Work supported by the National Institute of Health, Public Health Service, Grant GM-10397 and the United States Air Force Grant AFOSR-68-1415.

A SEQUENTIAL FIXED-WIDTH CONFIDENCE INTERVAL FOR THE MEAN OF A U-STATISTIC

by

RAYMOND NELSON SPROULE
Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 636

AUGUST 1969

RAYMOND NELSON SPROULE. A sequential fixed-width confidence interval for the mean of a U-statistic. (Under the direction of W. J. HALL.)

Preliminary properties of a U-statistic U_n, introduced by Hoeffding (*Ann. Math. Statist.* **19** (1948) 293-325), are developed along with an investigation of several related statistics.

The problem of estimating n $\mathrm{Var}\{U_n\}$ is considered in some detail, with emphasis placed on the asymptotic nature of the estimates. Two prime candidates emerge each possessing good asymptotic properties.

Hoeffding (Univ. of North Carolina, Inst. of Statist., Mimeo Series No. 302 (1961)) showed that a U-statistic may be expressed as an average of independent and identically distributed random variables plus a remainder term. A Kolmogorov-like inequality for this remainder term is developed and its (a.s.) convergence properties are examined. These properties are then related to the U-statistic. In addition, using a result of Anscombe (*Proc. Cambridge Philos. Soc.* **48** (1952) 600-607), the asymptotic normality of U_N, where N is a positive integer-valued random variable, is established under certain conditions.

Equipped with the preceding results, a sequential fixed-width confidence interval for the mean of a U-statistic, having coverage probability approximately equal to some preassigned α, is developed. It is also shown that the sequential procedure (or confidence interval) is asymptotically efficient, in the sense of Chow and Robbins (*Ann. Math. Statist.* **36** (1965) 457-462).

A SEQUENTIAL FIXED-WIDTH CONFIDENCE INTERVAL

FOR THE MEAN OF A U-STATISTIC

by

Raymond Nelson Sproule

A thesis 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 Statistics

Chapel Hill

1969

Approved by:

Adviser

TABLE OF CONTENTS

ACKNOWLEDGMENT

I wish to express my sincere appreciation to Professor W. J. Hall for his significant guidance, criticism and understanding during the course of this work.

I would also like to thank Professors W. Hoeffding, P. K. Sen and G. Simons for their helpful suggestions regarding the original manuscript and the Public Health Service for financing a Research Assistantship through grant GM-10397.

A very special thank you is due my wife Bonnie and Electra Sproule for their constant encouragement.

CHAPTER I

INTRODUCTION

Our primary goal is the development of a sequential confidence interval for the mean of a U-statistic, having fixed-width equal to 2d and coverage probability approximately equal to some preassigned α, where $0 < \alpha < 1$. The problem was solved by Chow and Robbins [3] for the special U-statistic, the sample mean. Starr [16] evaluated the Chow and Robbins [3] sequential procedure, assuming that the underlying distribution is normal, and concluded that the procedure is reasonably consistent and efficient for all values of the variance of the underlying distribution. Generally speaking a U-statistic is a generalization of a sample mean. This is clearly revealed by a decomposition due to Hoeffding [9] and which we refer to as the H-decomposition. By means of this decomposition a U-statistic U_n can be expressed as the sum of a sample mean and a remainder term R_n.

In Chapter II, notation, concepts and preliminary results used throughout this work are developed. In addition to an examination of the H-decomposition, two important statistics are introduced, the W-statistic and the Z-statistic.

In Chapter III we consider the problem of estimating the variance of a U-statistic. Two estimates, s_{wn}^2 and s_{zn}^2, of $n \operatorname{Var}\{U_n\}$ emerge as prime candidates - based on the W- and Z-statistics. It turns out that s_{wn}^2 is slightly superior theoretically, but that s_{zn}^2 can be calculated much more easily, in general. Therefore, both estimates are used in

Chapter V to define sequential procedures.

In Chapter IV a Kolmogorov-like inequality for U-statistics is established. Also, it is shown that the remainder term R_n converges to 0 in a strong sense. However, the main purpose of Chapter IV is to develop the asymptotic normality of a U-statistic based on a random number of observations, in the fashion of Anscombe [1]. This result leads to the asymptotic consistency of the sequential procedures offered in Chapter V.

In Chapter V, equipped with the results of the previous chapters, we present a sequential confidence interval for the mean of a U-statistic. The confidence interval has fixed-width and is asymptotically consistent. Making use of the fact that U-statistics are reverse martingales, the asymptotic efficiency of our sequential procedure is established, in much the same manner as Simons [15]. The theory is illustrated explicitly by obtaining sequential (non-parametric) fixed-width confidence intervals for (1) the population variance and (2) the probability of a concordant pair of observations in sampling from a bivariate population.

As a secondary goal, it is hoped that the techniques and results developed here can be used to extend other sequential procedures already available for the sample mean (where the variance is unknown) to the case of a U-statistic.

CHAPTER II

PRELIMINARY CONCEPTS AND RESULTS

2.1 Introduction. U-statistics, as well as some related statistics, are defined and some relevant properties are presented. A particular emphasis is placed on a decomposition of a U-statistic due to Hoeffding [9], and referred to as the H-decomposition.

2.2 Functionals. We now introduce some basic terminology following closely the lines of Hoeffding [7]. Let $\mathcal{D}$ be a subset of the set of all c.d.f.'s defined on a finite-dimensional Euclidean space. Suppose that the random variable X has c.d.f. $F \in \mathcal{D}$. We say that "$\theta(F)$ is a functional of F defined on $\mathcal{D}$" if for each $F \in \mathcal{D}$ a real number $\theta(F)$ is assigned. A functional $\theta(F)$ is "regular over $\mathcal{D}$" if there exists a positive integer n and a function $\phi(x_1,\cdots,x_n)$ such that

$$\int\cdots\int\phi(x_1,\cdots,x_n)dF(x_1)\cdots dF(x_n) = \theta(F)$$

for each $F \in \mathcal{D}$. In such a case $\phi(x_1,\cdots,x_n)$ is said to be "an unbiased estimate of $\theta(F)$ over $\mathcal{D}$". Let r be the smallest number of arguments required for a function to be an unbiased estimate of $\theta(F)$ over $\mathcal{D}$. Then r is said to be "the degree of $\theta(F)$ over $\mathcal{D}$" and the function $\phi(x_1,\cdots,x_r)$ is referred to as a "kernel of the regular functional $\theta(F)$". Clearly, for any regular functional $\theta(F)$, we can always find a kernel which is a symmetric function of its r arguments, namely,

$$(r!)^{-1}\Sigma\, \phi(x_{\alpha_1},\cdots,x_{\alpha_r})$$

where the summation is over the r! permutations $(\alpha_1,\cdots,\alpha_r)$ of the integers $\{1,2,\cdots,r\}$.

For example, suppose $\mathcal{D}$ is the set of all c.d.f.'s defined on the real line and having finite variance. Let $\theta(F)$ be the variance of the random variable X having c.d.f. $F \in \mathcal{D}$. Then

$$\theta(F) = \int (x_1^2 - x_1 x_2)\, dF(x_1)\, dF(x_2)$$

so that $\phi(x_1,x_2) = x_1^2 - x_1x_2$ is a kernel of $\theta(F)$ and $(x_1-x_2)^2/2$ is the corresponding symmetric kernel.

Hoeffding [7] shows that a polynomial in regular functionals is itself a regular functional. This useful result has a very straightforward proof.

2.3 U-statistics. U-statistics were introduced by Hoeffding [7]. Let $X_1,X_2,\cdots,X_n$ be independent and identically distributed random variables (henceforth referred to as I.I.D. random variables) having c.d.f. F. Let $f(x_1,\cdots,x_r)$ be a function of r arguments. Then a U-statistic is defined by

$$U_n = U(x_1,\cdots,x_n) = [n(n-1)\cdots(n-r+1)]^{-1}\Sigma\, f(x_{\alpha_1},\cdots,x_{\alpha_r})$$

where the summation is over all permutations $(\alpha_1,\cdots,\alpha_r)$ formed from the integers $\{1,2,\cdots,n\}$. We refer to $f(x_1,\cdots,x_r)$ as a "kernel of the U-statistic". Notice that U_n is symmetric in $x_1,x_2,\cdots,x_n$. Also, if $f(x_1,\cdots,x_r)$ is a kernel of a regular functional $\theta(F)$ over $\mathcal{D}$ then

U_n is an unbiased estimate of $\theta(F)$ over $\mathcal{D}$. Let $f_o(x_1,\cdots,x_r)$ be the symmetric function corresponding to $f(x_1,\cdots,x_r)$. We can then write

$$U_n = \binom{n}{r}^{-1} \Sigma^{(n,r)} f_o(x_{\alpha_1},\cdots,x_{\alpha_r})$$

where $\Sigma^{(n,r)}$ represents here, and in the sequel, the summation over all the combinations $(\alpha_1,\cdots,\alpha_r)$ formed from the integers $\{1,2,\cdots,n\}$. We refer to $f_o(x_1,\cdots,x_r)$ as the "symmetric kernel" of the U-statistic. Assume from this point on, without loss of generality, that $f(x_1,\cdots,x_r)$ is symmetric in $x_1,x_2,\cdots,x_r$.

