Chernoff-Savage theorems for dependent sequences of random variables and applications to asymptotic relative efficiency

Thomas Ray Fears
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Statistics and Probability Commons

Recommended Citation
Fears, Thomas Ray, "Chernoff-Savage theorems for dependent sequences of random variables and applications to asymptotic relative efficiency" (1972). Retrospective Theses and Dissertations. 5908.
https://lib.dr.iastate.edu/rtd/5908

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
FEARS, Thomas Ray, 1942-
CHERNOFF-SAVAGE THEOREMS FOR DEPENDENT
SEQUENCES OF RANDOM VARIABLES AND APPLICATIONS
TO ASYMPTOTIC RELATIVE EFFICIENCY.

Iowa State University, Ph.D., 1972
Statistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan
Chernoff-Savage theorems for dependent sequences of random variables and applications to asymptotic relative efficiency

by

Thomas Ray Fears

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of The Requirements for the Degree of DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1972
PLEASE NOTE:

Some pages may have indistinct print.
Filmed as received.

University Microfilms, A Xerox Education Company
TABLE OF CONTENTS

I. INTRODUCTION .......................... 1
   I.1. Historical background of the problem 1
   I.2. The method and notation 4
   I.3. The space D 6

II. THE P-DEPENDENT CASE ............... 9
   II.1. Introduction 9
   II.2. The space D x D 10
   II.3. The one-sample processes 14
   II.4. The two-sample process 30

III. THE MIXING CASE .................. 35
    III.1. Introduction 35
    III.2. The one-sample processes 36
    III.3. The two-sample process 57

IV. THE SCALE PROBLEM .................. 62
    IV.1. Introduction 62
    IV.2. The one-sample processes 65
    IV.3. The two-sample process 69

V. VARIATIONAL ARGUMENTS ............... 87
    V.1. Asymptotic relative efficiency 87
    V.2. The A.R.E. for the p-dependent case 88
    V.3. The A.R.E. for the mixing case 95

VI. BIBLIOGRAPHY ....................... 99

VII. ACKNOWLEDGEMENTS .................. 101
I. INTRODUCTION

I.1. Historical background of the problem

Let \( X_1, \ldots, X_m, \ldots \) and \( Y_1, \ldots, Y_n, \ldots \) be two sequences of random variables with each \( X_i, i = 1,2,\ldots \) (each \( Y_j, j = 1,2,\ldots \)) having a common continuous marginal distribution function \( F(x) \) (\( G(y) \)). Further, assume that with probability one, no two of the \( X \)'s or \( Y \)'s are equal. Taking the first \( m \) \( X \)'s and the first \( n \) \( Y \)'s, let \( N = m + n \) and define the statistic \( T_N^* \) by

\[
T_N^* = \sum_{k=1}^{N} Z_{Nk} C_{N,k}^*
\]

where \( C_{N,k}^* \) \( 1 \leq k \leq N \) are certain real numbers and \( Z_{Nk} = 1 \) if the \( k^{th} \) smallest of the combined \( N \) \( X \)- and \( Y \)-observations is an \( X \) and \( Z_{Nk} = 0 \) otherwise.

The asymptotic normality of such two sample rank-order statistics was studied by Chernoff and Savage (1958) under the assumption that all \( X \)'s and \( Y \)'s are mutually independent. Their paper is important in that it not only generalized earlier work on asymptotic normality of such statistics, but it prompted considerable activity in proving similar limit theorems and using them to evaluate asymptotic efficiencies in a variety of testing and estimation situations: e.g., Govindarajulu (1960), Puri (1964), Sen (1967, 1968), Hodges and Lehmann (1963). Subsequently, Govindarajulu et al. (1967), using deeper properties of empirical distribution functions,
weakened the conditions under which the Chernoff-Savage Theorem remained true.

In a majority of these papers the asymptotic normality is proved for statistics which are functions of relative ranks of observations that are mutually independent. Such statistics should be distinguished from those statistics based on the relative ranks of all observations in a problem. The latter class of statistics are called "joint ranking" statistics.

The reader is referred to the following papers: Sen (1967), Gastwirth et al. (1967), Gastwirth and Rubin (1969), dealing with asymptotic normality problems under dependence.

In this thesis joint ranking statistics under three types of dependence will be studied. In Chapter II we will consider the "p-dependent" case; in Chapter III we consider the "mixing" case; and in Chapter IV the case of double sequences specially constructed and suitable for the scale problem. Lastly, in Chapter V, expressions are derived for the asymptotic relative efficiency, in both the p-dependent and mixing cases, of the rank order tests for the two-sample location problem relative to the t-test.

**Definition I.1.1**

The sequence of vectors is said to be p-dependent if 

\(((X_i, Y_i), \ldots, (X_k, Y_k))\) and \(((X_k+l, Y_k+l), \ldots)\) are independent when \(l > p\).
Definition 1.1.2

Consider a sequence $X_1, X_2, \ldots$. For $a < b$, take $M_a^b$ to be the $\sigma$-field generated by $(X_a, \ldots, X_b)$. $\phi$ is a nonnegative function of positive integers. The sequence is $\phi$-mixing if for $k \geq 1$ and $n \geq 1$, $E_1 \in M_1^k$ and $E_2 \in M_{k+n}^\infty$ imply

$$|P(E_1 \cap E_2) - P(E_1)P(E_2)| \leq \phi(n)P(E_1).$$

The scale problem is defined as follows: Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be two independent random samples of observations based on populations with cumulative distribution functions $F(X-\xi)$ and $G(X-\eta) = F([X-\eta]/\delta)$ respectively, where $\xi$ and $\eta$ are unknown location parameters and $\delta$ the scale parameter. We consider the problem of testing that the populations differ only in their location parameters, i.e.,

$$H_0: \delta = 1$$
$$H_A: \delta \neq 1.$$

In this thesis the method employed for studying the asymptotic normality of $T_N$ is the same as employed by Pyke and Shorack (1968). Their approach, like that of Govindarajulu et al. (1967), is based on deeper properties of functions of empirical distribution functions.

Earlier, Fears and Mehra (1968) had used the original Chernoff and Savage method for $p$-dependent sequences. This approach proved cumbersome and it was therefore decided to
adopt the method of Pyke and Shorack. We believed that this method would be relatively simple when compared to the Chernoff and Savage technique for dependent sequences. Further, the Pyke and Shorack method requires assumptions that are comparable to those of Chernoff and Savage (see Theorem 5.1 of Pyke and Shorack (1968)). The following work will show that this belief was well founded.

1.2. The method and notation

The method of Pyke and Shorack will now be explained. Notation is also introduced which will be used throughout this work. Additional necessary notation will be introduced later and will be pertinent to the chapter in which it is introduced. The reader is warned that additional notation may have a different connotation in different chapters.

Pyke and Shorack (1968) consider two-sample rank-order statistics under the assumption that all X's and Y's are mutually independent. Let \( F_m \) be the empirical distribution function (e.d.f.) for the X-sample, \( G_n \) the e.d.f. for the Y-sample, and \( H_N \) the e.d.f. for the composite sample. \( F_m \), \( G_n \) and \( H_N \) will always refer to those observations on which \( T_N \) is based.

For later reference, let \( N = m+n \) and \( \lambda_N = m/N \) and assume that there exists \( \lambda_0 \leq \frac{1}{2} \) such that \( 0 < \lambda_0 \leq \lambda_N \leq 1-\lambda_0 < 1 \).
hold for all $N$. Let $\Delta = [\lambda_0, 1-\lambda_0]$. When $N$ is used as a subscript it denotes the pair $(m,n)$. $I_A$ is the indicator function of the set $A$. $K$ denotes a generic constant. For a distribution function $F$ take $F^{-1}(t) = \inf\{x : F(x) \geq t\}$. For any distribution functions $F$ and $G$ take $FG^{-1}$ to be the composed function $FG^{-1}(t) = F(G^{-1}(t))$. Let $\{U_m(t) : 0 \leq t \leq 1\}$ denote the one sample empirical process defined by

(I.2.1) \hspace{1cm} U_m(t) = m^{1/2} (F_mF^{-1}(t) - t).

Similarly, let $\{V_n(t) : 0 \leq t \leq 1\}$ denote the empirical process defined by

(I.2.2) \hspace{1cm} V_n(t) = n^{1/2}(G_nG^{-1}(t) - t).

Finally, define the two sample empirical process

(I.2.3) \hspace{1cm} L_N(t) = N^{1/2}(F_mH_N^{-1}(t) - FH^{-1}(t))

where $H = \lambda_NF + (1-\lambda_N)G$.

Pyke and Shorack show that

(I.2.4) \hspace{1cm} T_N = m^{-1} \sum_{k=1}^{N} Z_Nk C^*_N,k = \sum_{k=1}^{N} C_N,k F_m H_N^{-1}(k/N)

where $C_N,k = C^*_N,k - C^*_N,k+1$, $1 \leq k < N$, and $C^*_NN = C^*_NN$. 
If \( \nu_N \) denotes the signed measure which puts measure
\( \mathcal{C}_N,k \) on the point \( k/N \) for \( 1 \leq k \leq N \), and puts measure zero
elsewhere, then

\[
T_N = \int_0^1 F m H_N^{-1} \, d\nu_N.
\]

Take \( \mu_N = \int_0^1 F H^{-1} \, d\nu_N \) and set

\[
T^*_N = T_N - \mu_N = \int_0^1 (F m H_N^{-1} - F H^{-1}) \, d\nu_N.
\]

Let \( \nu \) be a Lebesgue-Stieltjes measure on \((0,1)\) for which
\( |\nu|([\varepsilon, 1-\varepsilon]) < \infty \) for \( \varepsilon > 0 \). By examining the one sample
processes and relating them to the two sample process, Pyke
and Shorack determine conditions to insure that \( \{\nu_N\} \) converges
to \( \nu \) in a manner which permits the substitution of \( \nu \) for \( \nu_N \) in
Equation I.2.6.

I.3. The space \( D \)

**Definition I.3.1**

\( D \) will denote the set of all right continuous real valued
functions on \([0,1]\) having only jump discontinuities.

There are several metrics defined on the space \( D \). The
uniform metric, \( \rho \), is defined by
\[ (I.3.1) \quad \rho(x,y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)|. \]

We will also use a complete metric defined in Chapter 3 of Billingsley (1968). If \( \tau \) is a nondecreasing function of \([0,1] \) with \( \tau(0) = 0 \) and \( \tau(1) = 1 \), take

\[ ||\tau|| = \sup_{s \neq t} \left| \log \frac{\tau(t) - \tau(s)}{t-s} \right|. \]

Let \( d \) denote the metric of Billingsley and let \( T \) be the class of strictly increasing, continuous mappings of \([0,1] \) onto itself.

For \( x, y \in D \), \( d(x,y) \) is defined to be the infimum of those positive \( \varepsilon \) for which \( T \) contains some \( \tau \) with

\[ \begin{cases} ||\tau|| < \varepsilon \\ \text{and} \\ \sup_t |x(t) - y(\tau(t))| < \varepsilon . \end{cases} \]

For other metrics on the space \( D \), the reader is referred to Prokhorov (1956) and Skorokhod (1956).

For \( n \geq 0 \) let \( \{W_n(t) : 0 \leq t \leq 1\} \) denote stochastic processes on the probability space \((\Omega, \beta, \mathbb{P})\) whose sample functions are elements of \((D, \delta)\) where \( \delta \) is one of the previously defined metrics.
The following definition of weak convergence will be used:

**Definition 1.3.1**

We write $W_n \xrightarrow{L} W_0$ relative to $(D, \delta)$ if $\lim_{m \to \infty} E(\psi(W_m)) = E(\psi(W_0))$ for all bounded functions $\psi$ defined on $D$ which are continuous in the $\delta$-metric and are such that $\psi(W_m), m \geq 0,$ are measurable with respect to $\beta$. Such convergence is called convergence in law or weak convergence.
II. THE P-DEPENDENT CASE

II.1. Introduction

Let $X_1, \ldots, X_m, \ldots$ and $Y_1, \ldots, Y_n, \ldots$ be two sequences of random variables with common continuous distribution functions $F$ and $G$ respectively. The following assumptions are made on the underlying probability distributions:

(II.1.1) \begin{align*}
(i) \quad & (X_1, Y_1), (X_2, Y_2), \ldots \text{ is a strictly stationary sequence of random vectors.} \\
(ii) \quad & (X_1, Y_1), (X_2, Y_2), \ldots \text{ is a p-dependent sequence of random vectors with } p \geq 1 \\
& \quad \text{(see Definition II.1.1).}
\end{align*}

In this chapter the following additional notation will be used:

(II.1.2) \begin{align*}
H_\lambda &= \lambda F + (1-\lambda)G \\
H &= H_{\lambda N} \\
K_N &= FH_N^{-1} \\
K_\lambda &= FH_\lambda^{-1} \\
K &= K_{\lambda N}.
\end{align*}

For later reference, let us define for $i = 1, \ldots, p+1$ and $m \geq 1$, ...
II.1.3 \[
\begin{aligned}
  m_i &= \left\lfloor \frac{m}{p+1} \right\rfloor \\ &\quad \text{if } \left\lfloor \frac{m}{p+1} \right\rfloor (p+1) + i > m \\
  &= \left\lfloor \frac{m}{p+1} \right\rfloor + 1 \quad \text{otherwise}
\end{aligned}
\]

\[
F_m^{(i)}(x) = m_i^{-1} \sum_{j=0}^{m_i-1} I_{(-\infty,x]}(X_j(p+1)+i)
\]

\[
U_m^{(i)}(t) = m_i^{1/2} (F_m^{(i)}F_m^{(-1)}(t) - t).
\]

Observe that \(F_m^{(i)}(x)\) is a sum of independent random variables and that

\[
\begin{aligned}
  F_m(x) &= \sum_{i=1}^{p+1} (m_i/m) F_m^{(i)}(x) \\
  U_m(t) &= \sum_{i=1}^{p+1} (m_i/m)^{1/2} U_m^{(i)}(t).
\end{aligned}
\]

II.2. The space \(D \times D\)

Consider the set \(D \times D\) where \(D\) is the set of right continuous, real valued functions on \([0,1]\), having only jump discontinuities. Let \((x',x'')\) and \((y',y'')\) be elements of \(D \times D\). Define a metric on \(D \times D\) by

\[
(II.2.1) \quad d'(x',y'),(y'',y'')) = \max\{d(x',y'),d(x'',y'')\}
\]

where \(d\) is Billingsley's metric discussed in Section I.3. As shown in Billingsley (1968, page 224), \(d'\) specifies the
product topology. Further, the space \((\mathbb{R}^d \times \mathbb{R}^d, d')\) is separable and complete.

For points \((s_1, t_1), \ldots, (s_k, t_k)\) in \([0,1] \times [0,1]\), define the projection \(\Pi^*_{(s_1, t_1, \ldots, s_k, t_k)}\) from \(\mathbb{R}^d \times \mathbb{R}^d\) to \(\mathbb{R}^{2k} = 2k\)-dimensional Euclidean space as

\[
(\Pi^*_{(s_1, t_1, \ldots, s_k, t_k)}(x, y)) = (\pi^S_{s_1}(x), \pi_{t_1}^y(y), \ldots, \pi_{t_k}^y(y))
\]

where \((x, y) \in \mathbb{R}^d \times \mathbb{R}^d\) and \(\pi^S_{s_1}\) is the natural projection from \(\mathbb{R}^d\) to \(\mathbb{R}^{1}\). We now show that \(\Pi^*_{(s_1, t_1, \ldots, s_k, t_k)}\) is measurable. To see this, observe that \(\pi^S_{s_1}\) is the natural projection from \(\mathbb{R}^d\) to \(\mathbb{R}^{1}\) and therefore measurable. Since a mapping into \(\mathbb{R}^{2k}\) is measurable if each component is, we have that \(\Pi^*_{(s_1, t_1, \ldots, s_k, t_k)}\) is measurable.

Define the finite dimensional sets as sets of the form \(\Pi^*_{(s_1, t_1, \ldots, s_k, t_k)}(H)\) with \(k \geq 1\) and \(H \in \mathbb{E}^{(2k)}\). Each finite dimensional set lies in \(\mathbb{R} \times \mathbb{R}\) by the definition of measurability.

If \(T_i, i = 1, 2\), are subsets of \([0,1]\), then let \(F_{T_1 \times T_2}\) be the class of sets \(\Pi^*_{(s_1, t_1, \ldots, s_k, t_k)}(H)\) where \(k\) is arbitrary, the \(s_i\) are arbitrary points of \(T_1\), the \(t_j\) are arbitrary points of \(T_2\) and \(H \in \mathbb{E}^{(2k)}\). \(F\) is a finitely additive field.

A subclass \(\mathcal{S}\) of \(\mathbb{R}^d \times \mathbb{R}^d\) is called a determining class, if two measures on \((\mathbb{R}^d \times \mathbb{R}^d, \mathcal{D} \times \mathcal{D})\) are identical whenever they agree
on $S$. Observe that any field that generates $\mathcal{D} \times \mathcal{D}$ will be a determining class.

It follows from Lemma II.2.1 below that if $P$ is a probability measure on $(\mathcal{D} \times \mathcal{D}, \mathcal{D} \times \mathcal{D})$ and its finite dimensional distributions are defined by $P \mathcal{H}^{*1}(s_1, t_1, \ldots, s_k, t_k)$, then $P$ is completely determined by its finite dimensional distributions for points in $T_1 \times T_2$ as defined in Lemma II.2.1.

Lemma II.2.1

If each $T_i$, $i = 1, 2$, contains $1$ and is dense on $[0,1]$, then $F_{T_1 \times T_2}$ is a determining class.

Proof:

Since $F_{T_1 \times T_2}$ is a finitely additive field, it will suffice to show that $F_{T_1 \times T_2}$ generates $\mathcal{D} \times \mathcal{D}$. Further, $\mathcal{D} \times \mathcal{D}$ is separable and it is therefore enough to show that each open $d'$-sphere $S_d((x,y), r)$ lies in the $\sigma$-field generated by $F_{T_1 \times T_2}$. Fix the center $(x,y)$ and the radius $r$. From the definition of $d'$ it follows that

$$(II.2.1) \quad S_d((x,y), r) = S_d(x,r) \times S_d(y,r).$$

Billingsley (1968 , page 121 and the proof of Theorem 14.5) shows that there exists in $T_1$ a sequence $s_1, s_2, \ldots, s_k, \ldots$ and in $E^{(k)}$ a double sequence of measurable sets $H_k(\varepsilon)$ for $k \geq 1$ and rational $\varepsilon$ in $[0,r]$ such that $S_d(x,r) = \bigcup_{\varepsilon \in S_1, \ldots, s_k} H_k(\varepsilon)$. Similarly, there exists in $T_2$ a
sequence $t_1,\ldots,t_k,\ldots$ and in $E^{(k)}$ a double sequence of measurable sets $H^*_k(\varepsilon)$ such that $S_d((x,y), r) = \bigcup \cap \prod_{t_1,\ldots,t_k}^{-1} (H^*_k(\varepsilon))$.

