., as given by (5.2.12), is $\geq 1/3$. Again since

$$\underset{\sim}{\Sigma}\underset{\sim}{T}^{-1} = \begin{pmatrix} \frac{1}{\tau_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\tau_{qq}} \end{pmatrix} \begin{pmatrix} \sigma_{11} & & 0 \\ & \ddots & \\ 0 & & \sigma_{qq} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_{11}}{\tau_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{\sigma_{qq}}{\tau_{qq}} \end{pmatrix}$$

which has eigenvalues $\sigma_{11}/\tau_{11}, \ldots, \sigma_{qq}/\tau_{qq}$, by Courant's theorem, the A.R.E. (5.2.12) lies between the q-th root of minimum and maximum values of $(\sigma_{11}/\tau_{11}, \ldots, \sigma_{qq}/\tau_{qq})$.

b) If $\underset{\sim}{B}_r = \lambda_r \underset{\sim}{I}_q$, $r = 0,1,2,\ldots$, where λ_r's are constants and $\varepsilon_{t,j}$'s are exchangeable random variables for $j = 1,..,q$, i.e., $\underset{\sim}{\Sigma}_\varepsilon = \sigma_\varepsilon^2[(1 - \rho_\varepsilon)\underset{\sim}{I}_q + \rho_\varepsilon \underset{\sim}{J}_q]$ then we can write for $h = 0,1,2,\ldots,$

$$\begin{aligned} E(\underset{\sim}{X}_t \underset{\sim}{X}'_{t+h}) &= \sum_{r=0}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\Sigma}_\varepsilon \underset{\sim}{B}'_r \\ &= \sigma_\varepsilon^2 \sum_{r=0}^{\infty} [(1 - \rho_\varepsilon)\underset{\sim}{B}_r \underset{\sim}{B}'_r + \rho_\varepsilon \underset{\sim}{B}_r \underset{\sim}{J}_q \underset{\sim}{B}'_r] \\ &= \sigma_\varepsilon^2 (\sum_{r=0}^{\infty} \lambda_r \lambda_{r+h})[(1 - \rho_\varepsilon)\underset{\sim}{I}_q + \rho_\varepsilon \underset{\sim}{J}_q] \end{aligned}$$

Hence in this case $X_{t,j}$'s are also exchangeable random variables for $j = 1,\ldots,q$. By (5.2.8) here we have

$$\begin{aligned} \underset{\sim}{\Sigma} = ((\sigma_{jj'})) &= \sum_{r=0}^{\infty} [\underset{\sim}{B}_r \underset{\sim}{\Sigma}_\varepsilon \underset{\sim}{B}'_r + 2 \sum_{h=1}^{\infty} \underset{\sim}{B}_r \underset{\sim}{\Sigma}_\varepsilon \underset{\sim}{B}'_{r+h}] \\ &= \sigma_\varepsilon^2 S[(1 - \rho_\varepsilon)\underset{\sim}{I}_q + \rho_\varepsilon \underset{\sim}{J}_q], \end{aligned} \tag{5.2.13}$$

where $S = \left(\sum_{r=0}^{\infty} \lambda_r^2 + 2 \sum_{h=1}^{\infty} \sum_{r=0}^{\infty} \lambda_r \lambda_{r+h} \right)$

The variables $X_{t,j}$ being also symmetrically distributed, from (2.3.23) and (2.4.1) we see that $\tau_{jj'}$ is same for all $j \neq j' = 1,\ldots,q$ and τ_{jj} is also same for $j = 1,\ldots,q$. Here for $p^{(j)} = 1/2$, $j = 1,\ldots,q$, let us write

$$\tau_{jj'} = \begin{cases} \tau \text{ for } j = j' \\ \mu \text{ for } j \neq j' \end{cases} \tag{5.2.14}$$

and define the function

$$J(F(x)) = \begin{cases} 1 \text{ for } F(x) \geq \frac{1}{2} \\ 0 \text{ otherwise.} \end{cases}$$

Then we can write the correlation coefficient between the medians for any two variates by $\rho_J = \mu/\tau$ and by Lemma 4.4 and Lemma 4.5 of Sen (1968 a) we get,

$$\rho_J \geq -(q-1)^{-1} \tag{5.2.15}$$

where the equality holds iff J(F(X)) is a linear function of X with probability one, in which case the distribution of X is degenerate. Hence in the non-degenerate case we can write

$$\rho_J > -(q-1)^{-1} \tag{5.2.16}$$

By use of (5.2.13) and (5.2.14), in this case the A.R.E. (5.2.11) can be written as

$$\{|\Sigma|/|\Sigma|\}^{1/q} = \left[\frac{(\sigma_\varepsilon^2 S)^q \{1+(q-1)\rho_\varepsilon\}(1-\rho_\varepsilon)^{q-1}}{\{\tau + (q-1)\mu\}(\tau - \mu)^{q-1}} \right]^{1/q} \tag{5.2.17}$$

$$= (\sigma_\varepsilon^2 S/\tau) \left[\frac{\{1 + (q-1)\rho_\varepsilon\}^{1/q} (1-\rho_\varepsilon)^{(q-1)/q}}{\{1 + (q-1)\rho_J\}^{1/q} (1-\rho_J)^{(q-1)/q}} \right]$$

We denote the efficiency (5.2.17) by $\tilde{e}$ and consider some particular cases.

i) Let $\rho_\varepsilon = 0$. Then

$$\tilde{e} = (\sigma_\varepsilon^2 S/\tau)[\{1 + (q-1)\rho_J\}^{1/q}(1-\rho_J)^{(q-1)/q}]^{-1}$$

$$= (\sigma_\varepsilon^2 S/\tau), \text{ if } \rho_J = 0$$

$$\geq (\sigma_\varepsilon^2 S/\tau q^{1/q}), \text{ if } 0 \leq \rho_J \leq 1$$

$$> (\sigma_\varepsilon^2 S/\tau)((q-1)/q)^{(q-1)/q}, \text{ if } -(q-1)^{-1} < \rho_J \leq 0$$

$$> (\sigma_\varepsilon^2 S/\tau q^{1/q})((q-1)/q)^{(q-1)/q}, \text{ if } -(q-1)^{-1} < \rho_J \leq 1.$$

Thus in this case we get different lower bound of the efficiency depending on the value of ρ_J.

ii) Let $\rho_\varepsilon = -(q-1)^{-1}$. Then as $\rho_J > -(q-1)^{-1}$ by (5.2.17), $\tilde{e} = 0$ so that our proposed procedure is not at all suitable for the degenerate case $\rho_\varepsilon = -(q-1)^{-1}$.

iii) In general for given ρ_ε and $-(q-1)^{-1} < \rho_J \leq 1$, we have

$$\tilde{e} > (\sigma_\varepsilon^2 S/\tau)\frac{\{1 + (q-1)\rho_\varepsilon\}^{1/q}(1-\rho_\varepsilon)^{(q-1)/q}(q-1)^{(q-1)/q}}{q}$$

$$= (\sigma_\varepsilon^2 S/\tau)[(q-1)^{(q-1)/q}\{1 + (q-1)\rho_\varepsilon\}^{1/q}(1-\rho_\varepsilon)^{(q-1)/q}]/q$$

so that whether we have a larger or smaller bound of the efficiency than in case (i) depends on the value of ρ_ε.