We now introduce regular functionals denoted by ρ which play a central role in what follows. Assume that the symmetric function $f(x_1,\cdots,x_r)$ has existing expectation for $F \in \mathcal{D}$. Write

$$\theta = \theta(F) = \mathcal{E}\{U_n\} = \mathcal{E}\{f(X_1,\cdots,X_r)\}.$$

Define

$$f_c(x_1,\cdots,x_c) = \mathcal{E}\{f(x_1,\cdots,x_c,X_{c+1},\cdots,X_r)\}$$

for $c = 1,2,\cdots,r$. Note that $f_r(x_1,\cdots,x_r) = f(x_1,\cdots,x_r)$. We interpret $\mathcal{E}\{f(x_1,\cdots,x_c,X_{c+1},\cdots,X_r)\}$ as the expected value of $f(X_1,\cdots,X_r)$ given that $X_1,\cdots,X_c$ are fixed at the values $x_1,\cdots,x_c$, respectively. Notice that $\theta = \mathcal{E}\{f_c(X_1,\cdots,X_c)\}$ for $c = 1,2,\cdots,r$. Define

$$\rho_c = \mathrm{Var}\{f_c(X_1,\cdots,X_c)\}$$

for $c = 1,2,\cdots,r$. In particular $f_1(x_1) = \mathcal{E}\{f(x_1,X_2,\cdots,X_r)\}$ and $\rho_1 = \mathrm{Var}\{f_1(X_1)\}$. Now $\rho_c = \rho_c(F)$ is a polynomial in regular functionals of F, and so, is itself a regular functional of F.

If $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, then Hoeffding [7] shows that the variance of U_n is given by

$$\text{(2.1)} \qquad \begin{aligned} \operatorname{Var}\{U_n\} &= \binom{n}{r}^{-1}\Sigma_{c=1}^{r}\binom{r}{c}\binom{n-r}{r-c}\rho_c \\ &= n^{-1}r^2\rho_1 + O(n^{-2}) \end{aligned}$$

for $n \geq r$. This result is generalized in Theorem 2.2 below. In Chapter III we consider the problem of estimation of the regular functional $\sigma^2 = r^2\rho_1$. The following lemma appears in Hoeffding [7].

LEMMA 2.1. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$. *Then*

(i) $0 \leq \rho_c/c \leq \rho_d/d$ *for* $1 \leq c < d \leq r$,

(ii) $r^2\rho_1/n \leq \operatorname{Var}\{U_n\} \leq r\rho_r/n$,

(iii) $n \operatorname{Var}\{U_n\}$ *is a non-increasing function of* n, *and*

(iv) $\operatorname{Var}\{U_r\} = \rho_r$ *and* $\lim_{n\to\infty} n \operatorname{Var}\{U_n\} = r^2\rho_1$.

We now introduce notation which is used to represent the covariance between two U-statistics. Let $X_1,X_2,\cdots,X_n$ be I.I.D. random variables, and let $f(x_1,\cdots,x_r)$ and $g(x_1,\cdots,x_s)$ be two symmetric functions with r and s arguments, respectively. Define

$$f_c(x_1,\cdots,x_c) = \mathcal{E}\{f(x_1,\cdots,x_c,X_{c+1},\cdots,X_r)\},\ c = 1,2,\cdots,r,$$

$$g_d(x_1,\cdots,x_d) = \mathcal{E}\{g(x_1,\cdots,x_d,X_{d+1},\cdots,X_s)\},\ d = 1,2,\cdots,s,$$

and

$$\xi_c = \operatorname{Cov}\{f_c(X_1,\cdots,X_c),g_c(X_1,\cdots,X_c)\},\ c = 1,2,\cdots,\min(r,s).$$

Define two U-statistics by

$$U_n = \binom{n}{r}^{-1}\Sigma^{(n,r)} f(x_{\alpha_1},\cdots,x_{\alpha_r})$$

and

$$V_m = \binom{m}{s}^{-1}\Sigma^{(m,s)} g(x_{\beta_1},\cdots,x_{\beta_s})$$

where $n \geq r$ and $n \geq m \geq s$.

<u>THEOREM 2.2</u>. <u>Assume</u> $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ <u>and</u> $\mathcal{E}\{g(X_1,\cdots,X_s)\}^2 < \infty$. <u>Then</u>, <u>for</u> $m \leq n$,

$$\begin{aligned}\mathrm{Cov}\{U_n,V_m\} &= \binom{n}{r}^{-1}\Sigma_{c=1}^{\min(r,s)}\binom{s}{c}\binom{n-s}{r-c}\xi_c \\ &= n^{-1}rs\xi_1 + O(n^{-2}).\end{aligned}$$

<u>PROOF</u>. First

$$\mathrm{Cov}\{U_n,V_m\} = \binom{n}{r}^{-1}\binom{m}{s}^{-1}\Sigma_{\alpha's}^{(n,r)}\Sigma_{\beta's}^{(m,s)}\mathrm{Cov}\{f(X_{\alpha_1},\cdots,X_{\alpha_r}),g(X_{\beta_1},\cdots,X_{\beta_s})\}.$$

Now, for each combination $(\beta_1,\cdots,\beta_s)$ formed from $\{1,2,\cdots,m\}$, the total number of combinations $(\alpha_1,\cdots,\alpha_r)$ formed from $\{1,2,\cdots,n\}$ and having exactly c suffixes in common with $(\beta_1,\cdots,\beta_s)$ is $\binom{s}{c}\binom{n-s}{r-c}$, where $c = 0,1,\cdots,\min(r,s)$. Notice also that $\mathrm{Cov}\{f(X_{\alpha_1},\cdots,X_{\alpha_r}), g(X_{\beta_1},\cdots,X_{\beta_s})\}$ is zero if $c = 0$, that is, if there are no suffixes in common. Thus

$$\mathrm{Cov}\{U_n,V_m\} = \binom{n}{r}^{-1}\binom{m}{s}^{-1}\left\{\binom{m}{s}\Sigma_{c=1}^{\min(r,s)}\binom{s}{c}\binom{n-s}{r-c}\xi_c\right\}$$

which completes the proof. Notice that the final expression for $\mathrm{Cov}\{U_n,V_m\}$ is free of m.

<u>COROLLARY</u>. <u>If</u> $r = s$ <u>and</u> $f(x_1,\cdots,x_r) = g(x_1,\cdots,x_r)$, <u>then for</u> $r \leq m < n$ <u>we have</u>

$$\mathrm{Cov}\{U_n,U_m\} = \binom{n}{r}^{-1}\Sigma_{c=1}^{r}\binom{r}{c}\binom{n-r}{r-c}\rho_c = \mathrm{Var}\{U_n\}$$

and

$$\text{Correlation}\{U_n, U_m\} = [\text{Var}\{U_n\}/\text{Var}\{U_m\}]^{1/2}$$

$$= n^{-1/2} m^{1/2} [1 + O(m^{-1})].$$

We next introduce notation used to represent the covariance between the squares of U_n and U_m. For $c = 0,1,\cdots,r$ define

(2.2) $$q^{(c)}(x_1,\cdots,x_{2r-c})$$

$$= \binom{2r-c}{r}^{-1} \binom{r}{c}^{-1} \Sigma^{(c)} f(x_{\alpha_1},\cdots,x_{\alpha_r}) f(x_{\beta_1},\cdots,x_{\beta_r})$$

where the summation $\Sigma^{(c)}$ is over all combinations $(\alpha_1,\cdots,\alpha_r)$ and $(\beta_1,\cdots,\beta_r)$ each formed from $\{1,2,\cdots,2r-c\}$ and such that there are exactly c integers in common. Put $\rho_o = 0$. Then $\mathcal{E}\{q^{(c)}(X_1,\cdots,X_{2r-c})\} = \rho_c + \theta^2$ for $c = 0,1,\cdots,r$. Now, for $c = 0,1,\cdots,r$ and $\ell = 1,2,\cdots,2r-c$ define

$$q_\ell^{(c)}(x_1,\cdots,x_\ell) = \mathcal{E}\{q^{(c)}(x_1,\cdots,x_\ell,X_{\ell+1},\cdots,X_{2r-c})\}$$

and

$$\rho_\ell^{(c)} = \text{Var}\{q_\ell^{(c)}(X_1,\cdots,X_\ell)\}.$$

In particular $q_1^{(0)}(x_1) = \theta f_1(x_1)$ so that $\rho_1^{(0)} = \theta^2 \rho_1$. For each $c = 0,1,\cdots,r$ define a U-statistic by

(2.3) $$U_n^{(c)} = \binom{n}{2r-c}^{-1} \Sigma^{(n,2r-c)} q^{(c)}(x_{\alpha_1},\cdots,x_{\alpha_{2r-c}}).$$

Put $t = \min(2r-c, 2r-d)$. Finally, for $\ell = 1,2,\cdots,t$ and $c,d = 0,1,\cdots,r$ define

$$\xi_{\ell}^{(c,d)} = \mathrm{Cov}\{q_{\ell}^{(c)}(X_1,\cdots,X_{\ell}), q_{\ell}^{(c)}(X_1,\cdots,X_{\ell})\}.$$

Notice that for $\ell = 1,2,\cdots,2r-c$ and $c = d = 0,1,\cdots,r$, $\xi_{\ell}^{(c,c)} = \rho_{\ell}^{(c)}$ so that, in particular, $\xi_1^{(0,0)} = \theta^2\rho_1$.

