ASYMPTOTIC RESULTS FOR STOPPING TIMES
BASED ON U-STATISTICS

By

GAUTE VIK

Candidatus Realium

University of Bergen

Bergen, Norway

1978

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1984

ASYMPTOTIC RESULTS FOR STOPPING TIMES

BASED ON U-STATISTICS

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to express special thanks to my major advisor, Dr. Nitis Mukhopadhyay. His professional competence, his comments, ideas, structured way of thinking, and interest and concern have been a major motivation and help for me to carry out my graduate studies.

My thanks also go to Dr. Lyle D. Broemeling, Dr. P. Larry Claypool, Dr. J. Leroy Folks and Dr. Hassan A. Melouk as members of my advisory committee, for their time and effort on my behalf. Special thanks go to Dr. J. Leroy Folks for the effort he put into arranging my stay here at O. S. U.

I also wish to thank the Norwegian Defense Research Establishment and the Royal Norwegian Council for Scientific and Industrial Research for their financial support.

Finally, I am most grateful to my wife, Lisbeth, and my two children for their patience and support to give me the opportunity to fulfill the requirements for a Ph.D.

TABLE OF CONTENTS

TABLE

CHAPTER I

INTRODUCTION AND PRELIMINARIES

1.1 Introduction

Let N be a positive integer valued random variable which says when to stop sampling when a sequential procedure is considered. This N will be referred to as the stopping time and will usually depend on the variability in the population from which we actually sample. Two situations may arise:

(i) No assumptions are made about the distribution from which we sample, other than having requirements for finiteness of certain moments.

(ii) The population is known to be a member of a certain family of distributions.

To estimate the variability, different estimators are usually used in these two situations. In this study, we will first consider a general nonparametric setting which will be useful for situation (ii). Earlier, similar approaches, utilizing Sen's (1960) estimator, had been developed for situation (i). This will become clear from the papers to which we have referred. In order to review the state of the art in the areas of sequential nonparametrics, one may also look at the books by Wald (1947), Puri and Sen (1971), and Sen (1981).

The ground work of this study will be a general stopping time N defined in terms of U-statistics. One major interest will be in deriving many asymptotic properties which are closely related to the asymptotic properties of U-statistics. In the first part of Chapter II we will investigate our proposed stopping time N along the lines of Sproule (1969, 1974), Ghosh and Mukhopadhyay (1979), and Sen and Ghosh (1981). The major finding will be the limiting distribution of the stopping time N, where our tools will be a slight generalization of a result in Ghosh and Mukhopadhyay (1975).

Recently, problems related to rates of convergences for randomly stopped U-statistics have received much attention. Some basic results are due to Landers and Rogge (1976), Ghosh and DasGupta (1980), and Ghosh (1980). In the last part of Chapter II we utilize some of these results and our new extensions to obtain various new rates of convergence. To do this we naturally need to impose conditions of finiteness of some moments.

Csenki (1980) used the results of Landers and Rogge (1976) for suitably normalized randomly stopped means to derive the rate of convergence of the coverage probability for Chow and Robbins' (1965) fixed-width sequential confidence interval procedure. Mukhopadhyay (1981) generalized this result by applying the results of Ghosh and DasGupta (1980) to Sproule's (1969, 1974) sequential procedure. In the first part of Chapter III we apply the same theorem of Ghosh and DasGupta (1980) to our stopping time, and obtain similar results relevant for situation (ii).

Motivated by the results of Mukhopadhyay (1980, 1982), we then define an alternative two-stage procedure in the second part of Chapter III. This procedure to construct a fixed-width confidence interval for the mean of a U-statistic is further compared with the purely sequential procedure discussed previously in Chapter III. In a certain sense, the purely sequential scheme comes out ahead of the proposed two-stage procedure. Our Chapter IV is devoted to some of the more technical proofs, and Chapter V deals with some examples where we consider sampling from Bernoulli (p), Poisson (δ), Gamma (δ,β) and $N(\mu,\sigma^2)$ populations, respectively.

1.2 Notation and Preliminary Results

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables having distribution function $F(\cdot)$, where $F(\cdot)$ belongs to a family G of distribution functions. Let $\phi(X_1,\ldots,X_r)$ be a symmetric kernel of degree r. Now for $n \geq r$, U-statistics are defined as follows:

$$U_n = \binom{n}{r}^{-1} \sum_{n,r} \phi(X_{\alpha_1},\ldots,X_{\alpha_r})$$

where $\sum_{n,r}$ is the summation over all combinations $\{\alpha_1 < \ldots < \alpha_r\}$ formed from the integers $\{1,2,\ldots,n\}$. Furthermore, we define

$$\theta = E\{\phi(X_1,\ldots,X_r)\},$$

$$\phi_c(x_1,\ldots,x_c) = E\{\phi(X_1,\ldots,X_r) \mid X_1 = x_1,\ldots,X_c = x_c\},$$

$$\xi_c = \mathrm{Var}\{\phi_c(X_1,\ldots,X_c)\}, \quad c = 1,2,\ldots,r.$$

As an illustration, we will give two examples on U-statistics. Let first G be a class of distribution functions F, for which the expectation (θ) and the variance (σ^2) exist. That is, $\theta = E_F(X_1)$ and $\sigma^2 = \mathrm{Var}_F(X_1)$ for all $F \in G$. Then

$$U_n = \binom{n}{1}^{-1} \sum_{n,1} X_{\alpha_i} = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$$

is a U-statistic, where $\phi(X_i) = X_i$, $\phi_1(X_1) = X_1$, and $\xi_c = \mathrm{Var}(\phi_1(X_1)) = \sigma^2$.

Next, let G be a class of distributions for which $E_F|X_1|^4$ exists. Consider $\phi(X_i,X_j) = \frac{1}{2}(X_i - X_j)^2$. This is a symmetric kernel of degree 2, and $E_F\{\phi(X_i,X_j)\} = \sigma^2$ for all $F \in G$. Now,

$$U(X_1,\ldots,X_n) = \binom{n}{2}^{-1} \sum_{n,2} \frac{1}{2}(X_{\alpha_1} - X_{\alpha_2})^2$$

$$= \frac{1}{2}\binom{n}{2}^{-1} \sum_{1 \le i < j \le n} (X_i - X_j)^2$$

$$= \frac{1}{(n-1)n}\left(n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i^2 + 2\sum_{1 \le i < j \le n} X_i X_j\right)\right)$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2,$$

which shows that the familiar sample variance is also a U-statistic. Here $\theta = \sigma^2$, with

$$\phi_1(x_1) = E\{\tfrac{1}{2}(x_1 - X_2)^2\} = \tfrac{1}{2}(x_1^2 - 2x_1 EX_2 + EX_2^2), \quad \text{and}$$

$$\xi_1 = \mathrm{Var}\{\phi_1(X_1)\} = \mathrm{Var}\{\tfrac{1}{2}(X_1^2 - 2\mu X_1)\}, \quad \text{where } \mu = EX_2.$$

Now, let us return to the general theory. Hoeffding (1948) proved that $n^{\frac{1}{2}}(U_n - \theta) \xrightarrow{L} N(0, r^2\xi_1)$ as $n \to \infty$, if $\xi_1 > 0$ and $E(\phi^2) < \infty$. For a similar result when the sample size is a random variable, we state the following theorem which is due to Sproule (1969, 1974).

<u>Theorem 1.1</u>: <u>Let $\{n_\nu\}$ be a nondecreasing sequence of positive integers tending to ∞ as $\nu \to \infty$, and let $\{N_\nu\}$ be a sequence of proper positive integer valued random variables. Assume that $\xi_1 > 0$ and $E\{\phi^2\} < \infty$. If</u>

$n_\nu^{-1} N_\nu \xrightarrow{P} 1$ as $\nu \to \infty$, then

$$N_\nu^{\frac{1}{2}}(U_{N_\nu} - \theta) \xrightarrow{L} N(0, r^2 \xi_1) \text{ as } \nu \to \infty.$$

Corollary 1.1: Theorem 1.1 still holds when $\{n_\nu\}$ is a sequence of non-decreasing positive real numbers.

Proof: Let $\varepsilon > 0$ be given, and let ν be big enough such that $n^{-1} < \min(\frac{\varepsilon}{2}, \frac{1}{2})$. Then

$$P\{|N_\nu(n_\nu - 1)^{-1} - 1| > \varepsilon\} = P\{|N_\nu n_\nu^{-1} + n_\nu^{-1} - 1| \cdot |1 - n_\nu^{-1}|^{-1} > \varepsilon\}$$

$$\leq P\{|N_\nu n_\nu^{-1} - 1| > \varepsilon|1 - n_\nu^{-1}| - \frac{\varepsilon}{2}\}$$

$$\leq P\{|N_\nu n_\nu^{-1} - 1| > \eta\}$$

for some $\eta > 0$. Now, since $N_\nu n_\nu^{-1} \leq N_\nu [n_\nu]^{-1} \leq N_\nu (n_\nu - 1)^{-1}$, $N_\nu [n_\nu]^{-1} \xrightarrow{P} 1$, and the corollary then follows from the theorem. (Here [x] means the largest integer smaller than x.)

In the sequel, we write

$$\xi^* = \begin{cases} \xi & \text{if } \xi \geq 1, \\ 2\xi - 1 & \text{if } \frac{1}{2} \leq \xi < 1 \end{cases} \qquad \ldots (1.1)$$

The following theorem is due to Sen and Ghosh (1981).

Theorem 1.2: Assume $E\{|\phi|^{2\xi}\} < \infty$ for some $\xi \geq \frac{1}{2}$. Then for all $n \geq r$, there exists k (> 0) such that

$$E\{|U_n - \theta|^{2\xi}\} \leq k n^{-\xi^*},$$

where k does not depend on n.

Corollary 1.2: *Let* $\{\varepsilon_n\}$ *be a sequence of positive real numbers converging to zero and* c (>0) *be an arbitrary constant. Assume* $E\{|\phi|^{2\xi}\} < \infty$ *for* $\xi \geq \frac{1}{2}$, *and let* α (>0) *be given. Then there exists* k (>0) *such that for all* $n \geq r$, *we have*

$$P\{|U_n^\alpha - \theta^\alpha| > c\varepsilon_n\} \leq k(n\varepsilon_n^2)^{-\xi^*},$$

where k *does not depend on* n, *and* ξ^* *is defined in (1.1)*.

Proof: We will need the following lemma to prove Corollary 1.2.

Lemma 1.1: *Let* α (>0) *and* δ (>0) *be constants, and let* $x \in (0,\infty)$. *For any* θ (>0) *we can find a constant* k (>0) *such that* $|x^\alpha - \theta^\alpha| > \varepsilon$ *implies that* $k|x-\theta| > \varepsilon$ *for all* $\varepsilon \in (0,\delta)$. *Here* k *may depend on* θ *and* δ, *but not on* ε.

The lemma is readily proved by the help of the Taylor expansion of x^α. Now, a proof of Corollary 1.2 can be established by using Lemma 1.1 and Theorem 1.2.

The following well-known Lemmas 1.2 and 1.3 are proved in Michel and Pfanzagl (1971). Let $\Phi(\cdot)$ be the distribution function of a standard normal random variable. For brevity we write only $\sup_x$ instead of $\sup_{x\in(-\infty,\infty)}$.