In this case if the number of variables q increases indefinitely and $\rho_J \geq 0$, then the efficiency

$$\tilde{e} = (\sigma_\varepsilon^2 S/\tau) \lim_{q \to \infty} \frac{\{(q-1)^{-1} + \rho_\varepsilon\}^{1/q}(1 - \rho_\varepsilon)^{(q-1)/q}}{\{(q-1)^{-1} + \rho_J\}^{1/q}(1 - \rho_J)^{(q-1)/q}}$$

$$= (\sigma_\varepsilon^2 S/\tau)(1-\rho_\varepsilon)/(1-\rho_J)$$

$$> (\sigma_\varepsilon^2 S/\tau)(1-\rho_\varepsilon)$$

and this lower bound is $(\sigma_\varepsilon^2 S/\tau)$ if $\rho_\varepsilon \to 0$. Hence if the number of variables increases indefinitely, the efficiency can be bounded below by a quantity which tends to a fixed limit irrespective of the value of $\rho_J (\geq 0)$.

c) If $\underset{\sim}{X}_t$ has a q-variate normal distribution then $X_{t,j}$ and $X_{t,j'}$ have jointly a bivariate normal distribution. If we write $\rho_{jj'}$ = correlation coefficient between $X_{t,j}$ and $X_{t,j'}$ and $\rho_{hjj'}$ = correlation coefficient between $X_{t,j}$ and $X_{t+h,j'}$, then using the relations

$$F_{[j,j']}(\xi_j,\xi_{j'}) = \frac{1}{4} + \frac{1}{2\pi}\,\mathrm{Sin}^{-1}\,\rho_{jj'} \quad \text{and} \quad F_{h[j,j']}(\xi_j,\xi_{j'}) = \frac{1}{4} + \frac{1}{2\pi}\,\mathrm{Sin}^{-1}\rho_{hjj'}$$

(See Cramér (1946), p. 290) we have the following simplifications. In fact then

$$\nu_{jj'} = \begin{cases} \frac{1}{2\pi}\left[\mathrm{Sin}^{-1}\rho_{jj'} + \sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_{hjj'} + \sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_{hj'j}\right], & \text{for } j \neq j' \\ \frac{1}{2\pi}\left[\frac{\pi}{2} + 2\sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_{hjj}\right], & \text{for } j = j' \end{cases}$$

and

$$f_{[j]}(\xi_j) = \frac{1}{\sqrt{2\pi}\,\sigma_j}, \quad \text{where } \sigma_j^2 = \mathrm{var}(X_{t,j})$$

5.3. Asymptotic Relative Efficiency for Univariate Stationary Autoregressive Processes

For univariate autoregressive processes of order k, we can simplify the results quite considerably. In this particular case, we can write the observation X_t as

$$(5.3.1) \qquad X_t = \sum_{r=0}^{\infty} b^{(r)} \varepsilon_{t-r}$$

where $b^{(r)} = \sum_{i=1}^{k} \gamma_i e_i^r$, $(e_1, \ldots, e_k)$ ($|e_i| < 1$ for $i = 1, \ldots, k$ by (2.2.6)) are the roots of the auxiliary equation of

$$(5.3.2) \qquad X_t + a_1 X_{t-1} + \ldots + a_k X_{t-k} = \varepsilon_t$$

and γ_i's are constants. Since by (2.2.6), $0 < e^* = \max_{1 \le i \le k} |e_i| < 1$, we can write

$$e^* = e^{-\rho^*}, \text{ where } \rho^* > 0.$$

Then

$$(5.3.3) \quad |b^{(r)}| = O(e^*)^r = O(e^{-r\rho^*}) \text{ or } |b^{(r)}| \le \lambda(e^*)^r$$

where λ is a finite constant independent of r. This shows that in this case the assumption (2.2.8) is obviously satisfied. Here all the other results can be obtained much more easily due to the simplicity of the expressions (5.3.1) and (5.3.3). But we shall study, in particular, the details of the A.R.E. values.

Let X_t, $t = 1, \ldots, n$ be the sample observations where X_t has the representation (5.3.1). We assume $\mathrm{var}(\varepsilon_t) = \sigma^2$ for all t and $0 < \sigma^2 < \infty$. Define $\bar{X}_n = n^{-1} \sum_{t=1}^{n} X_t$ and denote by t_n, the sample median and by $\xi_{1/2}$, the population median.

Since X_t has the representation (5.3.1), we have

$$E(X_t) = 0$$

$$\mathrm{var}(X_t) = \sum_{r,s=0}^{\infty} b^{(r)} b^{(s)} E(\varepsilon_{t-r}\, \varepsilon_{t-s}) = \sigma^2 \sum_{r=0}^{\infty} b^{(r)2}$$

$$\mathrm{cov}(X_t, X_{t+h}) = \sum_{r,s=0}^{\infty} b^{(r)} b^{(s)} E(\varepsilon_{t-r}\, \varepsilon_{t+h-s}) = \sigma^2 \sum_{r=0}^{\infty} b^{(r)} b^{(r+h)}$$

which imply

$$E(\overline{X}_n) = 0$$

and

$$(5.3.4) \quad n \operatorname{var}(\overline{X}_n) = n^{-1}[\sum_{t=1}^{n} \operatorname{var}(X_t) + 2\sum_{h=1}^{n}\sum_{t=1}^{n-h} \operatorname{cov}(X_t, X_{t+h})]$$

$$= \sigma^2[\sum_{r=0}^{\infty} b^{(r)2} + 2\sum_{h=1}^{n}(1 - \frac{h}{n})\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)}]$$

$$\to \sigma^2[\sum_{r=0}^{\infty} b^{(r)2} + 2\sum_{h=1}^{\infty}\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)}], \text{ as } n \to \infty,$$

since

$$\sum_{h=1}^{n}\frac{h}{n}\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)} \leq \lambda^2 \sum_{h=1}^{n} \frac{h}{n}(e^*)^h\{1 - (e^*)^2\}^{-1}$$

$$= (\lambda^2/n)\{1 - (e^*)^2\}^{-1}[e^*\{1-(e^*)^n\}(1 - e^*)^{-2} - n(e^*)^{n+1}(1 - e^*)^{-1}]$$

$$\to 0 \text{ as } n \to \infty.$$

By (5.3.3) the right hand side of (5.3.4) is of the order $(1 + e^*)(1 - e^*)^{-1}\{1-(e^*)^2\}^{-1}$ which is non-null for $0 < e^* < 1$. Also by (5.3.3),

$$\sum_{r=0}^{\infty} |b^{(r)}| \leq \lambda \sum_{r=0}^{\infty} (e^*)^r = \lambda(1 - e^*)^{-1} < \infty, \text{ for } 0 < e^* < 1.$$

Now taking $B^* = 0$ and $y_t = 1$ in the model (1.1.2), considered by Hannan (1961), we see that his (c3), (c4), and (c5) conditions are also satisfied. Hence by Hannan's (1961) Central Limit Theorem, $\sqrt{n}\,\overline{X}_n$ is asymptotically normal with mean 0 and variance $\sigma^2[\sum_{r=0}^{\infty} b^{(r)2} + 2\sum_{h=1}^{\infty}\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)}]$.

