SEQUENTIAL DECISION PROCEDURES FOR TESTING HYPOTHESES CONCERNING GENERAL ESTIMABLE PARAMETERS

By

Mahmoud Riad Mahmoud

Department of Biostatistics
The University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 881

AUGUST 1973
SEQUENTIAL DECISION PROCEDURES FOR TESTING HYPOTHESES
CONCERNING GENERAL ESTIMABLE PARAMETERS

by

Mahmoud Riad Mahmoud

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

Chapel Hill
1973

Approved by:

[Signatures]

Adviser

Reader

Reader
ABSTRACT

MAHMOUD RIAD MAHMOUD. Sequential Decision Procedures for Testing Hypotheses Concerning General Estimable Parameters. (Under the direction of DR. P. K. SEN.)

An invariance principle, due to Strassen, is extended here to continuous functions of several estimable parameters. A sequential procedure for discriminating between two close hypotheses concerning these functions is introduced. The invariance principle is used to justify the proposed procedure and to study its properties. The procedure is extended to discriminating among three hypotheses concerning the same functions is proposed and studied. A simulation study of an example is used to verify the proposed procedure.
ACKNOWLEDGMENTS

It is a pleasure to acknowledge my sincerest gratitude to
Professor P. K. Sen, my adviser, for proposing the problems in this
dissertation and giving me help in numerous ways. It was his patience
and encouragement that supported me all the time.

I wish also to thank Professor N. L. Johnson for going through
the manuscript and making many helpful suggestions. Thanks are also
due Dr. D. Quade and Dr. M. J. Symons for proofreading the manuscript.
I also thank Dr. R. P. Walker, the other member of my doctoral com-
mittee, and Dr. D. G. Kelley for sitting on my oral exam.

I wish also to thank Professor A. E. Sarhan of the Institute of
Statistical Studies and Research at Cairo University for his help in
many ways. It was through him that I was first introduced to Statistics
and he has been supporting me since.

I wish to acknowledge the Department of Statistics for financial
assistance during the first two years of my graduate program and the
Department of Biostatistics for financial assistance in the rest of my
graduate studies.

Finally, thankful appreciation is due Mrs. Gay Hinnant for her
excellent work in typing the manuscript.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>iv</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION AND REVIEW.</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Review of literature</td>
<td>2</td>
</tr>
<tr>
<td>1.3. Some definitions and previous results</td>
<td>5</td>
</tr>
<tr>
<td>1.4. The problem and results</td>
<td>9</td>
</tr>
<tr>
<td>II. ASYMPTOTIC PROPERTIES OF U-STATISTICS AND VON MISES' DIFFERENTIABLE STATISTICAL FUNCTIONS</td>
<td>12</td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>12</td>
</tr>
<tr>
<td>2.2. Notations</td>
<td>12</td>
</tr>
<tr>
<td>2.3. Weak convergence of U-statistics</td>
<td>14</td>
</tr>
<tr>
<td>2.4. Functions of several U-statistics</td>
<td>16</td>
</tr>
<tr>
<td>2.5. Estimation of $\gamma^2$</td>
<td>21</td>
</tr>
<tr>
<td>III. TESTING HYPOTHESES CONCERNING ESTIMABLE PARAMETERS</td>
<td>29</td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>29</td>
</tr>
<tr>
<td>3.2. Two hypotheses problem</td>
<td>30</td>
</tr>
<tr>
<td>3.3. OC function</td>
<td>32</td>
</tr>
<tr>
<td>3.4. ASN function</td>
<td>35</td>
</tr>
<tr>
<td>3.5. Three hypotheses problem</td>
<td>44</td>
</tr>
<tr>
<td>3.6. Probability of correct decision</td>
<td>45</td>
</tr>
<tr>
<td>3.7. Bounds for the ASN function</td>
<td>47</td>
</tr>
<tr>
<td>IV. EXAMPLES AND NUMERICAL INVESTIGATIONS</td>
<td>50</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>50</td>
</tr>
<tr>
<td>4.2. Correlation coefficients</td>
<td>50</td>
</tr>
<tr>
<td>4.3. Rank correlation</td>
<td>54</td>
</tr>
<tr>
<td>4.4. Comparison of two variances</td>
<td>55</td>
</tr>
<tr>
<td>4.5. Numerical results</td>
<td>56</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>65</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ASN (In Pairs) for the Correlation Coefficients $g_0 = -.1$ and $g_1 = .1$</td>
<td>58</td>
</tr>
<tr>
<td>2. ASN (In Pairs) for the Correlation Coefficients $g_0 = -.2$ and $g_1 = .2$</td>
<td>59</td>
</tr>
<tr>
<td>3. ASN (In Pairs) for the Correlation Coefficients $g_0 = -.1$</td>
<td>60</td>
</tr>
<tr>
<td>4. ASN (In Pairs) for the Correlation Coefficients $g_0 = -.2$</td>
<td>61</td>
</tr>
<tr>
<td>5. ASN (In Pairs) for the Correlation Coefficients $g_0 = .1$</td>
<td>62</td>
</tr>
<tr>
<td>6. ASN (In Pairs) for the Correlation Coefficients $g_0 = .2$</td>
<td>63</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION AND REVIEW

1.1. Introduction. Let $X$ be a random variable (r.v.) with distribution function (DF) $F(x)$. On the basis of a finite number of observations on $X$ we want to draw some inference concerning the (unknown) DF of $X$. Frequently we are interested in the values taken by some parameter $\theta(F)$ of $F(x)$. We shall be concerned with testing statistical hypotheses about the parameter $\theta = \theta(F)$. We are only interested in sequential procedures where the number of observations depends, in some specific way, on the data obtained in the course of study.

Suppose that $\theta$ takes values in some space $\Theta$, called the parameter space, and we want to test the hypothesis $H_0$: $\theta \in \omega_0$ against the alternative hypothesis $H_1$: $\theta \in \omega_1$, where $\omega_0$ and $\omega_1$ are subsets of $\Theta$. In sequential methods of testing statistical hypotheses, $H_0$ vs $H_1$, a rule is given for making one of the following decisions at any stage of the experiment: (1) to accept $H_0$ and terminate the procedure, (2) to accept $H_1$ and terminate the procedure, or (3) to continue the experiment by making an additional observation. With any sequential procedure for testing $H_0$ vs $H_1$ there will be associated two numbers $\alpha$ and $\beta$ between zero and one such that the probability is $\alpha$ that we reject $H_0$ when it is true and is $\beta$ that we accept $H_0$ when it is false.
We call the pair \((\alpha, \beta)\) the strength of the testing procedure.

Since the number of observations \(N\) at which we decide to stop sampling depends on the outcome of observations, it is a random variable. One particular property we desire in a testing procedure is that it will eventually terminate with probability one. If a testing procedure has this property we shall say that it "terminates." Also, two functions of the parameter \(\theta\) are of special interest. These are,

1. the operating characteristic (OC) function defined as the probability of accepting \(H_0\) when \(\theta\) is the true value of the parameter, and
2. the average sample number (ASN) function defined as the expected value of the number of observations \(N\) required by the procedure before stopping. A good discussion of the relevance of these functions is found in Wald (1947). Let \(S_1\) and \(S_2\) be two tests for testing \(H_0 vs H_1\) which have the same strength. If \(E_{S_1}(N; \theta) < E_{S_2}(N; \theta)\) for some \(\theta \in \Theta\), we say that \(S_1\) is more efficient than \(S_2\) at \(\theta\). We say that \(S^*\) is uniformly most efficient (UME) for testing \(H_0 vs H_1\) if

\[
E_{S^*}(N; \theta) = \inf_S E_S(N; \theta) \quad \text{for all } \theta \in \Theta,
\]

where the infimum is taken over all \(S\) of the same strength. Uniformly most efficient tests do not generally exist, and in such cases we confine ourselves to optimality within a class of tests.

1.2. Review of literature. Let the distribution of \(X\) be known except for the value taken by the unknown parameter \(\theta\). Wald (1947) gives a sequential test for testing \(H_0 vs H_1\) when both \(\omega_0\) and \(\omega_1\) contain only one point. The test is based on the log likelihood and is known as the sequential probability ratio test (SPRT). The test is rather simple, possesses some optimal properties and has a wide range of applicability.
However, in the majority of cases the DF involves unknown parameters other than the parameter under investigation. For example, if we are testing hypotheses about the mean of a normal population, the variance is usually unknown. On the other hand, the SPRT is designed for testing a simple hypothesis versus a simple alternative. So if the hypotheses do not specify completely all these parameters, the SPRT is not, at least directly, applicable. Furthermore, if the form of the DF is not known the SPRT is not at all applicable.

Wald (1947) suggested the use of some weight functions to integrate \( \theta \) over the ranges \( \omega_0 \) and \( \omega_1 \) or to integrate out the nuisance parameters. In principle, the choice of a weight function could be made to satisfy some kind of optimal properties, but there is no general method available for doing that. An alternative is to consider transforming the original sequence of observations; the transformation to be so chosen that the new sequence does not depend on the nuisance parameters. Then we apply the SPRT to the new sequence. Cox (1952) gave sufficient conditions for the existence of such transformation and gave an example on using it; see also Armitage (1947). Since the transformed observations are in general dependent, it is difficult to know what properties the OC and ASN functions have.

Cox (1963) developed a method due to Bartlett (1946) which follows from the asymptotic sufficiency of the maximum likelihood estimator (mle). For testing \( H_0: \theta = \theta_0 \) vs \( H_1: \theta = \theta_1 \) Cox (1963) suggested using the statistic \( T_n = n(\hat{\theta} - \frac{1}{2}(\theta_0 + \theta_1)) \), where \( \hat{\theta} \) is the mle of \( \theta \). It was assumed that \( \theta_0, \theta_1 \) and the value \( \theta \) of the parameter differ by amounts of order \( n^{-1/2} \). Although the validity of Cox's test depends on the convergence of \( T_n \) to a random walk, this result was rigorously
justified later by Breslow (1969). The statistic \( T_n \) is asymptotically a normal process with mean increment per unit time \( \theta - \frac{1}{2}(\theta_0 + \theta_1) \) and variance per step

\[
I^{\theta\theta} = (I^{\theta\theta} - I^{\theta\phi}/I^{\phi\phi})^{-1},
\]

where \( n I^{\theta\theta} = -\frac{\partial^2 L_n(x_n)}{\partial \theta^2} \), \( n I^{\theta\phi} = -\frac{\partial^2 L_n(x_n)}{\partial \theta \partial \phi} \) and

\( n I^{\phi\phi} = -\frac{\partial^2 L_n(x_n)}{\partial \phi^2} \); where \( L(x_n) \) is the likelihood function after \( n \) observations, and \( \phi \) is the nuisance parameter. Cox (1963) then suggested using Wald's SPRT for normally distributed observations as an approximate test for testing \( H_0 \) vs \( H_1 \).

In the methods mentioned above it was assumed that the DF \( F(x) \) of the r.v. \( X \) under consideration is known except for the value of some parameters \( (\theta, \phi) \). We shall be interested in cases where the parameter of interest is a measure of some characteristic of the DF without requiring the DF itself to be known. For example \( \theta \) may be the mean or the variance of the population.

So far we have been talking about testing \( H_0: \theta = \theta_0 \) vs \( H_1: \theta = \theta_1 \). But we may be interested in choosing one of three hypotheses about \( \theta \). Let \( a_1 \leq a_2 \) be two given numbers and suppose that we want to choose one of the three hypotheses

\[
(1.1) \quad H_1: \theta < a_1, \quad H_2: a_1 < \theta < a_2, \quad H_3: \theta > a_2.
\]

Sobel and Wald (1949) modified the problem a little by subdividing the parameter space into preference zones and indifference zones. Then they suggested using two SPRT's simultaneously. Armitage (1950)
suggested the use of three SPRT's simultaneously. Sobel and Wald (1949)
studied their procedure in detail when \( \theta \) is the mean of a normal pop-
ulation with known variances they gave an approximation to the OC func-
tion and some bounds for the ASN function. Also Simon (1967) studied
(1.1) when \( \theta \) is the mean of a normal population with known variance; he
used a Wiener process approximation for the discrete process \( S_n = \sum_{i=1}^{n} X_i \)
and treated it using a procedure similar to that of Armitage (1950).
This way he was able to get some better approximations to the OC and
ASN functions. However, as before, these procedures are applicable only
to cases where the DF is known except for the value of parameter \( \theta \).

1.3. Some definitions and previous results. Let \( X_1, X_2, \ldots \) be a
sequence of independent and identically distributed random variables
(iidrv) with each \( X_i \) having the same DF \( F(x), x \in \mathbb{R}^p \), the \( p(\geq 1) \)-
dimensional Euclidean space. Let \( F \) be a subset of the set of all DF's
in \( \mathbb{R}^p \). If to any \( F \in F \) a quantity \( \theta(F) \) is assigned then \( \theta(F) \) is called
a functional of \( F \) defined on \( F \). Suppose that for some sample size \( n \)
\( \theta \) admits an unbiased estimate for any DF \( F \in F \). That is, if
\((X_1, \ldots, X_n)\) is a random sample of size \( n \) there exists a function
\( f(x_1, \ldots, x_n) \) such that

\[
\theta(F) = \int_{\mathbb{R}^{np}} \ldots \int_{\mathbb{R}^{np}} f(x_1, \ldots, x_n) \, dF(x_1) \ldots dF(x_n),
\]
for every \( F \in F \). A functional \( \theta(F) \) of the form (1.2) is referred to
as regular over \( F \). A regular functional is also referred to as esti-
mable parameter.

Let \( m \geq 1 \) be the smallest sample size for which there exists an
unbiased estimate \( f(x_1, \ldots, x_m) \) of the regular functional \( \theta(F) \) over \( F \).
Then

\[(1.3) \quad \theta(F) = \int_{\mathbb{R}^{mp}} \ldots \int f(x_1, \ldots, x_m) \, dF(x_1) \ldots dF(x_m),\]

for every \( F \in \mathcal{F} \). We call \( m \) the degree over \( F \) of \( \theta(F) \). Any function \( f(x_1, \ldots, x_m) \) satisfying (1.3) is called a kernel of \( \theta(F) \). For any regular functional \( \theta(F) \) we can always find a kernel that is a symmetric function in its \( m \) arguments. Therefore, whenever we speak of a regular functional we shall assume that its kernel is symmetric. The simplest nontrivial example of a regular functional is the mean of a r.v. \( X \), where \( m = 1 \) and \( f(x) = x \).

Let \( X_1, \ldots, X_n \) be iid r.v.'s with DF \( F(x) \), \( x \in \mathbb{R}^p \). Let \( f(x_1, \ldots, x_m) \) be a symmetric function in its \( m \) arguments. A U-statistic is defined by

\[(1.4) \quad U_n = U_n(X_1, \ldots, X_n) = \binom{n}{m}^{-1} \sum_{c_{n,m}} f(x_{\alpha_1}, \ldots, x_{\alpha_m}),\]

where the summation is over all permutations \((\alpha_1, \ldots, \alpha_m)\) such that \( 1 \leq \alpha_1 < \ldots < \alpha_m \leq n \). The function \( f(x_1, \ldots, x_m) \) is referred to as the kernel of the U-statistic. If \( f(x_1, \ldots, x_m) \) satisfies (1.3) then \( U_n \) is the minimum-variance unbiased estimate of \( \theta(F) \).

For a random sample \((X_1, \ldots, X_n)\) of \( n \) iid r.v.'s with DF \( F(x) \), \( x \in \mathbb{R}^p \) the empirical DF is defined as

\[(1.5) \quad F_n(x) = n^{-1} \sum_{i=1}^{n} c(x - X_i), \quad x \in \mathbb{R}^p,\]

where \( c(u) \) is equal to one if all of its arguments are nonnegative and elsewhere it is equal to zero. The differentiable statistical function (DSF) of the regular functional \( \theta(F) \) is defined as
\begin{equation}
\theta(F_n) = \int_{\mathbb{R}^n} \cdots \int f(x_1, \ldots, x_n) \, dF_n(x_1) \cdots dF_n(x_n).
\end{equation}

Therefore, from (1.2), we have

\begin{equation}
\theta(F_n) = n^{-m} \sum_{\alpha_1 = 1}^{n} \cdots \sum_{\alpha_m = 1}^{n} f(x_{\alpha_1}, \ldots, x_{\alpha_m}).
\end{equation}

The U-statistics were introduced by Halmos (1946) and thoroughly studied by Hoeffding (1948). The DSP's were introduced by Von Mises (1947) and they are closely related to U-statistics. We may write

\begin{equation}
n^m \theta(F_n) = \binom{n}{m} U_n + \sum^{*} f(x_{\alpha_1}, \ldots, x_{\alpha_m}),
\end{equation}

where the summation $\sum^{*}$ extends over all m-tuples $(\alpha_1, \ldots, \alpha_m)$ in which at least one equality $\alpha_i = \alpha_j \quad i \neq j$ is satisfied. If the variance $\sigma^2(\theta(F_n))$ exists then $\theta(F_n) - U_n = \frac{1}{n} D_n$, where the expected value $\mathbb{E} D_n^2$ is bounded for $n$ tends to infinity. [See Hoeffding (1948)]. Then $\sqrt{n} \left[ \theta(F_n) - U_n \right]$ tends to zero in probability, as $n$ tends to infinity which implies that the limiting distribution of $\sqrt{n} \left[ \theta(F_n) - \theta(F) \right]$ is the same as that of $\sqrt{n} \left[ U_n - \theta(F) \right]$.