Lemma 1.2: *Let* $\{X_n\}$ *and* $\{Y_n\}$ *be sequences of random variables. Assume that for a sequence* $\{a_n\}$ *of real numbers tending to zero*, $\sup_x|P\{X_n \leq x\} - \Phi(x)| = 0(a_n)$, *and* $P\{|Y_n - 1| > a_n\} = 0(a_n)$. *Then we have*

$$\sup_x|P\{X_n \leq xY_n\} - \Phi(x)| = 0(a_n).$$

Lemma 1.3: *Let* $\{X_n\}$, $\{Y_n\}$ *and* $\{a_n\}$ *be as in Lemma 1.2. If*

$\sup_x |P\{X_n \le x\} - \Phi(x)| = O(a_n)$ and $P\{|Y_n| > a_n\} = O(a_n)$, then $\sup_x |P\{X_n + Y_n \le x\} - \Phi(x)| = O(a_n)$.

The following lemma serves as a tool in proving some of the results of section 2.

Lemma 1.4: *Let $\{Y_n\}$ be a sequence of random variables where $P\{Y_n = 0\} = 0$ for every n, and $\{\delta_n\}$ and $\{\varepsilon_n\}$ be sequences of real numbers tending to zero. If for arbitrary c_1 (> 0) we have $P\{|Y_n - 1| > c_1\varepsilon_n\} = O(\delta_n)$, then we can conclude that $P\{|Y_n^{-1} - 1| > c_2\varepsilon_n\} = O(\delta_n)$, where c_2 (> 0) is also arbitrary.*

Proof: Let n be large enough such that $c_1\varepsilon_n < \frac{1}{2}$. Then, $|Y_n - 1| < c_1\varepsilon_n$ implies that

$$|Y_n^{-1} - 1| < c_1\varepsilon_n|Y_n|^{-1} < c_1\varepsilon_n(1 - \varepsilon_n c_1)^{-1} < 2c_1\varepsilon_n.$$

Thus, for $c_2 = 2c_1$,

$$P\{|Y_n^{-1} - 1| > c_2\varepsilon_n\} \le P\{|Y_n - 1| > \tfrac{1}{2}c_2\varepsilon_n\} = O(\delta_n).$$

Now, c_2 (> 0) can be chosen arbitrarily since c_1 is arbitrary (> 0).

Remark 1.1: The Lemma 1.4 is easily generalized to the following: If $P\{|Y_n - \theta| > c_1\varepsilon_n\} = O(\delta_n)$ holds for a constant $\theta \ne 0$, then $P\{|Y_n^{-1} - \theta^{-1}| > c_2\varepsilon_n\} = O(\delta_n)$ holds under the conditions of Lemma 1.4.

CHAPTER II

A GENERAL STOPPING TIME AND ITS ASYMPTOTIC PROPERTIES

2.1 Definition of N_ν

Assume the following:

$\{U_n, n \geq r\}$ is a sequence of U-statistics and $U_n \geq 0$ w.p. 1;

$E\{U_n\} = \theta < \infty$ and $E\{\phi^2\} < \infty$;

$\alpha\ (>0)$, $\gamma\ (>0)$, $\delta\ (>0)$ and $t\ (\geq 0)$ are all constants;

$\{\psi_\nu, \nu = 1,2,\ldots\}$ is a sequence of nondecreasing real numbers (>0) where $\psi_\nu \to \infty$ as $\nu \to \infty$.

Now, we define our stopping time as

$$N_\nu = \inf\{n \geq n_o\ ;\ (n+t)^\delta \geq \psi_\nu (U_n^\alpha + n^{-\gamma})\}, \qquad \ldots (2.1)$$

$\nu = 1,2,\ldots$; n_o $(\geq r)$ is an integer and can be thought of as the starting sample size.

2.2 Some Important Convergence Results for N_ν

Some of the important convergence results are stated in the next four lemmas. The following lemma easily follows from the definition of N_ν in (2.1).

Lemma 2.1: *For the stopping time N_ν defined in (2.1), we have*

$P\{N_\nu < \infty\} = 1$ for $\nu = 1,2,\ldots;$

$N_{\nu_1} \le N_{\nu_2}$ w.p. 1 if $\nu_1 \le \nu_2;$

$N_\nu \to \infty$ w.p. 1 as $\nu \to \infty$.

Proof: From the definition (2.1) of N_ν, we get

$$P\{N_\nu = \infty\} = \lim_{n\to\infty} P\{N_\nu > n\}$$

$$\le \lim_{n\to\infty} P\{((n+t)^\delta \psi_\nu^{-1} - n^{-\gamma})^{\frac{1}{\alpha}} < U_n\}$$

$$= 0$$

since $((n+t)^\delta \psi_\nu^{-1} - n^{-\gamma})^{\frac{1}{\alpha}} \to \infty$ and $U_n \xrightarrow{P} \theta$ as $n \to \infty$. This proves the first part of the lemma. Next, for the second part of the lemma, $\nu_1 \le \nu_2$ implies that $\psi_{\nu_1}(U_n^\alpha + n^{-\gamma}) \le \psi_{\nu_2}(U_n^\alpha + n^{-\gamma})$. Now N_{ν_2} is one integer $\ge n_o$ such that $(n+t)^\delta \ge \psi_{\nu_1}(U_n^\alpha + n^{-\gamma})$. N_{ν_1} is the infimum of all such integers, which gives $N_{\nu_1} \le N_{\nu_2}$ w.p. 1. For the last part of the lemma, we will prove the almost sure convergence of N_ν by showing that $\lim_{\nu_1\to\infty} P\{\operatorname{Sup}_{\nu\le\nu_1} N_\nu \le k\} = 0$ for all $k > 0$. Utilizing the increasing property of N_ν and the definition (2.1), we can conclude the following:

$$\lim_{\nu_1\to\infty} P\{\operatorname{Sup}_{\nu\le\nu_1} N_\nu \le k\}$$

$$= \lim_{\nu_1\to\infty} P\{N_{\nu_1} \le k\}$$

$$= \sum_{i=n_o}^{[k]} \lim_{\nu_1\to\infty} P\{N_{\nu_1} = i\}$$

$$\le \sum_{i=n_o}^{[k]} \lim_{\nu_1\to\infty} P\{(i+t)^\delta \ge \psi_{\nu_1}(U_i^\alpha + i^{-\gamma})\}$$

$= 0.$

The following explains the last step: There exists an integer $\nu'(k)$ such that $\psi_{\nu_1} > (k+t)^{\delta} k^{\gamma}$ for all $\nu_1 > \nu'(k)$. Then, $(i+t)^{\delta} < \psi_{\nu_1} i^{-\gamma}$ for all $\nu_1 > \nu'(k)$, $i = n_o, \ldots, k$. Since $U_i \geq 0$ w.p. 1, this implies $(i+t)^{\delta} < \psi_{\nu_1}(U_i^{\alpha} + i^{-\gamma})$ for all $\nu_1 > \nu'(k)$, $i = n_o, \ldots, k$ and this gives the last equality

<u>Lemma 2.2</u>: <u>For the stopping time N_ν defined in (2.1), we have</u>

$$U_{N_\nu} \to \theta \quad \underline{\text{w.p.}}\ 1 \quad \underline{\text{as}} \quad \nu \to \infty;$$

$$U_{N_\nu - 1} \to \theta \quad \underline{\text{w.p.}}\ 1 \quad \underline{\text{as}} \quad \nu \to \infty.$$

<u>Proof</u>: Let a subset of the sample space be defined as $A = \{\lim_{n\to\infty} U_n = \theta,$ N_ν is non-decreasing in ν, $\lim_{\nu\to\infty} N_\nu = \infty\}$. Consider an element $w \in A$. For this w, the following arguments will hold: For all $\varepsilon > 0$, there exists an integer $n'(\varepsilon) > 0$ such that $|U_n - \theta| < \varepsilon$ for all $n > n'(\varepsilon)$. Also there exists a $\nu'(\varepsilon) > 0$ such that $P\{N_\nu \leq n'(\varepsilon)\} = 0$ for all $\nu > \nu'(\varepsilon)$. That is, $|U_{N_\nu} - \theta| < \varepsilon$ for all $\nu > \nu'(\varepsilon)$ which shows that $\lim_{\nu\to\infty} U_{N_\nu} = \theta$ for the fixed w. Since w was chosen arbitrarily in A, $\lim_{\nu\to\infty} U_{N_\nu} = \theta$ for all $w \in A$. Hoeffding (1961) showed that $U_n \to \theta$ w.p. 1, which together with Lemma 2.1 prove that $P\{A\} = 1$. This concludes the proof of the first part of the lemma. The second part follows by the same type of arguments.

<u>Lemma 2.3</u>: <u>For the stopping time N_ν defined in (2.1), we have</u>

$$N_\nu / (\psi_\nu^{1/\delta}\, \theta^{\alpha/\delta}) \to 1 \quad \underline{\text{w.p.}}\ 1 \quad \underline{\text{as}} \quad \nu \to \infty.$$

<u>Proof</u>: From the stopping time (2.1) we obtain the basic inequality,

namely,

$$\psi_\nu^{1/\delta}\ (U_{N_\nu}^{\alpha} + N_\nu^{-\gamma})^{1/\delta} - t \le N_\nu \le \psi_\nu^{1/\delta}\ (U_{N_\nu - 1}^{\alpha} + (N_\nu - 1)^{-\gamma})^{1/\delta} + (n_o + t - 1). \quad \ldots (2.2)$$

Now, dividing all over by $\psi_\nu^{1/\delta}\ \theta^{\alpha/\delta}$ and applying Lemmas 2.1 and 2.2, we can complete the proof of Lemma 2.3.

Lemma 2.4: Consider the stopping time N_ν defined in (2.1). If $E\{|\phi|^{2\xi}\} < \infty$ for some $\xi > m \ge 1$, then we have

$$\lim_{\nu \to \infty} E\{(N_\nu / (\psi_\nu^{1/\delta}\ \theta^{\alpha/\delta}))^m\} = 1.$$

A proof of this lemma is given in section 4.1.

2.3 Asymptotic Distribution of the Stopping Time N_ν

In this section, our main tool will be a slight generalization of a result in Ghosh and Mukhopadhyay (1975), stated as our Theorem 2.1. Theorem 2.2 and Corollary 2.1 are both used to show that our stopping time N_ν satisfies the assumptions of Theorem 2.1. An application of Theorem 2.1 then gives us our main result of this section which is stated as Theorem 2.3.

Theorem 2.1: Consider a sequence $\{N_\nu^*\}$ of positive integer valued random variables defined by $N_\nu^* = \inf\{n \ge n_o;\ n \ge b_\nu T_n - t\}$ where $n_o \ge 1$, $\{b_\nu\}$ is a sequence of positive real numbers tending to infinity as $\nu \to \infty$, $\{T_n;\ n \ge n_o\}$ is a sequence of statistics such that $P\{T_n > 0\} = 1$ for all $n \ge n_o$, and $t \in (-\infty, \infty)$ is a constant. Now, if for all $x \in (-\infty, \infty)$,

$$\lim_{\nu\to\infty} P\{N_\nu^{*\frac{1}{2}}(T_{N_\nu^*} - c) \le d\,x\} = \Phi(x), \text{ and}$$

$$\lim_{\nu\to\infty} P\{N_\nu^{*\frac{1}{2}}(T_{N_\nu^*-1} - c) \le d\,x\} = \Phi(x)$$

for some constants c (> 0) _and_ d (> 0), _then we have_

$$\lim_{\nu\to\infty} P\{c^{\frac{1}{2}}(N_\nu^* - cb_\nu) \le x\,d\,b_\nu^{\frac{1}{2}}\} = \Phi(x) \text{ for all } x \in (-\infty,\infty).$$

If we use $t = 0$ in the definition of N_ν^*, then we end up with the situation as in Ghosh and Mukhopadhyay (1975). The proof for the case $t \neq 0$ can be given along the same lines as in Ghosh and Mukhopadhyay (1975).