Also by (2.4.2), $\sqrt{n}(t_n - \xi_{1/2})$ is asymptotically normal with mean 0 and variance $\nu^2/f^2(\xi_{1/2})$, where $f(\xi_{1/2}) = F'(\xi_{1/2})$ and

$$(5.3.5) \qquad \nu^2 = \frac{1}{4} + 2 \sum_{h=1}^{\infty} \{F_h(\xi_{1/2},\xi_{1/2}) - \frac{1}{4}\} .$$

In the above model we have $E(\overline{X}_n) = 0$. Instead of 0, if $E(\varepsilon_t) = \mu_\varepsilon$ then $E(\overline{X}_n) = \sum_{r=0}^{\infty} b^{(r)}\mu_\varepsilon = \mu$, say. If however, $\mu = \xi_{1/2}$ then the sample mean and the sample median estimate the same parameter. In particular if the c.d.f. G is symmetric about 0, then $\mu = \xi_{1/2}$. In that case the A.R.E. of the sample median with respect to the sample mean is given by

$$(5.3.6) \qquad \mathrm{eff}(t/\overline{X}) = \frac{4\sigma^2 f^2(\xi_{1/2})\left[\sum_{r=0}^{\infty} b^{(r)^2} + 2\sum_{h=1}^{\infty}\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)}\right]}{[1 + 2\sum_{h=1}^{\infty}\{4F_h(\xi_{1/2},\xi_{1/2}) - 1\}]}$$

If the underlying parent distribution is normal then $f(\xi_{1/2}) = \left(\sqrt{2\pi}\ \sigma\sqrt{\sum_{r=0}^{\infty} b^{(r)^2}}\right)^{-1}$ and $F_h(\xi_{1/2},\xi_{1/2}) = \frac{1}{4} + (\mathrm{Sin}^{-1}\rho_h)/2\pi$ (See Cramér (1946), p. 290) where ρ_h is the correlation coefficient between X_t and X_{t+h}. Hence for a univariate Gaussian autoregressive process the A.R.E. of the sample median with respect to the sample mean is given by

$$(5.3.7) \qquad e_N(t/\overline{X}) = (2/\pi)\,\frac{1 + 2\sum_{h=1}^{\infty}\rho_h}{1 + (4/\pi)\sum_{h=1}^{\infty}\mathrm{Sin}^{-1}\rho_h}$$

For first order Gaussian autoregressive processes where

$$X_t = aX_{t-1} + \varepsilon_t,$$

$|a| < 1$ and ε_t are independent normal variables with mean 0 and variance σ^2, we have $\rho_h = a^h$ and $b^{(r)} = a^r$. Hence if we denote by e_{1N}, the A.R.E. of the sample median with respect to the sample mean for a first order normal autoregressive process then from (5.3.7) we get,

$$(5.3.9) \qquad e_{1N} = (2/\pi) \frac{[1 + 2 \sum_{h=1}^{\infty} a^h]}{1 + (4/\pi) \sum_{h=1}^{\infty} \text{Sin}^{-1}(a^h)}$$

$$= \frac{2[1 + 2a(1 - a)^{-1}]}{\pi + 4 \sum_{h=1}^{\infty} \text{Sin}^{-1}(a^h)}$$

$$= 2(1 + a)(1 - a)^{-1}\{\pi + 4 \sum_{h=1}^{\infty} \text{Sin}^{-1}(a^h)\}^{-1}$$

We now observe some limiting values and bounds of the efficiency e_{1N}. First, the series $\sum_{h=1}^{\infty} \text{Sin}^{-1}(a^h)$ being uniformly convergent for $|a| < 1$, $e_{1N} \to 2/\pi$ as $a \to 0$. Secondly, using the infinite series expansion for $\text{Sin}^{-1}(a^h)$ and summing term by term, we get

$$(5.3.10) \qquad \sum_{h=1}^{\infty} \text{Sin}^{-1}(a^h) = a(1-a)^{-1}[1 + (a^2/6)(1 + a + a^2)^{-1} + (3a^4/40)(1 + a + a^2 + a^3 + a^4)^{-1} + (5a^6/112)(1+a+a^2+a^3+a^4+a^5+a^6)^{-1} + \ldots]$$

which consequently imply as $a \to 1$,

$$(5.3.11) \qquad e_{1N} \to (1 + \frac{1}{18} + \frac{3}{200} + \frac{5}{784} + \frac{35}{10368} + \frac{63}{30976} + \ldots)^{-1} = .924$$

As $a \to -1$, the expression (5.3.10) also implies that $e_{1N} \to 0$. Again by using the inequality $(2/\pi)\ \text{Sin}^{-1}x \leq x$ for $x \geq 0$, we get for $a \geq 0$,

$$(5.3.12) \qquad e_{1N} \geq (2/\pi)(1 + a)(1 - a)^{-1}\{1 + 2 \sum_{h=1}^{\infty} a^h\}^{-1} = 2/\pi$$

so that the efficiency is bounded below by $(2/\pi)$ for non-negative values of a.

For a second order Gaussian autoregressive process where $X_t + a_1X_{t-1} + a_2X_{t-2} = \varepsilon_t$, a_1, a_2 are constants and ε_t are independent normal variables with mean 0 and variance σ^2, two cases may arise.

i) Case of unequal roots: If e_1 $(-1 < e_1 < 1)$ and e_2 $(-1 < e_2 < 1)$ are the distinct roots of the auxiliary equation $x^2 + a_1x + a_2 = 0$, then we can write

$$(5.3.13) \qquad X_t = (e_1 - e_2)^{-1} \sum_{r=0}^{\infty} (e_1^{r+1} - e_2^{r+1})\varepsilon_{t-r}$$

which gives us $b^{(r)} = (e_1 - e_2)^{-1}(e_1^{r+1} - e_2^{r+1})$. Therefore

$$\begin{aligned}(5.3.14) \qquad \sum_{r=0}^{\infty} b^{(r)^2} &= (e_1 - e_2)^{-2} \sum_{r=0}^{\infty} (e_1^{2r+2} + e_2^{2r+2} - 2e_1^{r+1}e_2^{r+1}) \\ &= (e_1 - e_2)^{-2}[e_1\{e_1(1-e_1^2)^{-1} - e_2(1-e_1e_2)^{-1}\} \\ &\quad + e_2\{e_2(1-e_2^2)^{-1} - e_1(1-e_1e_2)^{-1}\}] \\ &= (e_1-e_2)^{-1}[e_1(1-e_1^2)^{-1}(1-e_1e_2)^{-1} \\ &\quad - e_2(1-e_2^2)^{-1}(1-e_1e_2)^{-1}] \\ &= (1+e_1e_2)(1-e_1^2)^{-1}(1-e_2^2)^{-1}(1-e_1e_2)^{-1}\end{aligned}$$