<u>LEMMA 2.3</u>. <u>Assume</u> $\mathcal{E}\{f(X_1,\cdots,X_r)\}^4 < \infty$. <u>Then</u>, <u>for</u> $r \leq m \leq n$,

$$\mathrm{Cov}\{U_n^2,U_m^2\} = \binom{m}{r}^{-1}\binom{n}{r}^{-1}\Sigma_{c=o}^{r}\Sigma_{d=o}^{r}\binom{r}{c}\binom{r}{d}\binom{n-r}{r-c}\binom{m-r}{r-d}\binom{n}{2r-c}^{-1}$$

$$\Sigma_{\ell=1}^{t}\binom{2r-d}{\ell}\binom{n-2r+d}{2r-c-\ell}\xi_{\ell}^{(c,d)}$$

$$= n^{-1}4r^2\theta^2\rho_1 + 0(m^{-1}n^{-1}).$$

<u>PROOF</u>. First, from (2.2) and (2.3),

$$U_n^2 = \binom{n}{r}^{-2}\Sigma_{\alpha's}^{(n,r)}\Sigma_{\beta's}^{(n,r)}f(x_{\alpha_1},\cdots,x_{\alpha_r})f(x_{\beta_1},\cdots,x_{\beta_r})$$

$$= \binom{n}{r}^{-1}\Sigma_{c=o}^{r}\binom{r}{c}\binom{n-r}{r-c}U_n^{(c)}.$$

Then

$$\mathrm{Cov}\{U_n^2,U_m^2\} = \binom{m}{r}^{-1}\binom{n}{r}^{-1}\Sigma_{c=o}^{r}\Sigma_{d=o}^{r}\binom{r}{c}\binom{r}{d}\binom{n-r}{r-c}\binom{m-r}{r-d}\mathrm{Cov}\{U_n^{(c)},U_m^{(d)}\}.$$

But from Theorem 2.2

$$\mathrm{Cov}\{U_n^{(c)},U_m^{(d)}\} = \binom{n}{2r-c}^{-1}\Sigma_{\ell=1}^{t}\binom{2r-d}{\ell}\binom{n-2r+d}{2r-c-\ell}\xi_{\ell}^{(c,d)}$$

for $c,d = 0,1,\cdots,r$. This completes the proof.

2.4 <u>The H-decomposition.</u> In Hoeffding [9] a means for decomposing U_n is developed, one which has great value in establishing properties of U_n. We refer to this decomposition as the "H-decomposition". Continuing in our development of notation we define

$g_h(x_1,\cdots,x_h) = f_h(x_1,\cdots,x_h) - \theta$ for $h = 1,2,\cdots,r$. Also, let $g^{(1)}(x_1) = g_1(x_1)$ and

(2.4) $$g^{(h)}(x_1,\cdots,x_h)$$

$$= g_h(x_1,\cdots,x_h) - \Sigma_{j=1}^{h-1}\Sigma^{(h,j)} g^{(j)}(x_{\alpha_1},\cdots,x_{\alpha_j})$$

for $h = 2,3,\cdots,r$. For example, if $h = 2$,

$$g^{(2)}(x_1,x_2) = g_2(x_1,x_2) - g^{(1)}(x_1) - g^{(1)}(x_2).$$

From (2.4) it follows that

$$g_h(x_1,\cdots,x_h) = \Sigma_{j=1}^{h}\Sigma^{(h,j)} g^{(j)}(x_{\alpha_1},\cdots,x_{\alpha_j})$$

for $h = 1,2,\cdots,r$. Define $g_h^{(h)}(x_1,\cdots,x_h) = g^{(h)}(x_1,\cdots,x_h)$ and

$$g_c^{(h)}(x_1,\cdots,x_c) = \mathcal{E}\{g^{(h)}(x_1,\cdots,x_c,X_{c+1},\cdots,X_h)\}$$

for $h = 1,2,\cdots,r$ and $c = 1,2,\cdots,h-1$. For $n \geq r$ and $h = 1,2,\cdots,r$ define

(2.5) $$V_n^{(h)} = \binom{n}{h}^{-1}\Sigma^{(n,h)} g^{(h)}(x_{\alpha_1},\cdots,x_{\alpha_h}).$$

Note that $V_n^{(1)} = n^{-1}\Sigma_{i=1}^{n} g^{(1)}(x_i) = n^{-1}\Sigma_{i=1}^{n} f_1(x_i) - \theta$. Strictly speaking $V_n^{(h)}$ is not a U-statistic, as it may depend upon unknown functionals. Nevertheless, it does have most of the attributes of a U-statistic. The proof of the following lemma is simple and appears in Hoeffding [9].

<u>LEMMA 2.4</u>. <u>For</u> $h = 1,2,\cdots,r$ <u>the</u> <u>function</u> $g^{(h)}(x_1,\cdots,x_h)$ <u>is</u> <u>symmetric</u>. <u>If</u> $\mathcal{E}\{|f(X_1,\cdots,X_r)|\} < \infty$, <u>then</u> $\mathcal{E}\{g^{(h)}(X_1,\cdots,X_h)\} = 0$ <u>and</u> $g_c^{(h)}(x_1,\cdots,x_c) = 0$ <u>for</u> $h = 1,2,\cdots,r$ <u>and</u> $c = 1,2,\cdots,h-1$.

LEMMA 2.5. *Assume that* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and let* $\delta_h = \mathrm{Var}\{g^{(h)}(X_1,\cdots,X_h)\}$ *for* $h = 1,2,\cdots,r$. *Then*

(i) *for* $h = 1,2,\cdots,r$ *the mean of* $V_n^{(h)}$ *is* 0 *and the variance is given by*

$$\mathrm{Var}\{V_n^{(h)}\} = \binom{n}{h}^{-1}\delta_h = O(n^{-h}).$$

Also

(ii) for $r \leq m \leq n$ *we have*

$$\mathrm{Cov}\{V_n^{(h)}, V_m^{(\ell)}\} = \begin{cases} \mathrm{Var}\{V_n^{(h)}\} & h = \ell = 1,2,\cdots,r \\ 0 & h \neq \ell = 1,2,\cdots,r. \end{cases}$$

PROOF. First, since $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, then $\delta_h < \infty$ for $h = 1,2,\cdots,r$. Part (i) follows from Lemma 2.4 and the fact that

$$\mathrm{Var}\{V_n^{(h)}\} = \binom{n}{h}^{-1}\Sigma_{c=1}^{h}\binom{h}{c}\binom{n-h}{h-c}\,\mathrm{Var}\{g_c^{(h)}(X_1,\cdots,X_c)\}.$$

Part (ii) follows from Theorem 2.2.

We now introduce the H-decomposition by means of the following theorem given in Hoeffding [9].

THEOREM 2.6. *A U-statistic may be decomposed into a linear combination of U-statistics, specifically,*

$$U_n = \theta + \Sigma_{h=1}^{r}\binom{r}{h}V_n^{(h)} = \theta + rV_n^{(1)} + R_n \tag{2.6}$$

$$= \theta + rn^{-1}\Sigma_{i=1}^{n}(f_1(x_i)-\theta) + R_n$$

where $R_n = \Sigma_{h=2}^{r}\binom{r}{h}V_n^{(h)}$ *and* Correlation $\{V_n^{(1)}, R_n\} = 0$. *Further,* $\binom{n}{h}V_n^{(h)}$ *satisfies the martingale property, that is,*

$$\mathcal{E}\{ \binom{n}{h} V_n^{(h)}(x_1,\cdots,x_m,X_{m+1},\cdots,X_n) \} = \binom{m}{h} V_m^{(h)}$$

for $r \leq m < n$ and $h = 1,2,\cdots,r$.