The following theorem is a slightly modified version of a theorem stated in Rao (1973), and this can be proved in a similar way.

Theorem 2.2: _Let_ $\{Y_n\}$ _be a sequence of positive random variables such that_ $Y_n^{-1} \xrightarrow{P} 0$. _Let_ $g(\cdot)$ _be a real valued function defined on_ $(-\infty,\infty)$ _and assume_ $g'(\cdot)$ _exists_. _Let_ $\{T_n\}$ _be a sequence of random variables, and_ σ (> 0) _and_ $c \in (-\infty,\infty)$ _be constants_. _Now if_ $Y_n(T_n - c) \xrightarrow{L} N(0,\sigma^2)$ _as_ $n \to \infty$, _then we have_

$$Y_n(g(T_n) - g(c)) \xrightarrow{L} N(0,(\sigma g'(c))^2) \text{ as } n \to \infty, \text{ if } g'(c) \neq 0.$$

In Rao (1973) this result is given for $Y_n = n^{\frac{1}{2}}$. We have stated and proved a version giving the rate of convergence in our Theorem 2.5.

Corollary 2.1: _Let_ N_ν _be as in_ (2.1) _and_ $g(\cdot)$ _be as in Theorem 2.2_. _Now if_ $N_\nu^{\frac{1}{2}}(T_{N_\nu} - c) \xrightarrow{L} N(0,\sigma^2)$ _as_ $\nu \to \infty$, _then we have_

$$N_\nu^{\frac{1}{2}}(g(T_{N_\nu}) - g(c)) \xrightarrow{L} N(0,(\sigma g'(c))^2) \text{ as } \nu \to \infty.$$

Proof: The result follows immediately, recalling the properties of N_ν

given in Lemma 2.1.

Theorem 2.3: *Assume $\gamma > \frac{1}{2}$ in the definition of N_ν given in (2.1). Then we have*

$$(c/b_\nu)^{\frac{1}{2}} (N_\nu - cb_\nu) \xrightarrow{L} N(0,d^2),$$

as $\nu \to \infty$, *where* $c = \theta^{\alpha/\delta}$, $b_\nu = \psi_\nu^{1/\delta}$ *and* $d = r\xi_1^{\frac{1}{2}}\alpha\delta^{-1}\theta^{\alpha/\delta-1}$.

Proof: Our stopping time (2.1) can be written as $N_\nu = \inf\{n \geq n_o;\ n \geq \psi_\nu^{1/\delta}(U_n^\alpha + n^{-\gamma})^{1/\delta} - t\}$, which shows that N_ν is of the form considered in Theorem 2.1 with $T_n = (U_n^\alpha + n^{-\gamma})^{1/\delta}$, $b_\nu = \psi_\nu^{1/\delta}$. We will show that the assumptions in Theorem 2.1 are satisfied by this N_ν. From Corollary 1.1 we get $\{N_\nu^{\frac{1}{2}}(U_{N_\nu} - \theta)/(r\xi_1^{\frac{1}{2}})\} \xrightarrow{L} N(0,1)$ as $\nu \to \infty$, and thus by Corollary 2.1 we obtain

$$\{N_\nu^{\frac{1}{2}}(U_{N_\nu}^\alpha - \theta^\alpha)/(r\xi_1^{\frac{1}{2}}\alpha\theta^{\alpha-1})\} \xrightarrow{L} N(0,1) \text{ as } \nu \to \infty. \text{ Furthermore,}$$

$$\{N_\nu^{\frac{1}{2}}(U_{N_\nu}^\alpha + N_\nu^{-\gamma} - \theta^\alpha)/(r\xi_1^{\frac{1}{2}}\alpha\theta^{\alpha-1})\} \xrightarrow{L} N(0,1) \text{ as } \nu \to \infty, \text{ since}$$

$$\{N_\nu^{-\gamma}\ N_\nu^{\frac{1}{2}}/(r\xi^{\frac{1}{2}}\alpha\theta^{\alpha-1})\} \xrightarrow{P} 0 \text{ as } \nu \to \infty \text{ for } \gamma > \tfrac{1}{2}.$$

Applying the Corollary 1.1 once more, we now obtain

$$\{N_\nu^{\frac{1}{2}}((U_{N_\nu}^\alpha + N_\nu^{-\gamma})^{1/\delta} - \theta^{\alpha/\delta})/(r\xi_1^{\frac{1}{2}}\alpha\theta^{\alpha-1}\ \delta^{-1}\ \theta^{\alpha(1/\delta-1)})\} \xrightarrow{L} N(0,1), \text{ as}$$

$\nu \to \infty$, which actually translates to $\{N_\nu^{\frac{1}{2}}(T_{N_\nu} - c)/d\} \xrightarrow{L} N(0,1)$ as $\nu \to \infty$. Now, $\{(N_\nu - 1)/(\psi_\nu^{1/\delta}\ \theta^{\alpha/\delta})\} \xrightarrow{P} 1$ as $\nu \to \infty$, and using the same type of arguments, we conclude that

$$\{(N_\nu - 1)^{\frac{1}{2}}\ (T_{N_\nu - 1} - c)/d\} \xrightarrow{L} N(0,1) \text{ as } \nu \to \infty. \text{ Then we also get}$$

$$\{N_\nu^{\frac{1}{2}}(T_{N_\nu - 1} - c)/d\} \xrightarrow{L} N(0,1) \text{ as } \nu \to \infty, \text{ since}$$

$$\{N_\nu/(N_\nu - 1)\}^{\frac{1}{2}} \xrightarrow{P} 1 \text{ as } \nu \to \infty.$$

Now, Theorem 2.1 can be applied to conclude the desired result.

2.4 Rates of Convergences for the Stopping Time N_ν

The structure of this section will be the same as it was in Section 2.3. We first state a theorem that will serve as our main tool (Theorem 2.4). This gives the corresponding rate of convergence results for Theorem 2.1 and it is a slight generalization of a result of Ghosh (1980). Theorem 2.5 gives us the rate of convergence for Theorem 2.2. The next four theorems and the corollary give us more insight about N_ν. The main purpose of having these results, however, is that they enable us to show in the end that our stopping time N_ν satisfies the assumptions of Theorem 2.4. This in turn gives us Theorem 2.9 which is the main result of this section.

Theorem 2.4: *Let N_ν^*, b_ν and T_n be defined as in Theorem 2.1 and let $\{\varepsilon_\nu\}$ be a sequence of positive real numbers such that $\varepsilon_\nu \geq b_\nu^{-1}$. Now, if*

$$P\{|N_\nu^*(cb_\nu)^{-1} - 1| > \varepsilon_\nu\} = O(\varepsilon_\nu^{\frac{1}{2}}),$$

$$\sup_x |P\{N_\nu^{*\frac{1}{2}} (T_{N_\nu^*} - c) \leq d\,x\} - \Phi(x)| = O(\varepsilon_\nu^{\frac{1}{2}}), \text{ and}$$

$$\sup_x |P\{N_\nu^{*\frac{1}{2}} (T_{N_\nu^*-1} - c) \leq d\,x\} - \Phi(x)| = O(\varepsilon_\nu^{\frac{1}{2}}),$$

for some constants c (>0) *and* d (>0), *then we have*

$$\sup_x |P\{c^{\frac{1}{2}} (N_\nu^* - cb_\nu) \leq xdb_\nu^{\frac{1}{2}}\} - \Phi(x)| = O(\varepsilon_\nu^{\frac{1}{2}}).$$

This slight generalization does not call for any special difficulty in the proof, and thus the proof can follow the same lines as in Ghosh (1980). We omit any further details on this.

In the next theorem we will consider a real-valued function $g(\cdot)$ having the following properties:

(i) $g(\cdot)$ is defined on (a,b), $-\infty \leq a < b \leq \infty$,

(ii) $g(\cdot)$ is continuous,

(iii) $g'(c) \neq 0$ for some $c \in (a,b)$,

(iv) $g''(\cdot)$ exists and is bounded in a neighborhood around c.

Theorem 2.5: *Assume that a function $g(\cdot)$ has all the properties (i) - (iv) for an interval (a,b) and a constant $c \in (a,b)$. Let $\{Y_n\}$ be a sequence of positive random variables and let $\{\varepsilon_n\}$ be a sequence of decreasing positive real numbers. Furthermore, let $\{T_n\}$ be a sequence of random variables taking on values in (a,b). Now, if*

$$\sup_x |P\{Y_n(T_n - c) \leq \sigma x\} - \Phi(x)| = 0(\varepsilon_n), \text{ for some } \sigma > 0, \text{ and}$$

$$P\{Y_n(T_n - c)^2 > k\varepsilon_n\} = 0(\varepsilon_n),$$

for every $k > 0$, then we have

$$\sup_x |P\{Y_n(g(T_n) - g(c)) \leq x\sigma g'(c)\} - \Phi(x)| = 0(\varepsilon_n).$$

Proof: Let $R_n(y,c)$ be defined through

$$g(y) - g(c) = (y - c)(g'(c) + R_n(y,c)) \text{ for } y \neq c, \text{ and}$$

$R_n(c,c) = 0$. Then, we can write the following:

$$\frac{Y_n(g(T_n) - g(c))}{\sigma g'(c)} = \frac{Y_n(T_n - c)}{\sigma} + Y_n(T_n - c)R_n(T_n,c). \quad \ldots (2.3)$$

Using Taylor's expansion and the property (iv), one can see that there exists $\delta > 0$ such that for all $y \in [c-\delta, c+\delta]$, we can write

$$|R_n(y,c)| = |\tfrac{1}{2}g''(z)(y-c)| \leq k|y-c|$$

for some $k > 0$ and z lying between y and c. Now, we obtain

$$P\{|Y_n(T_n - c)R_n(T_n,c)| > \varepsilon_n\}$$

$$\leq P\{Y_n(T_n - c)^2 k > \varepsilon_n, |T_n - c| < \delta\}$$

$$+ P\{|Y_n(T_n - c)R_n(T_n,c)| > \varepsilon_n, |T_n - c| \geq \delta\},$$

$$\leq P\{Y_n(T_n - c)^2 > k^{-1} \varepsilon_n\} + P\{|T_n - c| \geq \delta\}.$$

Also,

$$P\{|T_n - c| \geq \delta\} = P\{Y_n(T_n - c)^2 \geq \delta|Y_n(T_n - c)|\}$$

$$\leq P\{Y_n(T_n - c)^2 \geq \delta\sigma\varepsilon_n, |Y_n(T_n - c)| \geq \varepsilon_n\sigma\}$$

$$+ P\{|Y_n(T_n - c)| < \varepsilon_n\sigma\}$$

$$= O(\varepsilon_n) + P\{|Y_n(T_n - c)| < \varepsilon_n\sigma\}.$$

Again, we have

$$P\{|Y_n(T_n - c)| < \varepsilon_n\sigma\}$$

$$= P\{Y_n(T_n - c) < \varepsilon_n\sigma\} - P\{Y_n(T_n - c) < -\varepsilon_n\sigma\}$$

$$\leq |P\{Y_n(T_n - c) < \varepsilon_n\sigma\} - \Phi(\varepsilon_n)|$$

$$+ |P\{Y_n(T_n - c) < -\varepsilon_n\sigma\} - \Phi(-\varepsilon_n)| + |\Phi(\varepsilon_n) - \Phi(-\varepsilon_n)|$$

$$= O(\varepsilon_n).$$

Hence, Theorem 2.5 now follows by an application of Lemma 1.3 and Equation (2.3).