and

$$\begin{aligned}(5.3.15) \qquad \sum_{h=1}^{\infty} \sum_{r=0}^{\infty} b^{(r)}b^{(r+h)} &= (e_1-e_2)^{-2} \sum_{h=1}^{\infty} \sum_{r=0}^{\infty} (e_1^{r+1}-e_2^{r+1})(e_1^{r+h+1}-e_2^{r+h+1}) \\ &= (e_1-e_2)^{-2} \sum_{h=1}^{\infty} [e_1^{h+1}\{e_1(1-e_1^2)^{-1} - e_2(1-e_1e_2)^{-1}\} \\ &\quad + e_2^{h+1}\{e_2(1-e_2^2)^{-1} - e_1(1-e_1e_2)^{-1}\}] \\ &= (e_1-e_2)^{-1} \sum_{h=1}^{\infty} [e_1^{h+1}(1-e_1^2)^{-1}(1-e_1e_2)^{-1} \\ &\quad - e_2^{h+1}(1-e_2^2)^{-1}(1-e_1e_2)^{-1}]\end{aligned}$$

$$= (e_1-e_2)^{-1}(1-e_1e_2)^{-1}[e_1^2(1-e_1)^{-1}(1-e_1^2)^{-1}$$

$$- e_2^2(1-e_2)^{-1}(1-e_2^2)^{-1}]$$

$$= (1-e_1e_2)^{-1}(1-e_1)^{-1}(1-e_1^2)^{-1}(1-e_2)^{-1}$$

$$\times (1-e_2^2)^{-1}(e_1+e_2-e_1e_2-e_1^2e_2^2)$$

Hence if we denote by e_{2N}, the A.R.E. of the sample median with respect to the sample mean for a second order Gaussian autoregressive process with unequal roots of auxiliary equation, then by (5.3.14) and (5.3.15) we get,

$$(5.3.16) \quad e_{2N} = \frac{2[1 + 2(e_1+e_2-e_1e_2-e_1^2e_2^2)(1-e_1)^{-1}(1-e_2)^{-1}(1+e_1e_2)^{-1}]}{\pi + 4\sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_h}$$

where

$$(5.3.17) \quad \rho_h = [e_1^{h+1}(1-e_2^2) - e_2^{h+1}(1-e_1^2)](e_1-e_2)^{-1}(1+e_1e_2)^{-1}$$

Since the series $\sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_h$ is uniformly convergent for $|e_1| < 1$, $|e_2| < 1$ and $\rho_h \to 0$ for $e_1 \to 0$, $e_2 \to 0$ with $e_1 \neq e_2$, $e_{2N} \to (2/\pi)$ as both $e_1 \to 0$ and $e_2 \to 0$ with the restriction $e_1 \neq e_2$. If ρ_h is given by (5.3.17) then

$$\sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_h = (e_1-e_2)^{-1}(1+e_1e_2)^{-1}\{e_1^2(1-e_2^2)(1-e_1)^{-1}-e_2^2(1-e_1^2)(1-e_2)^{-1}\}$$

$$+ (1/6)(e_1-e_2)^{-3}(1+e_1e_2)^{-3}\{e_1^6(1-e_2^2)^3(1-e_1^3)^{-1}$$

$$-3e_1^4e_2^2(1-e_2^2)(1+e_2) + 3e_1^2e_2^4(1-e_1^2)(1+e_1)$$

$$- e_2^6(1-e_1^2)^3(1-e_2^3)^{-1}\} + \ldots$$

and as $e_2 \to 0$,

$$\sum_{h=1}^{\infty} \mathrm{Sin}^{-1}\rho_h \to e_1(1-e_1)^{-1}[1 + (e_1^2/6)(1+e_1+e_1^2)^{-1} + \ldots]$$

which is exactly similar to (5.3.10). Therefore from above and (5.3.16) we conclude that the efficiency $e_{2N} \to e_{1N}$ as $e_2 \to 0$. The same conclusion can be arrived at for $e_1 \to 0$. This means that the A.R.E. of the second order Gaussian autoregressive process converges to that of first order process if any one of the roots of the auxiliary equation converges to zero. Again here $\rho_h \geq 0$ for all h, for $e_1 \geq 0$, $e_2 \geq 0$ and $e_1 - e_2 > 0$. Hence, similar to (5.3.12), for $e_1 \geq 0$, $e_2 \geq 0$ and $e_1 > e_2$,

$$(5.3.18) \qquad e_{2N} \geq \frac{2[1+2(e_1+e_2-e_1e_2-e_1^2e_2^2)(1-e_1)^{-1}(1-e_2)^{-1}(1+e_1e_2)^{-1}]}{\pi(1 + 2 \sum_{h=1}^{\infty} \rho_h)} = \frac{2}{\pi}$$

so that the A.R.E. e_{2N} is bounded below by $(2/\pi)$ if both the values of e_1 and e_2 are non-negative and $e_1 > e_2$.

i) Case of equal roots: If $e(-1 < e < 1)$ is the double root of the auxiliary equation $x^2 + a_1x + a_2 = 0$, then we can write

$$(5.3.19) \qquad X_t = \sum_{r=0}^{\infty} (r+1)e^r \varepsilon_{t-r}$$

which gives us $b^{(r)} = (r+1)e^r$ for $r = 0,1,\ldots$ Therefore,

$$(5.3.20) \qquad \sum_{r=0}^{\infty} b^{(r)2} = \sum_{r=0}^{\infty} (r+1)^2 e^{2r} = \sum_{r=1}^{\infty} r^2 e^{2r-2}$$

$$= \sum_{r=1}^{\infty} \{r(r-1)+r\}e^{2r-2} = 2e^2(1-e^2)^{-3} + (1-e^2)^{-2}$$

$$= (1 + e^2)(1 - e^2)^{-3}$$

and

$$(5.3.21) \quad \sum_{h=1}^{\infty} \sum_{r=0}^{\infty} b^{(r)} b^{(r+h)} = \sum_{h=1}^{\infty} \sum_{r=0}^{\infty} (r+1)(r+h+1)e^{2r+h}$$

$$= \sum_{h=1}^{\infty} \sum_{r=1}^{\infty} r(r+h)e^{2r+h-2}$$

$$= \sum_{h=1}^{\infty} \sum_{r=1}^{\infty} \{r(r-1) + (h+1)r\}e^{2r+h-2}$$

$$= \sum_{h=1}^{\infty} \{2e^2(1-e^2)^{-3} + (h+1)(1-e^2)^{-2}\}e^h$$

$$= \sum_{h=1}^{\infty} \{(1+e^2) + h(1-e^2)\}(1-e^2)^{-3}e^h$$

$$= e(1+e^2)(1-e)^{-1}(1-e^2)^{-3} + e(1+e)(1-e)^{-1}(1-e^2)^{-3}$$

$$= e(2 + e + e^2)(1 - e)^{-1}(1 - e^2)^{-3}$$

Hence if we denote by e_{2N}^*, the A.R.E. of the sample median with respect to the sample mean for a second order Gaussian autoregressive process with equal roots of auxiliary equation, then by use of (5.3.20) and (5.3.21) in (5.3.7) we get

$$(5.3.22) \quad e_{2N}^* = \frac{2[1 + 2e(2 + e + e^2)(1 - e)^{-1}(1 + e^2)^{-1}]}{\pi + 4 \sum_{h=1}^{\infty} \text{Sin}^{-1}\rho_h}$$

where

$$(5.3.23) \quad \rho_h = [1 + h(1 - e^2)(1 + e^2)^{-1}]e^h$$

Let us now observe some limiting values and bounds of the efficiency e_{2N}^*. First, we see that by (5.3.23), $\rho_h \to 0$ as $e \to 0$ and the series $\sum_{h=1}^{\infty} \text{Sin}^{-1}\rho_h$ is uniformly convergent for $|\rho_h| < 1$. Hence as $e \to 0$,