Let $\mathcal{C}_n$ be the $\sigma$-field generated by the unordered collection $(X_1, \ldots, X_n)$ together with $X_{n+1}$, $X_{n+2}$, $\ldots$, so that $\mathcal{C}_n$ is non-increasing in $n \geq 1$. The following lemma was proved by Berk (1966).

**Lemma 1.1.** If $\mathbb{E} \left| f(x_1, \ldots, x_m) \right| < \infty$, then $\{U_n, \mathcal{C}_n, n > m\}$ is a reverse martingale.

The following theorem was proved by Sen (1960) and Hoeffding (1961) by considering the representation of a U-statistic by a linear combination of uncorrelated random variables. In view of lemma 1.1 it was proved again by Berk (1966) using the martingale convergence theorem (cf Doob 1953).
Theorem 2.2. If \( E \{ |f(X_1, \ldots, X_m)| \} < \infty \) then

\[
U_n(X_1, \ldots, X_n) \to \theta(F)
\]

almost surely (a.s.) as \( n \to \infty \).

Sproule (1969) has shown that for \( \lambda > \frac{1}{2} \), \( (U_n - \theta(F)) = 0(n^{-1/2} (\log n)^\lambda) \) a.s. as \( n \to \infty \) and this result was improved upon by Ghosh and Sen (1970) who showed that \( (U_n - \theta(F)) = 0(n^{-1/2} (\log \log n)^{1/2}) \) a.s. as \( n \to \infty \). They also showed that if

\[
\zeta^* = \max_{1 < \alpha_1 < \ldots < \alpha_m < m} E[f(x_{\alpha_1}, \ldots, x_{\alpha_m})]^2 < \infty
\]

then

\[(1.9) \quad [\theta(F_n) - \theta(F)] = 0(n^{-1/2} (\log \log \log n)^{1/2}) \text{ a.s.} \]
as \( n \to \infty \). Miller and Sen (1972) proved that under the same condition for every \( \varepsilon > 0 \)

\[(1.10) \quad \lim_{n \to \infty} P\left( \max_{m < k < n} k \mid \theta(F_m) - U_n \mid > \varepsilon n^{1/2} \right) = 0, \]
as \( n \to \infty \). Grams and Serfling (1973) proved the following results:

(i) If \( f^{2r} < \infty, r \geq 1 \), then for every \( \varepsilon > 0 \)

\[(1.11) \quad P\{ \sup_{k > n} |U_k - \theta(F)| > \varepsilon \} = 0(n^{1-2r}). \]

(ii) If \( E\{ |f(x_{\alpha_1}, \ldots, x_{\alpha_m})|^r \} < A < \infty \) for all \( 1 \leq \alpha_1 \leq \ldots \leq \alpha_m \leq m \) and \( r \) a positive integer, then

\[(1.12) \quad E\{ |U_n - \theta(F_n)|^r \} = 0(n^{-r}). \]

In fact (ii) is also true for any \( r \geq 1 \).

These results show how closely related the U-statistics and
DSF's are. Therefore, the convergence of one can be obtained from the convergence of the other with the aid of the above results. With this in mind we shall develop our results in the following chapters for U-statistics where similar results for DSF's may be easily obtained if we replace conditions on the moments of $f(X_1, \ldots, X_m)$ by similar ones on $f(X_{\alpha_1}, \ldots, X_{\alpha_m}), 1 \leq \alpha_1 \leq \ldots \leq \alpha_m \leq n$.


1.4. The problem and results. Suppose that we want to test (1.1). If the analytic form of the DF of the random variable under consideration is not known, we cannot apply any of the methods mentioned above. We introduce a statistic similar to that of Cox (1963) and notice that Cox's statistic $T_n = n(\hat{\theta} - \frac{1}{2}(\theta_0 + \theta_1))$ is a measure of the discrepancy between the mle $\hat{\theta}$ of $\theta$ and the midpoint $\frac{1}{2}(\theta_0 + \theta_1)$, of the hypotheses $H_0$: $\theta = \theta_0$ and $H_1$: $\theta = \theta_1$. We define the statistics

$$M_n = n[\bar{\theta} - \frac{1}{2}(\theta_0 + \theta_1)],$$

and

$$M^*_n = n[\theta(F_n) - \frac{1}{2}(\theta_0 + \theta_1)],$$

where $\bar{\theta}$ is the U-statistic estimate of $\theta$ and $\theta(F_n)$ is the DSF corresponding to 0. As in $T_n$ the exact distributions of $M_n$ and $M^*_n$ are not known and we use their asymptotic distributions to develop a sequential
procedure for testing hypotheses about $\theta$. As we mentioned, the asymptotic distributions of $\bar{\theta}$ and $\theta(F_n)$ have been developed [cf. Miller and Sen (1972)]. But here we are more interested in testing hypotheses about continuous functions of one or several estimable parameters.

Let $g(\bar{\theta}) = g(\theta^{(1)}, \ldots, \theta^{(q)})$ be a continuous function in the regular functionals $\theta^{(1)}, \ldots, \theta^{(q)}$, $q \geq 1$. For example $g(\bar{\theta})$ may be the variance ratio $\sigma_1^2/\sigma_2^2$ of two random variables. If the function $g(\bar{\theta})$ is estimable then we have the same situation as in the previous paragraph. We are interested in cases where $g(\bar{\theta})$ may not be estimable. That is, there may not exist a function $f(x_1, \ldots, x_n)$ such that $g(\bar{\theta}) = Ef(X_1, \ldots, X_n)$. Now, we want to choose between the three hypotheses

\[(1.13) \quad H_1: \quad g(\bar{\theta}) < a_1, \quad H_0: \quad a_1 \leq g(\bar{\theta}) \leq a_2, \quad H_1: \quad g(\bar{\theta}) > a_2\]

where $a_1$ and $a_2$ are given.

In Chapter II we study the asymptotic distribution of the random variables $g(U_n^{(1)}, \ldots, U_n^{(q)})$ as $n \to \infty$, where $U_n^{(i)}$ is the U-statistic estimate of $\theta^{(i)}$, $i = 1, \ldots, q$. We prove that under some regularity conditions $g(U_n) = g(U_n^{(1)}, \ldots, U_n^{(q)})$ converges to a Wiener process with variance per unit step $\gamma^2$. Then we find a consistent estimate of $\gamma^2$. The invariance principle we prove for $g(U_n)$ is similar to that of Strassen (1967). We also show that same principle holds for $g(\bar{\theta}(F_n)) = g(\theta^{(1)}(F_n), \ldots, \theta^{(q)}(F_n))$, where $\theta^{(i)}(F_n)$ is the DSF corresponding to $\theta^{(i)}$, $i = 1, \ldots, q$.

It is to be pointed out here that we are not interested in transforming a problem from testing hypotheses about some parameter to testing corresponding hypotheses about a function of that parameter.
For one thing the problem is not invariant, in general, under continuous transformations. However, we are interested in testing hypotheses about a function that is of interest in itself, but the function is not estimable and it is a function of several estimable parameters.

In Chapter III we develop a sequential procedure similar to that of Sobel and Wald for the problem (1.13). We study the operating characteristic of the procedure and obtain a simple formula for some limiting process of the OC function. We also give some rough bounds for the ASN function.

In Chapter IV we give some examples of the proposed test and study, numerically, the ASN function of the test.
CHAPTER II

ASYMPTOTIC PROPERTIES OF U-STATISTICS AND VON MISES' DIFFERENTIABLE STATISTICAL FUNCTIONS

2.1. Introduction. As mentioned in Chapter I our main purpose is to develop sequential procedures for testing more than two hypotheses. We proposed to use the U-statistic estimates of some regular functionals as a basis for our test. For the development of the test we need certain mathematical tools. These include: (1) the asymptotic distribution theory of continuous functions of several U-statistics, (2) almost sure convergence of estimates of the variances of these functions, and (3) almost sure convergence of sequences of these functions to Wiener processes. These results will be studied in this chapter.

2.2. Notations. Let $X_1, X_2, \ldots, X_n$ be n iid r.v.'s having DF $F(x)$, $x \in \mathbb{R}^p$, $p \geq 1$. Let $f(x_1, \ldots, x_m)$ be symmetric in its arguments. Let $\theta(F)$ satisfy (1.3) and define $U_n$ as in (1.4). For $c = 1, 2, \ldots, m$ define

\begin{equation}
(2.1) \quad f_c(x_1, \ldots, x_c) = E f(x_1, \ldots, x_c, X_{c+1}, \ldots, X_m).
\end{equation}

Suppose that $E f^2(X_1, \ldots, X_m) < \infty$ and define

\begin{equation}
(2.2) \quad \zeta_0 = \zeta_0(F) = 0, \quad \zeta_c = \zeta_c(F) = E f_c^2(X_1, \ldots, X_c) - \theta^2.
\end{equation}
If for some \( F \in F = \{ F : |\theta(F)| < \infty \} \)

\begin{equation}
0 < \zeta_1(F) < \infty
\end{equation}

\( \theta(F) \) is said to be stationary of order zero at \( F \). If \( U_n^{(1)} \) and \( U_n^{(2)} \) are two \( U \)-statistics with kernels \( f^{(1)}(x_1, \ldots, x_{m_1}) \) and \( f^{(2)}(x_1, \ldots, x_{m_2}) \) define

\begin{equation}
\zeta^{(1,2)}_c = E \{ f^{(1)}_c(X_1, \ldots, X_c) f^{(2)}_c(X_1, \ldots, X_c) \},
\end{equation}

c = 1, 2, \ldots, \min(m_1, m_2).

For every \( h(1 \leq h \leq m) \) define

\begin{equation}
V_{n,h} = \int \cdots \int_{R^m} f_h(x_1, \ldots, x_h) \prod_{j=1}^m d[F_n(x_j) - F(x_j)],
\end{equation}

so that

\begin{equation}
\frac{n}{V_{n,1}} = \sum_{j=1}^n \{ f_1(X_j) - \theta(F) \}.
\end{equation}

Write

\begin{equation}
U_{n,1} = V_{n,1},
\end{equation}

and for \( 2 \leq h \leq m \) define

\begin{equation}
U_{n,h} = n^{-[h]} \sum_{p} \int \cdots \int_{R^m} f_h(x_1, \ldots, x_h) \prod_{j=1}^m d[c(x_j - x_{i_j}) - F(x_j)]
\end{equation}

where \( P_{n,h} = \{1 \leq i_1, \ldots, i_h \leq n\} \) and \( n^{-[h]} = [n \ldots (n - h + 1)]^{-1} \).

Then we can write

\begin{equation}
\theta(F_n) - \theta(F) = \sum_{h=1}^m \binom{m}{h} V_{n,h},
\end{equation}

and

\begin{equation}
U_n - \theta(F) = \sum_{h=1}^m \binom{m}{h} U_{n,h} + R_n,
\end{equation}

where \( R_n = m U_{n,1} + R_n \).
where $R_n = \sum_{h=2}^{m} \binom{m}{h} U_{n,h}$. The decomposition (2.10) was given by Hoeffding (1961) and he showed that $\{ \binom{n}{h} U_{n,h} \ n \geq m \}$ satisfies the martingale property. Miller and Sen (1972) showed that $\{ U_{n,h}, \ z_n \ n \geq m \}$ is a reverse martingale. For $h = 1, ..., m$ define

$$
\delta_h = \sum_{c=0}^{h-1} (-1)^c \binom{h}{c} \zeta_{h-c}.
$$

Sproule (1969) proved that if $E[f(X_1, ..., X_m)]^2 < \infty$ then

$$
\text{Var}(U_{n,h}) = \binom{n}{h}^{-1} \delta_h = o(n^{-h}), \quad h = 1, ..., m,
$$

and for $n_1 > n_2 \geq m$

$$
\text{Cov}(U_{n_1,h}, U_{n_2,h'}) = \begin{cases} 
\text{Var}(U_{n_1,h}) & h = h' = 1, ..., m \\
0 & h \neq h' = 1, ..., m 
\end{cases}
$$

2.3. **Weak convergence of U-statistics.** Let $S$ be a metric space and $L$ be the $\sigma$-field generated by all open sets of $S$. Let $P_n$ and $P$ be probability measures on $L$. If

$$
\int_S y dP_n \rightarrow \int_S y dP
$$

for every bounded, continuous real function $y$ on $S$, we say that $P_n$ converges weakly to $P$ and we write $P_n \Rightarrow P$. This concept of convergence is studied in detail by Billingsley (1968). On $\mathbb{R}^p$ consider the metric $\rho(x,y) = |x-y| = \left( \sum_{i=1}^{p} (x_i - y_i)^2 \right)^{1/2}$, and let $\mathcal{B}^p$ be the $p$-dimensional Borel sets in $\mathbb{R}^p$. On $\mathbb{R}^p$ weak convergence of probability measures and convergence of distribution functions are equivalent. But on general metric spaces this is not necessarily true.

Consider the space $C[0,1]$ of all continuous functions on
I = [0,1], and associate with it the uniform topology

\[ \rho(X(\cdot), Y(\cdot)) = \sup_{t \in I} |X(t) - Y(t)|, \]

where both X and Y belong to C[0,1]. Loynes (1970) obtained an invariance principle for reverse martingales and cited without proof U-statistics as an example to show that a process obtained by linear interpolation from \( \{n^{1/2}[U_k - \theta(F)]; k \geq n\} \) converges weakly to a Wiener process in C[0,1]. Miller and Sen (1972) have shown that under the same set of conditions, processes obtained by linear interpolation from \( \{n^{-1/2} k[U_k - \theta(F)], m \leq k \leq n\} \) converge weakly in C[0,1] to a Wiener process.

Let \( S = \{S(t) : 0 \leq t < \infty\} \) be a random process, where

\[
S(k) = S_k = \begin{cases} 
0, & 0 \leq k \leq m - 1 \\
[k[U_k - \theta(F)], & k \geq m, \end{cases}
\]

and \( S(t) = S_k \) for \( k \leq t < k + 1, k \geq 0 \). Consider a positive and real valued function \( y(t), t \in [0,\infty) \), such that

\[
y(t) \text{ is } + \text{ but } t^{-1} y(t) \text{ is } + \text{ in } t : 0 \leq t < \infty; \tag{2.16}
\]

and

\[
\sum_{n \geq 1} [y(cn)]^{-1} E[(f_1^{*}(X_n))^2 I((f_1^{*}(X_n))^2 > y(cn))] < \infty \tag{2.17}
\]

for every \( c > 0 \), where \( f_1^{*}(x) = f_1(x) - \theta(F) \) and \( I(A) \) denotes the indicator function of a set \( A \). Let \( \{\xi(t) : 0 \leq t < \infty\} \) be a standard Wiener process. Sen (1972) proved the following extension of Strassen's (1967) first a.s. invariance principle.