REMARK 1. Theorem 2.6 is extremely useful in establishing properties of U-statistics. In fact, it states that U_n is a linear combination of U-statistics, mutually uncorrelated (by Lemma 2.5) and each successive term having a variance of smaller order. It shows that a U-statistic is essentially the sum of an average of I.I.D. random variables $V_n^{(1)}$ and a zero-mean remainder term R_n, and that the two are uncorrelated. Of course, if $r = 1$ the remainder term R_n is zero. From Lemma 2.5 we see that $\mathrm{Var}\{R_n\} = O(n^{-2})$. In Chapter IV we show that under the assumption that $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, $n^{\gamma}V_n^{(h)}$ converges to zero almost surely as $n \to \infty$ for $\gamma < h/2$ and $h = 1,2,\cdots,r$. This implies that $n^{\gamma}R_n$ converges to zero almost surely as $n \to \infty$ for $\gamma < 1$. Hoeffding [9] uses the H-decomposition to show that, under the assumption that $\mathcal{E}\{|f(X_1,\cdots,X_r)|\} < \infty$, a U-statistic converges to its mean almost surely as $n \to \infty$. Sen [13] proves a somewhat weaker result in that he assumes that $\mathcal{E}\{|f(X_1,\cdots,X_r)|^{1+\epsilon}\} < \infty$ for some $\epsilon > 1 - r^{-1}$. Berk [2] contains a rather simple proof of the almost sure convergence of a U-statistic by recognizing that U-statistics are reverse martingales. More will be said about U-statistics as reverse martingales in Chapter V.

REMARK 2. Hoeffding [7] proves that if $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ and $\rho_1 > 0$ then $\sqrt{n}(U_n-\theta)$ has an asymptotic normal distribution $N(0,\sigma^2)$ where $\sigma^2 = r^2\rho_1$. The result follows directly from the H-decomposition by noticing that $r\sqrt{n}\, V_n^{(1)}$ is asymptotically $N(0,\sigma^2)$ by the Lindberg-

Lévy central limit theorem and that $\lim_{n\to\infty} \mathcal{E}\{\sqrt{n}\, R_n\}^2 = 0$.

REMARK 3. The H-decomposition along with Lemma 2.5 yields a second expression for the variance of a U-statistic, namely,

$$\text{Var}\{U_n\} = \Sigma_{h=1}^{r} \binom{n}{h}^{-1} \binom{r}{h}^{2} \delta_h \tag{2.7}$$

Compare (2.7) with (2.1). Equating these two expressions for the variance of a U-statistic enables us to obtain explicitly the relationship between the ρ's and the δ's. This leads us to the following lemma, also proved by Hoeffding [7], but in an entirely different manner.

LEMMA 2.7. The ρ's and the δ's are related for $h = 1,2,\cdots,r$ by

$$\rho_h = \Sigma_{c=1}^{h} \binom{h}{c} \delta_c \tag{2.8}$$

and

$$\delta_h = \Sigma_{c=o}^{h-1} (-1)^c \binom{h}{c} \rho_{h-c}. \tag{2.9}$$

PROOF. To prove (2.8) we proceed by induction. Putting $n = r$ and equating (2.7) with (2.1) we obtain $\rho_r = \Sigma_{c=1}^{r} \binom{r}{c} \delta_c$. Thus (2.8) holds for $h = r$. Now assume that (2.8) is true for $r - j < h \leq r$ for some $j > 1$. Put $n = r + j$ and equate (2.7) with (2.1) so as to obtain

$$\Sigma_{h=r-j}^{r} \binom{r+j}{r}^{-1} \binom{r}{h} \binom{j}{r-h} \rho_h = \Sigma_{h=1}^{r} \binom{r+j}{h}^{-1} \binom{r}{h}^{2} \delta_h.$$

Multiply through by $\binom{r}{j}^{-1} \binom{r+j}{r}$ and obtain

$$\rho_{r-j} + \Sigma_{h=r-j+1}^{r} \binom{r}{j}^{-1} \binom{r}{h} \binom{j}{r-h} \rho_h = \Sigma_{h=1}^{r} \binom{r+j}{h}^{-1} \binom{r}{j}^{-1} \binom{r}{h}^{2} \binom{r+j}{r} \delta_h.$$

Thus, from (2.8), we have

$$(2.10)\qquad \rho_{r-j} + \Sigma_{h=r-j+1}^{r}\binom{r}{j}^{-1}\binom{r}{h}\binom{j}{r-h}\Sigma_{c=1}^{h}\binom{h}{c}\delta_c = \Sigma_{h=1}^{r}\binom{r}{j}^{-1}\binom{r}{h}\binom{r+j-h}{j}\delta_h .$$

The double sum on the left side of (2.10) with $i = r - h$ becomes

$$\begin{aligned}
&\Sigma_{i=o}^{j-1}\binom{r}{j}^{-1}\binom{r}{i}\binom{j}{i}\Sigma_{c=1}^{r-i}\binom{r-i}{c}\delta_c \\
&= \Sigma_{i=o}^{j-1}\binom{r}{j}^{-1}\binom{j}{i}\Sigma_{c=1}^{r}\binom{r}{c}\binom{r-c}{i}\delta_c \\
&= \Sigma_{c=1}^{r}\binom{r}{j}^{-1}\binom{r}{c}\delta_c\Sigma_{i=o}^{j-1}\binom{j}{i}\binom{r-c}{i} \\
&= \Sigma_{c=1}^{r}\binom{r}{j}^{-1}\binom{r}{c}\delta_c\left\{\binom{r+j-c}{j} - \binom{r-c}{j}\right\} .
\end{aligned}$$

Thus (2.10) becomes

$$\begin{aligned}
\rho_{r-j} &= \Sigma_{h=1}^{r}\binom{r}{j}^{-1}\binom{r}{h}\binom{r-h}{j}\delta_h \\
&= \Sigma_{h=1}^{r-j}\binom{r-j}{h}\delta_h
\end{aligned}$$

which is (2.8) with $h = r - j$. This completes the proof of (2.8). In order to prove (2.9) we make use of (2.8) as follows:

$$\begin{aligned}
\Sigma_{c=o}^{h-1}(-1)^c\binom{h}{c}\rho_{h-c} &= \Sigma_{c=1}^{h}(-1)^{h-c}\binom{h}{c}\rho_c \\
&= \Sigma_{c=1}^{h}(-1)^{h-c}\binom{h}{c}\Sigma_{j=1}^{c}\binom{c}{j}\delta_j \\
&= \Sigma_{j=1}^{h}\Sigma_{c=j}^{h}(-1)^{h-c}\binom{h}{c}\binom{c}{j}\delta_j .
\end{aligned}$$

But

$$\Sigma_{c=j}^{h}(-1)^{h-c}\binom{h}{c}\binom{c}{j} = \Sigma_{i=o}^{h-j}(-1)^{h-i-j}\binom{h}{i+j}\binom{i+j}{j}$$
$$= \binom{h}{j}\Sigma_{i=o}^{h-j}(-1)^{h-i-j}\binom{h-j}{i}$$
$$= \binom{h}{j}(1-1)^{h-j}$$

which equals 1 whenever $j = h$, and 0 whenever $j < h$. This completes the proof of (2.9).

2.5 The W-statistic. For each $i = 1,2,\cdots,n$ define a U-statistic based on $x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n$ by

$$U_{(i)n} = \binom{n-1}{r}^{-1}\Sigma^{(n-1,r)} f(x_{\alpha_1},\cdots,x_{\alpha_r})$$

where the summation is over all combinations $(\alpha_1,\cdots,\alpha_r)$ formed from $\{1,\cdots,i-1,i+1,\cdots,n\}$. Define the W-statistics by

$$W_{in} = nU_n - (n-r)U_{(i)n} \tag{2.11}$$

for $i = 1,2,\cdots,n$ and notice that they are identically distributed. Furthermore, since $U_n = n^{-1}\Sigma_{i=1}^{n}U_{(i)n}$,

$$\overline{W}_n = n^{-1}\Sigma_{i=1}^{n}W_{in} = rU_n.$$

The W-statistics can be conveniently decomposed. To do so, for each $i = 1,2,\cdots,n$ and $h = 1,2,\cdots,r$, define

$$V_{(i)n}^{(h)} = \binom{n-1}{h}^{-1}\Sigma^{(n-1,h)} g^{(h)}(x_{\alpha_1},\cdots,x_{\alpha_h})$$

where the summation is over all combinations $(\alpha_1,\cdots,\alpha_h)$ formed from the integers $\{1,\cdots,i-1,i+1,\cdots,n\}$. Let

(2.12) $$W_{in}^{(h)} = nV_n^{(h)} - (n-r)V_{(i)n}^{(h)}$$

for $i = 1,2,\cdots,n$ and $h = 1,2,\cdots,r$. Then

(2.13) $$W_{in} = r\theta + \Sigma_{h=1}^{r}\binom{r}{h}W_{in}^{(h)}$$

for $i = 1,2,\cdots,n$. The following lemma gives us additional insight into the decomposition (2.13).