$$e^*_{2N} \to (2/\pi) \tag{5.3.24}$$

Secondly, using the infinite series expansion for $\text{Sin}^{-1}\rho_h$ with ρ_h given by (5.3.23) and summing up term by term, we get

$$(1 - e)(1 + e^2) \sum_{h=1}^{\infty} \text{Sin}^{-1}\rho_h \tag{5.3.25}$$

$$= e(1+e^2) + e(1+e) + (e^3/6)\{(1+e^2)(1+e+e^2)^{-1} + 3(1+e)(1+e+e^2)^{-2} + 3(1+e)^2(1+e^3)(1+e^2)^{-1}(1+e+e^2)^{-3} + (1+e)^3(1+4e^3+e^6)(1+e^2)^{-2}(1+e+e^2)^{-4}\} + \ldots$$

which consequently imply as $a \to 1$,

$$e^*_{2N} \to 4\left(4 + \frac{2}{9} + \frac{8}{81} + \frac{6}{100} + \frac{3}{125} + \frac{9}{625} + \frac{18}{3125} + \frac{1}{28} + \frac{3}{196} + \frac{15}{1372} + \frac{15}{2401} + \ldots\right)^{-1} \tag{5.3.26}$$

$$= 4/4.493$$

$$= .890$$

As $e \to -1$, the expression (5.3.17) also implies that $e^*_{2N} \to 0$. Again $e \geq 0$ implies $\rho_h \geq 0$ for all h. Therefore similar to (5.3.12) here we get for $e \geq 0$,

$$e^*_{2N} \geq \frac{2[1 + 2e(2+e+e^2)(1-e)^{-1}(1+e^2)^{-1}]}{\pi(1 + 2\sum_{h=1}^{\infty} \rho_h)} = 2/\pi \tag{5.3.27}$$

so that the efficiency e^*_{2N} is bounded below by $(2/\pi)$ for non-negative values of e.

For visualizing the nature of the efficiencies in details, the values of the efficiencies e_{1N}, e^*_{2N} and e_{2N} are tabulated at the end of this section (TABLE 5.3.1) for values of the roots lying between -.9 and .9.

For independent Gaussian processes the 27% mid-range estimate being optimum among the class of mid-range estimates, various authors have compared the performance of it with that of other similar statistics. Here for autoregressive processes we want to compare the performance of the 27% mid-range estimate with that of the process average.

For a multivariate autoregressive process if the parent distribution is symmetrical then the 27% mid-range estimate (vector) for the location parameter is $\underset{\sim}{U}_n = (U_n^{(1)},\ldots,U_n^{(q)})'$ where $U_n^{(j)} = (Z^{(j)}_{n,p_1^{(j)}} + Z^{(j)}_{n,p_2^{(j)}})/2$, $j = 1,\ldots,q$, with $p_1^{(j)} = .27$, $p_2^{(j)} = .73$, for each $j=1,\ldots,q$. By section 3.4, $\sqrt{n}\,\underset{\sim}{U}_n$ has a multivariate normal distribution with mean $\underset{\sim}{0}$ (population median vector) and dispersion matrix

$$\underset{\sim}{T}^* = ((\tau^*_{jj'})) \tag{5.3.28}$$

where

$$\begin{aligned}\tau^*_{jj'} = [&\nu_{11}^{(jj')} f^{-1}(\xi^{(j)}_{p_1^{(j)}}) f^{-1}(\xi^{(j')}_{p_1^{(j')}}) \\ &+ \nu_{12}^{(jj')} f^{-1}(\xi^{(j)}_{p_1^{(j)}}) f^{-1}(\xi^{(j')}_{p_2^{(j')}}) \\ &+ \nu_{21}^{(jj')} f^{-1}(\xi^{(j)}_{p_2^{(j)}}) f^{-1}(\xi^{(j')}_{p_1^{(j')}})\end{aligned} \tag{5.3.29}$$

$$+ \nu_{22}^{(jj')} f^{-1}(\xi_{p_2(j)}^{(j)}) f^{-1}(\xi_{p_2(j')}^{(j')})]$$

where $\nu_{ii'}^{(jj')}$ are given by (3.3.9).

Hence in the same situation described earlier in connection with A.R.E. of the sample median (vector) with respect to the sample mean (vector), the A.R.E. of the 27% mid-range estimate (vector) with respect to the sample mean (vector) for q-variate stationary auto-regressive processes is given by

(5.3.30) $$\{|\Sigma|/|T^*|\}^{1/q}$$

As in section 5.2, here also we can have some bounds of the efficiency (5.3.30) and we can consider those particular cases. However, to avoid repetition we omit the details.

For univariate autoregressive processes if the parent distribution is symmetrical then the 27% mid-range estimate for the location parameter is $U_n = (Z_{n,p_1} + Z_{n,p_2})/2$ with $p_1 = .27$, $p_2 = .73$. By section 3.4, $\sqrt{n}\, U_n$ is asymptotically normal with mean 0 (population median) and its variance is

(5.3.31) $$\tau^{*2} = [\nu_{11} f^{-2}(\xi_{p_1}) + 2\nu_{12} f^{-1}(\xi_{p_1}) f^{-1}(\xi_{p_2}) + \nu_{22} f^{-2}(\xi_{p_2})]/4$$

where ξ_{p_1} and ξ_{p_2} are the 27% and 73% quantiles respectively and

(5.3.32) $$\nu_{11} = p_1(1 - p_1) + 2 \sum_{h=1}^{\infty} \{F_h(\xi_{p_1}, \xi_{p_1}) - p_1^2\}$$

(5.3.33) $$\nu_{22} = p_2(1 - p_2) + 2 \sum_{h=1}^{\infty} \{F_h(\xi_{p_2}, \xi_{p_2}) - p_2^2\}$$

(5.3.34) $$\nu_{12} = F(\xi_{p_1}, \xi_{p_2}) - p_1 p_2 + \sum_{h=1}^{\infty} \{F_h(\xi_{p_1}, \xi_{p_2}) - p_1 p_2\} + \sum_{h=1}^{\infty} \{F_h(\xi_{p_2}, \xi_{p_1}) - p_1 p_2\}$$

Using the facts of symmetry the expression (5.3.31) can be simplified quite a lot. First we see that for symmetrical distributions $f(\xi_{p_1}) = f(\xi_{p_2})$. Again the variances of the sample quantiles of lower 27% and upper 27% being equal for such distributions we must have $\nu_{11} = \nu_{22}$. Using the above facts the expression (5.3.31) can be simplified to

$$\tau^{*2} = (\nu_{11} + \nu_{12})/2f^2(\xi_{p_1}). \tag{5.3.35}$$

Hence in the situation described earlier, the A.R.E. of the 27% mid-range estimate with respect to the sample mean for univariate stationary autoregressive processes with symmetric distribution is given by

$$\text{eff}(U/\bar{X}) = 2\sigma^2 f^2(\xi_{p_1})\left[\sum_{r=0}^{\infty} b^{(r)^2} + 2\sum_{h=1}^{\infty}\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)}\right]/(\nu_{11}+\nu_{12}) \tag{5.3.36}$$

where ν_{11} and ν_{12} are given by (5.3.32) and (5.3.34) respectively with $p_1 = .27$ and $p_2 = .73$.