Theorem 2.6: *Consider the stopping time* N_ν *defined in (2.1) and let* $\gamma \geq \frac{1}{2}$. *Also let* $\lambda \in (0,\frac{1}{2})$. *Now if* $E\{|\phi|^{2\xi}\} < \infty$ *for* $\xi \geq \max\{(5-2\lambda)/8\lambda, 1+(\delta+\gamma)(1-2\lambda)/4\delta\}$, *then we have*

$$P\{|N_\nu n_\nu^{-1} - 1| > c\psi_\nu^{-(\frac{1}{2}-\lambda)/\delta}\} = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}),$$

for arbitrary c (> 0) *and* $n_\nu = \psi_\nu^{1/\delta}\theta^{\alpha/\delta}$.

The proof of this theorem is given in Section 4.2. The following corollary is an immediate consequence of Theorem 2.6 and Lemma 1.4.

Corollary 2.2: *Consider the situation of Theorem 2.6. Then we have,*

$$P\{|n_\nu N_\nu^{-1} - 1| > c\psi_\nu^{-(\frac{1}{2}-\lambda)/\delta}\} = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}),$$

for arbitrary c (> 0).

The following theorem is due to Ghosh and DasGupta (1980). We state this result here for completeness.

Theorem 2.7: *Let* $\{\varepsilon_\nu\}$ *be a sequence of positive real numbers and* τ (> 0) *be a constant. Consider a sequence* $\{N_\nu^*\}$ *of positive integer valued random variables such that*

$P\{|N_\nu^*(\tau\nu)^{-1} - 1| > \varepsilon_\nu\} = O(\varepsilon_\nu^{\frac{1}{2}})$. *Now if* $E\{|\phi|^{2\xi}\} < \infty$ *for*

$\xi \geq 2$, $1 > \varepsilon_\nu > \nu^{-2\xi/(2\xi+1)}$, then we have

$$\sup_x \left|P\{(\tau\nu)^{\frac{1}{2}} (U_{N^*_\nu} - \theta) \leq xr\xi_1^{\frac{1}{2}}\} - \Phi(x)\right| = O(\varepsilon_\nu^{\frac{1}{2}}), \text{ and}$$

$$\sup_x \left|P\{N^{*\frac{1}{2}} (U_{N^*_\nu} - \theta) \leq xr\xi_1^{\frac{1}{2}}\} - \Phi(x)\right| = O(\varepsilon_\nu^{\frac{1}{2}}).$$

The next theorem is a slight modification of this result in our context.

Theorem 2.8: Let N_ν be the stopping time defined in (2.1) for $\gamma \geq \frac{1}{2}$. Let $\lambda \in (0,\frac{1}{2})$. Now if $E\{|\phi|^{2\xi}\} < \infty$ for $\xi \geq \max\{(5-2\lambda)/8\lambda, 1+(\delta+\gamma)(1-2\lambda)/4\delta, 2\}$, then we have

$$\sup_x \left|P\{n_\nu^{\frac{1}{2}} (U_{N_\nu} - \theta) \leq xr\xi_1^{\frac{1}{2}}\} - \Phi(x)\right| = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}), \quad \ldots (2.4)$$

$$\sup_x \left|P\{N_\nu^{\frac{1}{2}} (U_{N_\nu} - \theta) \leq xr\xi_1^{\frac{1}{2}}\} - \Phi(x)\right| = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}), \quad \ldots (2.5)$$

$$\sup_x \left|P\{(N_\nu - 1)^{\frac{1}{2}} (U_{N_\nu-1} - \theta) \leq xr\xi_1^{\frac{1}{2}}\} - \Phi(x)\right| = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}), \ldots (2.6)$$

$$\sup_x \left|P\{N_\nu^{\frac{1}{2}} (U_{N_\nu-1} - \theta) \leq xr\xi_1^{\frac{1}{2}}\} - \Phi(x)\right| = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}), \quad \ldots (2.7)$$

where $n_\nu = \psi_\nu^{1/\delta}\theta^{\alpha/\delta}$.

A proof of this theorem is given in Section 4.3.

Remark 2.1: In Theorem 2.7, no formal relationship between N_ν^* and U_n is really needed. This is also the case for Theorem 2.8. That is, the Theorems 2.7 and 2.8 remain true for U_n not necessarily being the same U-statistic as the one occurring in the definition of N_ν. We will utilize this fact in Chapter III.

Now, the results given in this section so far enable us to apply Theorem 2.4 for our stopping time in (2.1). So, in the sequel we in fact show that under certain conditions, N_ν satisfies the assumptions of Theorem 2.4. The final result is given in the following theorem which is our main finding in this section.

Theorem 2.9: Consider N_ν as defined in (2.1), and assume that $\gamma \geq \frac{3}{4}$. Let $\lambda \in (0,\frac{1}{2})$. If $E\{|\phi|^{2\xi}\} < \infty$ for

$\xi \geq \max\{(5-2\lambda)/(1+2\lambda)\ ,\ (5-2\lambda)/8\lambda\ ,\ 1+(\delta+\gamma)(1-2\lambda)/4\delta\}$, then

$$\sup_x \ |P\{c^{\frac{1}{2}}\ (N_\nu - cb_\nu) \leq xdb_\nu^{\frac{1}{2}}\} - \Phi(x)| = O(\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}),$$

where $c = \theta^{\alpha/\delta}$, $b_\nu = \psi_\nu^{1/\delta}$ and $d = r\xi_1^{\frac{1}{2}}\ \alpha\delta^{-1}\ \theta^{\alpha/\delta-1}$.

Proof: As in the proof of Theorem 2.3, the stopping time N_ν from (2.1) can be defined by

$$N_\nu = \inf\{n \geq n_o\ ;\ n \geq b_\nu T_n - t\}$$

where $b_\nu = \psi_\nu^{1/\delta}$ and $T_n = (U_n^\alpha + n^{-\gamma})^{1/\delta}$. Let $\varepsilon_\nu = \psi_\nu^{-(\frac{1}{2}-\lambda)/\delta}$. First we show that

$$\sup_x \ |P\{N_\nu^{\frac{1}{2}}\ (U_{N_\nu}^\alpha - \theta^\alpha) \leq xr\xi_1^{\frac{1}{2}}\ \alpha\theta^{\alpha-1}\} - \Phi(x)| = O(\varepsilon_\nu^{\frac{1}{2}}). \qquad \text{... (2.8)}$$

From Theorem 2.6, we have

$$P\{N_\nu \leq n_\nu(1-\eta)\} = O(\varepsilon_\nu^{\frac{1}{2}}), \qquad \text{... (2.9)}$$

where $n_\nu = \psi_\nu^{1/\delta}\theta^{\alpha/\delta}$, $0 < \eta < 1$. Let $a = [n_\nu(1-\eta)]+1$, where [x] is the largest integer smaller than x. For brevity, we will write k for a generic positive constant independent of ν. Then, we can write

$$P\{N_\nu^{\frac{1}{2}}(U_{N_\nu}-\theta)^2 > k\varepsilon_\nu^{\frac{1}{2}}\}$$

$$\leq \sum_{n=a}^{\infty} P\{n^{\frac{1}{2}}(U_n-\theta)^2 > k\varepsilon_\nu^{\frac{1}{2}}, N_\nu = n\} + P\{N_\nu \leq n_\nu(1-\eta)\}$$

$$\leq \sum_{n=a}^{\infty} P\{|U_n-\theta| > k(\varepsilon_\nu n^{-1})^{\frac{1}{4}}\} + 0(\varepsilon_\nu^{\frac{1}{2}})$$

$$\leq \sum_{n=a}^{\infty} k(n_\nu/\varepsilon_\nu)^{\xi/2} n^{-\xi} + 0(\varepsilon_\nu^{\frac{1}{2}})$$

$$\leq k\varepsilon_\nu^{-\xi/2} a^{-(\xi/2-1)}$$

$$= k\psi_\nu^{\xi(1-2\lambda)/4} \psi_\nu^{-(\xi-2)/2\delta}, \qquad \ldots (2.10)$$

if $\xi > 2$. This is of order $\varepsilon_\nu^{\frac{1}{2}}$ if $\xi \geq (5-2\lambda)/(1+2\lambda)$.

Theorem 2.5 now gives (2.8). Next, we have

$$P\{N_\nu^{\frac{1}{2}} N_\nu^{-\gamma} > \varepsilon_\nu^{\frac{1}{2}} r\xi_1^{\frac{1}{2}} \alpha\theta^{\alpha-1}\}$$

$$= P\{N_\nu < k\varepsilon_\nu^{1/(1-2\gamma)}\}$$

$$= 0(\varepsilon_\nu^{\frac{1}{2}}),$$

if $\{(1-2\lambda)/(4\gamma-2)\} < 1$, which follows from (2.9) since $k\varepsilon_\nu^{1/(1-2\gamma)} < n_\nu(1-\eta)$ for "big enough" ν in this case. Lemma 1.3 now implies that $U_{N_\nu}^\alpha$ can be replaced by $(U_{N_\nu}^\alpha + N_\nu^{-\gamma})$ in equation (2.8). Now,

$$P\{N_\nu^{\frac{1}{2}}(U_{N_\nu}^\alpha + N_\nu^{-\gamma} - \theta^\alpha)^2 < k\varepsilon_\nu^{\frac{1}{2}}\}$$

$$= P\{|U_{N_\nu}^\alpha + N_\nu^{-\gamma} - \theta^\alpha| > k(\varepsilon_\nu/N_\nu)^{\frac{1}{4}}\}$$

$$\leq P\{|U_{N_\nu}^\alpha - \theta^\alpha| > k(\varepsilon_\nu/N_\nu)^{\frac{1}{4}}\} + P\{N_\nu^{-\gamma} > k(\varepsilon_\nu/N_\nu)^{\frac{1}{4}}\}.$$

Note that

$$P\{N_\nu^{-\gamma} > k(\varepsilon_\nu/N_\nu)^{\frac{1}{4}}\}$$

$$= P\{N_\nu^{8\gamma-2} < k\psi_\nu^{(1-2\gamma)/\delta}\}$$

$$= 0(\varepsilon_\nu^{\frac{1}{2}}),$$

for $8\gamma - 2 \geq 1$. This follows from (2.9) since $k\psi_\nu^{(1-2\lambda)/\delta} < n_\nu(1-\eta)$ for "big enough" ν. Next we have

$$P\{|U_{N_\nu}^\alpha - \theta^\alpha| > k(\varepsilon_\nu/N_\nu)^{\frac{1}{4}}\}$$

$$\leq \sum_{n=a}^{\infty} P\{|U_n^\alpha - \theta^\alpha| > k(\varepsilon_\nu/n)^{\frac{1}{4}}\} + P\{N_\nu < a\}$$

$$= 0(\varepsilon_\nu^{\frac{1}{2}}),$$

if $\xi \geq (5-2\lambda)/(1+2\lambda)$. This follows from Corollary 1.2 and the same types of arguments used to obtain the bound for (2.10). From Theorem 2.5 it now follows that

$$\sup_x |P\{N_\nu^{\frac{1}{2}}((U_{N_\nu}^\alpha + N_\nu^{-\gamma})^{1/\delta} - \theta^{\alpha/\delta}) \leq xd\} - \Phi(x)| = 0(\varepsilon_\nu^{\frac{1}{2}}), \qquad \ldots (2.11)$$

where $d = r\xi_1^{\frac{1}{2}}\alpha\delta^{-1}\theta^{\alpha/\delta-1}$. Let $N_\nu^* = N_\nu - 1$. Then

$$P\{|(N_\nu^*/n_\nu) - 1| > k\varepsilon_\nu\}$$

$$\leq P\{|(N_\nu/n_\nu) - 1| > k\varepsilon_\nu - n_\nu^{-1}\}$$

$= O(\varepsilon_\nu^{\frac{1}{2}})$,

since for any $\delta > 0$, $n_\nu^{-1} < \delta\varepsilon_\nu$ for "big enough" ν. Thus, (2.9) also holds for N_ν^*. So, (2.6) allows us to use the same kind of arguments which led to (2.11) with N_ν being replaced by N_ν^*. Finally, Lemma 1.2 applied to this modified (2.11) and utilizing the fact that $P\{|N_\nu^* N_\nu^{-1} - 1| > \varepsilon_\nu\} = O(\varepsilon_\nu^{\frac{1}{2}})$ will show that the last property needed to apply Theorem 2.4 indeed holds. This concludes the proof of Theorem 2.9.