Now, by Equation II.2.2

$$S_{d'}((x,y), r) = \bigcup \cap \prod_{t_1,\ldots,t_k}^{-1} (H^*_k(\varepsilon)).$$

Let

$$A_k(\varepsilon) = \{(r_1, r_2, \ldots, r_{2k})/\{r_1, r_3, \ldots, r_{2k-1}\} \in H^*_k(\varepsilon)$$

$$B_k(\varepsilon) = \{(r_1, r_2, \ldots, r_{2k})/\{r_2, r_4, \ldots, r_{2k}\} \in H^*_k(\varepsilon)$$

Since $H^*_k(\varepsilon)$ and $H^*_k(\varepsilon)$ are open sets of $E^{(k)}$ we have that $A_k(\varepsilon)$ and $B_k(\varepsilon)$ are open sets of $E^{(2k)}$.

By Equations II.2.3 and II.2.4 and the definition of $\prod(s_1, t_1, \ldots, s_k, t_k)$ we have

$$S_{d'}((x,y), r) = \bigcup \cap \prod_{t_1,\ldots,t_k}^{-1} (A_k(\varepsilon)) \cap \{s_1, t_1, \ldots, s_k, t_k\}$$

$$\prod_{t_1,\ldots,t_k}^{-1} (B_k(\varepsilon)) \}.$$
generated by $F_{T_1 \times T_2}$. Hence, $F_{T_1 \times T_2}$ is a determining class and
the proof of the lemma is complete.

II.3. The one-sample processes

Take $P_{x,m}$ to be the measure on $(D,d)$ corresponding to the
random function $U_m$. Similarly, $P_{y,n}$ is the measure on $(D,d)$
corresponding to the random function $V_n$ and $P_N$ is the measure
on $(D \times D, d')$ corresponding to the two dimensional process
$\{(U_m(s), V_n(t)) : 0 \leq s,t \leq 1\}$.

Lemma II.3.1

There exists a vector of random functions $(U_o, V_o)$ such
that if $\lambda_N = \lambda_\times + 0(N^{-1/2})$ then $(U_m, V_n)^+ \rightarrow (U_o, V_o)$ where
$\{(U_o(s), V_o(t)) : 0 \leq s,t \leq 1\}$ is a two-dimensional Gaussian
process specified by

\[
\begin{align*}
(i) & \quad E[U_o(s)] = 0 \\
(ii) & \quad E[V_o(t)] = 0 \\
(iii) & \quad E[U_o(s) U_o(t)] = E(g_s(x_1) g_t(x_1)) \\
& \quad + \sum_{j=2}^{p+1} E(g_s(x_1) g_t(x_j)) + \sum_{j=2}^{p+1} E(g_s(x_j) g_t(x_1)) \\
(iv) & \quad E[V_o(s) V_o(t)] = E(h_s(y_1) h_t(y_1)) \\
& \quad + \sum_{j=2}^{p+1} E(h_s(y_1) h_t(y_j)) + \sum_{j=2}^{p+1} E(h_s(y_j) h_s(y_1))
\end{align*}
\]
\begin{align*}
(v) \quad & E(U_0(s)V_0(t)) = E(g_s(X_1)h_t(Y_1)) \\
& + \sum_{j=2}^{p+1} E(g_s(X_1)h_t(Y_j)) + \sum_{i=2}^{p+1} E(g_s(X_i)h_t(Y_i))
\end{align*}

where \( g_s(\alpha) = I(\alpha \leq F^{-1}(s)) - s \) and \( h_t(\beta) = I(\beta \leq G^{-1}(t)) - t \).

These series converge absolutely, and with probability one \( U_0 \) and \( V_0 \) are both continuous functions.

**Proof:**

Let \((s_1, t_1)\) and \((s_2, t_2)\) be elements of \([0, 1] \times [0, 1]\). It will be shown that \((U_m(s_1), V_n(t_1), U_m(s_2), V_n(t_2))\) has asymptotically a multivariate normal distribution with parameters as specified by Equation II.3.1.

If \( \lambda_N = 1/2 \) for all \( N \), take \( \eta_i = (g_{s_1}(X_i), h_{t_1}(Y_i), g_{s_2}(X_i), h_{t_2}(Y_i)). \) This sequence of random vectors is \( \phi \)-dependent and therefore \( \phi \)-mixing. From Billingsley (1968, page 177) if \( \xi_n = (\xi_n^{(1)}, \ldots, \xi_n^{(r)}) \) is \( \phi \)-mixing, \( \sum \phi_n^{1/2} < \infty \), and if the \( \xi_n^{(i)} \) have mean zero and finite variance then

\[
\sum_{k=1}^{n} \xi_k \text{ has asymptotically a normal distribution centered at the origin, with covariances}
\]

\[
\sigma_{ij} = E(\xi_1^{(i)} \xi_1^{(j)}) + \sum_{k=2}^{\infty} E(\xi_1^{(i)} \xi_k^{(j)}) + \sum_{k=2}^{\infty} E(\xi_k^{(i)} \xi_1^{(j)})
\]

where the series converges absolutely. Hence, \((U_m(s_1), V_n(t_1), U_m(s_2), V_n(t_2)) = n^{-1/2} \sum_{i=1}^{n} \eta_i\) will have asymptotically a
multivariate distribution with parameters as specified by Equation II.3.1.

If \( \lambda_* = 1/2 \) take \( m^* = \min (m, n) \), then \( \frac{m^*}{n} \rightarrow 1 \) as \( N \rightarrow \infty \) and

\[
\left| \frac{m-m^*}{n^{1/2}} \right| \rightarrow 0 \quad \text{as} \quad N \rightarrow \infty.
\]

Hence,

\[
U_m(s) = \left( \frac{m^*}{m} \right)^{1/2} U_{m^*}(s) + \left( \frac{1}{m} \right)^{1/2} \sum_{i=m^*+1}^{m} \frac{1}{(i-\infty, F^{-1}(s))(X_i)-s} \]

\[
= \left( \frac{m^*}{m} \right)^{1/2} U_{m^*}(s) + o(1).
\]

Similarly, \( V_n(t) = \left( \frac{m^*}{n} \right)^{1/2} V_{m^*}(t) + o(1). \) Therefore, by Slutsky's theorem, \( (U_m(s_1), V_n(t_1), U_m(s_2), V_n(t_2)) \) has asymptotically the same distribution as \( (U_{m^*}(s_1), V_{n^*}(t_1), U_{m^*}(s_2), V_{n^*}(t_2)) \). The result follows as in the previous case.

If \( \lambda_* < 1/2 \), there exists \( N^* \) such that \( m < n \) for \( N > N^* \). For \( N > N^* \) we have

\[
V_n(t) = \left( \frac{m}{n} \right)^{1/2} V_{m}(t) + \left( \frac{1}{n} \right)^{1/2} \sum_{j=m+1}^{m+p+1} \frac{1}{(i-\infty, G^{-1}(t))(Y_j) - t}
\]

\[
+ \left( \frac{1}{n} \right)^{1/2} \sum_{j=m+p+1}^{n} \frac{1}{(i-\infty, G^{-1}(t))(Y_j) - t}.
\]

Now, \( \left( \frac{1}{n} \right)^{1/2} \sum_{j=m+1}^{m+p+1} \frac{1}{(i-\infty, G^{-1}(t))(Y_j) - t} \frac{1}{n} = o(1). \) Since \( \lambda_N = \lambda_* + o(N^{-1/2}) \) we have that \( \frac{m}{n} \rightarrow \frac{\lambda}{1-\lambda} \) and \( \frac{n-m-p-1}{n} \rightarrow \frac{1-2\lambda}{1-\lambda} \) as \( N \rightarrow \infty \). It follows from Slutsky's theorem that \( (U_m(s_1), V_n(t_1), U_m(s_2), V_n(t_2)) \) has asymptotically the same
distribution as

\[
(U_m(s_1), \left\{ \frac{\lambda_*}{1-\lambda_*} \right\}^{1/2} V_m(t_1) + \left\{ \frac{1-2\lambda_*}{1-\lambda_*} \right\}^{1/2} \left( \frac{1}{n-m-p-l} \right)^{1/2} \sum_{j=m+p+1}^{n} I(\infty, G^{-1}(t_1)](Y_j-t_1)
\]

\[
U_m(s_2), \left\{ \frac{\lambda_*}{1-\lambda_*} \right\}^{1/2} V_m(t_2) + \left\{ \frac{1-2\lambda_*}{1-\lambda_*} \right\}^{1/2} \left( \frac{1}{n-m-p-l} \right)^{1/2} \sum_{j=m+p+1}^{n} I(-\infty, G^{-1}(t_2)](Y_j-t_2)\).
\]

Observe that \( Y_j \) and \( (U_m(s), V_m(t)) \) are independent if \( j \geq m+p+1 \). Applying the central limit for \( \phi \)-mixing processes we have that \( (U_m(s_1), V_n(t_1), U_m(s_2), V_n(t_2) \) has asymptotically a multivariate normal distribution with parameters as specified by Equation II.3.1.

A set \{\((s_1, t_1), \ldots, (s_k, t_k)\)\} of elements of \([0,1] \times [0,1]\), where \( k > 2 \), can be treated in the same way.

Since \( \{P_{X,m}\} \) and \( \{P_{Y,n}\} \) are each tight on \((D,d)\), (note that a \( p \)-dependent process is also \( \phi \)-mixing, tightness for \( \phi \)-mixing processes is shown in the proof of Theorem 22.1 of Billingsley (1968)) we have that \( P_N \) is tight on \((DxD, d')\) (This is exactly the result of problem 6 on page 41 of Billingsley (1968)). Since \( \{P_N\} \) is tight, it is also relatively
compact. Let \( \{P^*_n\} \) be a subsequence. Then, by relative compactness, there exists a further subsequence \( \{P^*_m\} \) which converges weakly to a limit, say \( P \). Since this limit must have the finite distributions described, it follows by Lemma II.2.1 that \( P \) is unique.

Let \( \{(U_\omega(s), V_\omega(t)) : 0 \leq s, t \leq 1\} \) be the vector of random functions corresponding to the measure \( P \). Observe that, marginally, \( U_\omega \) is a Gaussian random function specified by Equations II.3.1(i) and II.3.2(ii). Therefore, \( U_\omega \) has the same finite dimension distributions as the random function \( Y \), as would be specified by Theorem 22.1 of Billingsley (1968). Since a random function on \( D \) is determined by its finite dimensional distributions we have that \( U_\omega \) is equivalent to \( Y \). Now, \( Y \) is continuous with probability one. Thus, \( U_\omega \) is continuous with probability one. Similarly, \( V_\omega \) is continuous with probability one and is a Gaussian random function specified by Equations II.3.1(ii) and II.3.2(iv). Therefore, with probability one, \( U_\omega \) is continuous and \( V_\omega \) is continuous. The proof of Lemma II.3.1 is now complete.

**Lemma II.3.2**

There exist two dimensional processes \( \{(\tilde{U}_m(s), \tilde{V}_n(t)) : 0 \leq s, t \leq 1\} \), \( N > 0 \), having the same finite dimensional distributions as \( \{(U_m(s), V_n(t)) : 0 \leq s, t \leq 1\} \), \( N > 0 \), but which in addition satisfy \( d'((\tilde{U}_m, \tilde{V}_n), (\bar{U}_\omega, \bar{V}_\omega)) \to 0 \). Further, all
processes are defined on a single probability space \((\tilde{\Omega}, \tilde{U}, \tilde{P})\).

**Proof:**

The metric space \((D \times D, d')\) is complete and separable. If we consider the sequence of random variables \((U_m, V_n)\) and apply item 3.1.1 of Skorokhod (1956), the result follows.

**Lemma II.3.4**

\[ P(U_m, U_o) \to 0 \text{ as } m \to \infty. \]

**Proof:**

\(U_o\) is continuous with probability one. The proof is completed by observing that \(d\)-convergence to a continuous function is equivalent to \(p\)-convergence (see Billingsley (1968, page 112)).

**Lemma II.3.5**

Let \(q\) be any non-negative function which is nondecreasing on \([0, \theta]\) for \(\theta < 1\). Then there exists a constant \(C_\theta > 0\) such that

\[ (II.3.2) \quad P(|U_m(t)| \leq q(t), 0 \leq t \leq \theta) \geq 1 - C_\theta (p+1) \int_0^\theta (q(t))^{-2} dt \]

\[ (II.3.3) \quad P(|U_m(t)| \leq q(t), 1-\theta \leq t \leq 1) \geq 1 - C_\theta (p+1) \int_0^\theta (q(t))^2 dt \]

for all \(m \geq 0\). Moreover, \(C_\theta\) is nonincreasing as \(\theta \to 0\).
Proof:

Take \( m \geq 1 \). From Equation 2.4 and Lemma 2.2 of Pyke and Shorack (1968), it follows that

\[
\begin{align*}
\text{Proof:} \\
& \text{Take } m \geq 1. \text{ From Equation 2.4 and Lemma 2.2 of Pyke} \\
& \text{and Shorack (1968), it follows that} \\
& (II.3.4) \left\{ \begin{array}{l}
P\{ |U^i_m(t)| \leq q(t), 0 \leq t \leq \theta \} \\
\geq 1 - C_\theta \int_0^\theta (q(t))^{-2} \, dt
\end{array} \right.
\end{align*}
\]

for \( i = 1, \ldots, p + 1 \). Observe that \( C_\theta \) does not depend on \( m \) or \( p \) and that \( C_\theta \) is nonincreasing as \( \theta \to 0 \).

Since \( U^i_m(t) = \sum_{i=1}^{m_i} U^i_m(t) \),

\[
P\{ |U^i_m(t)| \leq q(t) : 0 \leq t \leq \theta \}
\geq 1 - \sum_{i=1}^{p+1} P\{ |U^i_m(t)| \geq q(t) : 0 \leq t \leq \theta \}.
\]

From Equations II.3.4 and II.3.5

\[
P\{ |U^i_m(t)| \leq q(t) : 0 \leq t \leq \theta \}
\geq 1 - \sum_{i=1}^{p+1} P\{ |U^i_m(t)| \geq q(t) : 0 \leq t \leq \theta \}
\]

For \( U_0 \) observe that Lemma II.2.1 can be extended from \( D_0^2 \) to \( D_0^{p+1} \). It will be shown that

\[
(U^1_m, U^2_m, \ldots, U^{p+1}_m) \overset{L}{\rightarrow} (U^1_0, U^2_0, \ldots, U_0^{p+1})
\]
where \((U_0^{(1)}, U_0^{(2)}, \ldots, U_0^{(p+1)})\) is a \((p+1)\)-dimensional process such that marginally each \(U_0^{(i)}\) is a tied down Wiener process and such that \((p+1)^{-1/2} \sum_{i=1}^{p+1} U_0^{(i)}\) and \(U_0\) are the same.

By Donsker's theorem, the measures corresponding to \(U_m^{(i)}, m = 1, 2, \ldots\), are tight. As observed, in the proof of Lemma II.3.1, this implies that the measures on \(D^{p+1}\) corresponding to \((U_m^{(1)}, U_m^{(2)}, \ldots, U_m^{(p+1)})\) are tight.

We will now consider the finite dimensional distributions. Let \(s_i, i = 1, 2, \ldots, 2(p+1)\), be elements of \([0,1]\). The asymptotic distribution of \((U_m^{(1)}(s_1), U_m^{(2)}(s_2), \ldots, U_m^{(1)}(s_{p+2}), \ldots, U_m^{(p+1)}(s_2(p+1))\) will be determined. For \(s \in [0,1]\) and any \(i\), we have

\[
U_m^i(s) = m_i^{-1/2} \sum_{j=1}^{m_i} (I(-\infty, F^{-1}(s)](X_j^{(p+1)}+i) - s)
\]

\[
= ([m/p]/m_i)^{1/2} ([m/p])^{-1/2} \sum_{j=1}^{[m/p]} (I(-\infty, F^{-1}(s)](X_j^{(p+1)}+i) - s) + o(1)
\]

\[
= ([m/p]/m_i)^{1/2} U_i^{(i)} (p[m/p]) (s) + o(1).
\]

Since for each \(i\), \([m/p]/m_i \to 1\) as \(N \to \infty\), we may, by Slutsky's theorem, assume \(m = kp\) for some \(k\). Let

\[
W_j = (g_{s_1}^{(1)}(X_j^{(p+1)}+1), \ldots, g_{s_{p+1}}^{(1)}(X_j^{(p+1)}(p+1)),
\]

\[
g_{s_{p+2}}^{(2)}(X_j^{(p+1)}+1), \ldots, g_{s_{2(p+1)}}(X_j^{(p+1)}(p+1)))
\]
where \( j = 0,1, \ldots \). Then by the multivariate central theorem
given by Billingsley (1968) on page 177, it follows that
\[
(U_{m_1}(s_1), \ldots, U_{m_2}(s_2(p+1))) = \frac{1}{K} \sum_{j=0}^{K-1} W_j
\]
has asymptotically a normal distribution, centered at the origin with covariances
\[
\sigma_{ij} = E(g_{s_i}(X_i) g_{s_j}(X_j))
+ E(g_{s_i}(X_{i+p+1}) g_{s_j}(X_j)) \text{ if } i \leq j \leq p+1
= E(g_{s_i}(X_{i-p-1}) g_{s_j}(X_{j-p-1}))
+ E(g_{s_i}(X_i) g_{s_j}(X_{j-p-1})) \text{ if } p+1 < i \leq j
= E(g_{s_i}(X_i) g_{s_j}(X_{j-p-1}))
+ E(g_{s_i}(X_i) g_{s_j}(X_j)) \text{ if } i \leq p+1 < j.
\]
Marginally \((U_{m_1_i}(s_i), U_{m_1_i}(s_{i+p+1}))\) has the asymptotic distribution appropriate for tied-down Wiener measure. Since
the sequence is tight, \( U_{m_i} \) converges weakly to a tied-down Wiener measure.