**Theorem 2.1.** If \( \theta(F) \) is stationary of order zero and \( \xi_m(F) < \infty \), then under (2.16) and (2.17), as \( t \to \infty \)
(2.18) \[ S(t) = \gamma \xi(t) + o(t) y(t)^{1/4} \log t \text{ a.s.,} \]

where \( \gamma^2 = m^2 \xi_1(F) (> 0 \text{ by stationarity of } \theta(F) \text{ of order 0}). \)

2.4. **Functions of several U-statistics.** Consider a function
g(x) = g(x^{(1)}, \ldots, x^{(2)}), q \geq 1, \text{ that does not depend on } n \text{ and in some}
nearhood of } \theta = (\theta^{(1)}, \ldots, \theta^{(q)}) \text{ is continuous and has continuous}
first and second order partial derivatives with respect to iidrv's with
\( x^{(1)}, \ldots, x^{(q)}. \) For \( i = 1, \ldots, q \) let \( \theta^{(i)} = \theta^{(i)}(F) \) be a regular
functional of degree \( m_i \) over \( F = \{ F : |\theta^{(i)}(F)| < \infty, i = 1, \ldots, q \}. \)
That is, for each \( i = 1, \ldots, q, \) if \( X_1, \ldots, X_{m_i} \) are iidrv's with
DF \( F \in F, \) then there exists a symmetric function \( f^{(i)}(x_1, \ldots, x_{m_i}) \)
such that

\[(2.19) \quad \theta^{(i)} = \theta^{(i)}(F) = \int_{\mathbb{R}^{m_i}} \cdots \int_{\mathbb{R}^{m_i}} f^{(i)}(x_1, \ldots, x_{m_i}) \prod_{j=1}^{m_i} dF(x_j). \]

For a sample \( (X_1, \ldots, X_n) \) \( n \geq m^* = \max(m_1, \ldots, m_q), \) let

\[(2.20) \quad U^{(i)}_n = U^{(i)}_n(X_1, \ldots, X_n) = \binom{n-1}{m_i} \sum_{C_{n,m_i}} f^{(i)}(X_{\alpha_1}, \ldots, X_{\alpha_{m_i}}), \]

\[ C_{n,m_i} = \{1 \leq \alpha_1 < \ldots < \alpha_{m_i} \leq n, 1 \leq i \leq q\}. \]

We define \( f^{(i)}_c(x_1, \ldots, x_c), \xi^{(i)}_1 \) and \( \xi^{(i,j)}_1 \) similar to (2.1), (2.2)
and (2.4) respectively. We assume that

\[(2.21) \quad 0 < \xi^{(i)}_1(F) < \infty, \quad \text{for all } F \in F, \text{ and } i = 1, \ldots, q. \]

\[(2.22) \quad \xi^{(i)}_{m_i}(F) < \infty \quad i = 1, \ldots, q. \]

Write \( U_n = (U_n^{(1)}, \ldots, U_n^{(q)}) \) and \( \theta = (\theta^{(1)}(F), \ldots, \theta^{(q)}(F)). \)
By almost sure convergence of a U-statistic to its expected value and (1.11) we can find \( n_0 \) large enough such that for \( n \geq n_0 \), \( U_n \) lies in the neighborhood of \( \theta(F) \) where \( g \) is continuous and has first and second order partial derivatives. We shall write \( \frac{\partial g(\theta)}{\partial x(i)} \) to mean \( \frac{\partial g(x)}{\partial x(i)} \bigg|_{x=\theta} \).

Then for \( n \geq n_0 \) we can expand \( g(U_n) \) about \( \theta \) to get

\[
(2.23) \quad g(U_n) - g(\theta) = \sum_{i=1}^{q} \left( U_n^{(i)} - \theta^{(i)} \right) \frac{\partial g(\theta)}{\partial x(i)} + \frac{1}{2} \sum_{i=1}^{q} \sum_{j=1}^{q} \left( U_n^{(i)} - \theta^{(i)} \right) \left( U_n^{(j)} - \theta^{(j)} \right) \frac{\partial^2 g(\theta')}{\partial x(i) \partial x(j)}
\]

where \( \theta' \) is some point in the neighborhood of \( \theta \) referred to above.

Then, by (2.10) we can write

\[
(2.24) \quad g(U_n) - g(\theta) = \sum_{i=1}^{q} m_i \frac{\partial g(\theta)}{\partial x(i)} U_{n,i} + \sum_{i=1}^{q} \sum_{j=1}^{q} \frac{m_i}{m_j} \frac{\partial g(\theta)}{\partial x(i)} U_{n,j} + \frac{1}{2} \sum_{i=1}^{q} \sum_{j=1}^{q} \left( U_n^{(i)} - \theta^{(i)} \right) \left( U_n^{(j)} - \theta^{(j)} \right) \frac{\partial^2 g(\theta')}{\partial x(i) \partial x(j)}
\]

Assume that there exists \( \varepsilon > 0 \) such that for every \( x \) such that \( |x - \theta| < \varepsilon \) we have

\[
(2.25) \quad \left| \frac{\partial g(x)}{\partial x(i)} \right| < M_1 < \infty \quad i = 1, \ldots, q,
\]

and

\[
(2.26) \quad \left| \frac{\partial^2 g(x)}{\partial x(i) \partial x(j)} \right| < M_2 < \infty \quad i = 1, \ldots, q, j = 1, \ldots, q,
\]

By (2.26) we have for all \( i, j = 1, \ldots, q \),
\[(2.27) \quad \left| \sum_{i=1}^{q} \sum_{j=1}^{q} (U_n^{(i)} - \theta^{(i)}) (U_n^{(j)} - \theta^{(j)}) \frac{\partial^2 g(\theta')}{\partial x^{(i)} \partial x^{(j)}} \right| \leq M_2 \left( \sum_{i=1}^{q} (U_n^{(i)} - \theta^{(i)}) \right)^2 \leq M_2 \left[ \sum_{i=1}^{q} |U_n^{(i)} - \theta^{(i)}|^2 \right] \]

Ghosh and Sen (1970) have shown that under (2.21) and (2.22) we have for \( i = 1, \ldots, q \).

\[(2.28) \quad |U_n^{(i)} - \theta^{(i)}| = O(n^{-1/2} (\log \log n)^{1/2}) \quad \text{a.s. as } n \to \infty. \]

Thus, as \( n \to \infty \)

\[(2.29) \quad \left| \sum_{i=1}^{q} \sum_{j=1}^{q} (U_n^{(i)} - \theta^{(i)}) (U_n^{(j)} - \theta^{(j)}) \frac{\partial^2 g(\theta)}{\partial x^{(i)} \partial x^{(j)}} \right| = O(n^{-1} \log \log n) \quad \text{a.s.} \]

Therefore, we have just proved the following lemma.

**Lemma 2.2.** If \( \theta^{(i)}(F) \), \( i = 1, \ldots, q \) are stationary of order zero and \( \xi_{m}^{(i)} < \infty \), \( i = 1, \ldots, q \), then under (2.26) we have as \( n \to \infty \)

\[(2.30) \quad g(U_n) - g(\bar{\theta}) = \sum_{i=1}^{q} m_i \frac{\partial g(\bar{\theta})}{\partial x^{(i)}} U_{n,1} + \sum_{i=1}^{q} \sum_{h=1}^{m_i} \frac{\partial g(\bar{\theta})}{\partial x^{(i)}} U_{n,h} + O(n^{-1} \log \log n) \quad \text{a.s.} \]

Now, recall the definition of the \( \sigma \)-fields \( \mathcal{C}_n \) cited in Section 1.3. The following lemma is due to Miller and Sen (1972).

**Lemma 2.3.** For every \( h \) \( (1 \leq h \leq m_i) \) \( \{ U_{n,h}^{(i)}, \mathcal{C}_n, n \geq h \} \) forms a reversed martingale sequence. This implies that
\{U_n^{(i)} = \sum_{h=2}^{m_i} \left( \frac{\partial g(\theta)}{\partial x(1)} \right)_{n,h} U_n^{(i)}, C_n, n \geq m_i \} forms a reversed martingale sequence. Let c_k be a sequence of positive numbers such that c_k is + in k and \( c_k^2 = 0(c_k^{3/2}) \). We prove the following.

**Lemma 2.4.** If \( E[f^{(i)}(X_1, \ldots, X_{m_c})] < \infty \), then for \( i = 1, \ldots, q \),

\[
\max_{k \geq n} c_k \left| \frac{U_n^{(i)}}{U_{n+k}^{(i)}} \right| \rightarrow 0 \text{ in probability as } n \rightarrow \infty.
\]

**Proof:** By (2.21) and (2.22) we have

\[
E[U_n^{(i)}]^2 = \sum_{h=2}^{m_i} \binom{m_i}{h} 2 \left( \frac{\partial g(\theta)}{\partial x(1)} \right) \binom{n-1}{h} \delta_h = o(n^{-2}),
\]

and

\[
E[U_n^{(i)} - U_{n+1}^{(i)}]^2 = \sum_{h=2}^{m_i} \binom{m_i}{h} 2 \left( \frac{\partial g(\theta)}{\partial x(1)} \right) \binom{n-1}{h} \left[ \delta_h - \binom{n+1-1}{h} \delta_h \right]
\]

\[
= \sum_{h=2}^{m_i} \binom{m_i}{h} 2 \left( \frac{\partial g(\theta)}{\partial x(1)} \right) \binom{n-1}{h} \left[ 1 - \frac{n+1-h}{n+1} \right]
\]

\[
= \sum_{h=2}^{m_i} \binom{m_i}{h} 2 \left( \frac{\partial g(\theta)}{\partial x(1)} \right) \frac{n}{n+1} \frac{n-1}{h} \delta_h
\]

\[
= o(n^{-3}).
\]

Then

\[
\lim_{n \rightarrow \infty} C_n^2 E[U_{n+k}^{(i)}]^2 = 0.
\]

If we now apply Theorem 1 of Chow (1960) [i.e. the Hájek-Rényi inequality for submartingales] we obtain for every \( \epsilon > 0 \)
\[ P_{k \geq n} \max \mathbb{C}_k |U_k^{(i)}| > \epsilon \leq \epsilon^{-2} \left\{ \sum_{k=n}^{\infty} c_k^2 E[U_k^{(i)} - U_{k+1}^{(i)}] \right\} \]

\[ = \epsilon^{-2} \sum_{k=n}^{\infty} 0(n^{-3/2}) \]

\[ = \epsilon^{-2} 0(n^{-1/2}) + 0 \quad \text{as } n \to \infty, \]

and the lemma follows.

From (2.30) and by lemma 2.4 we can write that

\[ g(U_n) - g(\Theta) = \sum_{i=1}^{q} \sum_{j=1}^{m_j} \frac{\partial g(\Theta)}{\partial x(i)} U_n^{(i)} + O(C_n^{-1}) + O\left(\frac{\log \log n}{n}\right) \]

\[ = \sum_{i=1}^{q} \sum_{j=1}^{m_j} \frac{\partial g(\Theta)}{\partial x(i)} U_n^{(i)} + O(C_n^{-1}), \quad \text{as } n \to \infty. \]

Write \( G(X_j) = \sum_{i=1}^{q} \sum_{j=1}^{m_j} \frac{\partial g(\Theta)}{\partial x(i)} f_1^{(i)}(X_j) - \Theta^{(i)}(x), \quad j = 1, \ldots, n. \)

Then \( \{G(X_j)\} \) are i.i.d r.v's with \( E G(X_j) = 0 \) and

\[ \gamma^2 = E[G(X_j)]^2 = \sum_{i=1}^{q} \sum_{j=1}^{m_j} \frac{\partial g(\Theta)}{\partial x(i)} \frac{\partial g(\Theta)}{\partial x(i)} \xi_{i,j}, \quad j = 1, \ldots, n. \]

Thus, as \( n \to \infty \)

\[ n [g(U_n) - g(\Theta)] = \sum_{j=1}^{n} G(X_j) + o(n C_n^{-1}). \]

We now prove a theorem similar to Theorem 2.1. Consider a positive and real valued function \( y(t), \quad t \in [0, \infty) \) such that

\[ y(t) \text{ is } \dagger \text{ but } t^{-1} y(t) \text{ is } \dagger \text{ in } t : 0 \leq t < \infty, \]

and

\[ \sum_{n=1}^{\infty} [y(cn)] E[\{G(X_1)\}^2 I\{G(X_1) > y(cn)\}] < \infty \]
for every $\varepsilon > 0$. Define a process $S(t) = \{S(t) : 0 \leq t < \infty\}$ by

\[
S(k) = \begin{cases} 0 & , 0 \leq k \leq m-1 = \max(m_1, \ldots, m_q) - 1 \\ k \left[ \sum_{j=1}^{k} G(X_j) \right] & k \geq m,
\end{cases}
\]

and $S(t) = S(k)$ for $k \leq t < k + 1$, $k \geq 0$.

**Theorem 2.5.** If $\theta^{(i)}(F)$, $i = 1, \ldots, q$ are stationary of order zero and $\xi_m^{(i)}(F) < \infty$, $i = 1, \ldots, q$, then under (2.35) and (2.36) as $n \to \infty$

\[
S(t) = \gamma \xi(t) + o((t \gamma(t))^{1/4} \log t) \text{ a.s.},
\]

where $\xi(t)$ is a standard Wiener process and $\gamma$ is defined in (2.33).

**Proof:** Take $c_n = n[(ny(n))^{1/4} \log n]^{-1}$, then from (2.34)

\[
n[g(U_n) - g(\theta)] = n \sum_{j=1}^{n} G_1(X_j) + o((ny(n))^{1/4} \log n).
\]

Since $G(X_j)$ and iid r.v.'s the theorem follows directly from Theorem 4.4 of Strassen (1967).

Now, if we choose $y(t) = t(\log \log t)^2/(\log t)^4$ and (2.36) holds, then we have

\[
S(t) = \gamma \xi(t) + O\left(t^{1/2} \sqrt{\log \log t}\right) \text{ a.s. as } n \to \infty.
\]

If $y(t) = t \left( \frac{\log \log t}{\log t} \right)^4$ and (2.38) holds, then

\[
S(t) = \gamma \xi(t) + O(t^{1/2}) \quad \text{a.s. as } t \to \infty.
\]

2.5. **Estimation of $\gamma^2$.** In Cox's (1963) formulation he substitutes the mle of $\theta$ in the quantities $I_{\theta\theta}$, $I_{\theta\phi}$ and $I_{\phi\phi}$. To use (2.38) we need to find an estimate for

\[
\gamma^2 = \sum_{i=1}^{q} \sum_{j=1}^{q} m_i m_j \frac{\partial g(\theta)}{\partial x(i)} \frac{\partial g(\theta)}{\partial x(i)} \xi^{(i,j)}.
\]
Since \( \gamma^2 \) depends on \( \theta \) we substitute for \( \theta \) its U-statistic estimate \( U_n \).

We can estimate \( \xi_1^{(1,j)} \) by its U-statistic but Sproule (1969) has shown that the U-statistic estimate for \( \xi_1^{(1,j)} \) is hard to compute. Sen (1960) decomposed a U-statistic as follows:

\[
(2.41) \quad U_n(X_1, \ldots, X_n) = \frac{1}{n} \sum_{\alpha=1}^{n} V_\alpha(X_1, \ldots, X_n),
\]

where \( V_\alpha \) is defined by

\[
(2.42) \quad V_\alpha = \left( \frac{n-1}{m-1} \right)^{-1} \sum_{\alpha_2}^{\alpha_m} f(X_{\alpha_2}, \ldots, X_{\alpha_m}), \quad \alpha = 1, \ldots, n.
\]

and the summation extends over \( 1 \leq \alpha_2 < \ldots < \alpha_m \leq n \) but \( \alpha_j \neq \alpha \) for \( j = 2, \ldots, m \). He utilized the decomposition to obtain a consistent estimate of \( \xi_1 \). We use a similar approach to obtain consistent estimates of \( \xi_1^{(i,j)} \). For \( i = 1, \ldots, q \) we define for \( \alpha = 1, \ldots, n \)

\[
(2.43) \quad V_\alpha^{(i)} = \left( \frac{n-1}{m_i-1} \right)^{-1} \sum_{\alpha_2}^{\alpha_m} f^{(i)}(X_{\alpha_2}, \ldots, X_{\alpha_m}).
\]

where the summation extends over \( 1 \leq \alpha_2 < \ldots < \alpha_m \leq n \) but \( \alpha_j \neq \alpha \) for \( j = 2, \ldots, m_i \). Thus \( U_n^{(i)} = \frac{1}{n} \sum_{\alpha=1}^{n} V_\alpha^{(i)}, \quad i = 1, \ldots, q \). Define

\[
(2.44) \quad \delta_{ij} = \frac{1}{n} (V_\alpha^{(i)} - U_n^{(i)})(V_\alpha^{(j)} - U_n^{(j)}).
\]

We propose to use

\[
(2.45) \quad \gamma_n^2 = \sum_{i=1}^{q} \sum_{j=1}^{q} \delta_{ij} \frac{\partial g(U_n)}{\partial x^{(i)}} \frac{\partial g(U_n)}{\partial x^{(j)}}
\]

as an estimate for \( \gamma^2 \).