Remark 2.2: The condition $\gamma \geq \frac{3}{4}$ is sufficient, but not necessary. It is enough for λ and γ, in addition to having $\lambda \in (0,\frac{1}{2})$, to satisfy the inequality given by $\gamma > (3 - 2\lambda)/4$.

CHAPTER III

RATES OF CONVERGENCE FOR THE COVERAGE PROBABILITY OF A FIXED-WIDTH CONFIDENCE INTERVAL FOR THE MEAN OF A U-STATISTIC

3.1 Introduction

Let $\{V_n, n \geq s\}$ be a sequence of U-statistics with a symmetric kernel $g(X_1,\ldots,X_s)$ in s arguments. Assume $E\{|g|^2\} < \infty$ and define

$$g_c(x_1,\ldots,x_c) = E\{g(X_1,\ldots,X_s) \mid X_1 = x_1,\ldots,X_c = x_c\},$$

$c = 1,2,\ldots,s$, and $\eta_c = \mathrm{Var}\{g_c(X_1,\ldots,X_c)\}$. Also let μ be defined by $\mu = E\{g(X_1,\ldots,X_s)\}$. Let $\{U_n, n \geq r\}$ be the U-statistics as defined and utilized in Sections 1 and 2. Given d (> 0) and $q \in (0,1)$, we would like to construct a confidence interval I_n for μ such that the length of I_n is 2d and $P\{\mu \in I_n\} \approx 1 - q$. Here we propose to consider the natural fixed-width confidence interval $I_n = [V_n - d, V_n + d]$ for the parameter μ. Now,

$$\begin{aligned} P\{\mu \in I_n\} &= P\{-d \leq V_n - \mu \leq d\} \\ &= P\{-n^{\frac{1}{2}} d \leq n^{\frac{1}{2}} (V_n - \mu) \leq n^{\frac{1}{2}} d\}. \end{aligned}$$

From Hoeffding's (1948) result we know that $n^{\frac{1}{2}} (V_n - \mu)$ is asymptotically $N(0, s^2\eta_1)$ if $\eta_1 > 0$. Therefore, for large n, $P\{\mu \in I_n\} \approx 2\Phi(n^{\frac{1}{2}} d/s\eta_1^{\frac{1}{2}}) - 1$. If we require this expression to be at least $1 - q$, we need to choose n such that $n^{\frac{1}{2}} d/s\eta_1^{\frac{1}{2}} \geq a$, where a is defined through the equation $\Phi(a) =$

$1-\frac{1}{2}q$. That is, we need

$$n \geq (as/d)^2 \eta_1. \qquad \ldots (3.1)$$

The problem now arises because η_1 is not known in many of the usual applications.

3.2 A Sequential Procedure

Assume that η_1 can be estimated by a U-statistic U_n in such a way that $(E\{U_n\})^{\alpha} = k_1^2 \eta_1$, k_1 (>0) being a known constant. Then, motivated by (3.1), we propose the following stopping time:

$$N_d = \inf\{n \geq n_o \,;\, nd^2 k_1^2 \geq s^2 a^2 (U_n^{\alpha} + n^{-\gamma})\}, \qquad \ldots (3.2)$$

where n_o ($\geq s$) is the starting sample size. By considering an arbitrary positive sequence $\{d_\nu, \nu = 1,2,\ldots\}$ tending to zero as $\nu \to \infty$, we recognize this stopping time (3.2) to have the same form as defined in (2.1). All our previous results will therefore be potentially applicable for this stopping time. We wish to use some of these properties to prove the following theorem.

<u>Theorem 3.1</u>: <u>Consider</u> N_d <u>defined by (3.2) and assume that</u> $\gamma \geq \frac{1}{2}$. <u>Let</u> $\lambda \in (0,\frac{1}{2})$. <u>Now if</u> $E\{|\phi|^{2\xi}\} < \infty$ <u>for</u> $\xi \geq \max\{(5-2\lambda)/8\lambda, 1+(1+2\gamma)(1-2\lambda)/4\}$ <u>and</u> $E\{|g|^4\} < \infty$, <u>then we have</u>

$$P\{\mu \in I_{N_d}\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

<u>Proof</u>: The proof will follow along the lines of Csenki's (1980) proof of his theorem. From Theorem 2.6, we obtain

$$P\{|\frac{N_\nu d_\nu^2}{s^2 a^2 \eta_1} - 1| > d_\nu^{2(\frac{1}{2}-\lambda)}\} = O(d_\nu^{\frac{1}{2}-\lambda}), \qquad \ldots (3.3)$$

where $\{d_\nu\}$ is a sequence of positive real numbers such that $d_\nu \to 0$ as $\nu \to \infty$. Theorem 2.8 together with our Remark 2.1 now give

$$\sup_x |P\{N_\nu^{\frac{1}{2}} (V_{N_\nu} - \mu) \leq xs\eta_1^{\frac{1}{2}}\} - \Phi(x)| = O(d_\nu^{\frac{1}{2}-\lambda}). \qquad \ldots (3.4)$$

From (3.3), we obtain

$$P\{|\frac{N_\nu^{\frac{1}{2}} d_\nu}{sa\eta_1^{\frac{1}{2}}} - 1| > d_\nu^{\frac{1}{2}-\lambda}\} = O(d_\nu^{\frac{1}{2}-\lambda}),$$

and now, Lemma 1.2 applied to (3.4) then gives

$$\sup_x |P\{V_{N_\nu} - \mu \leq d_\nu a^{-1} x\} - \Phi(x)| = O(d^{\frac{1}{2}-\lambda}). \qquad \ldots (3.5)$$

Now, from (3.5) we get the following for the coverage probability:

$$|P\{-d_\nu \leq V_{N_\nu} - \mu \leq d_\nu\} - (\Phi(a) - \Phi(-a))|$$

$$= |P\{V_{N_\nu} - \mu \leq d_\nu\} - P\{V_{N_\nu} - \mu \leq -d_\nu\} - (\Phi(a) - \Phi(-a))|$$

$$\leq |P\{V_{N_\nu} - \mu \leq -d_\nu\} - \Phi(-a)| + |P\{V_{N_\nu} - \mu \leq d_\nu\} - \Phi(a)|$$

$$= O(d_\nu^{\frac{1}{2}-\lambda}).$$

Since this is true for any sequence $\{d_\nu\}$ where $d_\nu \to 0$ as $\nu \to \infty$, the Theorem 3.1 is now proved.

3.3 A Two-Stage Procedure

Motivated by (3.1), (3.2) and the results of Mukhopadhyay (1980, 1981), we propose the following two-stage procedure: For $0 < \eta < 2$, let the starting sample size m_o be defined by

$$m_o = \max\{[(a/d)^{\eta}] + 1, s\}.$$

Then we define the two-stage sample size M_d by

$$M_d = \max\{[(\frac{as}{dk_1})^2 U_{m_o}^{\alpha}] + 1, m_o\}, \qquad \ldots (3.6)$$

where [x] is the largest integer smaller than x. If $M_d = m_o$, we do not take any more samples in the second stage. However, if $M_d > m_o$, we sample the difference $M_d - m_o$ in the second stage. We construct the confidence interval I_{M_d} for μ. We now state the following properties for M_d.

Lemma 3.1: *Consider the stopping time* M_d *defined in (3.6). Assume that* $E\{|\phi|^2\} < \infty$. *Then, for* $\eta_1 > 0$, *we have*

(a) $P\{M_d < \infty\} = 1$,

(b) $M_d \to \infty$ *w.p.* 1 *as* $d \to 0$,

(c) $E\{(M_d/m_d)^{\delta}\} \to 1$ *as* $d \to 0$ *if* $E\{|\phi|^{2\xi}\} < \infty$ *for* $\xi > \max\{\frac{\delta}{2}, \frac{1}{2}\}$, *where* $m_d = (as/d)^2 \eta_1$.

The proof of this lemma will follow along the same lines as for the equivalent results in the sequential case. We omit further details.

Now we will state two interesting results for a special case of the stopping time we have in (3.6). As examples of situations where these apply, we refer to our examples 5.1, 5.2 and 5.3 in Chapter V.

Lemma 3.2: *Let* $\alpha = 1$ *in the definition (3.6) of the stopping time* M_d. *If* $E\{|\phi|^2\} < \infty$, *then we have*

(a) $E\{\frac{M_d d^2}{a^2 s^2 \eta_1}\} = 1 + O(d^2)$,

(b) $\frac{Var\{M_d\}}{(as/dk_1)^4 Var\{U_{m_o}\}} = 1 + O(d^{2-\eta})$,

where k_1 is defined by $E\{U_{m_o}\} = k_1^2\eta_1$, and η appears in the definition of m_o in (3.6).

This lemma is proved in Section 4.4.

Now we consider the coverage probability when I_{M_d} is proposed as the confidence interval for μ. In order to do this, we will need the following result for M_d.

Theorem 3.2: Consider M_d as defined by (3.6). Assume that $E\{|\phi|^{2\xi}\} < \infty$ for $\xi \geq (1-2\lambda)/(2\eta-4+8\lambda)$, where $\lambda \in (\ell, \frac{1}{2})$ for $\ell = (2-\eta)/4$. Then

$$P\{|(M_d/m_d) - 1| > kd^{2(\frac{1}{2}-\lambda)}\} = O(d^{\frac{1}{2}-\lambda}),$$

where $m_d = (as/d)^2\eta_1$ and k (> 0) is arbitrary.

The proof of Theorem 3.2 is given in Section 4.4.

Finally, our main result for constructing the two-stage confidence interval procedure through M_d is the following.

Theorem 3.3: Consider M_d as defined by (3.6). Assume that $E\{|\phi|^{2\xi}\} < \infty$ for $\xi \geq (1-2\lambda)/(2\eta-4+8\lambda)$, $E\{|g|^4\} < \infty$ and $\lambda \in (\ell, \frac{1}{2})$ for $\ell = (2-\eta)/4$. Then we have

$$P\{\mu \in I_{M_d}\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

In view of Theorem 3.2, our Theorem 3.3 can be proved similarly as

Theorem 3.1.