Take \( s, t \in [0,1] \), then
\[
(p+1)^{-1/2} \sum_{i=1}^{p+1} U_{0i}(s),
(p+1)^{-1/2} \sum_{i=1}^{p+1} U_{0i}(t)
\]
has a bivariate normal distribution, centered at the origin and specified by
\[
E\left\{ \frac{1}{p+1} \left( \sum_{i=1}^{p+1} U_{oi}(s) \right) \left( \sum_{i=1}^{p+1} U_{oi}(t) \right) \right\} \\
= \frac{1}{p+1} \left( \sum_{i=1}^{p+1} E(g_s(X_i)g_t(X_i)) \right) + \\
+ \sum_{i=1}^{p+1} \sum_{j<i} (E(g_s(X_i)g_t(X_j)) + E(g_s(X_i)g_t(X_{j+p+1}))) \\
+ \sum_{i=1}^{p+1} \sum_{j>i} (E(g_s(X_i)g_t(X_j)) + E(g_s(X_{i+p+1})g_t(X_j))) \\
= \frac{1}{p+1} \left( (p+1) E(g_s(X_1)g_t(X_1)) \right) \\
+ \sum_{i=1}^{p+1} \sum_{j=1}^{i+p} E(g_s(X_i)g_t(X_j)) \\
+ \sum_{j=1}^{p+1} \sum_{i=j+1}^{p+1} E(g_s(X_i)g_t(X_j)) \\
+ \sum_{j=p+2}^{2(p+1)} \sum_{i=1}^{j-p-2} E(g_s(X_i)g_t(X_j)) \\
= \frac{1}{p+1} \left( (p+1) E(g_s(X_1)g_t(X_1)) \right) \\
+ \sum_{i=1}^{p+1} \sum_{j=i+1}^{i+p} E(g_s(X_i)g_t(X_j)) \\
+ \sum_{j=1}^{p+1} \sum_{i=j+1}^{j+p} E(g_s(X_i)g_t(X_j)) \\
= \frac{1}{p+1} \left( (p+1) E(g_s(X_1)g_t(X_1)) \right) \\
+ (p+1) \sum_{j=2}^{p+1} E(g_s(X_1)g_t(X_j)) + (p+1) \sum_{i=2}^{p+1} E(g_s(X_i)g_t(X_j)) \right) .
\]
Thus, the finite dimensional distributions of 
\[
\left\{(p+1)^{-1/2} \sum_{i=1}^{p+1} U_{oi}(s), 0 \leq s \leq 1\right\}
\]
are the same as those of 
\\{U_o(s), 0 \leq s \leq 1\}. Therefore, by Lemma II.2.1, the measure

on (D,d) determined by 
\[
(p+1)^{-1/2} \sum_{i=1}^{p+1} U_{oi}
\]
is the same as that determined by 
\(U_o\).

It now follows from lemma 2.2 of Pyke and Shorack that

\[
P\{\left|U_o(t)\right| \leq q(t), 0 \leq t \leq \theta\}
\]

\[
= P\left\{\left|\frac{1}{p+1} \sum_{i=1}^{p+1} U_{oi}(t)\right| \leq q(t), 0 \leq t \leq \theta\right\}
\]

\[
\geq 1 - \sum_{i=1}^{p+1} P\{\left|U_{oi}(t)\right| > q(t), 0 \leq t \leq \theta\}
\]

\[
\geq 1 - (p+1) C_\theta \int_0^\theta (q(t))^{-2} \, dt.
\]

The proof of Equation II.3.2 is now complete.

Since for all \(i\), the reversed process \(U_{m}^- = U_{m}^i((1-t)-)\) has the same finite dimensional distributions as the \(U_{m}^-\) process, we can in a similar manner prove Equation II.3.3.

The proof of Lemma II.3.5 is now complete.

**Definition II.3.1**

Let \(\rho_q(f,g) = \rho(f/q, g/q)\) and \(d_q(f,g) = d(f/q, g/q)\) whenever well defined.
**Definition II.3.2**

Let $Q'$ denote the class of all nonnegative functions defined on $[0,1]$ which for some $\varepsilon > 0$ are bounded away from zero on $(\varepsilon, 1-\varepsilon)$, are nondecreasing (nonincreasing) on $[0,\varepsilon]$ $([1-\varepsilon,1])$ and which have square integrable reciprocals. Let $Q = \{q \in D/q \geq q'$ some $q' \in Q'\}$.

**Theorem II.3.1**

For $q \in Q$, $\rho_q(U_m, U_o) \xrightarrow{p} 0$ and $d_q(U_m', U_o) \xrightarrow{p} 0$ as $m \to \infty$.

**Proof:**

The proof is similar to that of Theorem 2.1 of Pyke and Shorack (1968). For any $0 < a \leq \varepsilon \leq 1-\varepsilon \leq b < 1$,

\[
\rho_q(U_m, U_o) \leq \sup_{0 \leq t \leq a, b \leq t \leq 1} \frac{(U_m(t) + U_o(t))}{q(t)}
\]

(II.3.6)

\[+ \rho(U_m', U_o)/\inf_{a \leq t \leq b} q(t),\]

where $\varepsilon$ is determined by $q$. For each $a, b$, the second term on the right converges to zero in probability by Lemma II.3.4.

Fix $\delta > 0$. Since $q^{-2}$ is integrable, we can choose $a < \varepsilon$ and $1 > b > 1-\varepsilon$ so that $\int_{a}^{b} [q(t)]^{-2} dt < C_a^{-1} \delta^3/p+1$ and $\int_{a}^{b} [q(t)]^{-2} dt < C_{1-b}^{-1} \delta^3/p+1$. Then by Lemma II.3.5

\[P\{\sup_{0 < t \leq a} \frac{|U_m(t)/q(t)|}{\delta} \leq \delta\} > 1-\delta\]
and
\[ P\{ \sup_{0 \leq t \leq 1} |U_m(t)/q(t)| \leq \delta \} > 1-\delta. \]

Since these inequalities hold for \( m = 0 \) the first term on the right side of Equation II.3.6 does not exceed \( 4\delta \) on an event whose probability is greater than \( 1-4\delta \). This completes the proof for \( P_q \).

To see that the result for \( \rho \) implies the result for \( d \), observe that by the definition of \( d \), \( d(x,y) \leq \rho(x,y) \) for all \( x,y \in D \).

The proof of Theorem II.3.1 is now complete.

**Definition II.3.3**

Let \( U_m^*(K_N) \) equal \( U_m(K_N) \) for \( \frac{1}{N} \leq t \leq 1-\frac{1}{N} \) and equal zero otherwise.

**Theorem II.3.2**

For \( q \in Q \), \( \rho_q(U_m^*(K_N), U_0(K)) \not\rightarrow 0 \) uniformly in all continuous \( F \) and \( G \) and all \( \lambda_N \in \Delta \).

**Proof:**

The Glivenko-Cantelli Theorem implies that \( \rho(F^{(i)}_m, F) \) \( \not\rightarrow 0 \) and \( \rho(G^{(i)}_n, G) \) \( \not\rightarrow 0 \) uniformly in \( F \) and \( G \). Since

\[ H_N - H = (\lambda_N) \sum_{i=1}^{p+1} (F^{(i)}_m - F)(m_i/m) + (1-\lambda_N) \sum_{i=1}^{p+1} (G^{(i)}_n - G)(n_i/n), \]

it follows that \( \rho(H_N^*, H) \) \( \not\rightarrow 0 \) uniformly in \( F \) and \( G \) and all \( \lambda_N \in [0,1] \). It now follows (see Lemma 2.3 of Pyke and Shorack (1968)) that \( \rho(HH^{-1}_N, HH^{-1}) \) \( \not\rightarrow 0 \) uniformly in \( F \) and \( G \)
and \( \lambda_N \in [0,1] \).

Since \( \rho(HH^{-1}, HH^{-1}) \rightarrow 0 \) and \( H = \lambda_N F + (1-\lambda_N)G \) we have that \( \rho(K_N, K) \rightarrow 0 \); note that \( K_N - K \) and \( GH_N - GH \) have the same sign.

Applying the triangle inequality,

\[
\rho(U_m(K_N), U_\infty(K)) \leq \rho(U_m(K_N), U_m(K)) + \rho(U_m(K_N), U_\infty(K)).
\]

The first term on the right converges to zero almost surely, by Lemma II.3.4. Since \( \rho(K_N, K) \rightarrow 0 \) and \( U_\infty \) is continuous a.s. it follows that \( \rho(U_m(K_N), U_\infty(K)) \rightarrow 0 \) and therefore by Equation II.3.7 we have

\[
(II.3.8) \quad \rho(U_m(K_N), U_\infty(K)) \rightarrow 0
\]

uniformly in \( F \) and \( G \) and \( \lambda_N \in [0,1] \).

Using Lemma 8 of Govindarajulu et al. (1967), we have that for \( \epsilon > 0 \), there exist \( b_i, i = 1, \ldots, p+1 \), such that for each \( i \)

\[
P(A_m^{(i)}) > 1 - \epsilon/p+1
\]

where

\[
(II.3.9) \quad A_m^{(i)} = \{ F(t) \leq (b_i/2p+2) F_m^{(i)}(t) \text{ for all } t \}
\]

If \( b = \max_i b_i \), then for each \( i \)
The last inequality follows since

\[
\frac{m_i}{m} \geq \frac{[m/p+1]}{m} \geq \frac{[m/p+1]}{(p+1)([m/p+1]+1)} \geq \frac{1}{(2p+2)}.
\]

If

\[ A_m = \{ F(t) < b F_m(t) \text{ for all } t \text{ where } F_m(t) > 0 \}, \]

then by Equations II.3.9 and II.3.10 we have

\[
P(A_m^c) = 1 - P(A_m^c) \geq \sum_{i=1}^{p+1} P(A_m^c) \geq 1 - \varepsilon.
\]

It now follows, using the proof of Lemma 2.5 of Pyke and Shorack (1968), that for \( \varepsilon > 0 \) there exists \( b > 0 \) such that

\[
P(K_N(t) \leq 2b \lambda^{1/\kappa} t \text{ for } t > 1/N) \geq 1-\varepsilon.
\]
We can now use Equations II.3.8 and II.3.11 to prove that $\rho_q(U_m(K_N), U_0(K)) \to 0$ uniformly in continuous $F$ and $G$ and $\lambda_N \in [0,1]$. The proof is similar to that of Theorem 2.2, Pyke and Shorack (1968).

We need only consider the intervals $[0,\delta]$ and $[1-\delta, 1]$ where $\delta$ is small. To see this observe that on the interval $(\delta, 1-\delta), q(K)$ and with high probability, $q(K_N)$ are bounded away from zero. Hence, the supremum over $(\delta, 1-\delta)$ converges by Equation II.3.8. Also, we need only consider the interval $[0,\delta]$ since the interval $[1-\delta, 1]$ can be considered using the reverse process.

Assume without loss of generality that $q$ is nondecreasing on $[0,\delta]$. For $\varepsilon, \eta > 0$ choose $b$ to satisfy Equation II.3.11. Then use Lemma II.3.5 to choose $a > 0$ so that

$$P\{|U_m(t)| \leq \eta q(\lambda_o t/2b), 0 < t < a\} > 1-\varepsilon;$$

then choose $\delta > 0$ such that

$$P\{K_N(t) < a, 0 < t \leq \delta\} > 1-\varepsilon$$

for large $N$. Such a $\delta$ is possible since $\rho(HH_{-1}, HH_{-1}) \to 0$ and $K_N(t) \leq \lambda_o^{-1} HH_{-1}^{-1}(\delta)$ for $t \leq \delta$. Thus for $N$ large, we have

$$P\{\sup_{0 \leq t \leq \delta} |U^*_m(K_N(t))/q(t) | / q(t) \leq \eta\}$$

$$\geq P\{|U_m(t)| \leq \eta q(\lambda_o t/2b), 0 < t < a\}$$
\( \cap (K_N(t) < \alpha, 0 < t < \delta) \)

\( \cap (K_N(t) < 2b\lambda_o^{-1}(t), t \geq 1/N) \).

Since \( \epsilon \) and \( n \) are arbitrary this completes the proof of Theorem II.3.2.

II.4. The two-sample process

This section consists of results concerning the weak convergence of \( L_N \) to a limiting process when \( \lambda_N \to \lambda_o \) and the Chernoff and Savages Theorems which follow from the weak convergence of \( L_N \). These results are obtained using arguments given in sections 3, 4 and 5 of Pyke and Shorack (1968) and Pyke and Shorack (1969). The theorems of this section are therefore stated without proof.

**Lemma II.4.1**

With probability one,

\[
L_N(t) = (1-\lambda_o)^{-1/2} B_N(t) \Phi^{-1}(t) \]

\[
- (1-\lambda_o)^{-1/2} A_N(t) \Gamma^{-1}(t) + \delta_N(t)
\]

for all \( t \in [0,1] \) where

\( \delta_N(t) = A_N(t) N^{1/2} (H_N^{-1}(t) - t) \)

\( A_N(t) = (K(u_t) - K(t))/(u_t - t), u_t = H_N^{-1}(t) \)

and \( B_N \) is defined by
\[ \lambda_N A_N(t) + (1 - \lambda_N) B_N(t) = 1 \]

\( L_N \) is defined by left continuity at any otherwise undefined point.

**Proof:**

The proof is similar to that of Lemma 3.1 of Pyke and Shorack (1968).

Since \( \lambda_N F_{H^{-1}}(t) + (1 - \lambda_N) G_{H^{-1}}(t) = t \), \( F_{H^{-1}} \) and \( G_{H^{-1}} \) are absolutely continuous. Let \( a_N \) and \( b_N \) denote the derivatives of \( F_{H^{-1}} \) and \( G_{H^{-1}} \) respectively, which exist on \([0,1]\) except on a set of Lebesgue measure zero. Let \( a_0 \) and \( b_0 \) denote the derivatives of \( F_{H_0^{-1}} \), \( G_{H_0^{-1}} \) where \( H_0 = \lambda_0 F + (1 - \lambda_0) G \). Now, set

\[
L_0(t) = (1 - \lambda_*) \left\{ (1 - \lambda_*)^{-1/2} \left[ b_0(t) U_0(F_{H_0^{-1}}(t)) \right. \right.
\left. - (1 - \lambda_*)^{-1/2} a_0(t) V_0(G_{H_0^{-1}}(t)) \right\}.
\]

**Assumption II.4.1.**

The functions \( K_\lambda \) have derivatives \( a_\lambda \) for all \( t \in (0,1) \) and have one-sided limits at 0 and 1.

**Theorem II.3.**

(a) Suppose Assumption II.1 holds, \( \lambda_N + \lambda_0 \) and \( q \in \mathbb{Q} \). Then \( \rho_q(L'_N, L'_0) \rightarrow 0 \) so that \( L'_N \overset{p}{\rightarrow} L'_0 \) relative to \( (D^-, \rho_q) \), where \( D^- \) is the set of left continuous functions on \([0,1]\). The same statement holds for \( d_q \).
(b) If, in addition, the measures \( \{ \nu_N : N \geq 1 \} \) and \( \nu \) of Section 2.1 satisfy

\[
\begin{align*}
(ii) \quad & \int_0^1 q \, d|v| < \infty \\
(i) \quad & \int_0^1 \frac{1}{N} \, d(\nu_N - \nu) \to 0, \text{ and,}
\end{align*}
\]

then \( T_N^* \to \int_0^1 L_0 \, d|v| \), a \( N(0, \sigma_0^2) \) random variable where

\[
\sigma_0^2 = 2(1-\lambda_*)^2 \lambda_*^{-1} \int_0^1 \int_0^V b_0(u) b_0(v) FH_0^{-1}(u) (1-FH_0^{-1}(v)) \, dv(u) \, dv(v)
\]

\[
+ \lambda_*^{-1} \sum_{k=1}^d \int_0^1 \int_0^1 b_0(u) b_0(v) (F_k(H_0^{-1}(u), H_0^{-1}(v)) - FH_0^{-1}(u) FH_0^{-1}(v)) \, dv(u) \, dv(v)
\]

\[
+ (1-\lambda_*)^{-1} \int_0^1 \int_0^V a_0(u) a_0(v) GH_0^{-1}(u) (1-GH_0^{-1}(v)) \, dv(u) \, dv(v)
\]

\[
+ (1-\lambda_*)^{-1} \sum_{k=1}^d \int_0^1 \int_0^1 a_0(u) a_0(v) (G_k(H_0^{-1}(u), H_0^{-1}(v))
\]

\[
- GH_0^{-1}(u) GH_0^{-1}(v)) \, dv(u) \, dv(v).
\]

\[
+ (1-\lambda_*)^{-1} \sum_{k=0}^d \int_0^1 \int_0^1 a_0(u) b_0(v) (H_k(H_0^{-1}(u), H_0^{-1}(v))
\]

\[
- GH_0^{-1}(u) FH_0^{-1}(v)) \, dv(u) \, dv(v).
\]

where \( F_k, G_k \) and \( H_k \) denote respectively the joint distributions of \( (X_1, X_{k+1}), (Y_1, Y_{k+1}) \) and \( (X_1, Y_{1+k}) \).
Whenever well defined let
\[ ||f||_\nu = \int_0^1 |f(t)| \, d|\nu| (t). \]

**Theorem II.4.2**

(a) Suppose \( K_0 \) is differentiable a.e. \(|\nu|, \lambda_N = \lambda_* + O(N^{-1/2}) \) and \( \int_0^1 q \, d|\nu| < \infty \) for some \( q \in Q \). Then \( \|L_N - L_0\|_{\nu} \overset{p}{\to} 0 \) so that \( L_N \overset{d}{\to} L_0 \) relative to \((D^-, \|\cdot\|_\nu)\).

(b) Suppose in addition that the measures \( \{\nu_N : N \geq 1\} \) of Section 1 satisfy
\[ (i) \int_1^1 L_N \, d(\nu_N - \nu) \overset{p}{\to} 0. \]

Then \( T_N \overset{p}{\to} \int_0^1 L_0 \, d\nu \) which is a \( N(0, \sigma^2) \) random variable with \( \sigma^2 \) given in Theorem II.4.1.

**Remark:**

Proposition 5.1 of Pyke and Shorack (1968) shows that the conditions of Theorem II.4.1 are weak enough to cover most cases of interest.

The following conditions are used in the concluding theorem:

(C1) There exists \( q \in Q \) such that \( \int_0^1 q \, d|\nu| < M < \infty \).

(C2) \( N^{-1/2} \sum_{i=1}^N |C_{Ni}^* - J(i/N \cdot (1-1/N))| \leq \delta_N \)

where \( \delta_N = o(1) \).

(C3) \( N^{1/2} |\lambda_N - \lambda_*| \leq M_N \) where \( M_N = o(1) \) and \( F \circ H_{N^{-1}} \) is differentiable a.e. - \(|\nu|\)
(C4) \( F, G \) (which may depend on \( N \)) and \( \{\lambda_N\} \) are such that the functions \( F \circ H^{-1} \) have derivatives \( a_N \) which form a uniformly equicontinuous family and \( a_N \) converges uniformly to \( a_0 \) as \( N \to \infty \). (Given \( \varepsilon > 0 \) there exists \( \delta_\varepsilon, N_\varepsilon \) such that \( |a_N(s) - a_N(t)| < \varepsilon \) for all \( |s-t| < \delta_\varepsilon \) and for all \( N \), and \( N > N_\varepsilon \) implies \( |a_N(t) - a_0(t)| < \varepsilon \) for all \( t \).)