Now, by continuity of \( \frac{\partial g(\theta)}{\partial x^{(i)}} \), \( i = 1, \ldots, q \), for every \( \varepsilon > 0 \) there exists \( \delta_1 > 0 \) such that
\[
\left| \frac{\partial g(U_n)}{\partial x(i)} - \frac{\partial g(\theta)}{\partial x(i)} \right| \leq \varepsilon \text{ whenever } \|U_n - \theta\| \leq \delta,
\]
where \( \|U_n - \theta\| = \max_{1 \leq i \leq q} |U_n^{(i)} - \theta^{(i)}| \). Therefore, for every \( \varepsilon > 0 \) there exist \( \delta_1 > 0 \) and \( \delta_2 > 0 \) such that \( |\gamma_n^2 - \gamma^2| \leq \varepsilon \) whenever
\[
|U_n - \theta| \leq \delta_1 \text{ and } ||\delta_{ij} - \xi_{1}^{(i,j)}|| \leq \delta_2,
\]
where \( ||\delta_{ij} - \xi_{1}^{(i,j)}|| = \max_{1 \leq i, j \leq q} |\delta_{ij} - \xi_{1}^{(i,j)}| \). Thus
\[
(2.46) \quad \mathbb{P}\{|\gamma_n^2 - \gamma^2| \geq \varepsilon\} \leq \mathbb{P}\{||\delta_{ij} - \xi_{1}^{(i,j)}|| \geq \delta_1\} + \mathbb{P}\{|U_n - \theta| \geq \delta_2\}.
\]

The following theorem was proved by Grams and Serfling (1973).

**Theorem 2.6.** Let \( \{U_n\} \) be the sequence of U-statistics generated by a kernel \( f \) applied to a sequence of observations \( \{X_i\} \). Assume that \( E f^{2r} < \infty \) for a positive integer \( r \). Then for every \( \varepsilon > 0 \),
\[
(2.47) \quad \mathbb{P}\{\sup_{k \geq n} |U_k - \theta| > \varepsilon\} = O(n^{1-2r}).
\]
It can be easily shown [by Theorem 2.1 of Grams and Serfling (1973)] that if \( E f^{2r} < \infty \) then
\[
(2.48) \quad E(U_n - \theta)^{2r} = O(n^{-r}).
\]
We now utilize (2.47) and (2.48) to prove the following.

**Theorem 2.7.** Assume that for \( i = 1, \ldots, q \) \( E[f^{(i)}(X_1, \ldots, X_{2r})] \) is finite for some \( r > 1 \). Then for every \( \varepsilon > 0 \)
\[
(2.49) \quad \mathbb{P}\{|\delta_{ij} - \xi_{1}^{(i,j)}| > \varepsilon\} = O(n^{-r/2}).
\]
Proof: Let \( \xi^{(i,j)} = \mathbb{E} f^{(i)}(X_{\alpha}) f^{(j)}(X_{\alpha}) \) for any \( 1 \leq \alpha \leq n \).

Then \( \xi_1^{(i,j)} = \xi^{(i,j)} - \theta^{(i)}\theta^{(j)} \) and

\[
\delta_{ij} - \xi_1^{(i,j)} = (\frac{1}{n} \sum_{\alpha=1}^{n} \mathcal{V}_{\alpha}^{(i)} \mathcal{V}_{\alpha}^{(j)} - \xi_1^{(i,j)}) - (u_n^{(i)} u_n^{(j)} - \theta^{(i)}\theta^{(j)}).
\]

Therefore

\[
P(|\delta_{ij} - \xi_1^{(i,j)}| > \varepsilon) \leq P\left(\frac{1}{n} \sum_{\alpha=1}^{n} \mathcal{V}_{\alpha}^{(i)} \mathcal{V}_{\alpha}^{(j)} - \xi_1^{(i,j)} \right) \leq \varepsilon/2\) +

\[
P\left(\left|u_n^{(i)} u_n^{(j)} - \theta^{(i)}\theta^{(j)}\right| > \varepsilon/2\right)
\]

If \( \theta^{(i)} = 0 \) but \( \theta^{(j)} \neq 0 \), then

\[
P\left(\left|u_n^{(i)} u_n^{(j)}\right| > \varepsilon/2\right) \leq P\left(\left|u_n^{(i)} (u_n^{(j)} - \theta^{(j)})\right| > \varepsilon/4\right) +

\[
P\left(\left|u_n^{(i)}\right| > \varepsilon/\theta^{(j)}\right)
\]

\[
\leq \left(\frac{\varepsilon}{4}\right)^{-r} \mathbb{E}\left|u_n^{(i)} (u_n^{(j)} - \theta^{(j)})\right|^r + o(n^{1-2r})
\]

(by (2.48))

\[
\leq \left(\frac{\varepsilon}{2}\right)^{r} \mathbb{E}|u_n^{(i)}|^{2r} \mathbb{E}|u_n^{(j)} - \theta^{(j)}|^{2r} + o(n^{1-2r})
\]

\[
= \left(\frac{\varepsilon}{2}\right)^{r} (o(n^{-1/2}))^2 + o(n^{-2r})
\]

\[= o(n^{-r}).\]

Similarly if \( \theta^{(j)} = 0 \) but \( \theta^{(i)} \neq 0 \) then \( P\left(\left|u_n^{(i)} u_n^{(j)}\right| > \varepsilon/2\right) = o(n^{-r}). \)

If \( \theta^{(i)} = 0 \) and \( \theta^{(j)} = 0 \) then
\[
\mathbb{P}\{|U_{n1}^{(1)}U_{n2}^{(1)} - \theta^{(2)}(j)| < \varepsilon/2\} \leq (\varepsilon/2)^{-\frac{1}{2}}(\mathbb{E}|U_{n1}^{(1)}|^{2^{1/2}})(\mathbb{E}|U_{n2}^{(1)}|^{2^{1/2}})
= 0(n^{-r}).
\]

If \(\theta^{(1)} \neq 0\) and \(\theta^{(1)} \neq 0\), then
\[
\mathbb{P}\{|U_{n1}^{(1)}U_{n2}^{(1)} - \theta^{(1)}(j)| > \varepsilon/2\} \leq \mathbb{P}\{|(U_{n1}^{(1)} - \theta^{(1)})(U_{n2}^{(1)} - \theta^{(1)})| > \varepsilon/2\} \\
+ \mathbb{P}\{|U_{n1}^{(1)} - \theta^{(1)}| > \varepsilon/6\} \theta^{(1)}|j|) \\
+ \mathbb{P}\{|U_{n2}^{(1)} - \theta^{(1)}| > \varepsilon/6\} \theta^{(1)}|j|)
\]

We have
\[
\mathbb{P}\{|(U_{n1}^{(1)} - \theta^{(1)})(U_{n2}^{(1)} - \theta^{(1)})| > \varepsilon/2\} \leq (\varepsilon/2)^{-r}\mathbb{E}|(U_{n1}^{(1)} - \theta^{(1)})(U_{n2}^{(1)} - \theta^{(1)})|^{r} \\
\leq (\varepsilon/2)^{-r}\mathbb{E}|U_{n1}^{(1)} - \theta^{(1)}|^{2r}^{1/2}(\mathbb{E}|U_{n2}^{(1)} - \theta^{(1)}|^{2r}^{1/2}) \\
= (\varepsilon/2)^{-1} 0(n^{-r/2}) \times 0(n^{-r/2}) \quad \text{(by 2.56)} \\
= 0(n^{-r}).
\]

Also by (2.48) we have
\[
\mathbb{P}\{|U_{n1}^{(1)} - \theta^{(1)}| > \varepsilon\} = 0(n^{1-2r}).
\]

Therefore
\[
(2.52) \quad \mathbb{P}\{|U_{n1}^{(1)}U_{n2}^{(1)} - \theta^{(1)}(j)| > \varepsilon/2\} = 0(n^{-r}).
\]

By definition of \(v_{\alpha}^{(i)}\) and \(v_{\alpha}^{(j)}\) we have
\[
\frac{1}{n} \sum_{\alpha=1}^{n} v^{(i)} v^{(j)} = \frac{1}{n^{n-1}} \left( \begin{array}{c} n-1 \\ m_i-1 \end{array} \right) \sum_{\alpha=1}^{n} \sum_{c_i, c_j} f^{(i)}(X_{\alpha}, X_{\alpha_2}, \ldots, X_{\alpha_{m_i}}) f^{(j)}(X_{\alpha}, X_{\beta_2}, \ldots, X_{\beta_{m_j}})
\]

where the summation \( \Sigma \) extends over \( 1 \leq \alpha_2 < \ldots < \alpha_{m_i} \leq n \) but \( c_i \)
\( \alpha_{\ell} \neq \alpha, \ell = 2, \ldots, m_i \) and the summation \( \Sigma \) extends over \( c_j \)
\( 1 \leq \beta_2 < \ldots < \beta_{m_j} \leq n \) but \( \beta_{\ell} \neq \alpha, \ell = 2, \ldots, m_j \). For \( c = 0,1, \ldots, \min(m_i, m_j) - 1 \) write

\[
f^{(i,j)}(X_{\alpha}, X_{\alpha_2}, \ldots, X_{\alpha_{m_i}}) f^{(i,j)}(X_{\alpha}, X_{\beta_2}, \ldots, X_{\beta_{m_j}})
\]

where

\[
A_c: 1 \leq \alpha_2 < \ldots < \alpha_{m_i} \leq n \text{ but } \alpha_{\ell} \neq \alpha, \ell = 2, \ldots, m_i \text{ and }
1 \leq \beta_2 < \ldots < \beta_{m_j} \leq n \text{ but } \beta_{\ell} \neq \alpha, \ell' = 2, \ldots, m_j \text{ and }
\]
such that \((\alpha_2, \ldots, \alpha_{m_i})\) and \((\beta_2, \ldots, \beta_{m_j})\) have exactly \( c \) suffices in common.

The number of elements in \( A_c \), for \( c = 0,1, \ldots, \min(m_i, m_j) - 1 \) is equal to \( \left( \begin{array}{c} n-1 \\ m_i-1 \end{array} \right) \left( \begin{array}{c} n-m_i \\ m_j-1 \end{array} \right) \). Thus, assume \( m_i \leq m_j \),

\[
(2.53) \quad \frac{1}{n} \sum_{\alpha=1}^{n} v^{(i)} v^{(j)} = \sum_{i=0}^{m_i-1} \delta(c) u^{(i,j)}
\]
where \( \delta(c) = \binom{n-1}{m-j-1} \binom{n-m_i}{m_i-1} = O(n^{-c}) \), and \( U_{i,j}^{(i)} \) are U-statistics with kernels

\[
\xi_{i,j} = f^{(i)}(X_1, \ldots, X_{m_i}) f^{(j)}(X_1, \ldots, X_{m_i+1}, X_{m_i+1}, \ldots, X_{m_j-1}).
\]

Also \( E(f_{i,j}^{i,j}) < \infty \) if \( E(f^{(i)}) < \infty \) and \( E(f^{(j)}) < \infty \). Therefore

\[
\frac{1}{n} \sum_{\alpha=1}^{n} V_{\alpha}^{(i)} V_{\alpha}^{(j)} - \xi_{i,j}^{(i,j)} = \frac{m_i-1}{n} \sum_{i=0}^{m_i-1} \delta(i) \left[ U_{i,j}^{(i,j)} - EU_{i,j}^{(i,j)} \right] + \frac{m_i-1}{n} \sum_{c=0}^{m_i-1} \delta^*(c) EU_{i,j}^{(i,j)},
\]

where \( \delta^*(0) = \delta(0) - 1 = O(n^{-1}) \) and \( \delta^*(c) = \delta(c) \), \( c = 1, \ldots, m_i - 1 \).

Therefore

\[
(2.54) \quad E \left| \frac{1}{n} \sum_{\alpha=1}^{n} V_{\alpha}^{(i)} V_{\alpha}^{(j)} - \xi_{i,j}^{(i,j)} \right|^r \leq \left( m_{i+1} \right)^r \left[ E \left| U_{0}^{(i,j)} - EU_{0}^{(i,j)} \right|^r \right] (\delta(0))^{r-1}
\]

\[
+ \sum_{i=1}^{m_i-1} \delta(c) \left[ E \left| U_{i,j}^{(i,j)} - EU_{c}^{(i,j)} \right|^r + O(n^{-r}) \right]
\]

\[
= O(n^{-r/2}) + O(n^{-3r/2}) + O(n^{-r})
\]

Then, by (2.48), (2.52) and (2.54) we get (2.49).

Now, by (2.46), (2.47) and (2.49) we have the following.

**Theorem 2.8.** If for \( i = 1, \ldots, q \), \( E(f^{(i)})^{2r} < \infty \) for some \( r > 1 \) then
\begin{equation}
\begin{aligned}
(2.55) \quad & P\{|\gamma_n^2 - \gamma^2| > \varepsilon\} = 0(n^{-r/2}), \\
\text{where} \quad & \gamma_n^2 = \sum_{i=1}^{q} \sum_{j=1}^{q} m_i m_j \frac{\partial g(U_n)}{\partial x(i)} \frac{\partial g(U_n)}{\partial x(j)} \delta_{ij}. 
\end{aligned}
\end{equation}

Theorem 2.9. If for \( i = 1, \ldots, q \) \( E(f^{(i)})^2 < \infty \), then

\begin{equation}
(2.56) \quad \gamma_n^2 \rightarrow \gamma^2 \text{ a.s. as } n \rightarrow \infty.
\end{equation}

Proof: By continuity of \( \frac{\partial g(x)}{\partial x(i)} \) and convergence almost surely of U-statistics we have

\begin{equation}
(2.57) \quad \frac{\partial g(U_n)}{\partial x(i)} \rightarrow \frac{\partial g(\theta)}{\partial x(i)} \text{ a.s. as } n \rightarrow \infty.
\end{equation}

Also, from (2.50), (2.53) and convergence almost surely of U-statistics then

\begin{equation}
(2.58) \quad \delta_{ij} \xrightarrow{} \xi_{1(i,j)} \text{ a.s. as } n \rightarrow \infty.
\end{equation}

Hence the result.
CHAPTER III

TESTING HYPOTHESES CONCERNING ESTIMABLE PARAMETERS

3.1. Introduction. Let \( X_1, X_2, \ldots \) be a sequence of iid r.v.'s, with DF \( F(x), x \in \mathbb{R}^p, p \geq 1 \). For \( i = 1, \ldots, q \) let \( \theta^{(i)} = \theta^{(i)}(F) \) be a regular functional over some subset \( F = \{F: |\theta^{(i)}| < \infty, i = 1, \ldots, q\} \) of the set of all DF's in the p-dimensional Euclidean space. Consider a function \( g(x) = g(x^{(1)}, \ldots, x^{(q)}) \) that is continuous and has continuous first and second ordered partial derivatives with respect to \( x^{(1)}, \ldots, x^{(q)} \) in some neighborhood of \( \theta \). Given two real numbers \( g_0 \leq g_1 \) we want to choose one of the three hypotheses

\[
(3.1) \quad H_1: \ g(\hat{\theta}) < g_0 \quad H_2: \ g_0 \leq g(\hat{\theta}) \leq g_1 \quad H_3: \ g_1 < g(\hat{\theta}).
\]

The problem was studied by Sobel and Wald (1949), Armitage (1950) and Simon (1967) when \( g(\hat{\theta}) = \theta \), and \( \theta \) is the mean of a normal distribution with known variance. In this chapter we develop a similar procedure to that of Sobel and Wald (1949) for (3.1). First we develop a testing procedure for

\[
(3.2) \quad H_1': \ g(\hat{\theta}) = g_0 - \Delta \quad \text{vs} \quad H_2': \ g(\hat{\theta}) = g_0 + \Delta,
\]

where \( \Delta > 0 \) and \( g_0 \) are specified, and study its properties. Then we utilize this testing procedure to study (3.1). In fact we are going to study (3.1) only when \( g_0, g_1 \) and the true value of the parameter \( g(\hat{\theta}) \) differ by small amounts.
The sequential procedure for testing (3.1) may be simply described as follows. We establish a testing procedure \( R_1 \) for testing (3.2) and a similar procedure \( R_2 \) for testing

\[
(3.3) \quad H'_3: \ g(\theta) = g_1 - \Delta \ vs \ H'_4: \ g(\theta) = g_1 + \Delta.
\]

Both procedure \( R_1 \) and \( R_2 \) are computed at each stage of inspection until Either: one of \( R_1 \) and \( R_2 \) leads to a decision to stop before the other; then the former is no longer computed and the latter is continued until it leads to a decision to stop, Or: both \( R_1 \) and \( R_2 \) lead to a decision to stop at the same time; in this case both procedures are discontinued. We are going to show later that under some proper choice of constants a and b it is impossible for \( R_1 \) to accept \( H'_1 \) while \( R_2 \) accepts \( H'_4 \).

Then there are three possibilities for the final decision \( R \).