LEMMA 2.8. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and suppose that* $n > r$ *and* $i,j = 1,2,\cdots,n$. *Then, for* $h = 1,2,\cdots,r$ *we have that* $\mathcal{E}\{W_{in}^{(h)}\} = 0$ *and*

(2.14) $$\mathrm{Var}\{W_{in}^{(h)}\} = \binom{n-1}{h}^{-1}[r^2 + h(n-2r)]\delta_h = O(n^{-h+1}).$$

Also, for $h \neq \ell = 1,2,\cdots,r$ *we have that* $\mathrm{Cov}\{W_{in}^{(h)},W_{jn}^{(\ell)}\} = 0$, *whereas for* $h = \ell = 1,2,\cdots,r$,

(2.15) $$\mathrm{Cov}\{W_{in}^{(h)},W_{jn}^{(h)}\} = \binom{n-1}{h}^{-1}[(r^2-h)-(n-1)^{-1}h(r-1)^2]\delta_h$$
$$= O(n^{-h}).$$

PROOF. Lemma 2.4 and (2.12) imply that $\mathcal{E}\{W_{in}^{(h)}\} = 0$. Also, (2.14) follows from Lemma 2.5 and (2.12). If $h \neq \ell = 1,2,\cdots,r$, then $\mathrm{Cov}\{W_{in}^{(h)},W_{jn}^{(\ell)}\} = 0$ by Lemma 2.4. If $h = \ell = 1,2,\cdots,r$, then by Lemma 2.5 and (2.12), (2.15) holds.

LEMMA 2.9. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and suppose that* $n > r$ *and* $i,j = 1,2,\cdots,n$. *Then* $\mathcal{E}\{W_{in}\} = r\theta$,

$$\mathrm{Var}\{W_{in}\} = \Sigma_{h=1}^{r}\binom{n-1}{h}^{-1}\binom{r}{h}^2[r^2 + h(n-2r)]\delta_h = O(1)$$

and

$$\mathrm{Cov}\{W_{in},W_{jn}\} = \Sigma_{h=1}^{r} \binom{n-1}{h}^{-1} \binom{r}{h}^{2} [(r^2-h)-(n-1)^{-1}h(r-1)^2]\,\delta_h = O(n^{-1}).$$

The proof follows directly from (2.13) and Lemma 2.8. The W-statistic is closely related to a statistic introduced by Sen [13], as we shall see in the next section, and plays an important role in Chapter III.

2.6 The S-decomposition. Sen [13] defined a U-statistic by

$$V_{in} = \binom{n-1}{r-1}^{-1} \Sigma^{(n-1,r-1)} f(x_i, x_{\alpha_2}, \cdots, x_{\alpha_r})$$

for $i = 1,2,\cdots,n$ where the summation is over all combinations $(\alpha_2,\cdots,\alpha_r)$ formed from the integers $\{1,\cdots,i-1,i+1,\cdots,n\}$. The S-decomposition is

$$U_n = n^{-1}\Sigma_{i=1}^{n} V_{in}.$$

Notice that $W_{in} = rV_{in}$ for $i = 1,2,\cdots,n$ so that Lemma 2.9 may be used to determine variance and covariance expressions for the V_{in}'s. An estimate of $\sigma^2 = r^2\rho_1$ can be constructed from the W_{in}'s (or the V_{in}'s) and such an estimate is examined in Chapter III. Sen [13] proves that V_{in} converges in probability to $f_1(x_i)$ as $n \to \infty$ for any positive integer i. This also follows from Lemma 2.9.

Both the H-decomposition and the S-decomposition indicate that for large n a U-statistic behaves like a sample mean. However, the H-decomposition has the advantage that the remainder term R_n of Theorem 2.6 has a fairly explicit representation. There is a great bulk of theory in the literature developed for the sample mean. It appears then that this theory might, in certain cases, be extended to the case of a U-statistic by showing that, at least asymptotically, the remainder term is in some sense negligible. A specific result due to Chow and Robbins [3] is extended in Chapter V.

2.7 The Z-statistic. Let $Z_r = rU_r$. For $n > r$ define the Z-statistic by

$$Z_n = nU_n - (n-1)U_{n-1}. \tag{2.16}$$

Note that $U_n = n^{-1}\Sigma_{i=r}^{n} Z_i$. Now, define

$$Z_n^{(h)} = nV_n^{(h)} - (n-1)V_{n-1}^{(h)} \tag{2.17}$$

for $n > r$ and $h = 1,2,\cdots,r$. By (2.16) and (2.17), along with Theorem 2.6, we have

$$Z_n = \theta + \Sigma_{h=1}^{r} \binom{r}{h} Z_n^{(h)} \tag{2.18}$$

for $n > r$. An estimate of $\sigma^2 = r^2\rho_1$ can be constructed from the Z_n's and such an estimate is examined in Chapter III. The following lemma, an immediate consequence of (2.17) and Lemma 2.5, gives us some insight into the decomposition (2.18).

LEMMA 2.10. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and suppose that* $r < m < n$. *Then, for* $h = 1,2,\cdots,r$ *we have that* $\mathcal{E}\{Z_n^{(h)}\} = 0$ *and*

$$\operatorname{Var}\{Z_n^{(h)}\} = \binom{n-1}{h}^{-1}[(n-2)h+1]\delta_h = 0(n^{-h+1}). \tag{2.19}$$

Also, for $h \neq \ell = 1,2,\cdots,r$ *we have that* $\operatorname{Cov}\{Z_n^{(h)},Z_m^{(\ell)}\} = 0$, *whereas for* $h = \ell = 1,2,\cdots,r$, (2.20) $$\operatorname{Cov}\{Z_n^{(h)},Z_m^{(h)}\} = -\binom{n-1}{h}^{-1}(h-1)\delta_h = O(n^{-h}).$$

LEMMA 2.11. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$. *Then* $\mathcal{E}\{Z_r\} = r\theta$, $\mathcal{E}\{Z_n\} = \theta$ *for* $n > r$ *and*

(2.21) $$\operatorname{Var}\{Z_n\} = \begin{cases} r^2\rho_r & n = r \\ \Sigma_{h=1}^r \binom{n-1}{h}^{-1}\binom{r}{h}^2[(n-2)h+1]\delta_h = r^2\rho_1 + O(n^{-1}) & n > r. \end{cases}$$

Also, for $m < n$,

(2.22) $$\operatorname{Cov}\{Z_n,Z_m\} = \begin{cases} -r\Sigma_{h=2}^r \binom{n-1}{h}^{-1}\binom{r}{h}^2(h-1)\delta_h = O(n^{-2}) & m = r \\ -\Sigma_{h=2}^r \binom{n-1}{h}^{-1}\binom{r}{h}^2(h-1)\delta_h = O(n^{-2}) & m > r. \end{cases}$$

PROOF. The expectations follow from the definition of Z_r and (2.16). The variance expression (2.21) follows from (2.18), (2.19) and (2.20). Suppose $r < m < n$. Then, by (2.18),

$$\begin{aligned}\operatorname{Cov}\{Z_n,Z_m\} &= \Sigma_{h_1=1}^r\Sigma_{h_2=1}^r\binom{r}{h_1}\binom{r}{h_2}\operatorname{Cov}\{Z_n^{(h_1)},Z_m^{(h_2)}\} \\ &= \Sigma_{h=1}^r\binom{r}{h}^2\operatorname{Cov}\{Z_n^{(h)},Z_m^{(h)}\}\end{aligned}$$

which, upon applying (2.20), reduces to (2.22). Next, suppose $m = r < n$. Then, using (2.16), the corollary to Theorem 2.2, and (2.7) gives us

$$\begin{aligned}\operatorname{Cov}\{Z_n,Z_r\} &= rn\operatorname{Cov}\{U_n,U_r\} - r(n-1)\operatorname{Cov}\{U_{n-1},U_r\} \\ &= rn\operatorname{Var}\{U_n\} - r(n-1)\operatorname{Var}\{U_{n-1}\} \\ &= rn\,\Sigma_{h=1}^r\binom{n}{h}^{-1}\binom{r}{h}^2\delta_h - r(n-1)\Sigma_{h=1}^r\binom{n-1}{h}^{-1}\binom{r}{h}^2\delta_h\end{aligned}$$

which reduces to (2.22). This completes the proof.

REMARKS. If $r = 1$, notice that W_{in} and Z_i each reduce to $f(x_i)$. In general rU_n is the average of $W_{1n},\cdots,W_{nn}$, whereas U_n is a near average of $Z_r,\cdots,Z_n$, that is, $U_n = n^{-1}\Sigma_{i=r}^{n} Z_i$. In Chapter III we are concerned with the problem of estimating $\sigma^2 = r^2\rho_1$. One estimate is the sample variance of $W_{1n},\cdots,W_{nn}$, while a second estimate somewhat resembles the sample variance of $Z_r,\cdots,Z_n$. From the computational point of view, the Z-statistic is a little more suited to a sequential setting than is the W-statistic.

The next two lemmas are used in Chapter III to establish Theorem 3.3.