If the underlying parent distribution is normal, $\xi_{p_1} = -.6128$ and $f(\xi_{p_1}) = \left(\sqrt{2\pi}\,\sigma\sqrt{\sum_{r=0}^{\infty} b^{(r)^2}}\right)^{-1} \exp\{-\frac{1}{2}(.6128)^2\}$. Hence, for a univariate Gaussian autoregressive process, the A.R.E. of the 27% mid-range estimate with respect to the sample mean is given by

$$e_N(U/\bar{X}) = (1+2\sum_{h=1}^{\infty} \rho_h)\exp\{-(.6128)^2\}/\pi(\nu_{11} + \nu_{12}) \tag{5.3.37}$$

where

$$\rho_h = \sum_{r=0}^{\infty} b_r b_{r+h} / \sum_{r=0}^{\infty} b_r^2 \tag{5.3.38}$$

$$\nu_{11} = .1971 + 2\sum_{h=1}^{\infty}\{\Phi_h(-.6128,-.6128) - .0729\} \tag{5.3.39}$$

$$\nu_{12} = \{\Phi_o(-.6128,.6128)-.1971\} + 2\sum_{h=1}^{\infty}\{\Phi_h(-.6128,.6128) -.1971\} \tag{5.3.40}$$

and $\Phi_h(x,y)$ is the c.d.f. of standard bivariate normal distribution with correlation coefficient ρ_h.

For studying the nature of the efficiencies in details, the values of the efficiency (5.3.37) for first and second order Gaussian autoregressive processes are tabulated at the end of this section (TABLE 5.3.2) for values of the roots lying between -.9 and .9.

If the parent distribution is symmetrical about its location, the median of the mid-ranges provide another alternative estimator for the location parameter. The asymptotic distribution of the median of the mid-range estimator is studied in chapter 4. We now want to compare its performance with that of the process average.

For a multivariate autoregressive process if the parent distribution is symmetrical then the median of the mid-range estimator (vector) for the location parameter is $\underset{\sim}{\theta}_n = (\theta_n^{(1)},\ldots,\theta_n^{(q)})'$ where $\theta_n^{(j)} = \underset{1\leq t\leq t'\leq n}{\text{med}} \{\frac{1}{2}(X_{t,j} + X_{t',j})\}$, for $j = 1,\ldots,q$. By section 4.4, $\sqrt{n}\underset{\sim}{\theta}_n$ has multivariate normal distribution with mean $\underset{\sim}{0}$ (population median vector) and dispersion matrix

$$\underset{\sim}{T}^{**} = ((\tau^{**}_{jj'}))_{j,j'=1,\ldots,q} \tag{5.3.41}$$

where

$$\tau^{**}_{jj'} = (\delta_{jj'} + \sum_{h=1}^{\infty} \delta_{jj'h} + \sum_{h=1}^{\infty} \delta_{j'jh})/(\int_{-\infty}^{\infty} f^2_{[j]}(x)\,dx)(\int_{-\infty}^{\infty} f^2_{[j']}(x)dx) \tag{5.3.42}$$

where for $h = 0,1,2,\ldots$; $j,j' = 1,\ldots,q$, $\delta_{jj'h}$ is given by (4.4.2) and for $h = 0$ and $j = j' = 1,\ldots,q$, δ_{jj} is given by (4.4.3).

Hence in the same situation described earlier in connection with A.R.E. of the sample median (vector) with respect to the sample mean (vector), the A.R.E. of the median of the mid-range vector with respect to the sample mean vector for q-variate stationary autoregressive processes is given by

$$\{|\Sigma|/|\mathrm{T}^{**}|\}^{1/q} \tag{5.3.43}$$

Similar to section 5.2, here also we can have some bounds of the efficiency (5.3.43) and we can consider those particular cases. However, to avoid repetition we omit the details.

For univariate autoregressive processes with symmetrical distribution, the median of the mid-range estimate for the location parameter is $\hat{\theta}_n = \underset{1\le t\le t'\le n}{\text{med}} \{\frac{1}{2}(X_t + X_{t'})\}$. By section 4.4, $\sqrt{n}\,\hat{\theta}_n$ is asymptotically normally distributed with mean 0 (population median) and variance

$$\tau^{**2} = (1 + 24 \sum_{h=1}^{\infty} \delta_h)/12(\int_{-\infty}^{\infty} f^2(x)dx)^2 \tag{5.3.44}$$

where for h = 1,2,...,

$$\delta_h = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{F_h(x,y) - F(x)F(y)\}dF(x)dF(y) \tag{5.3.45}$$

Hence in the same situation described earlier in connection with A.R.E. of the sample median with respect to the sample mean, the A.R.E.of the median of the mid-range estimate with respect to the sample mean for univariate stationary autoregressive processes is given by

$$\text{eff}(\hat{\theta}|\overline{X}) = 12\sigma^2(\sum_{r=0}^{\infty} b^{(r)^2} + 2\sum_{h=1}^{\infty}\sum_{r=0}^{\infty} b^{(r)}b^{(r+h)})(\int_{-\infty}^{\infty} f^2(x)dx)^2 \Big/ (1+24\sum_{h=1}^{\infty}\delta_h) \tag{5.3.46}$$

where δ_h is given by (5.3.45).

If the underlying parent distribution is normal then $\int_{-\infty}^{\infty} f^2(x)dx =$ $(2\sqrt{\pi}\sigma_x)^{-1}$ and

(5.3.47) $$\delta_h = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\Phi_h(x,y) - \Phi(x)\Phi(y)\}d\Phi(x)d\Phi(y)$$

where $\Phi(x)$ is the c.d.f. of standard normal distribution and $\Phi_h(x,y)$ is the c.d.f. of standard bivariate normal distribution with correlation coefficient ρ_h. By Kendall (1948, p 118) and integration by parts, $\delta_h = (1/2\pi)\sin^{-1}(\rho_h/2)$. Consequently the expression (5.3.44) is simplified to

(5.3.48) $$\tau^{**2} = (\pi\sigma x^2/3)[1 + (12/\pi)\sum_{h=1}^{\infty} \text{Sin}^{-1}(\rho_h/2)].$$

Hence for a univariate Gaussian autoregressive process the A.R.E. of the median of the mid-ranges with respect to the sample mean is given by

(5.3.49) $$e_N(\hat{\theta}|\overline{X}) = (3/\pi)(1 + 2\sum_{h=1}^{\infty} \rho_h)\{1+(12/\pi)\sum_{h=1}^{\infty} \text{Sin}^{-1}(\rho_h/2)\}^{-1}$$

where

(5.3.50) $$\rho_h = \sum_{r=0}^{\infty} b^{(r)}b^{(r+h)} / \sum_{r=0}^{\infty} b^{(r)^2}.$$

If we denote by E_{1N}, the A.R.E. of the median of the mid-ranges with respect to the sample mean for a first order Gaussian autoregressive process, then similar to (5.3.9), here we get

(5.3.51) $$E_{1N} = 3(1 + a)(1 - a)^{-1}\{\pi + 12 \sum_{h=1}^{\infty} \text{Sin}^{-1}(a^h/2)\}$$

Similar to the case of A.R.E. of the sample mean with respect to the sample median for a univariate first order Gaussian autoregressive process, here also $E_{1N} \to 3/\pi$ as $a \to 0$, $E_{1N} \to 0$ as $a \to -1$ and $E_{1N} \to (1 + \frac{1}{72} + \frac{3}{3200} + \ldots)^{-1} = .985$ as $a \to 1$. In this case, from TABLE 5.3.3, we see that for $a \geq .5$, the efficiency E_{1N} is very close to its limiting value.