<table>
<thead>
<tr>
<th>( R_1 ) accepts</th>
<th>( R_2 ) accepts</th>
<th>( R ) accepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g_0 - \Delta ) and ( g_1 - \Delta )</td>
<td>then ( H_1 )</td>
<td></td>
</tr>
<tr>
<td>( g_0 + \Delta ) and ( g_1 - \Delta )</td>
<td>then ( H_2 )</td>
<td></td>
</tr>
<tr>
<td>( g_0 + \Delta ) and ( g_1 + \Delta )</td>
<td>then ( H_3 )</td>
<td></td>
</tr>
</tbody>
</table>

3.2. Two hypotheses problem. Suppose that \( \theta^{(i)}(F) \) is a regular functional of degree \( m_i, i = 1, \ldots, q \). Then there exist symmetric functions \( f^{(i)} = 1, \ldots, q \) such that (2.19) holds for \( i = 1, \ldots, q \).

Let \( U_n^{(i)} \) be defined by (2.20) and suppose that (2.21) and (2.22) hold.

We want to test

\[
(3.4) \quad H_{-1}: \ g(\theta) = g_0 - \Delta \ vs \ H_1: \ g(\theta) = g_0 + \Delta,
\]

where \( g_0 \) and \( \Delta (>0) \) are specified.
For a single parameter $\theta$ Sen (1973) gives a method for constructing sequential procedures for testing

$$H_0: \theta = \theta_0 \text{ vs } H_1: \theta = \theta_0 + \Delta \quad \Delta \text{ known}.$$  

His test is based on the U-statistic estimator of $\theta$, and it enjoys the optimality of Cox's (1963) sequential likelihood ratio test as $\Delta$ tends to zero. The same technique is used here to formulate our proposed sequential procedure and to study its various properties.

We start with initial sample size $n_0 = n_0(\Delta)$ moderately large for small $\Delta$. We shall take $n_0(\Delta) = \Delta^{-1}$; then for $n \geq n_0$ we use the following procedure.

**Procedure $P$**. We choose two real numbers $a$ and $b$ such that $b < 0 < a$. Define a stopping variable $N(\Delta)$ as the smallest integer $\geq n_0(\Delta)$ for which one of the following inequalities is violated.

$$b \gamma_n^2 < n \Delta[g(U_n) - g_0] < a \gamma_n^2. \tag{3.5}$$

Accept $H_0$ if for $N(\Delta) = n$, $n \Delta[g(U_n) - g_0] \leq b \gamma_n^2$, and accept $H_1$ if for $N(\Delta) = n$, $n \Delta[g(U_n) - g_0] \geq a \gamma_n^2$.

**Theorem 3.1.** Under conditions (2.21) and (2.22) we have

$$p_\theta\{N(\Delta) > n\} \to 0 \quad \text{as } n \to \infty. \tag{3.6}$$

**Proof:** For $n \geq n_0(\Delta)$ we have

$$p_\theta\{N(\Delta) > n\} \leq p_\theta\{b \gamma_n^2 < n \Delta[g(U_n) - g_0] < a \gamma_n^2\}. \tag{3.7}$$

If $g(\theta) = g_0$ then
\[ P_0 \{ N(\Delta) > n \} \leq P_0 \{ \frac{b}{\Delta} \hat{\gamma}_n n^{1/2} < \frac{\sqrt{n}}{\gamma_n} (g(U_n) - g(\hat{\theta})) < \frac{a}{\Delta} \hat{\gamma}_n n^{-1/2} \} \]

By Hoeffding (1948), the convergence of \( \hat{\gamma}_n \) to \( \gamma \) (in Theorem 2.9) and with the Slutsky theorem (cf Cramer (1946)) \( \frac{\sqrt{n}}{\gamma_n} [g(U_n) - g(\hat{\theta})] \) is asymptotically normally distributed with mean zero and unit variance. By continuity of the normal distribution and Theorem 2.9 we obtain that

\[ P_0 \{ \frac{b}{\Delta} \hat{\gamma}_n n^{-1/2} < \frac{\sqrt{n}}{\gamma_n} [g(U_n) - g(\hat{\theta})] < \frac{a}{\Delta} \hat{\gamma}_n n^{-1/2} \} \]

\( \to 0 \) as \( n \to \infty \). Then \( P_0 \{ N(\Delta) > n \} \to 0 \) as \( n \to \infty \). On the other hand, if \( g(\hat{\theta}) \neq g_0 \), then by (3.7) we have

\[ P_0 \{ N(\Delta) > n \} \leq P_0 \{ n^{-1} \hat{\gamma}_n^2 \frac{b}{\Delta} < g(U_n) - g_0 < n^{-1} \hat{\gamma}_n^2 \frac{a}{\Delta} \} \]

By Theorem 4 of Sen (1960) we have \( g(U_n) \overset{a.s.}{\to} g(\hat{\theta}) \) as \( n \to \infty \). Hence for \( g(\hat{\theta}) \neq g_0 \), \( [g(U_n) - g_0] \) tends, almost surely, to a value different from zero. Then by Theorem 2.9 and (3.8) we have \( P_0 \{ N(\Delta) > n \} \) converges to zero as \( n \to \infty \). Hence the theorem.

3.3. OC Function. We consider the asymptotic situation where we let \( \Delta \to 0 \). The results should provide a good approximation for practical situations where \( \Delta \) is fixed but small. A reasonably large initial sample size \( n_0(\Delta) \) is needed to insure the accuracy of \( \hat{\gamma}_n^2 \) as an estimate of \( \gamma^2 \) for all \( n \geq n_0(\Delta) \). We assume that

\[ \lim_{\Delta \to 0} n_0(\Delta) = \infty \quad \text{but} \quad \lim_{\Delta \to 0} \Delta^2 n_0(\Delta) = 0. \]
Let $F_{-1,0}(x)$ be the unknown DF of the $X_i$'s under the null hypothesis $H_{-1}: g(\Theta) = g_0 - \Delta$ and $F_{1,0}(x)$ be the DF under the alternative $H_1: g(\Theta) = g_0 + \Delta$. Let $L_F(\Theta)$ denote the OC function of the proposed test; i.e. $L_F(\Theta)$ is the probability of accepting $H_{-1}$ when $g(\Theta)$ is the true value. A rough approximation for $L_F(\Theta)$ would be

$$L_F(\Theta) = \frac{\frac{ah_1}{e^{\frac{1}{2}}} - 1}{\frac{ah_1}{e^{\frac{1}{2}}} - bh_1},$$

where

$$h_1 = \frac{g_0 - g(\Theta)}{\Delta}.$$

This approximation is rough because it involves three approximations; one is approximating $g(U_n) - g(\Theta)$ by a Wiener process, another is approximating $\gamma^2$ by $\hat{\gamma}^2_n$ and a third is neglecting shooting over the boundaries. Moreover, it can be shown that for every $g(\Theta) \neq g_0$, the OC function converges, as $\Delta \to 0$, to 0 or 1 according as $g$ is $>$ or $< g_0$.

To avoid such limiting degeneracy we let

$$g(\Theta) = g_0 + \phi \Delta, \quad \phi \in I = \{\phi : |\phi| < k\} \text{ for some } 1 < k < \infty.$$  

We write $L_F(\phi,\Delta)$ to denote the OC function of the proposed test when $g(\Theta) = g_0 + \phi \Delta$ and we write $F_{\phi,\Delta}$ to denote the DF in this case. We now study the limiting behavior of $L_F(\phi,\Delta)$ as $\Delta \to 0$.

**Theorem 3.2.** Under conditions in Theorem 3.1 we have, for every $\phi \in I$,

$$\lim_{\Delta \to 0} L_F(\phi,\Delta) = \begin{cases} \frac{e^{-2a\phi} - 1}{e^{-2a\phi} - e^{-2b\phi}}, & \phi \neq 0 \\ \frac{a}{a - b}, & \phi = 0. \end{cases}$$
Proof: We have

\[(3.12) \quad P_{\phi}^{(2)}(\phi, \Delta) = \mathbb{P} \left\{ \frac{b}{\Delta} \frac{\gamma^2}{n} \geq n[g(U_n) - g_0], \text{ for smaller } n \geq n_0(\Delta) \text{ than any other } n \geq n_0(\Delta) \text{ for which } \frac{a}{\Delta} \frac{\gamma^2}{n} \leq n[g(U_n) - g_0] \right\} \]

Since we do not know the DF of \(g(U_n)\) we now make use of its convergence to a Wiener process. By (2.40), for every \(\epsilon > 0\) and \(\eta > 0\), there exists a \(t_0 = t_0(\epsilon, \eta)\) such that

\[(3.13) \quad \mathbb{P}\left\{ \sup_{t \geq t_0} t^{-1/2} |s(t) - \gamma \xi(t)| > \epsilon/2 \right\} < \frac{n}{2}.\]

Also, by Theorem 2.9 there exists an \(n_0(\epsilon, \eta)\) such that

\[(3.14) \quad \mathbb{P}\left\{ \max_{n \geq n_0(\epsilon, \eta)} |\gamma^2_n - \gamma^2| > \frac{1}{2} \epsilon \right\} < \frac{n}{2}.\]

Thus we can take a \(\Delta_0 = \Delta_0(\epsilon, \eta)\) such that

\[(3.15) \quad n_0(\Delta_0) \geq \max\{t_0, n_0(\epsilon, \eta)\}.\]

Thus \(t \Delta^2 \leq \epsilon\) for all \(t \in [0, n_0(\Delta_1)]\), where \(\Delta \in [0, \Delta_0]\). Now, for every \(\delta > 0\), \(t_0 > 0\) and \(\phi \in I\), define

\[(3.16) \quad P(\phi, \delta, a, b, t_0) = \mathbb{P}\left\{ \xi(t) \leq \delta^{-1}b + t\delta \phi \text{ for smaller } t \geq t_0 \text{ than any other } t \geq t_0 \text{ for which } \xi(t) \geq \delta^{-1}a + t\delta \phi \right\} \]

Then, by Anderson (1960) we have

\[(3.17) \quad P(\phi, \delta, a, b, 0) =\begin{cases} \frac{e^{2\phi a} - 1}{2\phi a - 2b}, & \phi \neq 0 \\ \frac{a}{a - b}, & \phi = 0. \end{cases}\]
Define

\[(3.18) \quad s^{(1)}_\varepsilon = [1 + (-1)^i\varepsilon]a, \quad b^{(1)}_\varepsilon = [1 + (-1)^i\varepsilon]b, \quad i = 1, 2.\]

Thus, from (3.12) - (3.16) and (3.18) it follows that for every \(\Delta \in [0, \Delta_0]\) and \(g(\hat{\theta}) = g_0 + \phi \Delta, \phi \in \mathbb{I}.\)

\[(3.19) \quad P(-\omega, \frac{\Delta}{\gamma}, a^{(2)}_\varepsilon, b^{(1)}_\varepsilon, n_0(\Delta)) - \eta \leq L_{\varepsilon}(\phi, \Delta) \leq P(-\omega, \frac{\Delta}{\gamma}, a^{(1)}_\varepsilon, b^{(2)}_\varepsilon, n_0(\Delta)) + \eta.\]

By the Levy inequality, the probability that \(\{\xi(t); 0 \leq t \leq n_0(\Delta)\}\) crosses either of the two lines \((\frac{\gamma}{\Delta} b^{(j)}_\varepsilon - t \frac{\phi \Delta}{\gamma})\) or \((\frac{\gamma}{\Delta} a^{(j)}_\varepsilon - t \frac{\phi \Delta}{\gamma})\)

for \(i, j = 1, 2\) can be made smaller than \(\eta'(>0)\), where \(\eta' \rightarrow 0\) as \(\varepsilon \rightarrow 0.\)

Thus by (3.19) we get

\[(3.20) \quad P(-\phi, \frac{\Delta}{\gamma}, a^{(2)}_\varepsilon, b^{(1)}_\varepsilon, 0) - \eta - \eta' \leq L_{\varepsilon}(\phi, \Delta) \leq P(-\phi, \frac{\Delta}{\gamma}, a^{(1)}_\varepsilon, b^{(2)}_\varepsilon, 0) + \eta + \eta'.\]

Now, \(P(\phi, \delta, a, b, 0)\) is continuous in \(a\) and \(b\), so by choosing \(\varepsilon\) and \(\eta\)

sufficiently small, both sides of (3.20) can be made arbitrarily close to (3.17). Hence the theorem.

3.4. **ASN Function.** By (3.9) we have, as \(\Delta \rightarrow 0.\)

\[\Delta^2 E N(\Delta) - \Delta^2 \sum_{n \geq n_0(\Delta)} \{P_{\hat{\theta}} N(\Delta) > n\} = \Delta^2 n_0(\Delta) \rightarrow 0.\]

Now, for \(\varepsilon > 0,\)

\[(3.21) \quad P\{N(\Delta) > n\} \leq P\{\gamma^2_n b < n\Delta[g(U_n) - g_0] < \gamma^2_n a\}\]

\[= P\{\gamma^2_n b < n\Delta[g(U_n) - g_0] < \gamma^2_n a, \gamma^2_n - \gamma^2 > \gamma^2\varepsilon\}\]
\[ + P\{ \gamma_n^2 < n\Delta [g(U_n) - g_0] < \gamma_n^2 \} a, \gamma^2 - \gamma^2 \leq \gamma^2 e \} \]
\[ \leq P\{ \gamma^2 - \gamma^2 > \gamma^2 e \} + P\{by^2(l+e) < n\Delta [g(U_n) - g_0] \]
\[ < ay^2(l+e) \}. \]

Let \( g(\theta) = g_0 + \phi \Delta, \phi \in I. \) Then

\[(3.22) \quad P_{\theta} \{ N(\Delta) > n \} \leq P_{\theta} \{ \gamma^2 - \gamma^2 > \gamma^2 e \} \]
\[ + P_{\theta} \{ by^2(l+e) - n \phi \Delta^2 < n\Delta [g(U_n) - g(\theta)] \]
\[ < ay^2(l+e) - n \phi \Delta^2 \}. \]

Consider the case of \( \phi \neq 0 \) and let for every \( \Delta > 0, \)

\[(3.23) \quad n_{\phi, \Delta} = \frac{k^2 \gamma^2}{\Delta^2 \phi^2}, \quad 0 < k < \infty. \]

By proper choice of \( k, \gamma^2(a-b)(1+e) \) can be made smaller than \( \frac{n}{4}|\phi| \Delta. \)
Thus the second term on the right hand side of (3.22) is bounded above by

\[ P_{\theta} \{ |g(U_n) - g(\theta)| > \frac{3}{4} n \Delta | \phi | \}. \]

By continuity of \( g(\theta), \) for every \( \epsilon > 0 \) there exists \( \delta > 0 \) such that

\[ |g(U_n) - g(\theta)| < \epsilon \text{ whenever } ||U_n - \theta|| < \delta. \]

Therefore,

\[ P\{ |g(U_n) - g(\theta)| > \epsilon \} \leq P\{ ||U_n - \theta|| > \delta \}. \]

By Theorem 2.6 if we assume that \( E(g^{(1)})^{2r} < \infty, 1 = 1, 2, \ldots, q, r > 1 \)
then \( P\{ \sup_{k \geq n} |U_k - \theta| > \delta \} = o(n^{1-2r}). \)
\[ R_n = g(U_n) - g(\theta) - \sum_{i=1}^{q} m_i \frac{\partial g(\theta)}{\partial x(i)} U_n^{(i)} - U_{n,1}^{(i)} \]

\[ = g(U_n) - g(U_{n,1}) + (g(U_{n,1}) - g(\theta) - U_{n,1}^{(i)}) , \]

where \( U_{n,1} = (U_{n,1}^{(1)}, \ldots, U_{n,1}^{(r)}) \) and \( U_{n,1}^{(i)} = \sum_{i=1}^{q} m_i \frac{\partial g(\theta)}{\partial x(i)} U_n^{(i)} . \)

Write

\[ R_{n1} = g(U_n) - g(U_{n,1}) , \]

and

\[ R_{n2} = g(U_{n,1}) - g(\theta) - U_{n,1}^{(i)} . \]

Then

\[ R_{n1} = \sum_{i=1}^{q} \left( U_n^{(i)} - U_{n,1}^{(i)} \right) \frac{\partial g(\lambda)}{\partial x(i)} , \]

where \( \lambda \) is some point between \( U_n \) and \( U_{n,1}^{(i)} \). Also,

\[ R_{n2} = \frac{1}{2} \sum_{i=1}^{q} \sum_{j=1}^{q} \left( U_n^{(i)} - \theta(i) \right) \left( U_n^{(j)} - \theta(j) \right) \frac{\partial g(\lambda')}{\partial x(i)} , \]

where \( \lambda' \) is some point in the neighborhood of continuity of \( g(\cdot) \) around \( \theta \).