(C2') (a) \( \sum_{i=1}^{N-1} [C_{Ni}^* - J(i/N)] Z_{Ni} = o_p(N^{-1/2}). \)
(b) \( C_{NN}^* = o(N^{1/2}). \)
(c) \( J(1 - 1/N) = o(N^{1/2}). \)

Theorem II.4.3

(i) Under (C1), (C2') and (C3)

\[ T_N \to \int_0^1 L_0 \, dv \text{ as } N \to \infty. \]

(ii) Under (C1) (C2) (C3) and (C4)

\[ T_N \to \int_0^1 L_0 \, dv \text{ as } N \to \infty \]

and the convergence is uniform in the set of all \( F, G, \lambda_N, C_{Ni}^* \)'s and \( J \) such that the conditions hold for fixed \( q, M, \delta_N^* \)'s, \( M_N^* \)'s, \( \varepsilon_\delta^* \)'s and \( N_\varepsilon^* \)'s.
III. THE MIXING CASE

III.1. Introduction

Let $X_1, X_2, \ldots, X_m, \ldots$ and $Y_1, Y_2, \ldots, Y_n, \ldots$ be two independent sequences of random variables. The following assumptions are made on the underlying probability distributions:

\begin{align*}
\text{(i)} \quad & \text{The sequences, } \{X_i\} \text{ and } \{Y_j\}, \text{ are each strictly stationary.} \\
\text{(III.1.1)} \quad & \text{(ii)} \quad \text{Each sequence is mixing with a common mixing function } \phi \text{ (See Definition 1.2.2).} \\
& \text{(iii)} \quad \sum_{k=1}^{\infty} k^2 \phi_k^{1/2} < \infty.
\end{align*}

Define

\begin{align*}
\text{(III.1.2)} \quad & H_\lambda = \lambda F + (1-\lambda)G \\
& H = H_N \\
& K_N = F H_N^{-1} \\
& K_\lambda = F H_\lambda^{-1} \\
& K = F H^{-1}.
\end{align*}

Also, as before, we take

\begin{align*}
\text{(III.1.3)} \quad & L_N(t) = N^{1/2} (F H_N^{-1}(t) - F H^{-1}(t)) \text{ for } t \in [0,1].
\end{align*}
III.2. The one-sample processes

The results of this section rely heavily on two theorems. They are stated now as a matter of convenience.

Theorem III.2.1

Let $\xi_1, \ldots, \xi_n$ be random variables. Let $S_k = \xi_1 + \ldots + \xi_k$ ($S_0 = 0$) and put $M_m = \max_{0 \leq k \leq m} |S_k|$. If there exists $\gamma > 0$ and $\alpha > 1$ such that

$$E\{|S_j - S_i|^{\gamma}\} \leq \left( \sum_{i < \ell < j} u_{i,\ell} \right)^{\alpha}, \quad 0 \leq i \leq j \leq m,$$

where $u_1, \ldots, u_m$ are positive numbers, then,

$$P_H\{M_m \geq \lambda\} \leq (K/\lambda^{\gamma})(u_1 + \ldots + u_m)^{\alpha}$$

where $K$ depends on $\gamma$ and $\alpha$.

Proof:

See Theorem 12.2 of Billingsley (1968).

Theorem III.2.2

Let $\{\xi_i\}_{i=0}^\infty$ be stationary and $\phi$-mixing. Suppose that $|\xi_0| \leq 1$ with probability one, $E(\xi_0) = 0$ and $\sum k^2 \phi_k^{1/2} < \infty$.

Then,

$$E(S_n^4) \leq 288(n^2 E^2(\xi_0^2) + n E(\xi_0^2)) \left( \sum_{k=0}^{\infty} (k+1)^2 \phi_k^{1/2} \right)^2.$$

Proof:

This is Lemma 1 of section 22 of Billingsley (1968).
Lemma III.2.1

Let $q(t) = c(t(1-t))^{1/2-\delta}$, $0 \leq t \leq 1$, for a $\delta$ with $0 < \delta < 1/2$ and assume the conditions stated in III.1.1. Then, for $\theta$, $0 < \theta < 1/2$, such that $\int_0^\theta (q(t))^{-2} dt \leq 1$ there exists a constant $K_\phi$ such that

$$P \{ |U_m(t)| > q(t), 0 \leq t \leq \theta \} < \theta + K_\phi \int_0^\theta (q(t))^{-2} dt$$

for all $m > M_\theta = \max[(4/c\theta)^{2/\delta}, (4/\theta)^{2/\delta}]$. Moreover, $K_\phi$ does not depend on $\theta$.

Proof:

Let $q_t(x) = I_{(-\infty, F^{-1}(t)]}(x) - t$, and consider real points $0 < s_1 < s_2 < \ldots < s_m = \theta$, with $s_\ell = (\ell \theta/m), 1 \leq \ell \leq m$. Then, for every pair $(j,k)$ satisfying $1 < j < k \leq m$

$$(III.2.3) \quad E \left[ \frac{g_{s_k}(X_1)}{q(s_k)} - \frac{g_{s_j}(X_1)}{q(s_j)} \right]^2$$

$$= \frac{s_k(1-s_k)}{q(s_k)} + \frac{s_j(1-s_j)}{q(s_j)} - \frac{2s_j(1-s_k)}{q(s_j)q(s_k)}$$

$$= \frac{(1-s_k)}{q(s_k)} \left[ \frac{s_k}{q(s_k)} - \frac{s_j}{q(s_j)} \right] + \frac{s_j}{q(s_j)} - \frac{(1-s_j)}{q(s_j)}$$

To obtain a useful bound on the right hand side of Equation III.2.3 we show that $(t+a)/q(t)$ is nondecreasing in $t$ for $a < t < \theta$. To see this note, that the second factor on the
right hand side of \((t+a)/q(t) = c^{-2}(t^{1/2} + \alpha t^{-1/2})(t^\delta / (1-t)^{1/2})\) is clearly nondecreasing in \(t\) and the derivative of the first factor is positive, for \(\alpha < t < \theta\). Using \(\alpha = \theta/m\) and \(t = s_{l-2}\), we obtain that \(s_{l}/q(s_{l-1}) > s_{l-1}/q(s_{l-2})\) for \(2 < l \leq m\), so that the first term in Equation III.2.3, which equals

\[
\frac{1-s_k}{q(s_{k-1})} \sum_{j \leq k} \left( \frac{s_{l}}{q(s_{l-1})} - \frac{s_{l-1}}{q(s_{l-2})} \right) \leq \frac{1}{q(s_{k-1})} \sum_{j \leq k} \frac{(s_{l}-s_{l-1})}{q(s_{l-2})}
\]

(III.2.4)\] where we have used the monotonicity of \(q\) in \((0,1/2)\). The second term in Equation III.2.3 can be expressed as

\[
\frac{s_j}{q(s_j)} \sum_{j \leq k} \left( \frac{1}{q(s_{l-2})} - \frac{1}{q(s_{l-1})} \right)
\]

(III.2.5)\] + \[
\frac{s_j}{q(s_{j-1})} \sum_{j \leq k} \left( \frac{s_{l}}{q(s_{l-1})} - \frac{s_{l-1}}{q(s_{l-2})} \right) \right].
\]

Note that the monotonicity of \(t+\alpha/q(t)\), \(\alpha < t < 1/2\), as in Equation III.2.4, implies
\[
\sum_{j} \frac{s_j}{q(s_{j-1})} \leq \sum_{j} \frac{s_{\ell} - s_{\ell-1}}{q(s_{\ell-2})} \left( \frac{s_{\ell} - s_{\ell-1}}{q(s_{\ell-2})} \right)
\]

\[
\leq \frac{\theta}{m} \sum_{j} \frac{1}{q^2(s_{\ell-2})}
\]

Also, using the mean value theorem and the definition of \( q \), we have for each \( \ell, \ j < \ell < k \), a \( t_o \) with \( s_{\ell-2} < t_o < s_{\ell-1} \), such

\[
\frac{s_j}{q(s_{j-1})} \left( \frac{1}{q(s_{\ell-2})} - \frac{1}{q(s_{\ell-1})} \right)
\]

\[
\leq \frac{s_j(s_{\ell-1} - s_{\ell-2})}{q(s_{j-1})} \left( \frac{(1/2-\delta)(1-2t_o)}{c(t_o(1-t_o))^{3/2-\delta}} \right)
\]

\[
\leq \frac{\theta}{m} \cdot \frac{s_{\ell-1}}{q(s_{\ell-2})} \cdot \frac{(1/2-\delta)(1-2t_o)}{q(s_{\ell-2}) \cdot s_{\ell-2} \cdot (1-t_o)}
\]

\[
\leq \frac{\theta}{m} \frac{1}{q^2(s_{\ell-2})}
\]

from which it follows that
Using Equations III.2.7 and III.2.6 in III.2.5, it follows from Equations III.2.3 and III.2.4 that

\[
E \left( \left| \frac{g_{s_k}(X_1)}{q(s_{k-1})} - \frac{g_{s_j}(X_1)}{q(s_{j-1})} \right|^2 \right) \leq \frac{\theta}{m} \sum_{j \leq k} \frac{1}{q^2(s_{k-2})}.
\]

(III.2.8)

Further, for \(1 \leq k \leq m\)

\[
E \left( \left| \frac{g_{s_k}(X_1)}{q(s_{k-1})} \right|^2 \right) \leq \frac{s_k(1-s_k)}{q^2(s_{k-1})} \leq \frac{\theta}{m} \cdot \frac{k-1}{k-2} \cdot \frac{k-1}{q^2(s_{k-1})}
\]

\[
\leq \frac{2\theta}{m} \sum_{1 \leq k \leq k} \frac{1}{q^2(s_{k-1})}.
\]

(III.2.9)

Consider now for each pair \((j,k), 1 < j < k \leq m\), the sequence \(\{\eta_i\}_{i=1}^{m}\) with

\[
\eta_i = m^{-1/2+\delta/2} \left| \frac{g_{s_k}(X_1)}{q(s_{k-1})} - \frac{g_{s_j}(X_1)}{q(s_{j-1})} \right|
\]

(III.2.10)
and observe that for $m > (4/c\theta)^{2/\delta}$

$$|\eta_i| \leq m^{-1/2} + \frac{\delta}{2} \left( \frac{1}{q(s_{k-1}^j)} + \frac{1}{q(s_{j-1}^k)} \right)$$

$$\leq \frac{2m^{-1/2} + \delta/2}{q(s_1)}$$

(III.2.11)

$$\leq \frac{4m^{-\delta/2}}{c\theta^{1/2-\delta}}$$

$$\leq \frac{4}{c\theta} \cdot m^{-\delta/2}$$

$$\leq 1.$$

We can thus apply Theorem III.2.2 and use Equation III.2.8 to conclude the existence of a constant $K_\phi$ ($K_\phi$ is used as a generic constant below) such that

$$E \left| \sum_{i=1}^{m} \eta_i \right|^4 \leq K_\phi \left( m^2 E^2(\eta_1^2) + m E (\eta_1^2) \right)$$

(III.2.12)

$$\leq K_\phi \left( m^2 \delta \left( \sum_{j \leq k} \frac{1}{q^2(s_{k-1}^j)} \right) \right)^2$$

$$+ m^\delta \left( \sum_{j \leq k} \frac{1}{q^2(s_{k-1}^j)} \right).$$
From Equation III.2.12, we have for $1 < j < k \leq m$

$$E \left| \frac{U_m(s_k)}{q(s_{k-1})} - \frac{U_m(s_j)}{q(s_{j-1})} \right|^4 = m^{-2\delta} E \left| \sum_{i=1}^{m} \eta_i \right|^4$$ (III.2.13)

$$\leq K_\phi \left\{ \left( \frac{\theta}{m} \sum_{j \leq l < k} \frac{1}{q^2(s_{l-1})} \right)^2 + m^{-\delta} \left( \frac{\theta}{m} \sum_{j \leq l < k} \frac{1}{q^2(s_{l-1})} \right) \right\}.$$ 

Since $m > (4/\theta)^{2/\delta}$ we have $m^{-\delta/2} < \theta/4 < \theta^6/2/4$. Observe that $\int_0^\theta (q(t))^{-2} \, dt < 1$ so that from Equation III.2.13 we have

$$E \left| \frac{U_m(s_k)}{q(s_{k-1})} - \frac{U_m(s_j)}{q(s_{j-1})} \right|^4 \leq K_\phi \left\{ \left( \frac{\theta}{m} \sum_{j \leq l < k} \frac{1}{q^2(s_{l-1})} \right)^2 + \left( \frac{\theta}{m} \sum_{j \leq l < k} \frac{1}{q^2(s_{l-1})} \right)^{1+\delta/2} \right\}.$$ (III.2.14)

$$\leq K_\phi \left\{ \left( \frac{\theta}{m} \sum_{j \leq l < k} \frac{1}{q^2(s_{l-1})} \right)^{1+\delta/2} \right\}.$$
Similarly, using Equation III.2.9, Theorem III.2.2 and the same argument, it follows that

\[
(III.2.15) \quad E \left| \frac{U_m(s_k)}{q(s_{k-1})} \right|^4 \leq K \phi \left( \frac{\theta}{m} \sum_{1 \leq \ell < k} \frac{1}{q^2(s_{\ell-1})} \right)^{1+\delta/2}.
\]

Now let \( \xi_1 = \frac{U_m(s_2)}{q(s_1)} \),

\[ \xi_i = \frac{U_m(s_i)}{q(s_i)} - \frac{U_m(s_{i-1})}{q(s_{i-1})}, \]

\( 1 < i < m, \) (\( \xi_0 = 0 \)) and use Equations III.2.14 and III.2.15 and Theorem III.2.1 to obtain

\[
(III.2.16) \quad p\{ \max_{1 \leq i \leq m} \left| \frac{U_m(s_{i+1})}{q(s_i)} \right| \geq \frac{1}{2} \} \leq K \phi \left( \frac{\theta}{m} \sum_{\ell=1}^{m-1} \frac{1}{q^2(s_{\ell})} \right)^{1+\delta/2}.
\]

From Billingsley (1968, (22.17) page 119) we have for \( s < t \)

\[ |U_m(s)| \leq |U_m(t)| + \sqrt{m} (t-s), \]

so that for \( s_i < t \leq s_{i+1}, \) \( i = 1,2,\ldots,m \) we obtain, using the monotonicity of \( q(t) \) in \((0,\theta)\), that
\[(III.2.17) \quad \left| \frac{U_m(t)}{q(t)} \right| \leq \left| \frac{U_m(s_i+1)}{q(s_i)} \right| + \frac{\sqrt{m}}{q(\theta/m)} \]

\[\leq \left| \frac{U_m(s_i+1)}{q(s_i)} \right| + \frac{\sqrt{\theta/m}}{q(\theta/m)}.\]

Since \(m > (4/c\theta)^{2/\delta}\), we have

\[(III.2.18) \quad \frac{\sqrt{\theta/m}}{q(\theta/m)} \leq \frac{2}{c} (\theta/m)^{\delta} \leq \frac{1}{2}.\]

From Equations \(III.2.16\), \(III.2.17\) and \(III.2.18\), it follows that

\[(III.2.19) \quad P \left( \sup_{\frac{\theta}{m} \leq t \leq \theta} \left| \frac{U_m(t)}{q(t)} \right| \geq 1 \right) \]

\[\leq P \left( \max_{1 \leq i \leq m} \left| \frac{U_m(s_i+1)}{q(s_i)} \right| \geq \frac{1}{2} \right) \]

\[\leq K_\phi \left( \frac{\theta}{m} \sum_{s=1}^{m-1} (1/q(s))^2 \right)^{1+\delta/2} \]

\[\leq K_\phi \int_0^\theta \frac{dt}{q^2(t)} \]

where we have used the monotonicity of \(q\) and the assumption that \(\int_0^\theta (q(t))^{-2} dt < 1\).
If $F_m^{-1}(\theta/m) = 0$ and $m > (4/c8)^{2/8}$, then, for $0 < t < \theta/m$ we have

$$\left| \frac{U_m(t)}{q(t)} \right| = \frac{m^{1/2} t}{q(t)} \leq \frac{2}{c} m^{1/2} t^{1/2+\delta}$$

$$\leq \frac{2}{c} \theta^{1/2+\delta} \cdot \frac{\theta}{m}$$

$$\leq 1.$$

Thus,

$$P\{ |U_m(t)| \leq q(t), 0 < t < \theta/m \}$$

$$\geq P\{ F_m^{-1}(\theta/m) = 0 \}$$

(III.2.20)

$$= 1 - P\left\{ \bigcup_{i=1}^m (F(X_i) < \theta/m) \right\}$$

$$\geq 1 - m \cdot \frac{\theta}{m}$$

$$= 1 - \theta.$$

From Equations III.2.19 and III.2.20, we have for $m > M_\theta$

$$P\{ |U_m(t)| > q(t), 0 < t < \theta \}$$

$$< \theta + K_\phi \int_0^\theta \frac{dt}{q^2(t)}.$$ 

The proof is complete.
Theorem III.2.1

Suppose $X_1, X_2, \ldots, X_m, \ldots$ stationary and $\phi$-mixing with
\[ \sum k^2 \phi_k^{1/2} < \infty \] and suppose each $X_i$ has a continuous distribution function $F$. Then $U_m/q \rightarrow U_o/q$ relative to $(D,d)$ where $U_o$ is the Gaussian random function specified by

(III.2.21) \[ E\{U_o(t)\} = 0 \]

and

\[ E(U_o(s) U_o(t)) = E(g_s(X_1) g_t(X_1)) \]

(III.2.22) \[ + \sum_{k=2}^{\infty} E(g_s(X_1) g_t(X_k)) + \sum_{k=2}^{\infty} E(g_s(X_k) g_t(X_1)) \]

with $g_s(a) = I_{(-\infty, F^{-1}(s)]}(a) - s$. Further, $U_o$ is continuous with probability one.

Proof:

For any $0 < a < b < 1$,

III.2.23 \[ \omega_\delta(U_m/q) \leq 2 \sup_{0 < t < a} \frac{|U_m(t)|}{q(t)} \]

\[ + \frac{\omega_\delta(U_m)}{\min(q(a-2\delta), q(b+2\delta))} \]

\[ + \frac{\omega_\delta(q) \sup_{0 \leq t \leq 1} |U_m(t)|}{[\min(q(a-2\delta), q(b+2\delta))]^2} \]
For each such \( a \) and \( b \) the second term on the right converges to zero in probability by Equation 22.13 of Billingsley (1968).

Observe that \( f_0(q) \to 0 \) as \( \delta \to 0 \) since \( q \) is continuous.

That \( \sup_{0 \leq t \leq 1} |U_m(t)| = 0 \), follows from Equation 22.13 of Billingsley (1968) and the inequality

\[
|U_m(t)| \leq |U_m(0)| + \sum_{i=1}^{k} U_m(it/k) - U_m((i-1)t/k).
\]

Hence, the third term on the right converges to zero in probability.

Fix \( \eta > 0 \) and choose \( a < \frac{\eta}{2} \) so that \( \int_0^a (q(t))^{-2} \, dt < \min(K_\phi^{-1} \eta^{3/2}, 1) \). Then by Lemma III.2.1

\[
\text{III.2.24} \quad P \left( \sup_{0 \leq t \leq a} \left| \frac{U_m(t)}{q(t)} \right| < \frac{\eta}{2} + K_\phi \eta^{-2} \int_0^a (q(t))^{-2} \, dt \right) < \eta.
\]

In a similar manner, using the reverse process, we can choose \( b \) such that

\[
\text{III.2.25} \quad P \left( \sup_{b \leq t < 1} \left| \frac{U_m(t)}{q(t)} \right| > \eta \right) < \eta.
\]

By Equations III.2.24 and III.2.25, the first term on the right hand side of Equation III.2.23 does not exceed \( 2\eta \) on a set whose probability is greater than \( 1-2\eta \). Thus,
Hence, by Theorem I.3.2, \( \{U_m/q\} \) is tight relative to \((D,d)\).