Similar to (5.3.16), the efficiency E_{2N} of the median of the mid-ranges with respect to the sample mean for a second order Gaussian autoregressive process with unequal roots of auxiliary equation is given by

$$(5.3.52) \qquad E_{2N} = \frac{3[1 + 2(e_1+e_2-e_1e_2-e_1^2e_2^2)(1-e_1)^{-1}(1-e_2)^{-1}(1+e_1e_2)^{-1}]}{\pi + 12 \sum_{h=1}^{\infty} \text{Sin}^{-1}(\rho_h/2)}$$

where ρ_h is given by (5.3.17). Here also $E_{2N} \to 3/\pi$ as both $e_1 \to 0$ and $e_2 \to 0$ with the restriction $e_1 \neq e_2$.

Similar to (5.3.22), the efficiency E^*_{2N} for the case of equal roots of auxiliary equation is given by

$$E^*_{2N} = \frac{3[1 + 2e(2 + e + e^2)(1 - e)^{-1}(1 + e^2)^{-1}]}{\pi + 12 \sum_{h=1}^{\infty} \text{Sin}^{-1}(\rho_h/2)}$$

where ρ_h is given by (5.3.23). Similar to the corresponding case of A.R.E. of the sample mean with respect to the sample median, here also $E^*_{2N} \to 3/\pi$ as $e \to 0$, $E^*_{2N} \to 0$ as $e \to -1$ and $E^*_{2N} \to (1+\frac{1}{72}+\frac{1}{162}+\frac{3}{32000} + \ldots)^{-1} = .979$ as $e \to 1$. In fact, from TABLE 5.3.3, we see that for $e \geq .2$, the efficiency E^*_{2N} is very close to its limiting value.

For studying the nature of the efficiencies in details, the values of the efficiencies E_{1N}, E_{2N} and E^*_{2N} are tabulated in TABLE 5.3.3 for values of the roots lying between -.9 and .9.

TABLE 5.3.1. A.R.E. of Sample Median/Sample Mean for First and Second Order Gaussian Autoregressive Processes

	-.9	-.8	-.7	-.6	-.5	-.4	-.3	-.2	-.1	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
-.9	**.003**																		
-.8	.011	**.022**																	
-.7	.019	.034	**.051**																
-.6	.029	.049	.070	**.094**															
-.5	.041	.067	.093	.116	**.153**														
-.4	.056	.077	.120	.150	.189	**.228**													
-.3	.074	.101	.151	.190	.230	.280	**.298**												
-.2	.095	.143	.187	.230	.276	.331	.373	**.427**											
-.1	.132	.179	.230	.279	.326	.387	.430	.482	**.511**										
0	.152$^+$	.221$^+$	.280$^+$	.335$^+$	.387$^+$	.438$^+$	.490$^+$	.540$^+$	.589$^+$	**.637**									
.1	.189	.272	.338	.397	.451	.503	.552	.597	.641	.681$^+$	**.697**								
.2	.235	.332	.405	.467	.522	.571	.615	.655	.691	.724$^+$	.752	**.778**							
.3	.290	.401	.481	.544	.596	.639	.677	.709	.738	.762$^+$	.783	.802	**.834**						
.4	.358	.482	.564	.639	.670	.707	.736	.760	.780	.797$^+$	.724	.722	.853	**.957**					
.5	.438	.573	.653	.706	.743	.770	.790	.806	.818	.828$^+$	.835	.841	.847	.851	**.856**				
.6	.534	.671	.742	.783	.809	.846	.738	.846	.852	.855$^+$	.857	.859	.860	.860	.861	**.863**			
.7	.646	.772	.825	.851	.865	.873	.877	.879	.879	.878$^+$	.876	.875	.872	.870	.868	.866	**.865**		
.8	.773	.865	.893	.903	.907	.906	.905	.903	.900	.897$^+$	.893	.890	.792	.776	.876	.872	.868	**.866**	
.9	.899	.931	.934	.931	.928	.925	.921	.918	.914	.911$^+$	.907	.904	.900	.895	.890	.885	.879	.871	**.866**

+ → First order process.

[box] → Second order process with equal roots.

[double box] → Independent process.

TABLE 5.3.2. A.R.E. of 27% Mid-Range/Sample Mean for First and Second Order Gaussian Autoregressive Processes

	-.9	-.8	-.7	-.6	-.5	-.4	-.3	-.2	-.1	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
-.9	[.003]																		
-.8	.012	[.035]																	
-.7	.024	.054	[.081]																
-.6	.039	.078	.114	[.158]															
-.5	.058	.108	.155	.210	[.270]														
-.4	.080	.144	.204	.269	.335	[.401]													
-.3	.108	.189	.261	.333	.402	.467	[.529]												
-.2	.142	.242	.325	.401	.469	.530	.587	[.642]											
-.1	.184	.304	.395	.470	.535	.592	.644	.692	[.736]										
0	.235^{+}	.374^{+}	.468^{+}	.541^{+}	.601^{+}	.652^{+}	.698^{+}	.739^{+}	.776^{+}	[[.809]]									
.1	.297	.450	.542	.610	.664	.709	.748	.782	.812	.838^{+}	[.860]								
.2	.370	.529	.617	.678	.724	.762	.794	.821	.844	.863^{+}	.880	[.895]							
.3	.453	.610	.689	.742	.780	.811	.835	.855	.872	.886^{+}	.897	.908	[.917]						
.4	.544	.690	.757	.800	.831	.854	.871	.885	.896	.905^{+}	.913	.919	.925	[.931]					
.5	.638	.765	.820	.853	.875	.890	.901	.910	.916	.922^{+}	.926	.930	.934	.937	[.941]				
.6	.732	.833	.875	.897	.911	.920	.926	.931	.934	.936^{+}	.938	.940	.941	.943	.945	[.947]			
.7	.819	.893	.920	.933	.940	.944	.946	.947	.948	.949^{+}	.949	.949	.949	.949	.949	.949	[.950]		
.8	.897	.941	.954	.959	.960	.961	.961	.960	.960	.959^{+}	.959	.958	.957	.955	.954	.953	.951	[.951]	
.9	.973	.986	.987	.987	.986	.985	.984	.983	.982	.981^{+}	.980	.979	.977	.976	.974	.972	.970	.970	[.992]

+ → First order process.

[ ] → Second order process with equal roots.

[[ ]] → Independent process.