Let \( A'_{n,\varepsilon} = \{|U_n - \theta| < \varepsilon\} \), \( A_{n,\varepsilon} = \{|U_{n,1} - \theta| < \varepsilon\} \), and

\[ B_{n,\varepsilon} = A'_{n,\varepsilon} \cap A_{n,\varepsilon}^2 . \]

Then

\[ P(A_{n,\varepsilon}^2) \leq P(|U_{n,1} - \theta| > \varepsilon) \]

\[ P(B_{n,\varepsilon}) \leq P(|U_{n,1} - \theta| > \varepsilon) + P(|U_{n,1} - \theta| > \varepsilon) . \]
If we assume that $E(f^{(1)})^{2r} < \infty$, $r = 2 + \delta$, $\delta > 0$ then by Grams and Serfling (1973)

(3.33) \[ P\{|U_n - \theta| > \epsilon\} = o(n^{1-2r}), \]

and

(3.34) \[ P\{|U_n,1 - \theta| > \epsilon\} = o(n^{1-2r}). \]

Thus

(3.35) \[ P\{\tilde{A}_n, \epsilon\} = o(n^{1-2r}), \]

and

(3.36) \[ P\{\tilde{B}_n, \epsilon\} = o(n^{1-2r}). \]

Now,

\[
P\{k^{1/2}|R_{k1}| > \eta\} = P\{k^{1/2}|R_{k1}| > \eta, B_k, \epsilon\} \\
+ P\{k^{1/2}|R_k| > \eta, \bar{B}_k, \epsilon\} \\
\leq P\{k^{1/2}|R_{k1}| > \eta, B_k, \epsilon\} + P\{\bar{B}_k, \epsilon\} \\
= P\{k^{1/2}|R_{k1}| > \eta, B_k, \epsilon\} + O(k^{1-2r}).
\]

In $B_k, \epsilon \frac{\partial g(\lambda)}{\partial \theta^{(1)}}$ is bounded and we have

\[
k^{1/2}|R_{k1}| \leq C \epsilon k^{1/2} \sum_{i=1}^{q} (u^{(1)}_{k} - u^{(1)}_{k,1})
\]

and therefore
\[ p^{1/2} \left| R_{k,1} \right| > \eta, B_{k,\epsilon} \right\} \leq p^{1/2} \left( \sum_{i=1}^{q} (u_{k,1}^{(i)} - u_{i,1}^{(i)}) \right) > \eta' \]

\[ \leq \frac{c k^r}{\epsilon^{2r}} \sum_{i=1}^{q} E \left| u_{k,1}^{(i)} - u_{i,1}^{(i)} \right|^{2r} \]

\[ = \frac{c}{\epsilon^{2r}} k^r 0(k^{-2r}) \]

\[ = 0(k^{-r}), \quad r = 2 + \delta. \]

Thus

(3.37) \quad p^{1/2} \left| R_{k,1} \right| > \eta \right\} = 0(k^{-r}).

Similarly

(3.38) \quad p^{1/2} \left| R_{k,2} \right| > \eta \right\} \leq p^{1/2} \left| R_{k,2} \right| > \eta, A_{k,\epsilon} \right\} + p\left( A_{k,\epsilon}^2 \right) \]

\[ = p^{1/2} \left| R_{k,2} \right| > \eta, A_{k,\epsilon} \right\} + 0(n^{-2r}), \]

\[ \leq p^{1/2} \left( \sum_{i=1}^{q} (u_{k,1}^{(i)} - \theta^{(i)} \left)^2 > \eta' \right\}. \]

Now we can write

(3.39) \quad k^{1/2} \left| \sum_{i=1}^{q} (u_{k,1}^{(i)} - \theta^{(i)} \right)^2 = k^{1/2 - 1/2 + 2\delta} \left( \sum_{i=1}^{q} (u_{k,1}^{(i)} - \theta^{(i)} \right)^2 \]

\[ = k^{1/2 - 1 - 2\delta} \left( \sum_{i=1}^{q} (u_{k,1}^{(i)} - \theta^{(i)} \right)^2 \]

\[ = k \sum_{i=1}^{q} p\left( \left| u_{k,1}^{(i)} - \theta^{(i)} \right| > \eta' \right\}, k^{1/2 - 1/2 + 2\delta} \]

For large \( k, k \) can be made arbitrarily small and we have

(3.40) \quad p\left( \sum_{i=1}^{q} (u_{k,1}^{(i)} - \theta^{(i)} \right)^2 > \eta' \right\} \leq p\left( \sum_{i=1}^{q} (u_{k,1}^{(i)} - \theta^{(i)} \right)^2 > \eta' \right\} \]

\[ \leq k \sum_{i=1}^{q} p\left( \left| u_{k,1}^{(i)} - \theta^{(i)} \right| > \eta' \right\}, k^{1/2 - 1/2 + 2\delta} \]
\[ q \sum_{i=1}^{\infty} \sum_{j=1}^{n} \mathbb{P}\left\{ f^{(i)}(X_j) - \theta^{(1)} \right\} > \eta'' < 1^{1/2+\delta_1} \].

Since \( E(f^{(1)})^{4+2\delta} < \infty \), then we can apply Theorem 2 of Katz (1963) to get that

\[ \left(3.41\right) \quad \mathbb{P}\left\{ k^{1/2} \left| \sum_{i=1}^{n} U_{k,1}^{(1)} - \theta^{(1)} \right|^2 > \eta' \right\} = O(n^{1+\delta'}), \quad \delta' > 0. \]

Therefore, by \(3.38\)

\[ \left(3.42\right) \quad \mathbb{P}\left\{ k^{1/2} |R_{k,2}| > \eta \right\} = O(k^{-1-\delta'}), \quad \delta' > 0. \]

We now combine \(3.28\) - \(3.30\), \(3.37\) and \(3.42\) to get

\[ \left(3.43\right) \quad \mathbb{P}\left\{ k^{1/2} |R_k| > \eta \right\} = O(k^{-1-\delta'}), \quad \delta' > 0. \]

**Proof of Theorem 3.3.** Define two stopping variables \( N^{(1)}_\phi(\Delta) \) and \( N^{(2)}_\phi(\Delta) \) as the smallest positive integers \( (\geq n_0(\Delta)) \) for which

\[ n \sum_{i=1}^{\infty} m_1 \frac{\partial g(\theta)}{\partial \theta^{(1)}} U_{n,1}^{(i)} + \omega \Delta \]

is not contained in \( \left( \frac{\gamma^2(1+\epsilon)b}{\Delta}, \frac{\gamma^2(1-\epsilon)a}{\Delta} \right) \) and

\[ \left( \frac{\gamma^2(1+\epsilon)b}{\Delta}, \frac{\gamma^2(1-\epsilon)a}{\Delta} \right) \] respectively.

Then, similar to \(2.26\) we have

\[ \left(3.44\right) \quad |\Delta^2 E_{\phi}(n^{(i)}(\Delta) - \Delta^{2n^{(i)}(\Delta)} - \mathbb{P}_{\theta}(n^{(i)}(\Delta) > n)| < \epsilon/2, \quad i = 1, 2. \]

Now

\[ \left(3.45\right) \quad \mathbb{P}_{\theta}(N(\Delta) > n) \leq \mathbb{P}_{\theta}\left( k \sum_{i=1}^{\infty} m_1 \frac{\partial g(\theta)}{\partial x^{(i)}} U_{k,1}^{(i)} + k \omega \Delta \right) < \frac{\gamma^2 a(1+\epsilon)}{\Delta}, \quad n_0 \leq k \leq n \]
\[ + \mathbb{P} \left\{ \gamma_k^2 - \gamma^2 > \frac{\gamma^2 \varepsilon}{2} \text{ or } \Delta k | R_k | > \varepsilon/2 \right\} \]

for at least one \( k: \ n_0(\Delta) \leq k \leq n \),

\[ = \mathbb{P} \left\{ N^{(2)}_\phi(\Delta) > n \right\} + \mathbb{P} \left\{ \gamma_k^2 - \gamma^2 > \frac{\gamma^2 \varepsilon}{2} \right\} \]

for at least one \( k: \ n_0 \leq k \leq n \),

\[ + \mathbb{P} \left\{ \Delta k | R_k | > \varepsilon/2 \right\} \]

for at least one \( k: \ n_0 \leq k \leq n \).

Now, for \( n_0(\Delta) \leq k \leq n_{\phi, \Delta} \), \( \Delta k^{1/2} \) is bounded and therefore by (3.43) we have

\[ (3.46) \quad \mathbb{P} \{ \Delta k | R_k | > \varepsilon/2 \text{ for at least one } n_0 \leq k \leq n \} = 0(n_0^{-\delta'}) \quad \delta' > 0. \]

Also, by (2.55) we have

\[ (3.47) \quad \mathbb{P} \left\{ \gamma_k^2 - \gamma^2 > \frac{\gamma^2 \varepsilon}{2} \right\} \text{ for at least one } k: \ n_0(\Delta) \leq k \leq n \]

\[ = 0(n_0^{1-r/2}(\Delta)) = 0(n_0^{-\delta/2}(\Delta)), \quad \delta > 0. \]

Then, by (3.45) we have

\[ (3.48) \quad \mathbb{P}_\phi \{ N(\Delta) > n \} \leq \mathbb{P}_\phi \{ | N^{(2)}_\phi(\Delta) | > n \} + 0(n_0^{-\delta}(\Delta)), \quad \delta > 0. \]

Similarly,

\[ (3.49) \quad \mathbb{P}_\phi \{ N(\Delta) > n \} \geq \mathbb{P}_\phi \{ N^{(1)}_\phi(\Delta) > n \} - 0(n_0^{-\delta}(\Delta)). \]

Since \( \Delta^2 n_{\phi, \Delta} \) is bounded and \( n_0^{-\delta}(\Delta) \to 0 \) as \( \Delta \to 0 \), we have for \( \Delta \to 0 \)
\begin{align}
(3.50) \quad & \Delta^n \sum_{n=n_0(\Delta)}^{n=\Delta} P_{\phi}^{(1)}(\Delta) > n \quad - \quad \eta < \sum_{n=n_0(\Delta)}^{n=\Delta} P_{\Theta}^{(n-\Delta)} > n \quad - \\
& \quad < \Delta^2 \sum_{n=n_0(\Delta)}^{n=\Delta} P_{\phi}^{(2)}(\Delta) > n \quad + \quad \eta, \quad \eta > 0,
\end{align}

where \( \eta \) can be made arbitrarily small by choosing \( \Delta \) small.

Now, since \( N_{\phi}^{(1)}(\Delta) \) is based on a summation over independent random variables, then if we neglect the excess over the boundaries for small \( \Delta \), and use Theorem 3.1 then by Wald's (1947) fundamental identity we have

\begin{align}
(3.51) \quad & \lim_{\Delta \to 0} [\Delta^2 E_{\phi}^{(1)}(\Delta) - (1 + (-1)^j \varepsilon) \psi(\phi, \gamma)] = 0, \quad j = 1, 2,
\end{align}

where

\[ \psi(\phi, \gamma) = \begin{cases} 
\gamma^{-1} \left[ bP(-\phi) + a(1 - P(-\phi)) \right], & \phi \neq 0 \\
-\gamma^2 P'(0)(a-b), & \phi = 0.
\end{cases} \]

Then, the theorem follows for \( \Theta \neq 0 \) by choosing \( \Delta \) small to make, by (3.51)

\begin{align}
(3.52) \quad & |\Delta^2 E_{\phi}^{(1)}(\Delta) - \psi(\phi, \gamma)| < \eta \quad \text{for} \quad j = 1, 2.
\end{align}

For \( \phi = 0 \) we note that \( P(\phi) \) is continuous and differentiable in some neighborhood of \( \phi = 0 \), so that \( P'(0) \) exists. Hence, by considering a sequence of values of \( \phi \), say \( \phi = \varepsilon \Delta \), \( \Delta \geq 1 \) where \( \varepsilon \Delta \to 0 \) as \( \Delta \to \infty \), and then using the above proof, we obtain by L'Hopital rule that

\begin{align}
(3.53) \quad & \lim_{\Delta \to 0} \left[ \Delta^2 E_{0}^{(1)}(\Delta) \right] = \lim_{\Delta \to 0} \psi(\phi, \gamma) \\
& \quad = -\gamma^2 P'(0)(a-b).
\end{align}

Hence the theorem.
3.5. **Three hypotheses problem.** Suppose we want to choose one of the hypotheses

\[(3.54) \quad H_{-1}: \ g(\bar{c}) < g_0 \quad \quad H_0: \ g_0 \leq g(\bar{c}) \leq g_1 \quad \quad H_1: \ g_1 < g(\bar{c})\]

where \(g_0\) and \(g_1\) differ by a small amount. There is no loss in assuming that \(g_0 = -g_1\). Consider the following problems for \(\Delta > 0\).

\[(3.55) \quad \text{Testing } H'_1: \ g(\bar{c}) = g_0 - \Delta \quad \text{vs} \quad H'_2: \ g(\bar{c}) = g_0 + \Delta\]

and

\[(3.56) \quad \text{Testing } H'_3: \ g(\bar{c}) = g_1 - \Delta \quad \text{vs} \quad H'_3: \ g(\bar{c}) = g_1 + \Delta.\]

Let \(R_1\) and \(R_2\) be two procedures similar to procedure \(P\) described on page 31 with \((a,b)\) and \((a',b')\) respectively. Let \(N_1\) and \(N_2\) be the stopping variables corresponding to \(R_1\) and \(R_2\) respectively. We use \(R_1\) to test (3.55) and \(R_2\) to test (3.56). We define a procedure \(R\) as follows.

**Procedure R.** Both \(R_1\) and \(R_2\) are computed at each stage of the inspection until, Either: one of the procedures leads to a decision to stop before the other. Then the former is no longer computed and the latter is continued until it leads to a decision to stop. Or: both \(R_1\) and \(R_2\) lead to a decision to stop at the same stage. In this case both computations are discontinued. Then we take a final decision as in the following table.

<table>
<thead>
<tr>
<th>(R_1) accepts</th>
<th>(R_2) accepts</th>
<th>(final) (R) accepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>If (H'_1)</td>
<td>and</td>
<td>(H'_1)</td>
</tr>
<tr>
<td>If (H'_2)</td>
<td>and</td>
<td>(H'_3)</td>
</tr>
<tr>
<td>If (H'_1)</td>
<td>and</td>
<td>(H'_4)</td>
</tr>
</tbody>
</table>
If we take \( a \leq a' \) and \( b \leq b' \) it can be shown that it is impossible for \( R_1 \) to accept \( H_1 \) while \( R_2 \) accepts \( H_4' \). From Theorem 3.1, \( R_1 \) and \( R_2 \) both terminate with probability one. Then it follows that \( R \) will also terminate with probability one. Then

\[
(3.57) \quad P_{\theta}\{\text{accept } H_0 | R\} = 1 - P_{\theta}\{\text{accept } H_{-1} | R\} - P_{\theta}\{\text{accept } H_1 | R\}.
\]

There is no loss in assuming that \( g_0 = -g_1 = c \Delta \) for some \( c > 1 \).

We shall see later that if \( c \) is large we can simplify the probability of correct decision. Let \( g(\theta) = \phi \Delta, \phi \in I \). Write \( L(H_1 | \phi, \Delta; R) \) to denote the probability that the procedure \( R \) will accept \( H_1 \) when \( g(\theta) = \phi \Delta \) is the true parameter value. That is,

\[
L(H_1 | \phi, \Delta; R) = P_{\theta}\{\text{accept } H_1 | R\}, \quad g(\theta) = \phi \Delta. \quad \text{We can write } g(\theta) = \phi \Delta = g_0 + (\phi + c)\Delta = g_1 + (\phi - c)\Delta. \quad \text{Then, by Theorem 3.2 we obtain the following relations.}
\]

\[
(3.58) \quad L(H_{-1} | \phi, \Delta; R) = L(H_1 | \phi, \Delta; R_1) \rightarrow \begin{cases} \frac{e^{-2a(\phi+c)} - 1}{e^{-2a(\phi+c)} - e^{-2b(\phi+c)}}, & \phi \neq -c \\ \frac{a}{a - b}, & \phi = -c \end{cases}, \quad \phi \rightarrow \infty.
\]

\[
(3.59) \quad L(H_1 | \phi, \Delta; R) = L(H_4' | \phi, \Delta; R_2) \rightarrow \begin{cases} \frac{1 - e^{-2b'(\phi-c)}}{e^{-2a'(\phi-c)} - e^{-2b'(\phi-c)}}, & \phi \neq c, \\ \frac{-b'}{a' - b'}, & \phi = c \end{cases}, \quad \phi \rightarrow 0.
\]

3.6. **Probability of correct decision.** Let \( L(\phi, \Delta | R) \) be the probability of correct decision when the true value of the parameter is \( g(\theta) = \phi \Delta, \phi \in I \). In virtue of (3.57) we define \( L(\phi, \Delta | R) \) as follows.
L(\phi, \Delta | R) =

\begin{align*}
&L(H_{-1}\mid \phi, \Delta; R) \quad \phi \leq -(c+1) \\
&1 - L(H_{1}\mid \phi, \Delta; R) \quad -(c+1) < \phi \leq -(c-1) \\
&1 - L(H_{-1}\mid \phi, \Delta; R) - L(H_{1}\mid \phi, \Delta; R), \quad -(c-1) \leq \phi \leq (c-1) \\
&1 - L(H_{-1}\mid \phi, \Delta; R) \quad (c-1) < \phi < (c+1) \\
&L(H_{1}\mid \phi, \Delta; R) \quad \phi \geq (c+1)
\end{align*}

It is to be noted that we take \( L(\phi, \Delta | R) \) to be the smaller values at points of discontinuities. Let \( L(\phi | R) \) be the limit, as \( \Delta \to 0 \), of \( L(\phi, \Delta | R) \). Then

(i) \( L(\phi | R) \) is monotonically decreasing with negative curvature for \( \phi \leq -(c-1) \).