Theorem 22.1 of Billingsley states that \( U_m \) converges weakly (relative to \((D,d)\)) to a Gaussian process with parameters as specified in Equations III.2.15 and III.2.16, that is, \( U_m \overset{D}{\rightarrow} U_0 \) relative to \((D,d)\). Hence, by Slutzky's Theorem the finite dimensional distributions of \( U_m/q \) converge to those of \( U_0/q \).

Since \( \{U_m/q\} \) is tight relative to \((D,d)\) every subsequence contains a further subsequence which converges weakly to some limit. This limit must have the same finite dimensional distributions as \( U_0/q \). Hence, every subsequence of \( \{U_m/q\} \) contains a further subsequence which converges weakly to \( U_0/q \). Therefore, \( U_m/q \overset{L}{\rightarrow} U_0/q \) relative to \((D,d)\). Finally, \( U_0/q \) is continuous with probability one by Equation III.2.26 and Theorem 15.5 of Billingsley (1968). The proof of Theorem III.2.1 is complete.

Since \( U_m/q \overset{L}{\rightarrow} U_0/q \) relative to \((D,d)\) and \((D,d)\) is complete and separable, it is possible (see item 3.1.1 of Skorokhod (1956)) to construct processes \( \{\tilde{U}_m(t)/q(t) : 0 \leq t \leq 1\}, m \geq 0 \), with sample functions in \( D \) and having the same finite dimensional distributions as \( \{U_m(t)/q(t) : 0 \leq t \leq 1\}, m \geq 0 \), but which also satisfy

\[
(III.2.27) \quad d(\tilde{U}_m/q, \tilde{U}_0/q)_{a.s.} \to 0.
\]
We make an independent construction for the \( V_n \) processes. All processes are defined on a single probability space \((\tilde{\Omega}, \tilde{\beta}, \tilde{\mathbb{P}})\). We now drop the symbol \( \sim \) from the notation.

**Lemma III.2.2**

\[ \rho(U_m/q, U_\infty/q) \xrightarrow{a.s.} 0 \text{ as } m \to \infty. \]

**Proof:**

From Theorem III.2.1, \( U_\infty/q \) is continuous with probability one. The proof is completed by observing that \( d \)-convergence to a continuous function is equivalent to \( p \)-convergence (see Billingsley (1968, page 112)).

**Lemma III.2.3**

Take \( 0 < \delta < 1/2 \) and let \( q(t) = c(t(1-t))^{1/2-\delta} \) for \( t \in [0,1] \). For \( \theta < 1 \),

\[
P\{ \sup_{0 \leq t \leq \theta} |U_\infty(t)/q(t)| \leq 1 \} \\
\geq 1 - \theta - K_\phi \int_0^\theta \{q(t)\}^{-2} \, dt
\]

where \( K_\phi \) is the constant used in Lemma III.2.1.

**Proof:**

For all \( m \geq 1 \)

\[
\frac{U_\infty(t)}{q(t)} = \frac{U_m(t)}{q(t)} + \frac{U_\infty(t)}{q(t)} \frac{U_\infty(t) - U_m(t)}{q(t)}
\]
\[ \frac{U_m(t)}{q(t)} + \rho(U_m/q, U_0/q). \]

Hence, for \( m \) large, we have by Lemma III.2.1

\[
P\{ \sup_{0 < t \leq \theta} \left| \frac{U_0(t)}{q(t)} \right| \leq 1 \} \geq 1 - P\{ \sup_{0 < t \leq \theta} \left| \frac{U_m(t)}{q(t)} \right| > 1/2 \} - P\{\rho(U_m/q; U_0/q) > 1/2\}
\]

\[
\geq 1 - \int_0^\theta \{q(t)\}^{-2} \, dt - P\{\rho(U_m/q, U_0/q) > 1/2\}.
\]

Since \( m \) was arbitrary we have by Lemma III.2.2

\[
P\{ \sup_{0 < t \leq \theta} \left| \frac{U_0(t)}{q(t)} \right| \leq 1 \} \geq 1 - \int_0^\theta \{q(t)\}^{-2} \, dt.
\]

The proof of Lemma III.2.3 is complete.

**Lemma III.2.4**

As \( N \to \infty \), \( \rho(HH^{-1}_N, HH^{-1}) \) a.s. \( 0 \) uniformly in all continuous \( F \) and \( G \) and \( \lambda_N \in [0,1] \).

**Proof:**

The proof is similar to that of Lemma 2.3 of Pyke and Shorack (1968). Since \( H_N - H = \lambda_N(F_m - F) + (1-\lambda_N)(G_n - G) \), a generalization of the Glivenko-Cantelli Theorem by Tucker
(1959) implies that \( \rho(H_N', H) \to 0 \) uniformly in \( F \) and \( G \) and \( \lambda_N \in [0,1] \). Then

\[
\rho(H^{-1}_N, H^{-1}) \leq \rho(H^{-1}_N, H_N H^{-1}_N) + \rho(H_N H^{-1}_N, H^{-1})
\]

\[
\leq \rho(H^{-1}_N, H_N H^{-1}_N) + 1/N.
\]

Therefore, \( \rho(H^{-1}_N, H^{-1}) \to 0 \). The proof of Lemma III.2.4 is complete.

**Lemma III.2.5**

For each positive \( \varepsilon \) and \( t \), there exists \( b > 0 \) such that

\[
P(K_N(t) \leq b \lambda^{-1}_* t^{1-\tau} \text{ for } t > 1/N) \geq 1 - \varepsilon.
\]

**Proof:**

Since \( \rho(K_N', K) \to 0 \) and \( K(t) \leq \lambda^{-1}_* t \), the proof reduces to a study of the intervals \([0,\theta]\) and \([1-\theta,1]\) for sufficiently small \( \theta \). Further, it is enough to consider the interval \([0,\theta]\) since the interval \([1-\theta,1]\) may be studied using the reverse process.

For fixed \( \varepsilon, \tau > 0 \), take \( \delta = \tau/2(1-\tau) \) and \( q(t) = (t(1-t))^{1/2-\delta} \). Then choose \( \theta > 0 \) and \( M_{\theta} > 1 \) such that
(III.2.28) \( P(A_m) > 1 - \varepsilon \) for \( m > M_\theta \),

where

\[
A_m = \left\{ \sup_{0 < t < \theta} \left| \frac{U_m(t)}{q(t)} \right| < 1 \right\}.
\]

This choice of \( \theta \) and \( M_\theta \) is made using Lemma III.2.1.

Suppose \( t \geq 1/N \). If \( 1/m \leq F_m H_N^{-1}(t) \), then using

\[
F_m H_N^{-1}(t) \leq \lambda_*^{-1} t + 1/N \leq 2\lambda_*^{-1} t \quad \text{and the definition of } U_m,
\]

we have an \( A_m \)

\[
|K_N(t)| = |F_H^{-1}(t)| = |F_m H_N^{-1}(t) - m^{-1/2} U_m(F_H^{-1}(t))|
\]

(III.2.29) \[
\leq F_m H_N^{-1}(t) + (F_m H_N^{-1}(t))^{1/2} |q(F_H^{-1}(t))|.
\]

\[
\leq 2\lambda_*^{-1} t + (2\lambda_*^{-1} t)^{1/2} (F_H^{-1}(t))^{1/2-\delta}.
\]

For simplicity, take \( x = F_H^{-1}(t) \) and rewrite Equation III.2.29 as

(III.2.30) \[
x \leq 2\lambda_*^{-1} t + (2\lambda_*^{-1} t)^{1/2} x^{1/2-\delta}.
\]

Completing the square in Equation III.2.30, we have

\[
x + \left(\frac{x^{1/2-\delta}}{2}\right)^2 \leq \left( (2\lambda_*^{-1} t)^{1/2} + \frac{x^{1/2-\delta}}{2} \right)^2.
\]
Hence,
\[
\left( x + \left( \frac{x}{2} \right)^{1/2} \right)^{1/2} \leq (2\lambda_*^{-1} t)^{1/2} + \frac{x^{1/2-\delta}}{2}.
\]

It follows that

(III.2.31) \quad x^{1/2-\delta} \left( x^{2\delta} + \frac{1}{4} \right)^{1/2} - \frac{x^{1/2-\delta}}{2} \leq (2\lambda_*^{-1} t)^{1/2}.

Since \( x \leq 1 \), \( x^{2\delta} \geq (x^{2\delta})^2 \) and

(III.2.32) \quad \left( \frac{x^{2\delta}}{2} \right)^2 + \frac{x^{2\delta}}{2} + \frac{1}{4} \leq x^{2\delta} + \frac{1}{4}.

We combine Equations III.2.31 and III.2.32 to obtain

\[
x^{1/2-\delta} \left( \left( \frac{x^{2\delta}}{2} + \frac{1}{2} \right) - \frac{1}{2} \right) \leq (2\lambda_*^{-1} t)^{1/2},
\]

so that

(III.2.33) \quad x \leq (2\lambda_*^{-1} t)^{1/(1+2\delta)}.

Since, \( \delta = \tau/2(1-\tau) \), \( 1-\tau = 1/(1+2\delta) \) and by Equation III.2.33, we have

(III.2.34) \quad x \leq (2\lambda_*^{-1} t)^{1-\tau}.
Take \( b = 2^{1-\tau} \), observe that \( \lambda_*^{-1+\tau} < \lambda_*^{-1} \) and substitute \( \mathcal{F}_N^{-1}(t) \) for \( x \) in Equation III.2.34 to conclude that on \( A_m \) if \( \mathcal{F}_m^{-1}(t) \geq 1/m \).

\[
\mathcal{F}_N^{-1}(t) \leq b \lambda_*^{-1} t^{1-\tau}
\]

(III.2.35) if \( \mathcal{F}_m^{-1}(t) = 0 \) and \( t \geq \frac{1}{N} \), we see easily that \( K_N(t) \leq m^{-1/2} (K_N(t))^{1/2-\delta} \) so that on \( A_m \)

\[
\mathcal{F}_N^{-1}(t) \leq b \lambda_*^{-1} t^{1-\tau}
\]

(III.2.36) if \( \mathcal{F}_m^{-1}(t) = 0 \).

Combining Equations III.2.28, III.2.35 and III.2.36 we conclude that

\[
P \left( K_N(t) \leq b \lambda_*^{-1} t^{1-\tau} \text{ for } 1/N \leq t \leq \theta \right) \geq P(A_m) \geq 1-\varepsilon.
\]

The proof of Lemma III.2.5 is complete.
Definition III.2.1

Let \( U_m^*(K_N) \) equal \( U_m(K_N) \) for \( 1/N \leq t \leq 1 - 1/N \) and equal 0 otherwise.

Theorem III.2.2

Take \( 0 < \delta < 1/2 \) and let \( q(t) = K(t(1-t))^{1/2-\delta} \) for \( t \in [0,1] \), then

\[
\lim_{n \to \infty} \rho_q(U_m^*(K_N), U_0(K)) = 0
\]

uniformly in all continuous \( F \) and \( G \).

Proof:

The proof is similar to Theorem 2.2 of Pyke and Shorack (1968). As before we need only consider the interval for \([0,\delta]\) for small \( \delta \) since \( q(K) \) and with high probability \( q(K_N) \) are bounded away from zero on \([\delta,1-\delta]\). Also, the reverse process can be used to consider the interval \([1-\theta,1]\).

For given \( \epsilon, \eta > 0 \), take \( 0 < \tau < 2\delta \) and choose \( b > 0 \) by Lemma III.2.5 so that

\[
(III.2.37) \quad P\{K_N(t) \leq b\lambda_*^{-1} t^{1-\tau} \text{ for } t > 1/N\} > 1 - \epsilon/3.
\]

We assume without loss of generality that \( b\lambda_*^{-1} > 1 \).

Take \( q_1(t) = (1/2)^{1/2-\tau/2} c(t(1-t)) \) and observe that for \( t < 1/2 \)
\[ q_1(t^{1-\tau}) \leq \frac{1}{2^{1/2-\tau/2}} c t^{1/2-\tau/2} \]

\[ \leq \frac{1}{2^{1/2-\tau/2}} c(t(1-t))^{1/2-\tau/2} \]

(III.2.38) \[ \leq c(t(1-t))^{1/2-\delta} \]

\[ \leq q(t). \]

Since \( b\lambda_*^{-1} > 1 \), we have \( \lambda_* t^{1-\tau}/b < t^{1-\tau} \) and by Equation III.2.38 and the monotonicity we conclude that

(III.2.39) \[ q_1(\lambda_* t^{1-\tau} / b) \leq q_1(t) \leq q(t). \]

We can use Lemma III.2.1 to choose \( \theta > 0 \) and \( M_o > 1 \) to satisfy

(III.2.40) \[ P\{ |U_m(t)| \leq \eta q_1(\lambda_* t^{1-\tau} / b), 0 < t \leq \theta \} > 1 - \frac{\varepsilon}{3} \]

for \( m > M_o \).

Now \( \alpha > 0 \) so that

(III.2.41) \[ P\{ K_N(t) < \theta, 0 < t \leq \alpha \} > 1 - \frac{\varepsilon}{3}. \]

This is possible by Lemma III.2.3 since \( K_N(t) \leq \lambda_*^{-1} HH_N^{-1}(\alpha) \).

With these choices of \( b, \alpha, \) and \( \theta \) we have, by Equations III.2.37, III.2.39, III.2.40 and III.2.41, that
\[ P\{\mid U_m^*(K_N(t)) \mid \leq \eta q(t), \ 0 < t \leq \alpha \} \]

> \[ P\{\mid U_m(t) \mid \leq \eta q_1(\lambda_* t^{1-\tau}/b), \ 0 < t < \theta \} \]

\[ \lambda \{ K_N(t) < \theta, \ 0 < t < \alpha \} \]

\[ \lambda \{ K_N(t) \leq b\lambda_*^{-1} t^{1-\tau}, \ t \geq 1/N, \} \]

\[ < 1 - \varepsilon \]

for \( m > M_0 \).

Since \( K(t) \leq \lambda_*^{-1} t \) we can use Lemma III.2.3 to show that

\[ P\{\mid U_o(K) \mid \leq \eta q(t), \ 0 \leq t < \alpha \} > 1 - \varepsilon. \]

Since \( \varepsilon \) and \( \eta \) were arbitrary the proof of Theorem III.2.2 is complete.

III.3. The two-sample process

The results that follow are stated without proof. The results are similar to those of Sections 3 and 4 of Pyke and Shorack (1968) and Pyke and Shorack (1969).

Lemma III.3.1: With probability 1

\[ L_N(t) = (1-\lambda_N)\{\lambda_N^{-1/2} B_N(t) U_m(FH_N^{-1}(t)) \]

\[ - (1-\lambda_N)^{-1/2} A_N(t) V_N(GH_N^{-1}(t)) \} + \delta_N(t) \]
for all $t \in [0,1]$ where

$$
\delta_N(t) = A_N(t) N^{1/2} [H_N H_N^{-1}(t) - t],
$$

$$
A_N(t) = [K(\mu_t) - K(t)]/(\mu_t - t), \mu_t = HH_N^{-1}(t),
$$

and $B_N$ is defined by

$$
\lambda_N A_N(t) + (1 - \lambda_N) B_N(t) = 1.
$$

($L_N(t)$ is defined by left continuity at any otherwise undefined points.)

**Assumption III.3.1**

The functions $K_\lambda$ have derivatives $a_\lambda$ for all $t \in (0,1)$, and for some $\lambda'$, $a_\lambda'$ is continuous on $(0,1)$ and has one sided limits at 0 and 1.

**Definition III.3.1**

Let $a_\circ$, $b_\circ$ denote the derivatives of $FH_\circ^{-1}$, $GH_\circ^{-1}$ where $H_\circ = \lambda_\circ F + (1 - \lambda_\circ) G$. Now, set

$$
L_\circ(t) = (1 - \lambda_\circ)^{-1/2} \{ \lambda_\circ^{-1/2} b_\circ(t) U_\circ(FH_\circ^{-1}(t))

- (1 - \lambda_\circ)^{-1/2} a_\circ(t) V_\circ(GH_\circ^{-1}(t)) \}. 
$$
Theorem III.3.1

(a) Suppose Assumption III.3.1 holds and $\lambda_N \to \lambda_0$. Then $\rho(L^N_1, L_0) \to 0$ where $L^N_1$ equals $L_N$ on $[\frac{1}{N}, 1]$ and 0 elsewhere. Thus $L^N_1 \rightarrow L_0$ relative to $(D, \rho)$. The same statement holds for $d$.

(b) If in addition the measures $\{\nu_N : N \geq 1\}$ and $\nu$ satisfy

(i) $\int_0^1 L_N d(\nu_N - \nu) \to 0$, and,

(ii) $\int_0^1 d|\nu| < \infty$,

then $T^*_N \to \int_0^1 L_0 d\nu$. If $\sigma^2_0 < \infty$, where

$$\sigma^2_0 = 2(1-\lambda_0)^2 \left( \lambda_0^{-1} \int_0^1 \int_0^1 \nu(b_0(u)b_0(v)F_H^{-1}(u)) \right)$$

$$(1 - F_H^{-1}(v))d\nu(u)d\nu(v)$$

(III.3.1) $\quad + \lambda_0^{-1} \sum_{k=1}^\infty \int_0^1 \int_0^1 b_0(u)b_0(v)(F_K(H_0^{-1}(u), H_0^{-1}(v)))$

$\quad \quad - F_H^{-1}(u)F_H^{-1}(v))d\nu(u)d\nu(v)$

$\quad + (1-\lambda_0)^{-1} \int_0^1 \int_0^1 \alpha_0(u)\alpha_0(v)G_H^{-1}(u)(1-G_H^{-1}(v))$

$\quad \quad d\nu(u)d\nu(v)$
(1-\lambda_0)^{-1} \sum_{k=1}^{\infty} \int_0^1 \int_0^1 a_o(u)a_o(v)(G_k(GH_o^{-1}(u),
GH_o^{-1}(v))-FH_o^{-1}(u)FH_o^{-1}(v))d\nu(u)d\nu(v),

then \int_0^1 L_0 d\nu is a N(0,\sigma_o^2) random variable.

**Theorem III.3.3**

(a) Suppose \( K_0 \) is differentiable a.e. \( |v|, \lambda_N = \lambda_0 + O(N^{-1/2}) \), and \( \int_0^1 q d |v| < \infty \) for some q. Then \( ||L^1_N - L_0||_\nu \to 0 \)

so that \( L^1_N \sim L_0 \) relative to \( (D^-, ||\cdot||_\nu) \).

(b) If, in addition, the measures \( \{\nu_N : N \geq 1\} \) and \( \nu \)

satisfy

(i) \( \int_0^1 L_N d(\nu_N - \nu) \to 0, \)

then \( T_N \to \int_0^1 L_0 d\nu \) which is a \( N(0,\sigma_o^2) \) random variable if

\( \sigma_o^2 < \infty \), with \( \sigma_o^2 \) given by Equation III.3.1.