TABLE 5.3.3. A.R.E. of Median of the Mid-Ranges/Sample Mean for First and Second Order Gaussian Autoregressive Processes

	-.9	-.8	-.7	-.6	-.5	-.4	-.3	-.2	-.1	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
-.9	[.464]																		
-.8	.827	[.859]																	
-.7	.849	.855	[.850]																
-.6	.853	.852	.846	[.842]															
-.5	.855	.850	.844	.841	[.842]														
-.4	.855	.850	.845	.844	.847	[.854]													
-.3	.857	.852	.849	.850	.855	.865	[.877]												
-.2	.859	.856	.856	.860	.868	.879	.892	[.906]											
-.1	.862	.862	.866	.873	.882	.894	.907	.921	[.934]										
0	.867+	.871+	.878+	.888+	.899+	.911+	.923+	.935+	.946+	[[.955]]									
.1	.873	.883	.893	.904	.915	.927	.938	.947	.956	.963+	[.968]								
.2	.881	.896	.909	.921	.932	.942	.951	.958	.964	.969+	.973	[.976]							
.3	.891	.910	.925	.937	.947	.955	.962	.967	.971	.974+	.976	.978	[.979]						
.4	.902	.925	.941	.952	.961	.967	.971	.974	.976	.978+	.979	.979	.979	[.980]					
.5	.915	.941	.956	.966	.972	.976	.978	.980	.980	.981+	.981	.980	.980	.980	[.980]				
.6	.929	.956	.969	.976	.980	.982	.983	.983	.983	.983+	.982	.981	.981	.980	.980	[.979]			
.7	.946	.971	.980	.984	.986	.986	.986	.985	.985	.984+	.983	.982	.981	.981	.980	.979	[.979]		
.8	.965	.983	.988	.989	.988	.988	.987	.986	.986	.985+	.984	.983	.982	.982	.981	.980	.979	[.979]	
.9	.985	.990	.990	.989	.988	.987	.987	.986	.986	.985+	.985	.984	.984	.983	.982	.981	.981	.979	[.979]

\+ → First order process.

[ ] → Second order process with equal roots

[[ ]] → Independent process

5.4. Proposed Topics of Future Research.

In this dissertation we have solved several problems in connection with stationary autoregressive processes. However, there are still some other problems in connection with it which we did not consider here. In connection with a stationary autoregressive process the problems which come to our mind immediately and which are not considered in this dissertation are the following.

i) For stationary autoregressive processes we have derived in Chapter III, the asymptotic distribution of linear combination of a fixed number of order statistics but we did not consider the problem of finding the asymptotic distribution of a linear combination of all order statistics or a subset of that. This problem includes the derivation of asymptotic distribution of Trimmed mean, Winsorized mean, etc. for stationary autoregressive processes.

ii) For stationary autoregressive processes we have developed in Chapter IV, the asymptotic distribution of Wilcoxon scores which is a particular case of Hoeffding's (1948) U-statistics. The problem of deriving the asymptotic distribution for the more general form of U-statistics for stationary autoregressive processes remains still open.

iii) In Chapter IV, we have considered only a particular case of U-statistics, namely the Wilcoxon scores. Wilcoxon scores can also be expressed as a particular case of rank statistics of the following type:

$$T_n = \sum_{i=1}^{n} J_n\left(\frac{R_{n,i}}{n+1}\right) \mathrm{Sgn}(X_i) \tag{5.4.1}$$

where $\mathrm{Sgn}(X_i)$ is 1 or 0 according as X_i is >0 or ≤ 0, $R_{n,i}$ is the rank of $|X_i|$ among $|X_1|,\ldots,|X_n|$ and $J_n(\frac{i}{n+1})$, $1 \leq i \leq n$, are suitable rank scores. Such statistics are knwn as Chernoff-Savage statistics. Hence for stationary autoregressive processes similar extension, as referred in (ii), is also possible for more general form of Chernoff-Savage statistics, defined in (5.4.1).

REFERENCES

[1] BAHADUR, R. R. (1966). A note on quantiles in large samples. Ann. Math. Statist. 37, 577-580.

[2] BILLINGSLEY, P. (1968). Convergence of Probability Measures. John Wiley, New York.

[3] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Statist. 29, 972-994.

[4] CRAMÉR, H. (1946). Mathematical Methods of Statistics. Princeton Univ. Press.

[5] DOOB, J. L. (1953). Stochastic Processes. John Wiley, New York.

[6] EICKER, F. (1965). Limit theorems for regressions with unequal and dependent errors. Proc. 5-th Berkeley Symposium. Vol I, 59-82.

[7] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1967). Limit Distributions for Sums of Independent Random Variables. Addison-Wesley Publishing Company, California.

[8] GRENANDER, U. and ROSENBLATT, M. (1957). Statistical Analysis of Stationary Time Series. John Wiley, New York.

[9] HÁJEK, J. (1968). Asymptotic normality of sample linear rank statistics under alternatives. Ann. Math Statist. 39, 325-346.

[10] HANNAN, E. J. (1961). A central limit theorem for systems of regressions. Proc. Cambridge Philos. Soc. Vol 57, 583-588.

[11] HODGES, J. L. Jr. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. Ann. Math. Statist. 27, 324-335.

[12] HODGES, J. L. Jr. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. Ann. Math. Statist. 34, 598-611.

[13] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, 293-325.

[14] HOEFFDING, W. and ROBBINS, H. (1948). The central limit theorem for dependent random variables. Duke Math. J. 15, 773-780.

[15] KENDALL, M. G. (1948). Rank Correlation Methods. Charles Griffin & Company, London.

[16] KIEFER, J. (1967). On Bahadur's representation of sample quantiles. Ann. Math. Statist. 38, 1323-1342.

[17] KIEFER, J. (1970). Deviations between the sample quantile process and the sample DF. Nonparametric Techniques in Statistical Inference. 299-319, Cambridge University Press.

[18] LOÈVE, M. (1963). Probability Theory (3rd Edition) Van Nostrand, Princeton.

[19] MOSTELLER, F. (1946). On some useful "inefficient" statistics. Ann. Math. Statist. 17, 377-408.

[20] PURI, M. L. and SEN, P. K. (1971). Nonparametric Methods in Multivariate Analysis. John Wiley, New York.

[21] ROSENBLATT, M. (1956a). On the estimation of regression coefficients of a vector-valued time series with a stationary residual. Ann. Math. Statist. 27, 99-121.

[22] ROSENBLATT, M. (1956b). A central limit theorem and a strong mixing condition. Proc. Nat. Acd. Sc. Vol. 42, 43-47.

[23] SARHAN, A. E. and GREENBERG, B. G. (1962). Contribution to Order Statistics. John Wiley, New York.

[24] SEN, P. K. (1961). On some properties of the asymptotic variance of the sample quantiles and mid-ranges. JRSSB Vol 23, 453-459.

[25] SEN, P. K. (1963). On the properties of U-statistics when the observations are not independent; part one. Cal Statist. Assoc. Bull. 12, 69-92.

[26] SEN, P. K. (1968a). On a class of aligned rank order tests in two-way layouts. Ann. Math. Statist. 39, 1115-1124.

[27] SEN, P. K. (1968b). Asymptotic normality of sample quantiles for m-dependent processes. Ann. Math. Statist. 39, 1724-1730.

[28] SEN, P. K. and GHOSH, M. (1971). On bounded length sequential confidence intervals based on one-sample rank order statistics. Ann. Math. Statist. 41, 189-203.

[29] USPENSKY, J. V. (1937). Introduction to Mathematical Probability. McGraw Hill Book Company, New York.

[30] WHITTLE, P. (1953). The analysis of multiple stationary time series. JRSSB 15, 125-139.