(ii) \( L(\phi | R) \) is monotonically increasing with negative curvature for \( \phi \geq (c-1) \).

(iii) \( L(\phi | R) \) has negative curvature for \( -(c-1) < \phi < (c-1) \).

We now want to choose \((a, b)\) and \((a', b')\) to insure that

\( L(\phi | R) \geq 1 - \gamma_1 \) for \( \phi \leq -(c+1) \), \( L(\phi | R) \geq 1 - \gamma_2 \) for \( -(c+1) < \phi < (c+1) \) and \( L(\phi | R) \geq 1 - \gamma_3 \) for \( \phi \geq (c+1) \); where \( \gamma_1, \gamma_2 \) and \( \gamma_3 \) are positive constants. From the monotonic properties of \( L(\phi | R) \) it is enough to insure that

\begin{align*}
(3.61) \quad L(-(c+1) | R) &= 1 - \gamma_1, \quad L(-(c-1) | R) = L((c-1) | R) = 1 - \gamma_2 \\
&\text{and} \\
L((c+1) | R) &= 1 - \gamma_3.
\end{align*}

That is, we require

\begin{align*}
(3.62) \quad 1 - L(-(c+1) | R) &= \frac{1 - e^{2b}}{e^{2a} - e^{2b}} = \gamma_1,
\end{align*}
\[ (3.63) \quad 1 - L(-(c-1)|R) = \frac{e^{2b}(e^{2a} - 1)}{e^{2a} - e^{2b}} + \left[ \frac{1 - e^{2b'(2c-1)}}{e^{2a'(2c-1)} - e^{2b'(2c-1)}} \right] = \gamma_1. \]

\[ (3.64) \quad 1 - L((c-1)|R) = \frac{1 - e^{2b'}}{e^{2a'} - e^{2b'}} + \left[ \frac{e^{2b(2c-1)}(e^{2a(2c-1)} - 1)}{e^{2a(2c-1)} - e^{2b(2c-1)}} \right] = \gamma_2. \]

and

\[ (3.65) \quad 1 - L(c+1|R) = \frac{e^{2b}(e^{2a} - 1)}{e^{2a} - e^{2b}} = \gamma_3. \]

If \( c \) is sufficiently large that we can neglect the bracketed terms then we have

\[ (3.66) \quad e^{2a} = \frac{1 - \gamma_2}{\gamma_2}, \quad e^{2b} = \frac{\gamma_2}{1 - \gamma_1}, \]

and

\[ (3.67) \quad e^{2a'} = \frac{1 - \gamma_3}{\gamma_2}, \quad e^{2b'} = \frac{\gamma_3}{1 - \gamma_2}. \]

If \( \gamma_1 = \gamma_2 = \gamma_3 = \delta \) (say), then \( a = a' = -b = -b' = \log \left( \frac{1 - \delta}{\delta} \right)^{1/2} \).

3.7. **Bounds for the ASN function.** Let \( E(N, \phi|R) \) be the expected value of \( N \) when \( g(\theta) = \phi \Delta, \phi \in I \) is the true value of the parameter and \( R \) is the sequential procedure employed. From the definition of \( R \) it follows that for all \( \phi \in I \),

\[ (3.68) \quad E(N, \phi|R) > \max[E(N, \phi|R_1), E(N, \phi|R_2)]. \]

For \( \phi < -(c-1) \) the probability of coming to a decision to stop first with \( R_2 \) is high and therefore
(3.69) \[ E(N, \phi | R) \sim E(N, \phi | R_1) \text{ for } \phi < -(c-1). \]

Similarly

(3.70) \[ E(N, \phi | R) \sim E(N, \phi | R_2) \text{ for } \phi > (c+1). \]

Formula (3.68) gives a lower bound for \( E(N, \phi | R) \) for all \( \phi \in I \).

However, these bound by (3.69) and (3.70), will be close for \( |\phi| > (c-1) \) only. It is not easy to find close bounds for \( |\phi| < (c-1) \), but noticing that this corresponds to \( |g(\theta)| < (c-1)\Delta \), and that we study the case where \( \Delta \to 0 \), this is a small interval. Furthermore, we shall study the ASN over the whole range of \( \phi \) numerically in Chapter IV.

We now follow Sobel and Wald (1949) in deriving upper bounds for the ASN function. Let \( R_1^* \) be the following rule: "Continue sampling until \( R_1 \) accepts \( g_0 - \Delta \)." Since this implies that \( R_2 \) accepts \( g_1 - \Delta \) at the same or a previous stage, it follows that

(3.71) \[ E(N, \phi | R_1^*) \geq E(N, \phi | R). \]

Similarly if we define \( R_2^* \) to be the following rule: "Continue sampling until \( R_2 \) accepts \( g_1 + \Delta \). Then

(3.72) \[ E(N, \phi | R_2^*) \geq E(N, \phi | R). \]

Upon neglecting the excess over the boundary then we can prove, similar to Theorem 3.3, that

(3.73) \[ \Delta^2 E(N, \phi | R_1^*) \xrightarrow{\Delta \to 0} \frac{b \gamma^2}{\phi + \Delta} \text{ for } \phi < -(c+1) \]

and
(3.74) \[ \Delta^2 E(N, \phi | R^*_{(c)}) \rightarrow \frac{\alpha y^2}{\Delta \Delta} \quad \phi > (c+1). \]

Again, it is difficult to find upper bounds for $E(N, \phi | R)$ for $|\phi| \leq (c+1)$ and we refer this to the numerical studies in the next chapter.

In the following chapter we give some examples and make some numerical investigations of the proposed procedure.
CHAPTER IV

EXAMPLES AND NUMERICAL INVESTIGATIONS

4.1. Introduction. In the last chapter we developed a sequential procedure for discriminating between two hypotheses concerning continuous functions of several estimable parameters. The procedure is essentially valid for cases where the two hypothetical values and the true value of the function under consideration are close to each other. Based on this procedure we developed a test for discriminating among three hypotheses concerning the same function. We studied in some detail the OC function of the proposed test; we developed some approximating formula for the OC function. However, we were not able to study the ASN function in as much detail. We only gave some rough bounds for the ASN function over some intervals of the parameter space. This calls for some numerical investigations to see how close to actual values these approximating formula are. In this chapter we carry on a simulation study of one of the cases where the proposed test is applicable.

4.2. Correlation coefficients. Let $X_1, \ldots, X_n$ be a sequence of iid r.v.'s with DF $F(x)$, $x \in \mathbb{R}^2$. Let $X_1 = (X_{11}, X_{12})$, and $\theta = (\theta_1, \theta_2, \theta_3)$ where $\theta_1 = \sigma^2(X_{11})$, $\theta_2 = \text{Cov}(X_{11}, X_{12})$, and $\theta_3 = \sigma^2(X_{12})$. Then $\theta_1$, $\theta_2$ and $\theta_3$ are regular functionals of $F(x)$ such that:
\[
\begin{align*}
\theta_1 &= E(x_{11} - E X_{21})^2 = \frac{1}{2} \int \int (x_{11} - x_{21})^2 \, dF(x_1) \, dF(x_2) \\
\theta_2 &= E(x_{11} - E X_{21})(x_{12} - x_{22}) = \frac{1}{2} \int \int (x_{11} - x_{21})(x_{12} - x_{21}) \, dF(x_1) \, dF(x_2) \\
\theta_3 &= E(x_{12} - x_{22})^2 = \frac{1}{2} \int \int (x_{12} - x_{22})^2 \, dF(x_1) \, dF(x_2)
\end{align*}
\]

Let \( g(\theta) = g(\theta_1, \theta_2, \theta_3) = \frac{\theta_2}{[\theta_1 \theta_3]^{1/2}} \) be the correlation coefficient of \( X_{11} \) and \( X_{12} \). Suppose that \( \Delta > 0 \) and \( g_0 \) are given real values such that \( g_0 - \Delta > -1 \) and \( g_0 + \Delta < 1 \) and that we want to test

\[(4.2) \quad H_{-1}: \ g(\theta) = g_0 - \Delta \quad \text{vs} \quad H_1: \ g(\theta) = g_0 + \Delta.\]

Let \( U_{n1} \) be the U-statistic estimate of \( \theta_1 \), \( i = 1, 2, 3 \). Then

\[
\begin{align*}
U_{n1} &= \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_{11} - x_{j1})^2 \\
U_{n2} &= \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_{11} - x_{j1})(x_{12} - x_{j2}) \\
U_{n3} &= \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_{12} - x_{j2})^2
\end{align*}
\]

We base our test on the statistic

\[(4.4) \quad g(U_n) = g(U_{n1}, U_{n2}, U_{n3}) = \frac{U_{n2}}{[U_{n1} \ U_{n3}]^{1/2}}.\]

For \( \alpha = 1, \ldots, n \) we define
\[ V_{\alpha 1} = \frac{1}{2(n-1)} \sum_{i=1}^{n} (x_{\alpha 1} - x_{i1})^2 \]

\[ V_{\alpha 2} = \frac{1}{2(n-1)} \sum_{i=1}^{n} (x_{\alpha 1} - x_{i1})(x_{\alpha 2} - x_{i2}) \]

\[ V_{\alpha 3} = \frac{1}{2(n-1)} \sum_{i=1}^{n} (x_{\alpha 2} - x_{i2})^2. \]

For \( i, j = 1, 2, 3 \) we define

\[ \delta_{ij} = \frac{1}{n} \sum_{\alpha=1}^{n} V_{\alpha i} V_{\alpha j} - U_{i1} U_{nj}. \]

Then we define

\[ \gamma_n^2 = 4 \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial g(U_n)}{\partial \theta_i} \frac{\partial g(U_n)}{\partial \theta_j} \delta_{ij}, \]

where \( \frac{\partial g(U_n)}{\partial \theta_i} \bigg|_{\theta = U_n} = (-1)^i \frac{3 + (-1)^i}{4} \frac{g(U_n)}{U_n} \).

We choose two real numbers \( a \) and \( b \) such that \( b < 0 < a \) and for \( n \geq n_0 \) (depends on \( \Delta \)) we use the following sequential rule.

Rule \( R_1 \): Continue sampling until one of the following inequalities is violated.

\[ b \gamma_n^2 < n \Delta [g(U_n) - g_0] < a \gamma_n^2. \]

If the left inequality is violated first we accept \( H_{-1} \) and if the right inequality is violated first we accept \( H_1 \). If we desire that the test have a strength \((\alpha, \beta)\) then by Theorem 3.2 we approximate \( a \) and \( b \) by

\[ a = \frac{1}{2} \log \frac{1-\beta}{\alpha} \text{ and } b = \frac{1}{2} \log \frac{\beta}{1-\alpha}. \]
Now, suppose we want to choose one of the following hypotheses

\[(4.10) \quad H_{-1}: \ g(\theta) < 0 \quad H_0: \ g(\theta) = 0 \quad H_1: \ g(\theta) > 0.\]

Since the three hypotheses are very close to each other it would be difficult to discriminate among them with a high probability of making a correct decision. Instead we choose a small number \(g > 0\) and let \(g_0 = -g_1\) and find a testing procedure to discriminate among the three hypotheses

\[(4.11) \quad H_{-1}: \ g(\theta) < g_0 \quad H_0: \ g_0 \leq g(\theta) \leq g_1 \quad H_1: \ g(\theta) > g_1\]

Then if we accept \(H_i\) we accept the corresponding hypothesis \(\bar{H}_i\), \(i = 1, 2, 3.\)

Now, problem (4.11) is the same as (3.54) and we choose a small \(\Delta > 0\) such that \(\Delta < g_1\) and use procedure R in Chapter III. Since, in our case here, \(g(\theta)\) is a correlation coefficient then \(|g(\theta)| \leq 1\) and we choose \(g_1\) and \(\Delta\) such that \(|g_1 + \Delta| < 1\).

Going back to Chapters II and III, see Theorems 2.7, 2.8 and 3.3, we find that for the validity of our test we require \(E X_1^{4+\delta}\) to exist for some \(\delta > 0\). But since, in our case here, \(g(\theta)\) is bounded then possibly we may not need such a powerful condition. This condition is satisfied for many bivariate DF but it still puts some limitations to our procedure. On the other hand alternative measures of association such as Kendall's tau or Spearman's rho are valid for a much wider class of distributions. For normal distribution both product moments correlation and rank correlation are valid and we may like to compare their performances. Tests for rank correlations will be discussed in the following section.
4.3. **Rank correlation.** Let $X_1, \ldots, X_n$ be a sequence of iid r.v.'s with DF $F(x)$, $x \in \mathbb{R}^2$. We define the rank $R_{\alpha i}$ of $x_{\alpha i}$, $i = 1, 2$ by

$$R_{\alpha i} = \frac{1}{2} \sum_{\beta = 1}^{n} c(x_{\alpha i} - x_{\beta i}),$$

where

$$c(u) = \begin{cases} 
0 & \text{if } u < 0 \\
\frac{1}{2} & \text{if } u = 0 \\
1 & \text{if } u > 0.
\end{cases}$$

If in the sample $\{(x_{\alpha 1}, x_{\alpha 2})\}$, $\alpha = 1, \ldots, n$, all $x_{\alpha 1}$'s and all $x_{\alpha 2}$'s are different, the rank correlation coefficient is given by

$$K' = \frac{12}{n^3 - n} \sum_{\alpha = 1}^{n} \left( R_{\alpha 1} - \frac{n+1}{2} \right) \left( R_{\alpha 2} - \frac{n+1}{2} \right).$$

We have

$$K' = \frac{3}{n^3 - n} \sum_{\alpha = 1}^{n} \sum_{\beta = 1}^{n} \delta(x_{\alpha 1} - x_{\beta 1}) \delta(x_{\alpha 2} - x_{\beta 2})$$

where

$$\delta(u) = 2c(u) - 1 = \begin{cases} 
-1 & \text{if } u < 0 \\
0 & \text{if } u = 0 \\
1 & \text{if } u > 0.
\end{cases}$$

We can write

$$K' = \frac{(n-2) K + 3t}{n+1},$$

where $t = \frac{1}{n(n-1)} \sum_{\alpha \neq \beta} \delta(x_{\alpha 1} - x_{\beta 1}) \delta(x_{\alpha 2} - x_{\beta 2}).$
and $K = \frac{3}{n(n-1)(n-2)} \sum' \delta(x_{\alpha 1} - x_{\beta 1}) \delta(x_{\alpha 2} - x_{\gamma 2})$, and the summation

$\Sigma'$ being over all different subscripts $\alpha$, $\beta$ and $\gamma$. The quantities $K$ and $t$ are U-statistics of degrees 3 and 2 respectively, and

\begin{align*}
\tau &= \mathbb{E}t = \int \int \delta(x_{11} - x_{21}) \delta(x_{12} - x_{22}) \, dF(x_1) \, dF(x_2) \\
\kappa &= \mathbb{E}K = 3 \int \int \delta(x_{11} - x_{21}) \delta(x_{12} - x_{32}) \, dF(x_1) \, dF(x_2) \, dF(x_3)
\end{align*}

Then $\tau$ is the product moment correlation of the difference signs, and $\kappa$ is the grade correlation. In the continuous case $K'$ is a biased estimate of the grade correlation $\kappa$ for finite $n$ but it is unbiased in the limit.