The conditions used in the concluding theorem are the same as those employed in Theorem II.4.3.

**Theorem III.3.3**

(i) Under (C1), (C2) and (C3)

\( T_N \to \int_0^1 L_0 d\nu \) as \( N \to \infty \).
(ii) Under (Cl), (C2), (C3) and (C4)

\[ T_N \to \int_0^1 L_0 \, d\nu \quad \text{as } N \to \infty \]

and the convergence is uniform in the set of all \( F, G, \lambda_N, \]
\( C_{N_i} \)'s and \( J \) such that the conditions hold for fixed \( q, M, \delta_N 's, \]
\( M_N 's, \delta_\varepsilon 's, \) and \( N_\varepsilon 's. \)
IV. THE SCALE PROBLEM

IV.1. Introduction

The scale problem will be approached by considering tests based on the observation centered at the respective estimated location parameters. Conditions will be given under which the asymptotic distribution of this modified rank order test will not depend on the estimated location parameters.

Sukhatme (1958) considered the special case of bounded U-statistics. He specified sufficient conditions under which the modified U-statistic and the original U-statistic have the same asymptotic distribution.

Raghavachari (1965) considered the normal score test. He assumed that the underlying distribution functions, $F$ and $G$, possess density functions $f$ and $g$, respectively which are symmetric about their location parameters. He also assumed that

$$f(x)/\phi^{-1}\{F(x)\}, \ g(x)/\phi^{-1}\{G(x)\}$$

are bounded where $\phi(x) = (2\pi)^{-1/2} \exp\{-x^2/2\}$.

Sen (1967) applied the Chernoff-Savage technique using a general $J$ function. He assumed that:

1. $F$ and $G$ have continuous densities ($f$ and $g$ respectively) which are symmetric about their location parameters
2. $J(u)$ is symmetric about $u = 1/2$
(3) $J'(F(x-t))f(x-t) < KT(x)$ and $J'(G(x-t))g(x-t) < KT(x)$ for all $|t| < c$ (c a constant) where $T(x)$ is square integrable with respect to $F$ and $G$.

In this chapter we generalize the work of Pyke and Shorack (1968) and improve upon the results of Raghavachari (1965) and Sen (1967). Assume, without loss of generality, that $\xi = \eta = 0$. Let $\hat{\xi}$ and $\hat{\eta}$ be suitable estimates of $\xi$ and $\eta$ respectively. Further, we assume that $|\hat{\xi}| = |\hat{\eta}| = O_p(N^{-1/2})$. Define

$$\begin{align*}
F^*_m(x) &= \frac{\text{(number of } X_i \text{ such that } x_i \leq x)}{m}, \\
G^*_n(x) &= \frac{\text{(number of } Y_j \text{ such that } y_j \leq x)}{n}.
\end{align*}$$

Therefore,

$$\begin{align*}
F_m(x) &= F^*_m(x + \hat{\xi}) \\
G_n(x) &= G^*_n(x + \hat{\eta}) \\
H_N(x) &= \lambda_N F^*_m(x + \hat{\xi}) + (1-\lambda_N) G^*_n(x + \hat{\eta}).
\end{align*}$$

Let

$$\begin{align*}
H_N(x) &= \lambda_N F(x + \hat{\xi}) + (1-\lambda_N) G(x + \hat{\eta}), \\
H_0(x) &= \lambda_N F(x) + (1-\lambda_N) G(x), \\
K_N(t) &= F(H_N^{-1}(t) - \hat{\xi}) \\
K_\lambda(t) &= F(H_\lambda^{-1}_N(t) - \hat{\xi}) \\
K(t) &= F H_0^{-1}(t).
\end{align*}$$
Also, take

\begin{align}
U_m^*(t) &= m^{1/2} (F_m^* F_{1-m}^{-1}(t) - t) \\
V_n^*(t) &= n^{1/2} (G_n^* G_{1-n}^{-1}(t) - t).
\end{align}

(IV.1.4)

$U_m^*$ and $V_n^*$ are one-sample empirical processes under independence. Hence, the results of Section 2 of Pyke and Shorack (1968) can be used. With this in mind, let $U_0^*$ and $V_0^*$ be independent tied down Wiener processes such that $U_m^* \overset{L}{\to} U_0^*$ and $V_n^* \overset{L}{\to} V_0^*$ relative to $(D,d)$. The reader will note that the results of Pyke and Shorack (1968) may apply only to specially constructed sequences (see item 2.2 of Pyke and Shorack). The $\overset{\sim}{\to}$ notation is again not used.

We define the two sample empirical process $\{L_N^*(t) : 0 \leq t \leq 1\}$ by

$$L_N^*(t) = N^{1/2} \{F_N^* H_N^{-1}(t) - F H_{1-N}^{-1}(t)\}.$$ 

Write

$$T_N^* = \int_0^1 L_N^* \, d\nu_N$$

where $\nu_N$, $N \geq 0$, is defined in Section 1.2. We will find conditions under which the asymptotic distribution of $T_N^*$ depends on $F$, $G$, and $\nu$.

It is natural to propose the modified 2-sample empirical process $\{L_N^*(t) : 0 \leq t \leq 1\}$ defined by
\[ L_N^*(t) = N^{1/2} \{ F_{m N}^{-1}(t) - F(H_{\lambda_N}^{-1}(t) + \xi) \}. \]

The asymptotic behavior of \( L_N^* \) will be studied in order to determine the asymptotic distribution of \( T_N^* \).

IV.2. The one-sample processes

Lemma IV.2.1

\[ \rho(H_{\lambda_N}, H_{\lambda_N}^{-1}) \rightarrow 0 \text{ uniformly in } F \text{ and } G \text{ and } \lambda_N \in \Delta. \]

Proof:

We have by IV.1.2

\[
\rho(H_{\lambda_N}, H_{\lambda_N}^{-1}) = \sup_{x} |\lambda_N (F_{m}^{*}(x+\xi) - F(x+\xi)) \]

\[ + (1-\lambda_N) (G_{n}^{*}(x+\hat{\eta}) - G(x+\hat{\eta}))| \]

\[ \leq \lambda_N \rho(F_{m}^{*}, F) + (1-\lambda_N) \rho(G_{n}^{*}, G). \]

Hence, by the Glivenko-Cantelli Theorem,

\[ \rho(H_{\lambda_N}, H_{\lambda_N}^{-1}) \rightarrow 0. \]

Further,

\[ \rho(H_{\lambda_N}, H_{\lambda_N}^{-1}) \leq \rho(H_{\lambda_N}, H_{\lambda_N}^{-1}) \]

\[ + \rho(H_{\lambda_N}, H_{\lambda_N}^{-1}). \]

The first term on the right hand side of Equation IV.2.3 converges a.s. to zero as \( N \rightarrow \infty \) by Equation IV.2.2. The second term on the right hand side is less than or equal to
Hence, $p(H_N^{-1}, H_N^{-1}) \rightarrow 0$. The proof of Lemma IV.2.1 is complete.

**Lemma IV.2.2**

$$p(U_N^*(K_N), U_N(K_N)) \rightarrow 0 \text{ uniformly in } F \text{ and } G \text{ and } \lambda_N \in \Delta.$$  

**Proof:**

Using the definition of $p$ and the triangle inequality, we have

$$p(U_N^*(K_N), U_N(K_N))$$

(IV.2.4)

$$\leq p(U_N^*, U_N) + p(U_N(K_N), U_N(K_N)).$$

Since

$$H_N^{-1}(t) - H_N^{-1}(t) = \lambda_N(F(H_N^{-1}(t) + \xi) - F(H_N^{-1}(t) + \xi))$$

$$+ (1-\lambda_N)(G(H_N^{-1}(t) + \hat{\eta}) - G(H_N^{-1}(t) + \hat{\eta}))$$

we have, from Lemma IV.2.1 and the fact that $K_N - K_N$ and $G(H_N^{-1} + \hat{\eta}) - G(H_N^{-1} + \hat{\eta})$ are of the same sign, that

(IV.2.5)

$$p(K_N, K_N) \rightarrow 0.$$

Thus, since $U_N$ is a.s. continuous on $[0,1]$, it follows from Equation IV.2.5 that

(IV.2.6)

$$p(U_N(K_N), U_N(K_N)) \rightarrow 0.$$
Lemma 2.1 of Pyke and Shorack states that \( \rho(U_m', U_o') \rightarrow 0 \).

Hence, by Equations IV.2.4 and IV.2.6 we have

\[
\rho(U_m'(K_N'), U_o'(K_N')) \rightarrow 0.
\]

The proof of Lemma IV.2.2 is complete.

**Lemma IV.2.3**

Given \( \epsilon > 0 \), there exists \( b > 0 \) such that

\[
P\{K_N(t) \leq 2b \lambda_*^{-1} t \text{ for all } t \geq 1/N \} > 1 - \epsilon.
\]

**Proof:**

For \( \epsilon > 0 \), we have, using Lemma 8 of Govindarajulu et al. (1967), that there exists \( b > 0 \) such that

\[
P(A_m) > 1 - \epsilon
\]

where

\[
A_m = \{ F(x) \leq b F_m^*(x) \text{ for all } x \text{ where } F_m^*(x) > 0 \}.
\]

Thus, if \( t \geq 1/N \) and \( 0 < F_m(H_N^{-1}(t)) \), then on \( A_m \)

\[
K_N(t) = F(H_N^{-1}(t) + \xi) \leq b F_m^*(H_N^{-1}(t) + \xi)
\]

\[
\leq b \lambda_*^{-1} H_N H_N^{-1}(t) \leq 2b \lambda_*^{-1} t.
\]

Also, if \( t \geq 1/N \) and \( F_m H_N^{-1}(t) = 0 \) then we have

\[
H_N^{-1}(t) < F_m^{-1}(1/m) = F_m^*{-1}(1/m) - \xi.
\]
Hence, if \( t > 1/N \) and \( F_{mN}^{-1}(t) = 0 \), then on \( A_m \)

\[
K_N(t) = F(H_N^{-1}(t) + \hat{\varepsilon}) \leq F(F_m^{*-1}(1/m))
\]

(IV.2.9)

\[
< b \frac{F^* F_m^{*-1}(1/m)}{m} \leq b/m.
\]

Thus, from Equations IV.2.7, IV.2.8 and IV.2.9

\[
P\{K_N(t) < 2b\lambda^{-1}_N \text{ for all } t > 1/N\} \geq 1-\varepsilon.
\]

The proof of Lemma IV.2.3 is complete.

**Definition IV.2.1**

Let \( U_{m}^{**}(K_N) \) equal \( U_{m}^{*}(K_N) \) for \( 1/N \leq t \leq 1-1/N \) and equal zero otherwise.

**Theorem IV.2.1**

For \( q \in Q \)

\[
\rho_q(U_{m}^{**}(K_N), U_{o}(K_{N})) \to 0
\]

as \( N \to \infty \) uniformly in \( F \) and \( G \) and \( \lambda_N \in \Lambda \).

**Proof:**

The proof is similar to that of Theorem 2.2 of Pyke and Shorack (1968). The details are therefore left to the reader.
IV.3. The two-sample process

Lemma IV.3.1

With probability one,

\[ L^*_N(t) = (1-\lambda_N)^{-1/2} B_N(t) \ U^*_m(F(H^{-1}_N(t) + \xi)) \]

\[ \quad - (1-\lambda_N)^{-1/2} A_N(t) \ V^*_n(G(H^{-1}_N(t) + \eta)) \] + \delta_N(t)

for all \( t \in [0,1] \) where

\[ \delta_N(t) = A_N(t) N^{1/2} \{ H_N H^{-1}_N(t) - t \}, \]

\[ A_N(t) = \{ K_{\lambda N} (u_t) - K_{\lambda N} (t) \} / (u_t - t), \ u_t = H_{\lambda N} H^{-1}_N(t) \]

and \( B_N \) is defined by

\[ \lambda_N A_N(t) + (1-\lambda_N) B_N(t) = 1. \]

Proof:

The proof is similar to that of Lemma 3.1 of Pyke and Shorack (1968) and is therefore omitted.

Definition IV.3.1

Whenever well defined, set

\[ L^*_O(t) = (1-\lambda_O)^{-1/2} b_O(t) \ U^*_o(FH^{-1}_O(t)) \]

\[ \quad - (1-\lambda_O)^{-1/2} a_O(t) \ V^*_o(GH^{-1}_O(t)) \]
where \( a_0 \) and \( b_0 \) denote the derivatives of \( F_{H_0}^{-1} \) and \( G_{H_0}^{-1} \).

We now consider the convergence of the \( L_N^* \)-process to the limiting \( L_0 \)-process when \( \lambda_N + \lambda_* \) and \( \int_0^1 q \, d|\nu| < \infty \) for some \( q \in Q \) (see Definition II.3.2). Let \( \delta_N^* \) be equal to \( \delta_N \) on \([1/N, 1-1/N]\) and 0 elsewhere, and let \( L_N^* \) equal \( L_N^* \) on \([1/N, 1]\) and 0 elsewhere. Since \( |\delta_N^*| \leq \lambda_*^{-1} \), we have

\[
\rho_q(\delta_N^*, 0) = o(1). \tag{IV.3.2}
\]

Also,

\[
\sup_{1-1/N < t < 1} \frac{|L_N(t)/q(t)|}{|L_N(t)/q(t)|} = o(1), \tag{IV.3.3}
\]

since \( |L_N(t)| = N^{1/2} |1-F(H^{-1}_{H_N}(t)+\hat{\xi})| \leq N^{1/2} \lambda_*^{-1}(1-t) \) in this interval.

Recall that whenever well defined

\[
||f||_\nu = \int_0^1 |f(t)| \, d|\nu|(t).
\]

Also, if \( \int_0^1 q \, d|\nu| < \infty \) then

\[
||f||_\nu \leq \rho_q(f, 0) \int_0^1 q \, d|\nu|. \tag{IV.3.4}
\]

Hence, by Equation IV.3.2

\[
||\delta_N^*||_\nu = o(1)
\]
and by Equation IV.3.3
\[
\int_{1-\frac{1}{N}}^{1} |I_N| d|\nu| = o(1).
\]

Thus, in order to prove that
\[
||I_N' - L_o||_\nu = 0
\]

it will suffice to show that

(IV.3.5) \[ ||b^*_m(F(H^{-1}_N(t)+\xi)) - b_o(t) U_o(K_o)||_\nu \]

and the similar expression in A_n, a_o, V^**, V_o and G converge in probability to zero. By the definition of \[ ||f||_\nu \], Equation IV.3.5, and the triangle inequality, we must show that

\[
\rho(B_N,0)\{||U^*_m(K_N) - U_o(K_o)||_\nu
\]

(IV.3.6) \[ + ||U_o(K_{n}) - U_o(K_o)||_\nu \}

\[ + \rho_q(U_o(K_o),0)||B_N - b_o||_q \]

converges to zero. Since |B_N(t)| \leq (1-\lambda_N)^{-1}, we have that |B_N(t)| = o(1). Apply Theorem IV.2.1 and Equation IV.3.4 to get that \[ ||U^*_m(K_N) - U_o(K_o)||_\nu = 0 \]. By Lemma 2.2 of Pyke and Shorack (1968) we have that \[ \rho_q(U_o(K_o),0) = O_p(1) \]. Hence, in order to establish
it suffices to find conditions under which

\[ (IV.3.8) \quad \|U_0(K_{\lambda_N}^\perp) - U_0(K_0)\|_\Psi^p \not\to 0 \]

and

\[ (IV.3.9) \quad \|B_N - b_o\|_\Psi^p \not\to 0. \]

We now make an assumption which will imply IV.3.8 and IV.3.9.

**Assumption IV.3.1**

We assume that \( F \) and \( G \) possess densities \( f \) and \( g \) such that the boundary of the set of common zeroes of \( f \) and \( g \) is finite. Further, in any compact subset \( E_1 \) of \( S_1 = \{x: 0 < F(x) < 1\} \), \( f \) has at most a finite number of discontinuities. Similarly, in any compact subset \( E_2 \) of \( S_2 = \{x: 0 < G(x) < 1\} \), \( g \) has at most a finite number of discontinuities.

**Lemma IV.3.2**

For all \( t \in [0,1] \) we have

(a) \( H_{o,\lambda_N}^{1-1}(t) - \tau_N \leq H_{\lambda_N}^{1-1}(t) \leq H_{o,\lambda_N}^{1-1}(t) + \tau_N \),

(b) \( H_{N}^{*1}(t) - \tau_N \leq H_{N}^{1}(t) \leq H_{N}^{*1}(t) + \tau_N \)

and (c) \( H_{o}^{1}(t - \theta_N) \leq H_{o,\lambda_N}^{1}(t) \leq H_{o}^{1}(t + \theta_N) \).
where \( \tau_N = \max(|\hat{\xi}|, |\hat{\eta}|) = o_p(N^{-1/2}) \) and 

\[ \theta_N = 2 \max(|\lambda_N - \lambda_o|, N^{-1/2}) = o(N^{-1/2}). \]

**Proof:**

(a) Clearly, we have

\[ (IV.3.10) \quad H_0,\lambda_N (x - \tau_N) \leq H,\lambda_N (x) \leq H_0,\lambda_N (x + \tau_N). \]

For convenience let \( H_0,\lambda_N (x + \tau_N) = R(x) \). We have \( R^{-1}(t) = H_0,\lambda_N (x + \tau_N) - t \) \( N^{-1} \) and using Equation IV.3.10 \( R^{-1}(t) \leq H^{-1}_\lambda (t) \).

Hence, we have \( H^{-1}_0,\lambda_N (t) - \tau_N \leq H^{-1}_\lambda (t) \). Similarly,

\[ H^{-1}_\lambda (t) \leq H^{-1}_0,\lambda_N (t) - \tau_N. \]

The inequalities involving \( H^{-1}_N \) are obtained similarly.

(b) Since \( H_0,\lambda_N = H_0 + (\lambda_N - \lambda_*) (F - G) \), we have

\[ H_0 (x) - \theta_N/2 \leq H_0,\lambda_N (x) \leq H_0 (x) + \theta_N/2. \]

Hence,

\[ (IV.3.11) \quad H_0 (x) - \theta_N < H_0,\lambda_N (x) < H_0 (x) + \theta_N. \]

Let \( y = H^{-1}_0,\lambda_N (t) \). Using Equation IV.3.11

\[ H_0 (y) - \theta_N < H_0,\lambda_N H^{-1}_0,\lambda_N (t) = t \]

since \( F \) and \( G \) are continuous. Thus,
\(H_O(y) < t + \theta_N\) so that \(y = H_O^{-1}(t) < H_O^{-1}(t + \theta_N)\). Similarly, \(H_O^{-1}(t - \theta_N) < H_O^{-1}(t)\) and the proof is complete.

Let \(C_O \subset (0,1)\) be the set on which \(FH_O^{-1}\) is differentiable, \(H_O^{-1}(t)\) is a continuity point of \(f\) and \(g\), and \(H_O^{-1}\) is continuous. By Assumption IV.2.1 and the fact that \(\lambda_O FH_O^{-1}(t) + (1 - \lambda_O) GH_O^{-1}(t) = t\), the Lebesgue measure of \(C_O\) is one.