LEMMA 2.12. Assume $\mathcal{E}\{f(X_1,\cdots,X_r)\}^4 < \infty$ and suppose that $n > r$. Let $\beta_1 = \mathrm{Var}\{[g^{(1)}(X_1)]^2\}$. Then

$$\text{(2.23)} \qquad \mathrm{Var}\{(Z_n-\theta)^2\} = r^4\beta_1 + O(n^{-1}).$$

PROOF. For convenience write $z_n = Z_n^{(1)} = g^{(1)}(x_n)$ (see (2.17)) and set $P_n = \Sigma_{h=2}^{r}\binom{r}{h} Z_n^{(h)}$. Then, by (2.18),

$$\text{(2.24)} \qquad (Z_n-\theta)^2 = r^2z_n^2 + 2rz_nP_n + P_n^2$$

and so

$$\text{(2.25)} \quad \mathrm{Var}\{(Z_n-\theta)^2\} = r^4\beta_1 + 4r^2\mathrm{Var}\{z_nP_n\} + \mathrm{Var}\{P_n^2\}$$
$$+ 4r^3\mathrm{Cov}\{z_n^2,z_nP_n\} + 2r^2\mathrm{Cov}\{z_n^2,P_n^2\} + 4r\mathrm{Cov}\{z_nP_n,P_n^2\}.$$

Our major task is to show that

$$\text{(2.26)} \qquad \mathrm{Var}\{P_n^2\} = O(n^{-2}).$$

Once this has been accomplished, it is a simple matter to show that each of the remaining variance and covariance terms in (2.25) is of order n^{-1}. To prove (2.26), first notice that

$$(2.27)\quad P_n^2 = \Sigma_{h=2}^{r}\binom{r}{h}^2 Z_n^{(h)2} + \Sigma_{h_1=2}^{r}\Sigma_{\substack{h_2=2\\ h_1\neq h_2}}^{r}\binom{r}{h_1}\binom{r}{h_2} Z_n^{(h_1)} Z_n^{(h_2)}.$$

Now, using (2.17) and (2.15), we can write

$$(2.28)\qquad Z_n^{(h)} = W_{nn}^{*(h)} - (h-1)V_{n-1}^{(h)}$$

where

$$(2.29)\quad W_{nn}^{*(h)} = h\binom{n-1}{h-1}^{-1}\Sigma_{\alpha's}^{(n-1,h-1)} g^{(h)}(x_n, x_{\alpha_2}, \cdots, x_{\alpha_h})$$

for $h = 2,3,\cdots,r$ and $n > r$. (Notice that $W_{nn}^{*(h)}$ is related to $V_n^{(h)}$ in much the same way as W_{nn}, defined in (2.11), is related to U_n, and that $W_{nn}^{*(h)}$ differs slightly from $W_{nn}^{(h)}$, defined in (2.12).) Therefore, from (2.28)

$$(2.30)\quad Z_n^{(h)2} = W_{nn}^{*(h)2} - 2(h-1)W_{nn}^{*(h)}V_{n-1}^{(h)} + (h-1)^2V_{n-1}^{(h)2}$$

for $h = 2,3,\cdots,r$. We now are equipped to prove that

$$(2.31)\qquad \mathrm{Var}\{Z_n^{(h)2}\} = 0(n^{-2})$$

for $h = 2,3,\cdots,r$. From Lemma 2.3 (with $m = n$) and the fact that $\mathcal{E}\{V_{n-1}^{(h)}\} = 0$, we have

$$(2.32)\qquad \mathrm{Var}\{V_{n-1}^{(h)2}\} = 0(n^{-2})$$

for $h = 2,3,\cdots,r$. Although $W_{nn}^{*(h)}$ is not a U-statistic, the proof of Lemma 2.3 can be adapted to show that

$$(2.33) \qquad \mathrm{Var}\{W_{nn}^{*(h)2}\} = O(n^{-2})$$

for $h = 2,3,\cdots,r$. Notice that, by Lemma 2.4, $\mathcal{E}\{W_{nn}^{*(h)}V_{n-1}^{(h)}\} = 0$ for $h = 2,3,\cdots,r$. Then, by the Schwarz inequality

$$(2.34) \quad \mathrm{Var}\{W_{nn}^{*(h)}V_{n-1}^{(h)}\} \leq [\mathcal{E}\{W_{nn}^{*(h)}\}^4 \mathcal{E}\{V_{n-1}^{(h)}\}^4]^{1/2}$$

$$= [\mathrm{Var}\{W_{nn}^{*(h)2}\} + \mathrm{Var}^2\{W_{nn}^{*(h)}\}]^{1/2}[\mathrm{Var}\{V_{n-1}^{(h)2}\} + \mathrm{Var}^2\{V_{n-1}\}]^{1/2}$$

for $h = 2,3,\cdots,r$. Thus, (2.32), (2.33) and (2.34), along with Lemmas 2.5 and 2.8, imply that

$$(2.35) \qquad \mathrm{Var}\{W_{nn}^{*(h)}V_{n-1}^{(h)}\} = O(n^{-2})$$

for $h = 2,3,\cdots,r$. Also, (2.32), (2.33), (2.35) and the Schwarz inequality imply that the three covariances involving $W_{nn}^{*(h)2}$, $V_{n-1}^{(h)2}$ and $W_{nn}^{*(h)}V_{n-1}^{(h)}$ are each of order n^{-2} for $h = 2,3,\cdots,r$. This proves (2.31).

An argument similar to that in (2.34), along with (2.31) gives us

$$(2.36) \qquad \mathrm{Var}\{Z_n^{(h_1)} Z_n^{(h_2)}\} = O(n^{-2})$$

for $h_1 \neq h_2 = 2,3,\cdots,r$. Again, by the Schwarz inequality, the covariances between the terms in (2.27) are each of order n^{-2}. This proves (2.26).

We now tackle the remaining terms in (2.25). From Lemma 2.4, $\mathcal{E}\{z_n^k P_n\} = 0$ for $k = 1,3$ so that

$$(2.37) \qquad \mathrm{Cov}\{z_n^2, z_n P_n\} = \mathrm{Cov}\{z_n^3, P_n\} = 0.$$

By Lemma 2.10, $\operatorname{Var}\{P_n\} = O(n^{-1})$; also $\operatorname{Var}\{z_n^k\} = O(1)$ for $k = 1,2$. Hence, by an argument similar to that in (2.34) we obtain

$$\operatorname{Var}\{z_n P_n\} = O(n^{-1}). \tag{2.38}$$

An application of (2.26) and the Schwarz inequality proves that $\operatorname{Cov}\{z_n^2, P_n^2\}$ and $\operatorname{Cov}\{z_n P_n, P_n^2\}$ are each of order n^{-1}. This, along with (2.26), (2.37), (2.38) and (2.25), completes the proof of (2.23).

LEMMA 2.13. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^4 < \infty$ *and suppose that* $r < m < n$. *Then*

$$\text{(2.39)}\qquad \operatorname{Cov}\{(Z_n-\theta)^2,(Z_m-\theta)^2\}$$

$$= r^4[(r-1)\beta_2 + (r-1)^2\beta_3]n^{-1} + O(n^{-1}m^{-1})$$

where $\beta_2 = \mathcal{E}\{g^{(1)2}(X_1)g^{(1)}(X_2)g^{(2)}(X_1,X_2)\}$ *and* $\beta_3 = \mathcal{E}\{g^{(1)}(X_1)g^{(1)}(X_2)g^{(2)}(X_1,X_3)g^{(2)}(X_2,X_3)\}$.

PROOF. From (2.24), writing z_n for $g^{(1)}(x_n)$,

$$\text{(2.40)}\quad \operatorname{Cov}\{(Z_n-\theta)^2,(Z_m-\theta)^2\} = r^2\operatorname{Cov}\{z_n^2,(Z_m-\theta)^2\}$$

$$+ 2r^3\operatorname{Cov}\{z_nP_n,z_m^2\} + 4r^2\operatorname{Cov}\{z_nP_n,z_mP_m\} + 2r\operatorname{Cov}\{z_nP_n,P_m^2\}$$

$$+ r^2\operatorname{Cov}\{P_n^2,z_m^2\} + 2r\operatorname{Cov}\{P_n^2,z_mP_m\} + \operatorname{Cov}\{P_n^2,P_m^2\}.$$

We treat the seven terms in (2.40) one at a time. Since $m < n$, z_n^2 is independent of $(Z_m-\theta)^2$, so that

$$\operatorname{Cov}\{z_n^2,(Z_m-\theta)^2\} = 0. \tag{2.41}$$

From (2.26) and the Schwarz inequality

(2.42) $$\mathrm{Cov}\{P_n^2, P_m^2\} = 0(n^{-1}m^{-1}).$$

We now consider the term $\mathrm{Cov}\{z_n P_n, z_m^2\}$ in (2.40). From (2.28), (2.29) and the fact that z_n is independent of $z_m^2 V_{n-1}^{(h)}$ for $m < n$,

(2.43) $$\mathrm{Cov}\{z_n P_n, z_m^2\} = \Sigma_{h=2}^{r} \binom{r}{h} \mathcal{E}\{z_n z_m^2 Z_n^{(h)}\} = \Sigma_{h=2}^{r} \binom{r}{h} \mathcal{E}\{z_n z_m^2 W_{nn}^{*(h)}\}$$
$$= \Sigma_{h=2}^{r} \binom{r}{h} \binom{n-1}{h-1}^{-1} \Sigma_{\alpha' s}^{(n-1,h-1)} \mathcal{E}\{z_n z_m^2 g^{(h)}(X_n, X_{\alpha_2}, \cdots, X_{\alpha_h})\}.$$