Remark 3.1: If we compare this result about the rate of convergence of the coverage probability with the corresponding result for the sequential procedure (Theorem 3.1), we readily see that for the two-stage procedure, λ is bounded below by a positive constant. This gives us a slower rate of convergence for the two-stage procedure than that for the sequential one. In the terminology of Mukhopadhyay (1981), this two-stage procedure is only "first-order asymptotically consistent", while the sequential procedure is also "second-order asymptotically consistent". Note that the rate for the two-stage procedure gets better as $\eta \in (0,2)$ gets larger in the definition of m_o.

Remark 3.2: Here, we will take the opportunity to correct the requirements for Theorem 5 in Mukhopadhyay (1981). In the context of that paper, Theorem 5 holds for $(1-\eta)/4 < \gamma < \frac{1}{2}$. Thus, the sharper rate of convergence of the coverage probability is obtained in Mukhopadhyay's (1981) Theorem 5, if we choose larger η in $[0,1)$. One may note that this observation is completely in agreement with our present Theorem 3.2 and Remark 3.1.

CHAPTER IV

SOME PROOFS OF LEMMAS AND THEOREMS

4.1 Proof of Lemma 2.4

Let k be a generic constant and $0 < \varepsilon < 1$ be arbitrary. Define $n_\nu = \psi_\nu^{1/\delta}\theta^{\alpha/\delta}$ and $a = [n_\nu(1+\varepsilon)]$ where $[x]$ is the largest integer smaller than x. Now,

$$E\{(N_\nu/n_\nu)^m\}$$

$$\leq (1+\varepsilon)^m P\{N_\nu \leq a\} + \sum_{j=a+1}^{\infty} (j/n_\nu)^m P\{N_\nu = j\}$$

$$= (1+\varepsilon)^m P\{N_\nu \leq a\} + n_\nu^{-m} \sum_{j=a+1}^{\infty} j^m(P\{N_\nu \geq j\} - P\{N_\nu \geq j+1\})$$

$$= (1+\varepsilon)^m P\{N_\nu \leq a\} + n_\nu^{-m}(1+a)^m P\{N_\nu > a\}$$

$$+ n^{-m} \sum_{j=a+1}^{\infty} P\{N_\nu > j\}((j+1)^m - j^m).$$

Now, we have

$$(j+1)^m - j^m = j^{m-1} \sum_{\ell=1}^{\infty} \binom{m}{\ell} j^{1-\ell} \leq 2^m j^{m-1},$$

which implies

$$E\{(N_\nu/n_\nu)^m\}$$

$$\leq (1+\varepsilon)^m + n_\nu^{-m}(a+1)^m P\{N_\nu > a\} + 2^m n_\nu^{-m} \sum_{j=a+2}^{\infty} j^{m-1} P\{N_\nu > j\} \ldots \quad (4.1)$$

Next, from the definition of N_ν in (2.1), for $j \geq a$, we get

$$P\{N_\nu > j\}$$

$$\leq P\{j^\delta < \psi_\nu(U_j^\alpha + j^{-\gamma})\}$$

$$\leq P\{U_j^\alpha > (a^\delta/\psi_\nu) - a^{-\gamma}\}$$

$$\leq P\{|U_j^\alpha - \theta^\alpha| > ((1+\varepsilon)^\delta - 1 - n_\nu^{-\delta})\theta^\alpha - a^{-\delta}\}$$

$$\leq kj^{-\xi},$$

for "big enough" ν. This follows from Corollary 1.2. In the special case for $j = a$, this gives $P\{N_\nu > a\} \leq ka^{-\xi}$, which tends to zero as $\nu \to \infty$. Then,

$$n_\nu^{-m}(a+1)^m P\{N_\nu > a\} \to 0,$$

as $\nu \to \infty$, since $n_\nu^{-m}(a+1)^m$ is bounded as $\nu \to \infty$. Furthermore,

$$2^m n_\nu^{-m} \sum_{j=a+1}^{\infty} j^{m-1} P\{N_\nu > j\}$$

$$\leq 2^m n_\nu^{-m} \sum_{j=a+1}^{\infty} kj^{m-1-\xi}$$

$$\leq kn_\nu^{-m},$$

since $\xi > m$. Equation (4.1) now gives

$$\limsup_{\nu\to\infty} E\{(N_\nu/n_\nu)^m\} \leq (1+\varepsilon)^m,$$

for all $\varepsilon > 0$. Therefore we have

$$\limsup_{\nu\to\infty} E\{(N_\nu/n_\nu)^m\} \le 1.$$

Finally, from Fatou's Lemma and our Lemma 2.3 we can conclude that

$$\liminf_{\nu\to\infty} E\{(N_\nu/n_\nu)^m\} \ge E\{\liminf_{\nu\to\infty} (N_\nu/n_\nu)^m\} = 1.$$

This completes the proof of Lemma 2.4.

4.2 Proof of Theorem 2.6

Let k be a generic constant and $0 < \varepsilon < 1$ be arbitrary. Then,

$$P\{|(N_\nu/n_\nu) - 1| > k\psi_\nu^{-(\frac{1}{2}-\lambda)/\delta}\}$$

$$= P\{N_\nu \le n_\nu(1-\varepsilon)\} + P\{n_\nu(1-\varepsilon) < N_\nu < n_\nu - kn_\nu^{\frac{1}{2}+\lambda}\}$$

$$+ P\{N_\nu > n_\nu + kn_\nu^{\frac{1}{2}+\lambda}\}. \qquad \ldots (4.2)$$

We will show that each of these three expressions in (4.2) has the specific desired order. From the definition of the stopping time N_ν, we get

$$P\{N_\nu = n\} \le P\{(n+t)^\delta/\psi_\nu - n^{-\gamma} > U_n^\alpha\}.$$

Since $P\{U_n^\alpha \ge 0\} = 1$, we therefore obtain

$$P\{N_\nu = n\} = 0 \text{ for } n \le \psi_\nu^{1/(\delta+\gamma)}. \qquad \ldots (4.3)$$

Also, for $n \le n_\nu(1-\varepsilon) = \psi_\nu^{1/\delta}\theta^{\alpha/\delta}(1-\varepsilon)$, we have

$$P\{N_\nu = n\}$$

$$\le P\{n+t > \psi_\nu^{1/\delta}\, U_n^{\alpha/\delta}\}$$

$$\leq P\{\theta^{\alpha/\delta} - U_n^{\alpha/\delta} > \theta^{\alpha/\delta}\varepsilon - t\psi_\nu^{-1/\delta}\}$$

$$\leq kn^{-\xi},$$

for "big enough" ν. This follows from Corollary 1.2 under the condition that $E\{|\phi|^{2\xi}\} < \infty$ for $\xi \geq 1$. Let $a_1 = [\psi_\nu^{1/(\delta+\gamma)}] + 1$ and $a_2 = [n_\nu(1-\varepsilon)]$. Then, we obtain

$$P\{N_\nu \leq a_2\} = \sum_{n=a_1}^{a_2} P\{N_\nu = n\} \leq \sum_{n=a_1}^{\infty} kn^{-\xi} \leq k\psi_\nu^{(1-\xi)/(\delta+\gamma)}, \qquad \ldots (4.4)$$

for $\xi > 1$. This expression in (4.4) is of order $\psi_\nu^{-(\frac{1}{2}-\lambda)/2\delta}$ if $\xi \geq 1 + (\delta+\gamma)(1-2\lambda)/4\delta$. Next, for $n \leq n_\nu - kn_\nu^{\frac{1}{2}+\lambda} = \psi_\nu^{1/\delta}\theta^{\alpha/\delta} - k\psi_\nu^{(\frac{1}{2}+\lambda)/\delta}$, we have

$$P\{N_\nu = n\}$$

$$\leq P\{n \geq \psi_\nu^{1/\delta}\, U_n^{\alpha/\delta} - t\}$$

$$\leq P\{\theta^{\alpha/\delta} - U_n^{\alpha/\delta} \geq k\psi_\nu^{(\lambda-\frac{1}{2})/\delta} - t\psi_\nu^{-1/\delta}\}$$

$$\leq kn^{-\xi}\,\psi_\nu^{-\xi(2\lambda-1)/\delta},$$

for ν "big enough". This follows from Corollary 1.2, given that $E\{|\phi|^{2\xi}\} < \infty$ for $\xi \geq 1$. Then, for $a_3 = [n_\nu - kn_\nu^{\frac{1}{2}+\lambda}]$, we have

$$P\{n_\nu(1-\varepsilon) < N_\nu \leq n_\nu - kn_\nu^{\frac{1}{2}+\lambda}\}$$

$$\leq \sum_{n=a_2+1}^{a_3} k\psi_\nu^{-\xi(2\lambda-1)/\delta}\, n^{-\xi}$$

$$\leq k\psi_\nu^{(1-2\xi\lambda)/\delta}, \qquad \ldots (4.5)$$

which is easily seen to be of the specific desired order if $\xi \geq (5-2\lambda)/8\lambda$.

Finally, let $a_4 = [n_\nu + kn_\nu^{\frac{1}{2}+\lambda}]$. Then, we obtain

$$P\{N_\nu > a_4\}$$

$$\leq P\{a_4^\delta < \psi_\nu U_{a_4}^\alpha + \psi_\nu a_4^{-\gamma}\}$$

$$\leq P\{\psi_\nu \theta^\alpha + k\psi_\nu^{1+(\lambda-\frac{1}{2})/\delta} < \psi_\nu U_{a_4}^\alpha + \psi_\nu \psi_\nu^{-\gamma/\delta}\}$$

$$\leq P\{U_{a_4}^\alpha - \theta^\alpha > k\psi_\nu^{(\lambda-\frac{1}{2})/\delta}\},$$

if $\lambda - \frac{1}{2} > -\gamma$ and ν "big enough". The requirement for λ and γ will always hold if $\gamma \geq \frac{1}{2}$. Now, Corollary 1.2 gives

$$P\{N_\nu > a_4\} \leq ka_4^{-\xi}\ \psi_\nu^{-2\xi(\lambda-\frac{1}{2})/\delta} \leq k\psi_\nu^{-2\xi\lambda/\delta}, \qquad \ldots (4.6)$$

which has the specific desired order if $\xi \geq (1-2\lambda)/8\lambda$. This requirement is clearly satisfied if $\xi \geq (5-2\lambda)/8\lambda$, which was the condition for the terms in (4.5) to satisfy the specific desired order. The Theorem 2.6 now follows by combining (4.2), (4.4), (4.5), and (4.6).