**Lemma IV.3.3**

For \(t \in C_O\),

\[K_{\lambda_N} - K_O \xrightarrow{P} 0.\]

Further, for every subsequence of \(\{N\}\) there is a further subsequence \(\{N'\}\) such that

\[P\{K_{\lambda_N'} \rightarrow K_O \text{ on } C_O\} = 1.\]

**Proof:**

Let \(t \in C_O\). Using Lemma IV.3.2 and the triangle inequality, we have

\[|K_{\lambda_N}(t) - K_O(t)| = |F(H_O^{-1}(t) + \xi) - FH_O^{-1}(t)|.\]

(IV.3.12)

\[\leq |F(H_O^{-1}(t + \theta_N) + 2\tau_N) - FH_O^{-1}(t)| + |F(H_O^{-1}(t - \theta_N) - 2\tau_N) - FH_O^{-1}(t)|.\]
Consider $F(H^{-1}_O(t + \theta_N) + 2\tau_N) - FH^{-1}_O(t)$. By an application of the triangle inequality,

$$|F(H^{-1}_O(t + \theta_N) + 2\tau_N) - FH^{-1}_O(t)|$$

$$\leq |F(H^{-1}_O(t + \theta_N) + 2\tau_N) - F(H^{-1}_O(t + \theta_N))|$$

$$+ |F(H^{-1}_O(t + \theta_N)) - FH^{-1}_O(t)|.$$ 

Thus, using the definition of $C_0$ and a Taylor expansion we have for $N$ large,

$$|F(H^{-1}_O(t + \theta_N) + 2\tau_N) - F(H^{-1}_O(t))|$$

(IV.3.13) $\leq 2\tau_N f(H^{-1}_O(t + \theta_N) + 2\Delta_1 \tau_N)$

$$+ \theta_N a_o(t + \Delta_2 \theta_N)$$

where $|\Delta_1|,|\Delta_2| < 1$. Since $\tau_N, \theta_N = O_p(N^{-1/2}),$

we have by Equation IV.3.13 and the continuity of $f$ at $H^{-1}_O(t)$

(IV.3.14) $|F(H^{-1}_O(t + \theta_N) + 2\tau_N) - FH^{-1}_O(t)|_p \rightarrow 0.$

Similarly,

(IV.3.15) $|F(H^{-1}_O(t - \theta_N) - 2\tau_N) - FH^{-1}_O(t)|_p \rightarrow 0.$

Combining Equations IV.3.12, IV.3.14 and IV.3.15 we have

$$|K_{\lambda N}(t) - K_o(t)|_p \rightarrow 0$$

for all $t \in C.$
For any subsequence of \(\{N\}\) there exists a further subsequence \(\{N'\}\) such that \(\tau_{N'}\), \(\theta_{N'}\) a.s. 0. Hence, by Equation IV.3.13
\[
P\{F(H^{-1}_O(t + \theta_{N'})) - FH^{-1}_O(t) + 0, t \in C_o\} = 1.
\]
Similarly,
\[
P\{F(H^{-1}_O(t - \theta_{N'})) - FH^{-1}_O(t) + 0, t \in C_o\} = 1,
\]
so that
\[
P\{K_{\lambda N'} \to K_{O}, t \in C_o\} = 1.
\]
The proof of Lemma IV.3.3 is now complete.

**Lemma IV.3.4**

If
\[
\int_0^1 |q|d\nu \ < \ \infty,
\]
then
\[
||U_0(K_{\lambda N}) - U_0(K_{O})||_{\nu} \to 0.
\]

**Proof:**

First, we will show that for \(\delta > 0\) there exists a constant \(c > 0\) such that \(P(A_{\delta}) > 1-\delta\) where

\[
(IV.3.16) \ A_{\delta} = \{ \omega: |U_0(K_{\lambda N}(t)) - U_0(K_{O}(t))| < c \ q(t) \ \text{for all} \ t \in (0,1) \}.
\]
Let $\varepsilon$ be the value associated with $q$ as in Definition II.3.2.

We have

\[
\sup_{0 \leq t \leq 1} \left| \frac{U_0(K_{\lambda N}(t))}{q(t)} - \frac{U_0(K_0(t))}{q(t)} \right| 
\]

\[
\leq \sup_{0 \leq t \leq \varepsilon} \left| \frac{U_0(K_{\lambda N}(t))}{q(t)} + \frac{U_0(K_0(t))}{q(t)} \right| 
\]

\[
+ \sup_{1-\varepsilon \leq t \leq 1} \left| \frac{U_0(K_{\lambda N}(t))}{q(t)} + \frac{U_0(K_0(t))}{q(t)} \right| 
\]

\[
+ 2\rho(U_0,0) + \inf_{\varepsilon \leq t \leq 1-\varepsilon} \frac{2p(U_0,0)}{q(t)}. 
\]

(IV.3.17)

Consider the first term on the right. Since $t = \lambda N^{-1}(F_{H^\lambda(t)} + \xi)$

\[
+ (1-\lambda N)G(H_{\lambda N}^{-1}(t)+\eta), \text{ we have } K_{\lambda N}(t) = F(H_{\lambda N}^{-1}(t)+\xi) \leq \lambda_*^{-1}t. 
\]

Hence, for $0 \leq t \leq a$ (with $a \leq \varepsilon \lambda_*$)

(IV.3.18) $U_0(K_{\lambda N}(t))/q(t)) < U_0(K_{\lambda N}(t))/q(\lambda_* K_{\lambda N}(t))$.

Similarly, $F_{H_0^{-1}(t)} \leq \lambda_*^{-1}t$ and for $0 \leq t \leq a$,

(IV.3.19) $\frac{U_0(K_0(t))}{q(t)} \leq \frac{U_0(K_0(t))}{q(\lambda_* K_0(t))}$. 
By Equations IV.3.18 and IV.3.19 we have
\[
\sup_{0 \leq t \leq a} \left| \frac{U_0(K(t))}{q(t)} - \frac{U_0(K(t))}{q(t)} \right|
\]
(IV.3.20)
\[
\leq 2 \sup_{0 \leq t \leq \varepsilon} \left| \frac{U_0(t)}{q(\lambda_o t)} \right|.
\]

The second term on the right side of Equation IV.3.17 may be dealt with in a similar manner by considering the reverse process. Thus, we have
\[
\sup_{0 \leq t \leq 1} \left| \frac{U_0(K(t))}{q(t)} - \frac{U_0(K(t))}{q(t)} \right|
\]
(IV.3.20)
\[
\leq 4 \sup_{0 \leq t \leq \varepsilon} \left| \frac{U_0(t)}{q(\lambda_o t)} \right| + \frac{\rho(U_0,0)}{\inf_{\varepsilon \leq t \leq 1-\varepsilon} q(t)}.
\]

By Lemma 2.2 of Pyke and Shorack (1968) and the definition of \(q\), each term on the right is bounded in probabilities. Hence,
\[
\sup_{0 \leq t \leq 1} \left| \frac{U_0(K(t))}{q(t)} - \frac{U_0(K(t))}{q(t)} \right| = O_p(1)
\]
and for \(\delta > 0\), \(c\) and \(A_\delta\) can be chosen.

On \(A_\delta\), \(|U_0(K(t)) - U_0(K(t))|\) is bounded by a \(|v|\) integrable function. By Lemma IV.3.3 and the a.s. continuity
of $U_\omega$, we have for every subsequence of \{N\}, a further subsequence \{N'\} such that (same argument as Lemma IV.3.6)

$$|U_\omega(K_{\lambda N}^n(t)) - U_\omega(K_0^n(t))| \to 0 \text{ a.e. } P \times |\nu|.$$ 

Hence, by the dominated convergence theorem, for $\delta > 0$

$$P\{\int_0^1 |U_\omega(K_{\lambda N}^n(t)) - U_\omega(K_0^n(t))| d|\nu| \to 0/A_\delta \} = 1.$$ 

Thus

$$P\{\int_0^1 |U_\omega(K_{\lambda N}^n(t)) - U_\omega(K_0^n(t))| d|\nu| \to 0/A_\delta \} = 1$$

and

$$\int_0^1 |U_\omega(K_{\lambda N}^n(t)) - U_\omega(K_0^n(t))| d|\nu| \to 0.$$

**Lemma IV.3.5**

For every subsequence of \{N\} there is a further subsequence \{N'\} such that

$$P\{\lim_{N \to \infty} H_{N'}^{-1}(t) = \lim_{N \to \infty} H_{\lambda N'}^{-1}(t) = H_0^{-1}(t) \text{ for all } t \in C\} = 1.$$ 

**Proof:**

First, we show that

$$(IV.3.21) \ P\{\omega: \lim_{N \to \infty} H_{N}^{-1}(t) = H_0^{-1}(t) \text{ for all } t \in C_O\} = 1.$$ 

Continuity of $H_0^{-1}$ at $t$ implies (see Hajek and Sidak (1967) and page 33) that for $\varepsilon > 0$
It follows that for \( \varepsilon' > 0 \), there is an \( \varepsilon'' > 0 \) such that
\[
|y - H^{-1}_O(t)| > \varepsilon' \text{ implies } |t - H^{-1}_O(y)| > \varepsilon''.
\]
Thus, setting \( Y_N = H^{-1}_N(t) \), it is enough to show that \( H^{-1}_O(H^{-1}_N(t)) - t \) a.s. \( 0 \) uniformly in \( t \). This follows using arguments similar to those of Lemma 2.3 of Pyke and Shorack (1968). For almost all \((P)\)

\[\omega \in \Omega,\]

\[
\sup_{0 < t < 1} |H^{-1}_O H^{-1}_N(t) - t| \to 0 \text{ as } N \to \infty
\]

which shows that for almost all \((P)\) \( \omega \in \Omega \),

\[H^{-1}_N(t) \to H^{-1}_O(t) \text{ for all } t \in C_0.\]

Thus, Equation IV.3.21 is proved.

We have, from Lemma IV.3.2

\[
|H^{-1}_N(t) - H^{-1}_O(t)| \leq \tau_N + |H^{-1}_N(t) - H^{-1}_O(t) |
\]

where \( \tau_N \) is as defined in Lemma IV.3.2. From Equation IV.3.21 and the definition of \( \tau_N \), there exists a further subsequence \( \{N'\} \) such that

\[
P\{\lim_{N' \to \infty} H^{-1}_N(t) \to H^{-1}_O(t) \text{ for all } t \in C_0\} = 1.
\]

Similarly, by Lemma IV.3.2, there exists further subsequence \( \{N'\} \) such that

\[
P\{\lim_{N' \to \infty} H^{-1}_O, \lambda_N(t) = H^{-1}_O(t) \text{ for all } t \in C_0\} = 1.
\]
Since
\[ |H_{\lambda N}^{-1}(t) - H_0^{-1}(t)| \leq |H_{\lambda N}^{-1}(t) - H_{\lambda N,0}^{-1}(t)| \]
\[ + |H_{\lambda N,0}^{-1}(t) - H_0^{-1}(t)| \]
\[ \leq \tau_N + |H_{\lambda N,0}^{-1}(t) - H_0^{-1}(t)|, \]
we have by Lemma IV.3.2 that
\[ P\left( \lim_{N \to \infty} H_{\lambda N}^{-1}(t) \to H_0^{-1}(t) \text{ for all } t \in C_0 \right) = 1. \]

The proof of the lemma is now complete.

**Lemma IV.3.6**

Under Assumption IV.3.1, if \( \lambda_N = \lambda_* + O(N^{-1/2}) \), then
\[ \| (A_N - a_0)q \|_v \to 0, \quad (\| (B_N - b_0)q \|_v \to 0) \]
as \( N \to \infty \).

**Proof:**

We prove the result for \( A_N \); for \( B_N \) the proof is similar.

We have
\[ (IV.3.22) \quad A_N(t) = A_{N1}(t) A_{N2}(t), \]
where
\[ A_{N1}(t) = \frac{K_{H_0^{-1}(t)+\hat{\xi}} - K_{H_0^{-1}(t)+\hat{\xi}}}{H_0^{-1}(t)+\hat{\xi}} \]
and

\[ A_{N2}(t) = \frac{H_o(H^{-1}_N(t) + \hat{\xi}) - H_o(H^{-1}_N(t) + \hat{\xi})}{H^{-1}_N(t) - t} \].

If we let

\[ D_N^{(1)} = F(H^{-1}_N(t) + \hat{\xi}) - F(H^{-1}_N(t) + \hat{\xi}) \]

(IV.3.23)

\[ - G(H^{-1}_N(t) + \hat{\xi}) + G(H^{-1}_N(t) + \hat{\xi}) \]

and

\[ D_N^{(2)} = G(H^{-1}_N(t) + \hat{\xi}) - G(H^{-1}_N(t) + \hat{\xi}) \]

(IV.3.24)

\[ - G(H^{-1}_N(t) + \hat{\eta}) + G(H^{-1}_N(t) + \hat{\eta}), \]

then we obtain

\[ H_o(H^{-1}_N(t) + \hat{\xi}) - H_o(H^{-1}_N(t) + \hat{\xi}) \]

(IV.3.25)

\[ = \lambda_o D_N^{(1)} + G(H^{-1}_N(t) + \hat{\xi}) - G(H^{-1}_N(t) + \hat{\xi}) \]

and

\[ H^{-1}_N(t) - t = \lambda N D_N^{(1)} + D_N^{(2)} \]

(IV.3.26)

\[ + G(H^{-1}_N(t) + \hat{\eta}) - G(H^{-1}_N(t) + \hat{\eta}). \]

From Equations IV.3.24, IV.3.25 and IV.3.26, we have
\[ H_0(H_N^{-1}(t)+\hat{\xi}) - H_0(H_N^{-1}(t)+\hat{\xi}) \]

(IV.3.27)

\[ = H_N^{-1}(t) - t + (\lambda_0 - \lambda_N) D_N^{(1)} + (1 - \lambda_N) D_N^{(2)}. \]

Also,

\[ H_N^{-1}(t) - t = H_N^{-1}(t) - H_N^{-1}(t) + H_N^{-1}(t) - t \]

(IV.3.28)

\[ = d_N(t) - N^{-1/2} w_N(t) \]

where \(0 \leq d_N(t) = H_N^{-1}(t) - t \leq 1/N\) and

\[ w_N(t) = N^{1/2}(H_N^{-1}(t) - H_N^{-1}(t)). \]

Further,

\[ w_N(t) = N^{1/2}(\lambda_N(F_m^{-1}(t) - F(H_N^{-1}(t)+\hat{\xi})) + (1-\lambda_N)(G_N^{-1}(t) - G(H_N^{-1}(t)+\hat{\eta}))) \]

(IV.3.29)

\[ = \lambda_N^{1/2} u_m(K_N(t)) + (1-\lambda_N)^{1/2} v_N(K_N^*(t)) \]

where \(K_N^*(t) = G(H_N^{-1}(t)+\hat{\eta}).\) From Equations IV.3.27 and IV.3.28 we have

(IV.3.30)

\[ \lambda_{N2}(t) = 1 + \frac{N^{1/2} (\lambda_0 - \lambda_N) D_N^{(1)}(t) + N^{1/2} (1-\lambda_N) D_N^{(2)}(t)}{N^{1/2} d_N(t) - w_N(t)}. \]
From Lemma IV.2.3, we have

\[(IV.3.31) \quad \rho(W_{N}, W_{ON}) \rightarrow 0 \text{ as } N \rightarrow \infty\]

where

\[(IV.3.32) \quad W_{ON}(t) = \lambda_{N}^{1/2} U_{O}(K_{N}^{*}(t)) + (1-\lambda_{N})^{1/2} V_{O}(K_{N}^{*}(t))\]

with \(K_{N}^{*} = G(H_{N}^{-1}(t) + \eta)\). From Lemma IV.3.3 we have for every subsequence of \(\{N\}\) a further subsequence \(\{N^{*}\}\) such that

\[P\{K_{N}^{*}(t) \rightarrow e. K_{O}(t) \text{ and } K_{N}^{*}(t) \rightarrow e. G_{O}^{-1}(t) \text{ for } t \in C\} = 1.\]

Using these convergences, Equation IV.3.31, and the a.s. continuity of \(U_{O}\) and \(V_{O}\), we have for every subsequence of \(\{N\}\) a further subsequence \(\{N^{*}\}\) such that

\[(IV.3.33) \quad P\{W_{N^{*}}(t) \rightarrow W_{O}(t) \text{ for all } t \in C\} = 1.\]

Using a Taylor expansion we have

\[D_{N}^{(2)}(t) = \hat{\xi}(g(H_{N}^{-1}(t) + \Delta_{1} \hat{\xi}) - g(H_{N}^{-1}(t) + \Delta_{2} \hat{\xi}))\]

\[- \hat{\eta}(g(H_{N}^{-1}(t) + \Delta_{1} \hat{\eta}) - g(H_{N}^{-1}(t) + \Delta_{2} \hat{\eta}))\]

for \(t \in C\), where \(|\Delta_{1}|, |\Delta_{2}| \leq 1\). Using Lemma IV.3.5 and the continuity of \(g\) at \(H_{O}^{-1}(t)\), we have for every subsequence of \(\{N\}\) a further subsequence \(\{N^{*}\}\) such that

\[(IV.3.34) \quad P\{N^{*1/2} D_{N}^{(2)}(t) \rightarrow 0 \text{ for all } t \in C\} = 1.\]
By similarity,

\[(IV.3.35)\quad P\{N'^{1/2}D_N(1)(t) \to 0 \text{ for all } t \in C_o\} = 1.\]

Now, for each \(\omega \in \Omega\), let \(C_{\omega, \omega} = \{t \in (0,1) : \omega(t) \neq 0\}\). It follows from Equations IV.3.30, IV.3.33, IV.3.34 and IV.3.35 that every subsequence of \(\{N\}\) contains a further subsequence \(\{N'\}\) such that

\[(IV.3.36)\quad P\{\omega \in \Omega : A_{N'^{1/2}}(t) \to 1 \text{ for all } t \in C_{\omega, \omega} \cap C_{o}\} = 1.\]

To study the limit of \(A_{N_1}\), we use the same argument as that of Lemma 5.1 of Pyke and Shorack (1968). For this, note that for \(t \in C_{o, \omega} \cap C_o\) and almost all \(\omega\)

\[
\begin{align*}
(H_{\lambda_N}^{-1}(t) - H_{\lambda_N}^{-1}(t + \hat{n}))(H_{\lambda_N}^{-1}(t) + \hat{\xi}) - H_{\lambda_N}^{-1}(t + \hat{n}) - H_{\lambda_N}^{-1}(t + \hat{\xi}) & = \{N^{1/2}(\lambda_N - \lambda_*) (F(H_{\lambda_N}^{-1}(t) + \hat{\xi}) - G(H_{\lambda_N}^{-1}(t) + \hat{n}))

+ N^{1/2}(1 - \lambda_*) (\hat{n} \ g(H_{\lambda_N}^{-1}(t) + \Delta_2 \hat{n}) - \hat{\xi} \ g(H_{\lambda_N}^{-1}(t) + \Delta_2 \hat{\xi}))

\{N^{1/2}d_N(t) - W_N(t) + (\lambda_* - \lambda_N)N^{1/2}D_N(1)(t) + (1 - \lambda_N)N^{1/2}D_N(2)(t)\}^{-1}
\end{align*}
\]
remains bounded (see Equations IV.3.27, IV.3.28, IV.3.33, IV.3.34, and IV.3.35). It follows as in Pyke and Shorack (1968) that

\[(IV.3.37)\quad P\{\omega \in \Omega : A_{N'_1}(t) \to a_0(t) \text{ for all } t \in C_{o, \omega} \cap C_o\} = 1.\]
Note that \( \nu(C_0) = 0 \). By combining Equations IV.3.36 and IV.3.37 we have

\[
A_N(t) \rightarrow a_0(t) \quad \text{a.e.} \quad P \times |\nu|.
\]

Hence, applying the dominated convergence theorem (as in Pyke and Shorack (1968)),

\[
| |(A_N - a_0)| | \underset{a.s.}{\rightarrow} 0,
\]

and from this we have

\[
| |(A_N - a_0)| | \underset{a.s.}{\rightarrow} 0.
\]

Convergence in probability follows and the proof of Lemma IV.3.6 is complete.