For convenience we shall write $g_{\alpha_1 \cdots \alpha_h} = g^{(h)}(x_{\alpha_1}, \cdots, x_{\alpha_h})$ for any combination $(\alpha_1, \cdots, \alpha_h)$. Now, for $h = 3,4,\cdots,r$, by Lemma 2.4,

(2.44) $$\mathcal{E}\{z_n z_m^2 g_{n\alpha_2 \cdots \alpha_h}\} = 0$$

for any combination $(\alpha_2, \cdots, \alpha_h)$ chosen from $\{1,2,\cdots,n-1\}$. Also, for $h = 2$

(2.45) $$\mathcal{E}\{z_n z_m^2 W_{nn}^{*(2)}\} = (n-1)^{-1} \Sigma_{i=1}^{n-1} \mathcal{E}\{z_n z_m^2 g_{ni}\}$$
$$= (n-1)^{-1} \beta_2 .$$

Putting (2.44) and (2.45) into (2.43) gives us

(2.46) $$\mathrm{Cov}\{z_n P_n, z_m^2\} = \binom{r}{2} \beta_2 n^{-1} + 0(n^{-2}).$$

Next, we consider the term $\mathrm{Cov}\{z_n P_n, z_m P_m\}$ in (2.40). Since $\mathcal{E}\{z_n P_n\} = 0$

(2.47) $$\mathrm{Cov}\{z_n P_n, z_m P_m\}$$
$$= \Sigma_{h_1=2}^{r} \Sigma_{h_2=2}^{r} \binom{r}{h_1} \binom{r}{h_2} \mathcal{E}\{z_n z_m Z_n^{(h_1)} Z_m^{(h_2)}\}.$$

From (2.28) and the independence of z_n from $z_m V_{n-1}^{(h_1)} Z_m^{(h_2)}$, we have

$$(2.48) \qquad \mathcal{E}\{z_n z_m Z_n^{(h_1)} Z_m^{(h_2)}\} = \mathcal{E}\{z_n z_m W_{nn}^{*(h_1)} W_{mm}^{*(h_2)}\}$$
$$- (h_2-1)\mathcal{E}\{z_n z_m W_{nn}^{*(h_1)} V_{m-1}^{(h_2)}\}$$

for $h_1, h_2 = 2,3,\cdots,r$. For $h_1 = h_2 = 2$

$$\mathcal{E}\{z_n z_m W_{nn}^{*(2)} W_{mm}^{*(2)}\} = (n-1)^{-1}(m-1)^{-1}\Sigma_{i=1}^{n-1}\Sigma_{j=1}^{m-1}\mathcal{E}\{z_n z_m g_{ni} g_{mj}\}$$

where $\mathcal{E}\{z_n z_m g_{ni} g_{mj}\} = 0$ except possibly when $i = j = 1,2,\cdots,m-1$. Thus

$$(2.49) \quad \mathcal{E}\{z_n z_m W_{nn}^{*(2)} W_{mm}^{*(2)}\} = (n-1)^{-1}(m-1)^{-1}\Sigma_{j=1}^{m-1}\mathcal{E}\{z_n z_m g_{nj} g_{mj}\}$$
$$= (n-1)^{-1}\beta_3.$$

Higher values of h_1 and h_2 lead to terms of order n^{-2} or higher, that is

$$(2.50) \qquad \mathcal{E}\{z_n z_m W_{nn}^{*(h_1)} W_{mm}^{*(h_2)}\} = 0(n^{-2}).$$

A close examination yields

$$(2.51) \qquad \mathcal{E}\{z_n z_m W_{nn}^{*(h_1)} V_{m-1}^{(h_2)}\} = 0(n^{-4})$$

for $h_1, h_2 = 2,3,\cdots,r$. (More specifically, $\mathcal{E}\{z_n z_m W_{nn}^{*(h_1)} V_{m-1}^{(h_2)}\} = 0$ for $h_1 = 2,3$ and $h_2 = 2,3,\cdots,r$.) Putting (2.49), (2.50) and (2.51) into (2.48), and then, (2.48) into (2.47) yields

$$(2.52) \qquad \mathrm{Cov}\{z_n P_n, z_m P_m\} = \binom{r}{2}^2 \beta_3 n^{-1} + 0(n^{-2}).$$

We now consider the term $\mathrm{Cov}\{z_n P_n, P_m^2\}$ in (2.40). First, from (2.28)

$$\text{(2.53)} \qquad \mathrm{Cov}\{z_n P_n, P_m^2\} = \Sigma_{h=2}^{r} \binom{r}{h} \mathcal{E}\{z_n Z_n^{(h)} P_m^2\}$$

$$= \Sigma_{h=2}^{r} \binom{r}{h} \mathcal{E}\{z_n W_{nn}^{*(h)} P_m^2\}$$

since z_n is independent of $V_{n-1}^{(h)} P_m^2$ for $m < n$. For $h = 2,3,\cdots,r$

$$\text{(2.54)} \quad \mathcal{E}\{z_n W_{nn}^{*(h)} P_m^2\} = \Sigma_{h_1=2}^{r} \binom{r}{h_1}^2 \mathcal{E}\{z_n W_{nn}^{*(h)} Z_m^{(h_1)2}\}$$

$$+ \Sigma_{\substack{h_1=2 \\ h_1 \neq h_2}}^{r} \Sigma_{h_2=2}^{r} \binom{r}{h_1} \binom{r}{h_2} \mathcal{E}\{z_n W_{nn}^{*(h)} Z_m^{(h_1)} Z_m^{(h_2)}\}.$$

For $h = h_1 = 2$, by (2.28)

$$\text{(2.55)} \qquad \mathcal{E}\{z_n W_{nn}^{*(2)} Z_m^{(2)2}\} = \mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)2}\}$$

$$- 2\mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)} V_{m-1}^{(2)}\} + \mathcal{E}\{z_n W_{nn}^{*(2)} V_{m-1}^{(2)2}\}.$$

But

$$\mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)2}\} = (n-1)^{-1}(m-1)^{-2} \Sigma_{i=1}^{n-1} \Sigma_{\ell_1=1}^{m-1} \Sigma_{\ell_2=1}^{m-1} \mathcal{E}\{z_n g_{ni} g_{m\ell_1} g_{m\ell_2}\}$$

where $\mathcal{E}\{z_n g_{ni} g_{m\ell_1} g_{m\ell_2}\} = 0$ except possibly when $i = m$ and $\ell_1 = \ell_2 = 1,2,\cdots,m-1$ and when $i = \ell_1 = \ell_2 = 1,2,\cdots,m-1$. Thus

$$\text{(2.56)} \qquad \mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)2}\} = 0(n^{-1} m^{-1}).$$

Also, it is not difficult to show that

$$\text{(2.57)} \qquad \mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)} V_{m-1}^{(2)}\} = 0$$

and

(2.58) $$\mathcal{E}\{z_n W_{nn}^{*(2)} V_{m-1}^{(2)2}\} = O(n^{-1}m^{-2}).$$

Putting (2.56), (2.57) and (2.58) into (2.55) gives us

(2.59) $$\mathcal{E}\{z_n W_{nn}^{*(2)} Z_m^{(2)2}\} = O(n^{-1}m^{-1}).$$

We now show that

(2.60) $$\mathcal{E}\{z_n W_{nn}^{*(2)} Z_m^{(2)} Z_m^{(3)}\} = O(n^{-1}m^{-1}).$$

For $h = h_1 = 2$ and $h_2 = 3$

$$\mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)} W_{mm}^{*(3)}\}$$
$$= (n-1)^{-1}(m-1)^{-1}\binom{m-1}{2}^{-1}\Sigma_{i=1}^{n-1}\Sigma_{\ell=1}^{m-1}\Sigma_{(\ell_1,\ell_2)}^{(m-1,2)}\mathcal{E}\{z_n g_{ni} g_{m\ell} g_{m\ell_1\ell_2}\}$$

where $\mathcal{E}\{z_n g_{ni} g_{m\ell} g_{m\ell_1\ell_2}\} = 0$ except possibly when (i,ℓ) equals (ℓ_1,ℓ_2). Thus

(2.61) $$\mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)} W_{mm}^{*(3)}\} = O(n^{-1}m^{-1}).$$

It is a simple matter to show that

(2.62) $$\mathcal{E}\{z_n W_{nn}^{*(2)} W_{mm}^{*(2)} V_{m-1}^{(3)}\} = 0,$$

(2.63) $$\mathcal{E}\{z_n W_{nn}^{*(2)} V_{m-1}^{(2)} V_{m-1}^{(3)}\} = O(n^{-1}m^{-2})$$

and

(2.64) $$\mathcal{E}\{z_n W_{nn}^{*(2)} V_{m-1}^{(2)} W_{mm}^{*(3)}\} = O(n^{-1}m^{-2}).$$

Then (2.60) follows from (2.61), (2.62), (2.63) and (2.64). Higher values of h, h_1 and h_2 yield terms of order $n^{-1}m^{-2}$ or higher. Therefore, from (2.53), (2.54), (2.59), (2.60) and the extensions of (2.59) and (2.60) to higher values of h, h_1 and h_2, we obtain

$$\text{Cov}\{z_n P_n, P_m^2\} = O(n^{-1}m^{-1}). \tag{2.65}$$

Further application of the techniques already used gives us

$$\text{Cov}\{P_n^2, z_m^2\} = O(n^{-2}) \tag{2.66}$$

and

$$\text{Cov}\{P_n^2, z_m P_m\} = O(n^{-2}). \tag{2.67}$$

We have now completed our treatment of each of the seven terms in (2.40). Combining (2.41), (2.42), (2.46), (2.52), (2.65), (2.66) and (2.67) yields (2.39). This completes the proof of the lemma.