4.3 Proof of Theorem 2.8

We will first prove (2.5) for a subsequence $\{\nu_i\}$ of $\{1,2,\ldots\}$ for which

$$[\psi_{\nu_1}^{1/\delta}] < [\psi_{\nu_2}^{1/\delta}] < \ldots. \qquad \ldots (4.7)$$

For brevity, we will use the following notations: $N_i = N_{\nu_i}$, $b_i = \psi_{\nu_i}^{1/\delta}$ and $\varepsilon_i = \psi_{\nu_i}^{-(\frac{1}{2}-\lambda)/\delta}$. Now, let i be "large enough" such that $[b_i] > 0$. Then we have

$$P\{|(N_i/(\theta^{\alpha/\delta}[b_i])) - 1| > \varepsilon_i\}$$

$$\leq P\{|(N_i/(\theta^{\alpha/\delta}[b_i])) - (N_i/(\theta^{\alpha/\delta}b_i))| > \tfrac{1}{2}\varepsilon_i\}$$

$$+ P\{|(N_i/(\theta^{\alpha/\delta}b_i)) - 1| > \tfrac{1}{2}\varepsilon_i\}$$

$$= P\{(N_i/n_i)(b_i - [b_i])/[b_i] > \tfrac{1}{2}\varepsilon_i\} + O(\varepsilon_i^{\frac{1}{2}}),$$

by Theorem 2.6. Let i be sufficiently large for the following to be true:

$$(b_i - [b_i])/[b_i] < 1/(b_i - 1) < 2/b_i.$$

Then, by the help of Theorem 2.6, we get

$$P\{(N_i/n_i)(b_i - [b_i])/[b_i] > \tfrac{1}{2}\varepsilon_i\}$$

$$\leq P\{(N_i/n_i) - 1 > \tfrac{1}{4}b_i\varepsilon_i - 1\}$$

$$\leq O(\varepsilon_i^{\frac{1}{2}}),$$

which now, combined with the above, gives

$$P\{|(N_i/(\theta^{\alpha/\delta}[b_i])) - 1| > \varepsilon_i\} = O(\varepsilon_i^{\frac{1}{2}}). \qquad \ldots (4.8)$$

By (4.8) and Theorem 2.7 we get (2.5) for the sequence $\{\nu_i\}$. Next,

$$P\{|(N_i/n_i)^{\frac{1}{2}} - 1| > \varepsilon_i^{\frac{1}{2}}\}$$

$$\leq P\{|((N_i/n_i)^{\frac{1}{2}} - 1)((N_i/n_i)^{\frac{1}{2}} + 1)| > \varepsilon_i\}$$

$$= O(\varepsilon_i^{\frac{1}{2}}), \qquad \ldots (4.9)$$

and considering the sequence $\{\nu_i\}$, (2.4) now follows from our Lemma 1.2. Furthermore, since we have

$$P\{|((N_i - 1)/n_i) - 1| > k\varepsilon_i\} \leq P\{|(N_i/n_i) - 1| > k\varepsilon_i - n_i^{-1}\} = O(\varepsilon_i^{\frac{1}{2}}),$$

the same arguments as used above will be applicable to conclude (2.5) and (2.4) when N_i is replaced by $N_i - 1$. This gives us (2.6) and (2.7) for the sequence $\{\nu_i\}$. Finally we need to prove that equations (2.4) - (2.7) hold for $\nu = 1,2,\ldots$. We define

$$V = \{\text{All sequences } \{\nu_i\} \text{ such that (4.7) holds}\},$$

$$f_\nu(x) = P\{n_\nu^{\frac{1}{2}}(U_{N_\nu} - \theta) \le xr\xi_1^{\frac{1}{2}}\},$$

$\nu = 1,2,\ldots$, and $x \in (-\infty,\infty)$. So, we can write (2.4) as

$$\sup_x |f_\nu(x) - \Phi(x)| = O(\varepsilon_\nu^{\frac{1}{2}}).$$

Also there exists a positive number k_1 $(<\infty)$ such that

$$\sup_x |f_{\nu_i}(x) - \Phi(x)| \le k_1 \varepsilon_{\nu_i}^{\frac{1}{2}}, \qquad \ldots (4.10)$$

$i = 1,2,\ldots$, for all $\{\nu_i\} \in V$.

[If this was not the case, we could find a sequence $\{\nu_{i*}\} \in V$ such that (2.4) is violated, and this would contradict what we have just shown.] Now, (2.4) follows from (4.10) since $\nu \in \{\nu_i\}$ for at least one sequence $\{\nu_i\} \in V$, $\nu = 1,2,\ldots$. The same arguments hold for (2.5) - (2.7), and the Theorem 2.8 is thus proved.

4.4 Proof of Lemma 3.2

Let k be a generic positive constant which is independent of d, and let $c = (as/k_1)^2$. From (3.6) we get

$$cd^{-2}U_{m_o} \le M_d \le cd^{-2}U_{m_o} + 1 + m_o I\{cd^{-2}U_{m_o} \le m_o\} \qquad \ldots (4.11)$$

where $I\{x \le y\} = \begin{cases} 1 \text{ if } -\infty < x \le y < \infty, \\ 0 \text{ if } -\infty < y < x < \infty. \end{cases}$

Taking the expectations in (4.11) leads to

$$cd^{-2}\theta \le E(M_d) \le cd^{-2}\theta + 1 + m_o P\{cd^{-2}U_{m_o} \le m_o\}, \qquad \dots (4.12)$$

where $E\{U_{m_o}\} = \theta = k_1^2\eta_1$. Now,

$$P\{cd^{-2}U_{m_o} \le m_o\}$$

$$\le P\{\theta - U_{m_o} \ge \theta - kd^{2-\eta}\}$$

$$\le km_o^{-\xi}, \qquad \dots (4.13)$$

where it is assumed that $E\{|\phi|^{2\xi}\} < \infty$ for some $\xi \ge 1$. In this case, the first part of the lemma then follows from (4.12). Next, we have from (3.6),

$$M_d^2 \le \max\{(cd^{-2}U_{m_o} + 1)^2, m_o^2\},$$

and thus,

$$c^2d^{-4}E\{U_{m_o}^2\} \le E(M_d^2) \le c^2d^{-4}E(U_{m_o}^2) + 2cd^{-2}\theta + 1$$

$$+ m_o^2 P\{cd^{-2}U_{m_o} \le m_o\}. \qquad \dots (4.14)$$

Now, combining (4.12) and (4.14), and utilizing (4.13) we get

$$c^2d^{-4}\operatorname{Var} U_{m_o} - k(d^{\eta(\xi-1)-2} + d^{2\eta(\xi-1)} + d^{\eta(\xi-1)} + d^{-2} + 1)$$

$$\le \operatorname{Var} M_d$$

$$\le c^2d^{-4}\operatorname{Var} U_{m_o} + k(d^{-2} + d^{\eta(\xi-2)} + 1). \qquad \dots (4.15)$$

From Hoeffding (1948) we get the variance of U_{m_o} which gives

$$c^2d^{-4}\operatorname{Var} U_{m_o} = c^2d^{-4}(r^2\xi_1 m_o^{-1} + 0(m_o^{-2})) = 0(d^{\eta-4}). \qquad \dots (4.16)$$

Observing the coefficients of k in (4.15) and assuming as before that $\xi \geq 1$, the dominating term will be the one involving d^{-2} as $d \to 0$. Thus (4.16) and (4.15) will lead to the second part of the lemma.

4.5 Proof of Theorem 3.2

Let k be a generic positive constant which is independent of d, and let $c = (as/k_1)^2$. From (3.6) we get the basic inequality

$$cd^{-2}U^{\alpha}_{m_o} \leq M_d \leq cd^{-2}U^{\alpha}_{m_o} + 1 + I\{cd^{-2}U^{\alpha}_{m_o} \leq m_o\}m_o .$$

Then, we have

$$P\{M_d > m_d + km_d^{\frac{1}{2}+\lambda}\}$$

$$\leq P\{cd^{-2}U^{\alpha}_{m_o} + 1 > m_d + km_d^{\frac{1}{2}+\lambda},\ cd^{-2}U^{\alpha}_{m_o} > m_o\}$$

$$+ P\{cd^{-2}U^{\alpha}_{m_o} \leq m_o\}. \qquad \ldots (4.17)$$

Now, we also have

$$P\{cd^{-2}U^{\alpha}_{m_o} > m_d + km_d^{\frac{1}{2}+\lambda} - 1\}$$

$$= P\{U^{\alpha}_{m_o} - \theta^{\alpha} > km_d^{\lambda-\frac{1}{2}} - \theta^{\alpha}m_d^{-1}\}$$

$$= 0(d^{\xi(4(\lambda-\frac{1}{2}) + \eta)}), \qquad \ldots (4.18)$$

by Corollary 1.2. For this to have the specific order, we need $\xi \geq (1-2\lambda)/(2\eta - 4 + 8\lambda)$ and $\lambda > (2-\eta)/4$. As in (4.13), we can also write

$$P\{cd^{-2}U^{\alpha}_{m_o} \leq m_o\} \leq km_o^{-\xi} = 0(d^{\frac{1}{2}-\lambda}),$$

if $\xi \geq (1-2\lambda)/2\eta$. This will always hold if the conditions for (4.18) to have the specific order hold. So far we have shown that under the conditions of Theorem 3.2,

$$P\{(M_d/m_d) - 1 > kd^{2(\frac{1}{2}-\lambda)}\} = P\{M_d > m_d + km_d^{\frac{1}{2}+\lambda}\} = O(d^{\frac{1}{2}-\lambda}). \quad \ldots (4.19)$$

Next, we have

$$P\{1 - M_d/m_d > kd^{2(\frac{1}{2}-\lambda)}\}$$

$$= P\{M_d < m_d - km_d^{\frac{1}{2}+\lambda}\}$$

$$\leq P\{cd^{-2}U_{m_o}^{\alpha} < m_d - km_d^{\frac{1}{2}+\lambda}\}$$

$$= P\{\theta^{\alpha} - U_{m_o}^{\alpha} > km_d^{\lambda-\frac{1}{2}}\}. \quad \ldots (4.20)$$

By comparing with equation (4.18), we see that the same kind of arguments that led to (4.19) will also be applicable to obtain the specific order in (4.20). This concludes the proof of Theorem 3.2.

CHAPTER V

EXAMPLES

The examples are mostly given for the purely sequential setting. The applications of the two-stage procedure will follow similarly. In example 5.4, the two-stage procedure is given for the purpose of comparing it to the existing procedure of Graybill and Connell (1964). For simplicity, we will write N,M instead of N_d and M_d, respectively.

5.1 Bernoulli (p)

Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli (p) random variables, $p \in (0,1)$. Having recorded $X_1, \ldots, X_n$, the usual estimator for p is taken as $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$. $\bar{X}_n$ is a U-statistic for which $\eta_1 = \mathrm{Var}\, X_1 = p(1-p)$, where η_1 is as defined in Section 3.1, and $S_n^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 = n(n-1)^{-1} \bar{X}_n (1 - \bar{X}_n)$ is a U-statistic estimating η_1 for $n \geq 2$. In order to obtain a $(1-q)$-confidence interval for p having the preassigned width 2d, (3.2) suggests the following stopping time:

$$N = \inf\{n \geq n_o \,;\, n \geq a^2 d^{-2} (S_n^2 + n^{-\gamma})\}.$$

S_n^2 is based upon the kernel ϕ defined by $\phi(X_1, X_2) = \tfrac{1}{2}(X_1 - X_2)^2$. Following the notations of Section 1.3, it is straightforward to show that

$$\xi_1 = \mathrm{Var}\{\phi_1(X_1)\} = \tfrac{1}{4}(p - 5p^2 + 8p^3 - 4p^4).$$

Now, for $\lambda \in (0, \tfrac{1}{2})$, Theorems 2.6 and 2.9 lead to

$$P\{|(Nd^2/a^2p(1-p))-1| > d^{2(\frac{1}{2}-\lambda)}\} = O(d^{\frac{1}{2}-\lambda}),$$

for $\gamma \geq 1/2$, and if $\gamma \geq 3/4$ and ξ_1 is as given above, then we also have

$$\sup_x |P\{(p(1-p))^{\frac{1}{2}}(N-a^2p(1-p)d^{-2}) \leq xad^{-1}2\xi_1^{\frac{1}{2}}\} - \Phi(x)| = O(d^{\frac{1}{2}-\lambda}).$$

We estimate p by the confidence interval $I_N = [\bar{X}_N - d, \bar{X}_N + d]$. Then from Theorem 3.1, we obtain

$$P\{p \in I_N\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

5.2 Poisson (δ)

Let $X_1, X_2, \ldots$ be i.i.d. Poisson (δ) random variables, $\delta \in (0,\infty)$. Our interest is in estimating δ. We propose the interval $I_N = [\bar{X}_N - d, \bar{X}_N + d]$, where

$$N = \inf\{n \geq n_o ; n \geq a^2d^{-2}(\bar{X}_n - n^{-\gamma})\}.$$

Then for any $\lambda \in (0,\frac{1}{2})$, we get (similar to the previous example) for $\gamma \geq 3/4$,

$$\sup_x |P\{N - a^2d^{-2}\delta \leq xad^{-1}\} - \Phi(x)| = O(d^{\frac{1}{2}-\lambda}),$$

and for $\gamma \geq 1/2$,

$$P\{\delta \in I_N\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

5.3 Gamma (δ,β)

Let $X_1, X_2, \ldots$ be i.i.d. Gamma (δ,β) random variables for $\delta \in (0,\infty)$ and $\beta \in (0,\infty)$, that is the probability density function of the X's is assumed to be proportional to $x^{\delta-1}\exp(-\beta x)$ for $x > 0$.