**Theorem IV.3.1**

(a) Suppose Assumption IV.3.1 holds and \( K_0 \) is differentiable a.e. \( |\nu| \), \( \lambda_N = \lambda_0 + O(N^{1/2}) \) and \( \int_0^1 q \, d|\nu| < \infty \) for some \( q \in \mathbb{Q} \). Then \( | |L_N^r - L_0^r| | \underset{P}{\rightarrow} 0 \) so that \( L_N^r \overset{p}{\rightarrow} L_0^r \) relative to \( (D,|\ast|,|\nu|) \).

(b) Suppose in addition that

(i) \( \int_0^1 \frac{L_N}{N} \, d(\nu_N - \nu) \overset{p}{\rightarrow} 0 \)

(ii) \( \lim_{N \to \infty} N^{1/2} \int_0^1 (F(H_N^{-1}(t) + \xi) - FH_0^0(t)) \, d\nu \overset{p}{\rightarrow} 0. \)

Then, \( T_N^* \overset{p}{\rightarrow} \int_0^1 L_0 \, d\nu \) which is a \( N(0,c_o^2) \) random variable where \( c_o^2 \) is as defined in Pyke and Shorack (1968).
V. VARIATIONAL ARGUMENTS

V.1. Asymptotic relative efficiency

We have established that the limiting distribution of Chernoff-Savage statistics, $T_N$, is asymptotically normal under $p$-dependence or a mixing condition (see Definitions I.1.1 and I.1.2). We will now study the efficiency properties of the statistics $T_N$ relative to the corresponding normal theory statistics.

Consider a sequence $X_1, X_2, \ldots, X_m, \ldots$ of identically distributed random variables. They may be independent or not. Denote by $L(X)$ the probability law of $X = (X_1, X_2, \ldots, X_m, \ldots)$. Also, consider a double sequence $Y^{(n)} = (Y_{1n}, Y_{2n}, \ldots, Y_{nn}, \ldots)$ and let $L(Y^{(n)})$ be the probability law of $Y^{(n)}$. Let $\theta_N$ and $\phi_N$ be real numbers and assume that $\Delta_N = \theta_N - \phi_N = cN^{-1/2}$. We also assume that

\[(V.1.1) \quad L(X - \theta_N) = L(Y^{(n)} - \phi_N),\]

where $\theta_N = (\theta_N, \theta_N, \ldots)$ and $\phi_N = (\phi_N, \phi_N, \ldots)$.

Further,

\[(V.1.2) \quad 0 < \lim_{N \to \infty} \lambda_N = \lambda_0 < 1.\]
Under these assumptions, one can evaluate the asymptotic relative efficiency, i.e. Pitman efficiency (see Hodges and Lehman (1956)), of two-sample rank-order tests relative to the t-test for testing $H_0: \Delta = 0$ vs. a sequence of near alternatives $K_N: \Delta_N = cN^{-1/2}$. Let $T_N$ be a test statistic such that there are functions $a(N(\Delta_N))$ and $b(N(\Delta_N))$ such that

$$L \left( \frac{T_N - a(N(\Delta_N))}{b(N(\Delta_N))} \right) \to N(0,1),$$

(V.1.3)

$$\lim_{N \to \infty} \frac{b(N(\Delta_N))}{b(N(0))} = 1,$$

(V.1.4)

and

$$E_T = \lim_{N \to \infty} \frac{a(N(\Delta_N)) - a(N(0))}{b(N(\Delta_N)) b(N(0))}^2$$

(V.1.5)

exists and is independent of $c$. $E_T$ is the efficiency of the sequence $T_N$. If $T^*_N$ is another test statistic appropriate for the same near alternative and $E_{T^*}$ exists, then $E_{T,T^*} = E_T / E_{T^*}$ is the asymptotic relative efficiency (A.R.E.) of $T_N$ with respect to $T^*_N$.

V.2. The A.R.E. for the p-dependent case

In this section we consider two sequences $(X_1, X_2, \ldots)$ and $(Y_1, Y_2, \ldots)$ which satisfy the conditions for the Pitman criteria
described in Section V.1. Further, let these sequences be p-dependent with $p \geq 1$. In addition we assume that $X_i$ and $Y_j$ are independent for any $(i,j)$ combination. Observe that the notation here is the same as that in Chapter IV.

**Lemma V.2.1**

If Assumption IV.1 is satisfied, $\lambda_N = \lambda_0 + O(N^{-1/2})$, and $\mu$ is a Lebesgue-Stieltjes measure on $(0,1)$ for which $\mu(C^C_0) = 0$, then $A_N + a_0$ and $B_N + b_0$ in $\mathbb{P} \times \mu$ measure as $N \to \infty$.

**Proof:**

We have $F(x - \theta_N) = G(x - \phi_N)$ where without loss of generality, $\theta_N = O(N^{-1/2})$ and $\phi_N = O(N^{-1/2})$. Thus, Lemma V.2.1 is a special case of Lemma IV.2.1. The proof is therefore omitted.

**Theorem V.2.1**

If

(i) Assumption IV.4.1 is satisfied

and

(ii) $\lambda_N = \lambda_0 + O(N^{-1/2})$, then

$$T_N \xrightarrow{\mathbb{P}} \int_0^1 L_0 \, d\nu,$$

a $N(0, c_0^2)$ random variable, with
\[ \sigma^2_0 = \int_0^1 J^2(u) \, du - \left( \int_0^1 J(u) \, du \right)^2 \]

\[ + 2 \sum_{i=1}^{p} \left( \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} J(F(u)) J(F(v)) \, dF_i(u,v) - \left( \int_0^1 J(u) \, du \right)^2 \right). \]

Proof (1) We use the results of Lemma V.1.1 and the argument of Theorem II.4.2 so that \( T_N \xrightarrow{p} \int_0^1 L_0 \, dv \) which is a \( N(0, \sigma^2_0) \). To see that the expression for \( \sigma^2_0 \) is the above theorem is equivalent to that of Theorem II.4.2 observe that \( dv = -dJ \), (see Pyke and Shorack (1969)), apply integration by parts, and use the arguments of Corollary 2 of Chernoff and Savage (1958).

Now, in the manner of Sen (1967), define

\[ A^2(F,J) = \int_0^1 J^2(u) \, du - \left( \int_0^1 J(u) \, du \right)^2; \]

\[ \rho_i(F,J) = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} J(F(x)) J(F(y)) \, dF_i(x,y) - \left( \int_0^1 J(u) \, du \right)^2}{A^2(F,J)} \]

for \( i = 1, \ldots, p \); and

\[ B(F,J) = \lim_{N \to \infty} N^{1/2} \int_{-\infty}^{+\infty} J(\lambda_N F(X + \theta_N) + (1-\lambda_N) F(X + \theta_N)) \]

\[ - J(F(x)) \, dF(x). \]
Then, by Equation V.1.5 and Theorem V.2.1, the efficacy of the procedure based on $T_N$ is

$$B^2(F, J)/A^2(F, J) \{1 + 2 \sum_{i=1}^{p} \rho_i(F, J)\}.$$  

Let $\rho_i = \text{Cov}(X_i, X_{i+1})/\sigma^2$, then the efficacy of the two sample t-statistic is $c[\sigma^2(1 + 2 \sum_{i=1}^{p} \rho_i)]^{-1}$. Thus, under $p$-dependence, the A.R.E. of $T_N$ with respect to the two-sample t-statistic is

$$B^2(F, J) \left( \frac{1 + 2 \sum_{i=1}^{p} \rho_i}{1 + 2 \sum_{i=1}^{p} \rho_i(F, J)} \right).$$  

Note that the first factor depends on the univariate marginal c.d.f. $F(x)$, while the second depends on the bivariate distributions. The values for the first term are generally known and therefore will not be given. The second factor cannot, in general, be bounded below (see Sen (1967)). Thus, for application of the above expression the reader must consider his special case of interest. The following comments are, however, of some importance.

(i) The Normal Score Test: Take $J$ equal to the inverse of the standard normal c.d.f. $\Phi$. If the parent univariate distributions are normal, then by Theorem 3 of Chernoff and Savage
(1958) the first term in Equation V.2.1 is one. If \( F \) is normal, say mean \( \mu \) and variance \( \sigma^2 \), then \( F(x) = \phi\left(\frac{x-\mu}{\sigma}\right) \) so that \( J(F(x)) = \frac{1}{\sigma}(x-\mu) \). Thus, if the parent bivariate distributions are also normal \( \rho_i(F,J) = \rho_i \) and Equation V.2.1 reduces to unity.

(ii) Median Test: Take

\[
J(v) = \begin{cases} 
0 & \text{for } 0 \leq v \leq 1/2 \\
1 & \text{for } 1/2 < v \leq 1,
\end{cases}
\]

and assume that the underlying bivariate distributions are normal. It is known (see Hajek and Sidak (1967) and problem 8 on page 278) that the first term of Equation V.2.1 reduces to \( 2/\pi \).

To simplify the second term, we have

\[
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dF_i(x,y) = P\{X_1 > 0, X_2 > 0\}.
\]

It follows from Cramer (1945) and page 290 that

\[
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dF_i(x,y) = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \rho.
\]

Thus,

\[
\rho_i(F,J) = 4 \cdot \frac{1}{2\pi} \sin^{-1} \rho,
\]

and the efficiency of the rank sum test with respect to the t-test is
\[ E = \frac{2}{\pi} \left( \frac{1 + 2 \sum_{i=1}^{p} \rho_i}{1 + 4 \sum_{i=1}^{p} \sin^{-1} \rho_i} \right) \]

or taking \( \rho_i = \sin u_i \pi \) we have

\[ E = \frac{2}{\pi} \left( \frac{1 + 2 \sum_{i=1}^{p} \sin u_i \pi}{1 + 4 \sum_{i=1}^{p} u_i} \right). \]

Since

\[ \sin u_i \pi \geq 2 u_i \text{ for } 0 \leq u_i < 1/2 \]

we have

\[ E \geq \frac{2}{\pi} \left( \frac{1 + 4 \sum_{i=1}^{p} u_i}{1 + 4 \sum_{i=1}^{p} u_i} \right) \text{ for } 0 \leq u_i < 1/2, \quad i = 1, \ldots, p. \]

Since \( \rho_i = \sin u_i \pi \), it follows that \( E \geq 2/\pi \) if \( \rho_i > 0 \), \( i = 1, \ldots, p \), and using a similar argument that \( E < 2/\pi \) if \( \rho_i < 0 \) for \( i = 1, \ldots, p \).

(iii) Rank Sum Test: Take \( J(u) = u \) and assume that the underlying bivariate distributions are normal. It is known (see Hajek and Sidak (1967) and problem 8 on page 278) that the first term of Equation V.2.1 reduces to \( 3/\pi \).
To simplify the second term, we follow the proof of Theorem 5.2 in Bickel (1964). We have

\[ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(x) F(y) \, dF_i(x,y) \]

\[ = \mathbb{P} \{ X-X' > 0, Y-Y' > 0 \} \]

where \( X' \) and \( Y' \) are independent, identically distributed as \( F \) and \((X,Y)\) is distributed as \( F_i(x,y) \). Since these distributions are normal and the processes are stationary, we have that \((X-X', Y-Y')\) has a \( N(0, 2\sigma^2, 2\sigma^2, \rho_i/2) \) distribution. It follows from Cramer (1956) and page 290 that

\[ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(x)F(y)dF_i(x,y) = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \rho_i/2. \]

Thus,

\[ \rho_i(F,J) = 12 \cdot \frac{1}{2\pi} \sin^{-1} \rho_i/2 \]

and efficiency of the rank sum test with respect to the \( t \) test is

\[ E = \frac{3}{\pi} \left( \frac{1 + 2 \sum_{i=1}^{P} \rho_i}{1 + \frac{12}{\pi} \sum_{i=1}^{P} \sin^{-1} \frac{\rho_i}{2}} \right). \]
Since \( \sin^{-1} \frac{y}{y} \) is non-decreasing for \( 0 < y \leq \frac{1}{2} \), we have

\[
\frac{\sin^{-1} \frac{y}{y}}{y} \leq \frac{\sin^{-1} \frac{1}{2}}{1/2} = \frac{\pi}{3} \quad 0 < y \leq \frac{1}{2}
\]

or

\[
\sin^{-1} y < \frac{\pi}{3} y \quad 0 < y \leq \frac{1}{2}.
\]

It follows that

\[
E > \frac{3}{\pi} \left( \frac{1 + 2 \sum_{i=1}^{p} \rho_i}{1 + 2 \sum_{i=1}^{p} \rho_i} \right) \quad \text{for } 0 < \rho_i \leq 1
\]

\[
i = 1, \ldots, p.
\]

In a similar manner it can be shown that \( E < \frac{3}{\pi} \)

if \( \rho_i < 0 \) for \( i = 1, \ldots, p \).

V.3. The A.R.E. for the mixing case

Assume that \((X_1, X_2, \ldots)\) and \((Y_1, Y_2, \ldots)\) are \( \phi \)-mixing (see Definition I.2.2) and that they satisfy the conditions described in section III.1. The notation used here is the same as that in Chapter III.

Lemma V.3.1

If Assumption IV.1 is satisfied, \( \lambda_N = \lambda_0 + O(N^{-1/2}) \), \( \sum n^2 \phi^{1/2}(n) < \infty \), and \( \mu \) is a Lebesque-Sieltjes measure on \((0,1)\) for which \( \mu(C_0^C) = 0 \), then \( A_n \to a_0 \) and \( B_n \to b_0 \) in \( P \times \mu \) measure.
as $N \to \infty$.

Proof:

As in Lemma V.2.1 the proof is similar to that of Lemma IV.4.1.

**Theorem V.3.1**

Suppose

(i) Assumption IV.4.1 is satisfied

(ii) $\lambda_N = \lambda_0 + O(N^{-1/2})$

(iii) $\sum n^2 \phi^{1/2}(n) < \infty$.

Then

(1) $T_N \overset{p}{\to} \int_0^1 L_0 \, d\nu$, a $\mathcal{N}(0, \sigma^2_0)$

r.v. with

$$
\sigma^2_0 = \int_0^1 J^2(u) \, du - \left( \int_0^1 J(u) \, du \right)^2
$$

$$
+ 2 \sum_{i=1}^{\infty} \left( \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} J(F(u)) J(F(v)) \, dF_i(u,v) \right.

- \left. \int_0^1 J(u) \, du \right)^2.
$$

Proof:

The proof is similar to that of Theorem V.2.1 and the details are therefore omitted.
Define

\[ A^2(F,J) = \int_0^1 J^2(u) \, du - \left( \int_0^1 J(u) \, du \right)^2; \]

\[ \rho_i(F,J) = \frac{\int_0^1 \int_0^1 J(F(x)) \cdot J(F(y)) \, dF(x,y) - \left( \int_0^1 J(u) \, du \right)^2}{A^2(F,J)} \]

for \( i = 1, 2, \ldots; \)

\[ B(F,J) = \lim_{N \to \infty} N^{1/2} \int_{-\infty}^{+\infty} [J(\lambda_N F(x+\theta_N) + (1-\lambda_N) F(x + \phi_N)) - J(F(x))] \, dF(x). \]

The efficacy of the procedure based on \( T_N \) is

(V.3.1) \[ \left( \frac{B^2(F,J)}{A^2(F,J)} \right) \left( 1 + \sum_{i=1}^{\infty} \rho_i(F,J) \right)^{1/2}. \]

Let \( \rho_i = \frac{\text{Cov}(X_i, X_{i+1})}{\sigma^2} \), then the efficacy of the t-statistic is \( c[\sigma^2(1+\Sigma \rho_i)]^{-1} \). Thus, the A.R.E. of \( T_N \) with respect to the t-statistic is

(V.3.2) \[ \left( \frac{\sigma^2 B^2(F,J)}{A^2(F,J)} \right) \left( \frac{1 + 2 \Sigma \rho_i}{1 + 2 \Sigma \rho_i(F,J)} \right). \]

The comments made following Equation V.2.3 are again relevant. The following comments are derived in a manner similar to those of the previous section.
(i) The Normal Score Test: If the underlying bivariate distributions are normal then Equation V.3.2 reduces to unity.

(ii) The Median Test: If the underlying distribution of \( \{X_i\}_{i=1}^{\infty} \) and \( \{Y_i\}_{i=1}^{\infty} \) are normal, Equation V.3.2 reduces to

\[
E = \frac{2}{\pi} \left( \frac{1 + 2 \sum_{i=1}^{\infty} \rho_i}{1 + 4 \sum_{i=1}^{\infty} \sin^{-1} \rho_i} \right).
\]

E is bounded below by \( \frac{2}{\pi} \) if \( \rho_i > 0 \) for \( i = 1, 2, \ldots \) and E is bounded above by \( \frac{2}{\pi} \) if \( \rho_i < 0 \) for \( i = 1, 2, \ldots \).

(iii) The Rank Sum Test: Again, under normality of the processes \( \{X_i\}_{i=1}^{\infty} \) and \( \{Y_j\}_{j=1}^{\infty} \) then Equation V.3.2 reduces to

\[
E = \frac{3}{\pi} \left( \frac{1 + 2 \sum_{i=1}^{\infty} \rho_i}{1 + \frac{12}{\pi} \sum_{i=1}^{\infty} \sin^{-1} (\pi_i / 2)} \right).
\]

E is bounded below by \( \frac{3}{\pi} \) if \( \rho_i > 0 \) for \( i = 1, 2, \ldots \) and E is bounded above by \( \frac{3}{\pi} \) if \( \rho_i < 0 \) for \( i = 1, 2, \ldots \).
VI. BIBLIOGRAPHY


VII. ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. K. L. Mehra for introducing me to the topic and advising me throughout the research and preparation of this thesis.

I would also like to thank my parents and my wife Marty for their patience and encouragement. Finally, I wish to express my appreciation to the National Institutes of Health and Iowa State University for their financial assistance.