If we write $\theta_1 = \kappa$ and $\theta_2 = t$, then $g(\theta) = [(n-1) \theta_1 + 3\theta_2]/(n+1)$ is a continuous function in $\theta_1$ and $\theta_2$ but it depends on $n$. However, as $n$ becomes large, $\frac{n-1}{n+1} \theta_1$ becomes a dominating factor and in testing hypotheses about $g(\theta)$ we may only consider testing similar hypotheses concerning $\theta_1$. We use the procedure R of Chapter III to test two or three hypotheses concerning $\theta_1$. All moments of the kernel of $x$ exist and the procedure would have a wider range of applicability than for the product moment correlation coefficient.

4.4. Comparison of two variances. Let $X$ and $Y$ be two random variables with variances $\sigma_1^2$ and $\sigma_2^2$ respectively. It is desired to discriminate between the hypotheses

\begin{align*}
\overline{H}_1: \frac{\sigma_1^2}{\sigma_2^2} < 1 & \quad \quad \overline{H}_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 & \quad \quad \overline{H}_1: \frac{\sigma_1^2}{\sigma_2^2} > 1.
\end{align*}
We have $\sigma_1^2$ and $\sigma_2^2$ as two regular functionals where

$$\sigma_1^2 = \frac{1}{2}E(X_1 - X_2)^2$$

and

$$\sigma_2^2 = \frac{1}{2}E(Y_1 - Y_2)^2.$$  

As in the correlation coefficient problem, instead of (4.18) we study the problem of discriminating among the hypotheses

$$(4.19) \begin{align*} 
H_{-1}: \frac{\sigma_1^2}{\sigma_2^2} < 1 - g_1 \\
H_0: 1 - g_1 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 1 + g_1 \\
H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1 + g_1
\end{align*}$$

where $g_1$ is some constant $> 0$. We choose $\Delta > 0$ such that $1 - g_1 - \Delta > 0$ and write $g(\theta) = \frac{\theta}{\theta_2}$. We let $\theta_1 = \sigma_1^2$ and $\theta_2 = \sigma_2^2$ and apply the procedure $R$ of Chapter III to study (4.19).

4.5. **Numerical results.** In Chapter III we gave some bounds for the ASN function of the proposed procedure over some interval of the parameter space. Namely we showed that

$$E(N, \phi|R_1^*) \geq E(N, \phi|R),$$

$$E(N, \phi|R_2^*) \geq E(N, \phi|R),$$

where

$$\Delta^2 E(N, \phi|R_1^*) \rightarrow \frac{b}{\phi + c} \text{ for } \phi + c \leq -1$$

and

$$\Delta^2 E(N, \phi|R_2^*) \rightarrow \frac{c}{\phi - c} \text{ for } \phi - c \geq 1.$$  

We were not able to obtain such upper bounds for $|\phi| \leq c + 1$. However, since the expected value of the maximum of two random variables is greater than or equal to the maximum of their expected values, so we can obtain some lower bounds of $E(N, \phi|R)$ as

$$E(N, \phi|R) \geq \max(E(N, \phi|R_1), E(N, \phi|R_2)).$$
We tested these bounds numerically to see how close they are. We took the correlation coefficient example of section 4.2. We took X to be a bivariate normal with mean vector zero and covariance matrix
\[
\begin{bmatrix}
1 & \rho \\
\rho & 1
\end{bmatrix}.
\]

We used a computer program to generate random numbers with the above distribution. We applied the proposed sequential procedure for different values of \( \Delta, g_1 \) and \( \rho \). In each case the initial sample size was taken to be equal to \( 1/\Delta \). Although our main interest as we developed the procedure was for cases where the two hypothetical values \( g_0 \) and \( g_1 \) and the true value of \( \rho \) are close to each other we also considered some cases where the true value was far from the hypothetical values.

For each case considered we applied the procedure 40 times and took the average of the sample sizes. This average is what is referred to in the tables as "actual." The results are summarized in Tables 1 and 2. Among the cases considered there are two cases where 40 samples were too costly and our results there are based on 25 samples in one and 20 in the other. There are 4 other cases where the expected sample sizes are too large and extensive costs prohibited us from getting results concerning these cases. These cases are marked in the tables by question marks and we hope that we can investigate these cases in a future time. The tables show also the number of cases where a wrong decision was made.

For cases where \( \Delta = .025 \) or \( \Delta = .05 \) the results obtained by simulation are "fairly close" to the expected results. However, we may
need more than just 40 samples in each case to decide how close they are. Also we hope that in the future we may be able to improve upon the expected values to get better approximations.

As for $\Delta = .1$ the number of wrong decisions made is larger than expected. This is more apparent in Tables 3 and 4 where the problem of choosing one of two hypotheses was considered. At first we thought that this may be considered a random fluctuation and that a larger number of samples may reveal different results. However, it was suggested that the difference between actual and expected values is too large. We investigated these cases by taking more samples and we obtained the same type of results. As the expected values were obtained by letting $\Delta$ tend to zero it is our belief that those large differences are due to the fact that $\Delta = .1$ may not be small enough to obtain good approximations.

Another problem now open is to study the ASN function through its moments. In our simulation we took only one case and we used U-statistics estimators. In the future we hope to tackle other cases using U-statistics as well as Von Mises' differentiable statistical functions. Although U-statistics cover a wide range of problems of interest, they leave others that are as much important uncovered. This suggests that other estimators may be studied to see if some justifications could be proved to base similar procedures upon them. However, this may require a completely different approach than we used in proving Wiener process approximations if they hold. Nevertheless they should be considered.
TABLE 1
ASN (IN PAIRS) FOR THE CORRELATION COEFFICIENTS
\(\alpha = \beta = .05, g_0 = -.1\) and \(g_1 = .1\)

<table>
<thead>
<tr>
<th>(\Delta)</th>
<th>(\rho)</th>
<th>Expected</th>
<th>Actual</th>
<th>SE</th>
<th>No. of wrong decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>*588.89</td>
<td>††837.96</td>
<td>41.41</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.025</td>
<td>*3398.78</td>
<td>†††</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>542.72</td>
<td>512.25</td>
<td>45.25</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>*294.44</td>
<td>424.08</td>
<td>32.14</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>*849.69</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>.05</td>
<td>271.36</td>
<td>197.00</td>
<td>23.96</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.3</td>
<td>121.36</td>
<td>85.35</td>
<td>8.73</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>69.25</td>
<td>59.80</td>
<td>7.30</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>41.40</td>
<td>33.98</td>
<td>3.38</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>*147.22</td>
<td>144.63</td>
<td>13.51</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>*212.42</td>
<td>159.13</td>
<td>20.39</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>2</td>
<td>135.68</td>
<td>84.08</td>
<td>13.95</td>
<td>6</td>
</tr>
<tr>
<td>.3</td>
<td>60.96</td>
<td>44.82</td>
<td>6.46</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>33.33</td>
<td>30.38</td>
<td>4.28</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>20.70</td>
<td>22.05</td>
<td>2.46</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

* upper bound
†† based on 25 samples
<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>$\rho$</th>
<th>Expected</th>
<th>Actual</th>
<th>SE</th>
<th>No. of wrong decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>*294.44</td>
<td>360.70</td>
<td>12.18</td>
<td>0</td>
</tr>
<tr>
<td>.025</td>
<td>.1</td>
<td>*530.15</td>
<td>†533.05</td>
<td>45.11</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.2</td>
<td>*3198.97</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>*147.22</td>
<td>197.63</td>
<td>9.40</td>
<td>0</td>
</tr>
<tr>
<td>.1</td>
<td>.2</td>
<td>*799.00</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>.3</td>
<td></td>
<td>243.80</td>
<td>263.73</td>
<td>34.55</td>
<td>0</td>
</tr>
<tr>
<td>.4</td>
<td></td>
<td>103.88</td>
<td>96.88</td>
<td>9.89</td>
<td>0</td>
</tr>
<tr>
<td>.5</td>
<td></td>
<td>55.20</td>
<td>52.85</td>
<td>5.93</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>.1</td>
<td>*144.29</td>
<td>90.35</td>
<td>10.72</td>
<td>6</td>
</tr>
<tr>
<td>.1</td>
<td>.2</td>
<td>*199.75</td>
<td>161.77</td>
<td>28.31</td>
<td>0</td>
</tr>
<tr>
<td>.3</td>
<td></td>
<td>121.90</td>
<td>80.10</td>
<td>10.98</td>
<td>0</td>
</tr>
<tr>
<td>.4</td>
<td></td>
<td>51.94</td>
<td>42.53</td>
<td>6.25</td>
<td>0</td>
</tr>
<tr>
<td>.5</td>
<td></td>
<td>27.60</td>
<td>19.75</td>
<td>2.06</td>
<td>0</td>
</tr>
</tbody>
</table>

* upper bound
† based on 20 samples
### TABLE 3
ASN (IN PAIRS) FOR THE CORRELATION COEFFICIENTS

\( \alpha = \beta = .05, g_0 = -.1 \)

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>( \rho )</th>
<th>Expected</th>
<th>Actual</th>
<th>SE</th>
<th>No. of wrong decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>588.89</td>
<td>622.44</td>
<td>58.27</td>
<td>0</td>
</tr>
<tr>
<td>.025</td>
<td>.1</td>
<td>288.58</td>
<td>302.92</td>
<td>16.35</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.2</td>
<td>180.91</td>
<td>165.68</td>
<td>10.17</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>.1</td>
<td>294.44</td>
<td>310.73</td>
<td>26.36</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.2</td>
<td>144.29</td>
<td>145.60</td>
<td>12.40</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.3</td>
<td>90.45</td>
<td>57.80</td>
<td>5.40</td>
<td>0</td>
</tr>
<tr>
<td>.05</td>
<td>.3</td>
<td>60.96</td>
<td>40.85</td>
<td>3.94</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.4</td>
<td>41.55</td>
<td>35.65</td>
<td>4.07</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.5</td>
<td>33.33</td>
<td>20.10</td>
<td>2.09</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>.1</td>
<td>147.22</td>
<td>118.52</td>
<td>13.69</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>.2</td>
<td>72.15</td>
<td>59.42</td>
<td>12.08</td>
<td>2</td>
</tr>
<tr>
<td>.10</td>
<td>.3</td>
<td>45.23</td>
<td>28.54</td>
<td>4.51</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>.4</td>
<td>30.48</td>
<td>22.78</td>
<td>2.43</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.5</td>
<td>20.76</td>
<td>18.85</td>
<td>1.95</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.5</td>
<td>16.67</td>
<td>14.93</td>
<td>1.31</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^{\dagger\dagger}\text{based on 25 samples}\)

SE = Standard Error
TABLE 4
ASN (IN PAIRS) FOR THE CORRELATION COEFFICIENTS

\[ \alpha = \beta = .05, \ g_0 = -.2 \]

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>( \rho )</th>
<th>Expected</th>
<th>Actual</th>
<th>SE</th>
<th>No. of wrong decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>294.44</td>
<td>292.85</td>
<td>17.68</td>
<td>0</td>
</tr>
<tr>
<td>.025</td>
<td>.1</td>
<td>192.39</td>
<td>188.95</td>
<td>14.76</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.2</td>
<td>135.68</td>
<td>127.00</td>
<td>7.51</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>147.22</td>
<td>152.47</td>
<td>12.09</td>
<td>0</td>
</tr>
<tr>
<td>.1</td>
<td></td>
<td>96.19</td>
<td>78.36</td>
<td>5.09</td>
<td>0</td>
</tr>
<tr>
<td>.05</td>
<td>.2</td>
<td>60.95</td>
<td>56.83</td>
<td>3.93</td>
<td>0</td>
</tr>
<tr>
<td>.3</td>
<td></td>
<td>48.86</td>
<td>44.18</td>
<td>3.57</td>
<td>0</td>
</tr>
<tr>
<td>.4</td>
<td></td>
<td>34.63</td>
<td>28.43</td>
<td>1.75</td>
<td>0</td>
</tr>
<tr>
<td>.5</td>
<td></td>
<td>23.66</td>
<td>27.75</td>
<td>1.66</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>73.61</td>
<td>55.35</td>
<td>7.15</td>
<td>2</td>
</tr>
<tr>
<td>.1</td>
<td></td>
<td>45.23</td>
<td>40.33</td>
<td>5.01</td>
<td>0</td>
</tr>
<tr>
<td>.2</td>
<td></td>
<td>30.48</td>
<td>30.48</td>
<td>2.79</td>
<td>0</td>
</tr>
<tr>
<td>.10</td>
<td>.3</td>
<td>24.38</td>
<td>18.28</td>
<td>1.66</td>
<td>0</td>
</tr>
<tr>
<td>.4</td>
<td></td>
<td>17.32</td>
<td>16.00</td>
<td>1.38</td>
<td>0</td>
</tr>
<tr>
<td>.5</td>
<td></td>
<td>11.83</td>
<td>12.05</td>
<td>.62</td>
<td>0</td>
</tr>
</tbody>
</table>
### TABLE 5

ASN (IN PAIRS) FOR THE CORRELATION COEFFICIENTS

\[ \alpha = \beta = .05, \ g_0 = .1 \]

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>( \rho )</th>
<th>Expected</th>
<th>Actual</th>
<th>SE</th>
<th>No. of wrong decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>588.89</td>
<td>†† 596.96</td>
<td>53.20</td>
<td>0</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>3398.78</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>542.72</td>
<td>288.17</td>
<td>45.25</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0</td>
<td>294.44</td>
<td>266.38</td>
<td>30.30</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>849.71</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>271.36</td>
<td>197.00</td>
<td>23.96</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>121.90</td>
<td>85.55</td>
<td>8.73</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>69.25</td>
<td>59.80</td>
<td>7.30</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>41.98</td>
<td>33.98</td>
<td>4.39</td>
<td>0</td>
</tr>
<tr>
<td>0.10</td>
<td>0</td>
<td>147.22</td>
<td>81.83</td>
<td>10.49</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>212.24</td>
<td>153.00</td>
<td>20.15</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>135.68</td>
<td>79.38</td>
<td>14.22</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>60.96</td>
<td>44.30</td>
<td>2.43</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>34.63</td>
<td>30.38</td>
<td>4.28</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>20.70</td>
<td>22.05</td>
<td>2.46</td>
<td>0</td>
</tr>
</tbody>
</table>

††based on 25 samples
### TABLE 6

ASN (IN PAIRS) FOR THE CORRELATION COEFFICIENTS

\( \alpha = \beta = .05, \theta_0 = .2 \)

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>( \rho )</th>
<th>Expected</th>
<th>Actual</th>
<th>SE</th>
<th>No. of wrong decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>294.44</td>
<td>277.73</td>
<td>14.97</td>
<td>0</td>
</tr>
<tr>
<td>.025</td>
<td>.1</td>
<td>577.16</td>
<td>530.15</td>
<td>45.95</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>.2</td>
<td>3196.01</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>.3</td>
<td>487.60</td>
<td>426.50</td>
<td>32.68</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>.1</td>
<td>147.22</td>
<td>151.08</td>
<td>9.28</td>
<td>0</td>
</tr>
<tr>
<td>.1</td>
<td>.2</td>
<td>288.58</td>
<td>275.73</td>
<td>30.94</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>.3</td>
<td>799.00</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>.05</td>
<td>.3</td>
<td>243.80</td>
<td>263.73</td>
<td>34.55</td>
<td>0</td>
</tr>
<tr>
<td>.4</td>
<td>.4</td>
<td>103.90</td>
<td>96.88</td>
<td>9.89</td>
<td>0</td>
</tr>
<tr>
<td>.5</td>
<td>.5</td>
<td>55.21</td>
<td>52.85</td>
<td>5.93</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>.1</td>
<td>73.61</td>
<td>63.18</td>
<td>8.35</td>
<td>3</td>
</tr>
<tr>
<td>.1</td>
<td>.2</td>
<td>144.29</td>
<td>79.83</td>
<td>11.20</td>
<td>6</td>
</tr>
<tr>
<td>.2</td>
<td>.3</td>
<td>199.75</td>
<td>158.65</td>
<td>28.65</td>
<td>0</td>
</tr>
<tr>
<td>.10</td>
<td>.3</td>
<td>121.90</td>
<td>80.10</td>
<td>10.98</td>
<td>0</td>
</tr>
<tr>
<td>.4</td>
<td>.4</td>
<td>51.95</td>
<td>42.53</td>
<td>6.25</td>
<td>0</td>
</tr>
<tr>
<td>.5</td>
<td>.5</td>
<td>27.60</td>
<td>19.75</td>
<td>2.06</td>
<td>0</td>
</tr>
</tbody>
</table>
REFERENCES