5.3.1 Estimating δ When β is Known

Define $\mu = E(\bar{X}_n) = \delta/\beta$ and $I_n = [\beta\bar{X}_n - d, \beta\bar{X}_n + d]$. Estimating δ by I_n is equivalent to estimating μ by $I_n^* = [\bar{X}_n - (d/\beta), \bar{X}_n + (d/\beta)]$ which is really in the context of our Section 3. We then propose I_N as an interval for δ, where we write

$$N = \inf\{n \geq n_o ; n \geq a^2\beta d^{-2}(\bar{X}_n + n^{-\gamma})\}.$$

Now, as we did in the previous examples, we obtain and state, as illustrations, the following results for any $\lambda \in (0,\frac{1}{2})$:

$$\sup_x |P\{N - a^2\delta d^{-2} \leq xad^{-1}\} - \Phi(x)| = O(d^{\frac{1}{2}-\lambda}),$$

for $\gamma \geq 3/4$, while for $\gamma \geq 1/2$ we have

$$P\{\delta \in I_N\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

5.3.2 Estimating β^{-1} When δ is Known

Define $I_n = [\delta^{-1}\bar{X}_n - d, \delta^{-1}\bar{X}_n + d]$. Estimating β^{-1} by I_n is equivalent to estimating μ by $I_n^* = [\bar{X}_n - \delta d, \bar{X}_n + \delta d]$. Now we propose I_N as a confidence interval for β^{-1}, where we write

$$N = \inf\{n \geq n_o ; n \geq a^2 d^{-2}\delta^{-3}(\bar{X}_n^2 + n^{-\gamma})\}.$$

In this example, we then readily obtain for any $\lambda \in (0,\frac{1}{2})$ and $\gamma \geq 3/4$,

$$\sup_x |P\{N - a^2\beta^{-2}d^{-2}\delta \leq 2a(d\delta\beta)^{-1}x\} - \Phi(x)| = O(d^{\frac{1}{2}-\lambda}),$$

while for $\gamma \geq 1/2$, we have

$$P\{\beta^{-1} \in I_N\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

5.4 Normal (μ,σ^2)

Let $X_1, X_2, \ldots$ be i.i.d. $N(\mu,\sigma^2)$ random variables $\mu \in (-\infty,\infty)$, $\sigma \in (0,\infty)$. We consider the problem of estimating σ^2 when μ is unknown. For $n \geq 2$, let us use $S_n^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$, which is a U-statistic, as an estimator for σ^2. Now, by a result from Hoeffding (1948), we get $4\eta_1 = \lim_{n\to\infty} n \operatorname{Var} S_n^2$, that is $\eta_1 = \sigma^4/2$. We then suggest $I_N = [S_N^2 - d\, ,\, S_N^2 + d]$ as a confidence interval for σ^2, where

$$N = \inf\{n \geq n_o\, ;\, n \geq 2a^2 d^{-2}(S_n^4 + n^{-\gamma})\}.$$

Now, for $\gamma \geq 3/4$ and any $\lambda \in (0,\frac{1}{2})$, our Theorem 2.9 gives

$$\sup_x \left| P\{N - 2a^2\sigma^4 d^{-2} \leq 4\sigma^2 a d^{-1} x\} - \Phi(x) \right| = O(d^{\frac{1}{2}-\lambda}), \text{ while for } \gamma \geq 1/2$$

we have

$$P\{\sigma^2 \in I_N\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

Next, let us consider the two-stage procedure defined through (3.6). In the context of this example, we get

$$m_o = \max\{[(a/d)^{\eta}] + 1\, ,\, 2\},$$

$$M = \max\{[2a^2 d^{-2} S_{m_o}^4] + 1\, ,\, m_o\},$$

for $\eta \in (0,2)$. Now, Lemma 3.1 gives

$$E\{Md^2/2a^2\sigma^4\} \to 1 \qquad \ldots (5.1)$$

as $d \to 0$, and for $\lambda \in (\frac{1}{4}(2-\eta)\, ,\, 1/2)$, Theorem 3.3 gives

$$P\{\sigma^2 \in I_M\} = 1 - q + O(d^{\frac{1}{2}-\lambda}).$$

Let us now consider the two-stage procedure leading to the sample size M^* proposed by Graybill and Connell (1964). This M^* is defined by

$$M^* = m_1 + M_1,$$

where m_1 (≥ 2) is the starting sample size and

$$M_1 = [2 + \pi(q^{-2/(m_1-1)} - 1)^2 (m_1 - 1)^2 S^4_{m_1} /4d^2] + 1.$$

Then $P\{\sigma^2 \in I_{M_1}\} \geq 1 - q$, and

$$E(M^*) \simeq m_1 + 2 + h(q,m_1)\sigma^4/d^2,$$

where $h(q,m_1) = \pi(q^{-2/(m_1-1)} - 1)^2 (m_1^2 - 1)/4$.

Let us now compare our M and Graybill and Connell's (1964) M^* by considering

$$e(q,m_1) = \lim_{d\to 0} \{E(M^*)/E(M)\} = h(q,m_1)/2a^2.$$

As an illustration, let $q = .05$. Now, the following table gives the values of $e(.05,m_1)$ for some values of m_1. The quantity $e(.05,m_1)$ being larger than unity will signify the superiority of our two-stage procedure. In the context of our particular problem, the Table I shows that the procedure through M^* will require more than 3.5 times the sample size required by the procedure through our M, over the range of m_1 considered here. Notice that this remark holds for m_1 being as large as 10^5. However, the larger average sample size required by the procedure through M^* is expected to provide us with higher coverage probability than our target, namely, $(1 - q)$ as $d \to 0$.

TABLE I

VALUES OF $e(.05,m_1)$

m_1	$e(.05,m_1)$
10	9.05
50	4.32
100	3.98
500	3.73
1000	3.70
10000	3.67
100000	3.67

CHAPTER VI

CONCLUSIONS

This study considers a general stopping time based on U-statistics. The purpose is to develop a general setting that will apply to many situations where the family of distributions from which we sample, is known apart from some unknown parameters.

Our interest is in studying asymptotic properties of the stopping time N. First, some convergence results are given where the limiting normality of standardized N is stated in the main theorem in that section (Theorem 2.3). Next, we give the rate of convergence for these results under some additional moment conditions (Theorem 2.9).

These results are then utilized for the problems of estimating the mean of a U-statistic by a fixed-width confidence interval of length 2d. The convergence rate of the coverage probability to a pre-determined confidence coefficient is shown to be of order $O(d^{\frac{1}{2}-\lambda})$, $\lambda \in (\ell, \frac{1}{2})$, where $\ell = 0$ in the sequential case (Theorem 3.1), and $\ell > 0$ for the two-stage procedure (Theorem 3.3). Some examples are given where sampling is carried out from a Bernoulli (p), Poisson (δ), Gamma (δ, β), and $N(\mu, \sigma^2)$ population.

SELECTED BIBLIOGRAPHY

Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. Ann. Math. Statist., 36, 457-462.

Csenki, A. (1980). On the convergence rate of fixed-width sequential confidence intervals. Scand. Actuarial J., 107-111.

Ghosh, M. and DasGupta, R. (1980). Berry-Esseen theorems for U-statistics in the non iid case. To appear in the proceedings of the conference in nonparametric inference held in Budapest, Hungary.

Ghosh, M. and Mukhopadhyay, N. (1975). Asymptotic normality of stopping times in sequential analysis. Unpublished manuscript.

Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. Commun. in Statist., Series A, 8, No. 7, 637-652.

Ghosh, M. (1980). Rate of convergence to normality for random means: Applications to sequential estimation. Sankhyā A., 42, 231-240.

Graybill, F. A. and Connell, T. L. (1964). Sample size for estimating the variance within d units of the true value. Ann. Math. Statist., 35, 438-440.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist., 19, 293-325.

Hoeffding, W. (1961). The strong law of large numbers of U-statistics. Institute of Statistics Mimeo Series No. 302, UNC, Chapel Hill.

Landers, D. and Rogge, L. (1976). The exact approximation order in the central-limit-theorem for random summation. Z. Wahr. Ver. Geb., 36, 269-283.

Michel, R. and Pfanzagl, J. (1971). The accuracy of the normal approximation for the minimum contrast estimates. Z. Wahr. Ver. Geb., 18, 73-84.

Mukhopadhyay, N. (1980). A consistent and asymptotically efficient two stage procedure to construct fixed-width confidence intervals for the mean. Metrika, 27, 281-284.

Mukhopadhyay, N. (1981). Convergence rates of sequential confidence intervals and tests for the mean of a U-statistic. Commun. in Statist., Series A, 10, No. 21, 2231-2244.

Mukhopadhyay, N. (1982). Stein's two-stage procedure and exact consistency. Skandinavisk Aktuarietidskrift, 110-112.

Puri, M. L. and Sen, P. K. (1971). Nonparametric Methods in Multivariate Analysis. John Wiley & Sons, Inc., New York.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications, 2nd Edition. John Wiley & Sons, Inc., New York.

Sen. P. K. (1960). On some convergence properties of U-statistics. Cal. Statist. Assoc. Bul., 10, 1-18.

Sen, P. K. (1981). Sequential Nonparametrics. John Wiley & Sons, Inc., New York.

Sen, P. K. and Ghosh, M. (1981). Sequential point estimation of estimable parameters based on U-statistics. Sankhyā A., 43, 331-344.

Sproule, R. N. (1969). A sequential fixed-width confidence interval for the mean of a U-statistic. Ph.D. Dissertation, UNC, Chapel Hill.

Sproule, R. N. (1974). Asymptotic properties of U-statistics. Trans. Amer. Math. Soc., 199, 55-64.

Wald, A. (1947). Sequential Analysis. Dover Publications, Inc., New York.

γ

VITA

Gaute Vik

Candidate for the Degree of

Doctor of Philosophy

Thesis: ASYMPTOTIC RESULTS FOR STOPPING TIMES BASED ON U-STATISTICS

Major Field: Statistics

Biographical:

Personal Data: Born in Stavanger, Norway, June 26, 1950.

Education: Graduated from St. Svithun High School, Stavanger, Norway, in June 1969; received Cand. Real. degree in Applied Mathematics from Bergen University, Bergen, Norway, in November, 1978; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1984.

Professional Experience: Statistician, Norwegian Defense Research Establishment, Kjeller, Norway, 1978-1982.

Professional Organizations: Vice Chair, Norwegian Operations Research Society, 1981-1982; Member of Norwegian Statistical Association, American Statistical Association, Institute of Mathematical Statistics.