CHAPTER III

ESTIMATION OF THE VARIANCE OF A U-STATISTIC

3.1 Introduction. Our chief purpose in this chapter is to obtain an estimate for the variance of a U-statistic having certain desirable properties. Assume $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ and that $\rho_1 > 0$. Then, from (2.1), recall that

$$\mathrm{Var}\{U_n\} = n^{-1}r^2\rho_1 + O(n^{-2})$$

so that we may confine our attention to estimation of $\sigma^2 = r^2\rho_1$, since any good estimate of σ^2 will be a good estimate of n $\mathrm{Var}\{U_n\}$, if second order terms are negligible. More specifically, we are concerned with obtaining an estimate for $\sigma^2 = r^2\rho_1$ that has several basic properties. (1) It converges to $\sigma^2 = r^2\rho_1$ almost surely as $n \to \infty$. (2) For each n, it is positive almost surely. (3) The variance of the estimate may be evaluated for large n. (4) The nature of the estimate is such that, for it, we can establish the asymptotic efficiency of the sequential procedure appearing in Chapter V. (5) In addition, we would hope that the sequential calculation of the estimate is not too tedious. In this chapter three suitable candidates are examined. The two leading candidates are compared when U_n is the unbiased estimate of the population variance.

3.2 The U-statistic estimate of ρ_1. Recall that $\rho_1 = \mathrm{Var}\{f_1(X_1)\} = \mathcal{E}\{f_1(X_1)\}^2 - \theta^2$. Now ρ_1 is a regular functional, and so, has a U-statistic estimate. From definition (2.2) in section 2.3, notice that ρ_1 has a kernel

$$(3.1)\quad q^{(1)}(x_1,\cdots,x_{2r-1}) - q^{(0)}(x_1,\cdots,x_{2r})$$

$$= r^{-1}\binom{2r-1}{r}^{-1}\Sigma^{(1)} f(x_{\alpha_1},\cdots,x_{\alpha_r})f(x_{\beta_1},\cdots,x_{\beta_r})$$

$$- \binom{2r}{r}^{-1}\Sigma^{(0)} f(x_{\alpha_1},\cdots,x_{\alpha_r})f(x_{\beta_1},\cdots,x_{\beta_r})$$

where the summation $\Sigma^{(1)}$ is over all combinations $(\alpha_1,\cdots,\alpha_r)$ and $(\beta_1,\cdots,\beta_r)$ each formed from $\{1,2,\cdots,2r-1\}$ and such that there is exactly one integer in common, and the summation $\Sigma^{(0)}$ is over all combinations $(\alpha_1,\cdots,\alpha_r)$ and $(\beta_1,\cdots,\beta_r)$ each formed from $\{1,2,\cdots,2r\}$ and such that there are no integers in common. From (2.3), the U-statistic estimate of ρ_1 is given by

$$(3.2)\qquad Q_{1n} = U_n^{(1)} - U_n^{(0)}$$

$$= \binom{n}{2r-1}^{-1}\Sigma^{(n,2r-1)} q^{(1)}(x_{\alpha_1},\cdots,x_{\alpha_{2r-1}})$$

$$- \binom{n}{2r}^{-1}\Sigma^{(n,2r)} q^{(0)}(x_{\alpha_1},\cdots,x_{\alpha_{2r}}).$$

We next evaluate the variance of this estimate of ρ_1. From (2.1) and Theorem 2.2,

$$(3.3)\quad \mathrm{Var}\{Q_{1n}\} = \mathrm{Var}\{U_n^{(1)}\} - 2\,\mathrm{Cov}\{U_n^{(1)},U_n^{(0)}\} + \mathrm{Var}\{U_n^{(0)}\}$$

$$= \Delta_1 n^{-1} + 0(n^{-2})$$

where

$$(3.4) \qquad \Delta_1 = 4r^2\theta^2\rho_1 - 4r(2r-1)\xi_1^{(0,1)} + (2r-1)^2\rho_1^{(1)}$$

and the functionals $\xi_1^{(0,1)}$ and $\rho_1^{(1)}$ are defined in section 2.3. Then, from (3.4),

$$(3.5) \quad \Delta_1 = 4r^2\theta^2\mathrm{Var}\{g^{(1)}(X_1)\} - 4r(2r-1)\mathrm{Cov}\{q_1^{(0)}(X_1), q_1^{(1)}(X_1)\}$$

$$+ (2r-1)^2\mathrm{Var}\{q_1^{(1)}(X_1)\}$$

$$= \mathrm{Var}\{(2r-1)q_1^{(1)}(X_1) - 2rq_1^{(0)}(X_1)\}.$$

We would now like to express $(2r-1)q_1^{(1)}(x_1) - 2rq_1^{(0)}(x_1)$ in terms of the $g^{(h)}$ functions introduced in section 2.4. First $q_1^{(0)}(x_1) = \theta f_1(x_1) = \theta g^{(1)}(x_1) + \theta^2$. Define $f_o(x_1) = \mathcal{E}\{f_1(X_2)f_2(x_1,X_2)\}$. Then, from section 2.3,

$$(3.6) \quad q_1^{(1)}(x_1)$$

$$= \mathcal{E}\{q^{(1)}(x_1,X_2,\cdots,X_{2r-1})\}$$

$$= r^{-1}\binom{2r-1}{r}^{-1}\mathcal{E}\{\Sigma^{(1)}f(X_{\alpha_1},\cdots,X_{\alpha_r})f(X_{\beta_1},\cdots,X_{\beta_r})\,|X_1=x_1\}$$

$$= r^{-1}\binom{2r-1}{r}^{-1}\left[\binom{2r-2}{r-1}f_1^2(x_1) + 2(2r-2)\binom{2r-3}{r-2}f_o(x_1)\right]$$

$$= (2r-1)^{-1}[f_1^2(x_1) + 2(r-1)f_o(x_1)].$$

Now, put $g_o(x_1) = \mathcal{E}\{g^{(1)}(X_2)g^{(2)}(x_1,X_2)\}$ and note that $\mathcal{E}\{g_o(X_1)\} = 0$. From Lemma 2.4

$$(3.7)\quad f_o(x_1) = \mathcal{E}\{(g^{(1)}(X_2)+\theta)(g^{(2)}(x_1,X_2) + \theta + g^{(1)}(x_1) + g^{(1)}(X_2))\}$$

$$= g_1(x_1) + \theta g^{(1)}(x_1) + \rho_1 + \theta^2.$$

Thus, putting (3.7) into (3.6) yields

$$q_1^{(1)}(x_1)$$

$$= (2r-1)^{-1}[g^{(1)2}(x_1) + 2r\theta g^{(1)}(x_1) + 2(r-1)g_o(x_1) + 2(r-1)(\rho_1+\theta^2) + \theta^2]$$

and so

$$(3.8)\quad (2r-1)q_1^{(1)}(x_1) - 2rq_1^{(0)}(x_1)$$

$$= g^{(1)2}(x_1) + 2(r-1)g_o(x_1) + 2(r-1)\rho_1 - \theta^2.$$

Before we put (3.8) into (3.5), notice that

$$\beta_1 = \mathrm{Var}\{g^{(1)2}(X_1)\},$$

$$(3.9)\quad \beta_2 = \mathcal{E}\{g^{(1)2}(X_1)g^{(1)}(X_2)g^{(2)}(X_1,X_2)\}$$

$$= \int g^{(1)2}(x_1)\{\int g^{(1)}(x_2)g^{(2)}(x_1,x_2)dF(x_2)\}dF(x_1)$$

$$= \int g^{(1)2}(x_1)g_o(x_1)dF(x_1)$$

$$= \mathrm{Cov}\{g^{(1)2}(X_1),g_o(X_1)\}$$

and

$$(3.10)\quad \beta_3 = \mathcal{E}\{g^{(1)}(X_1)g^{(1)}(X_2)g^{(2)}(X_1,X_3)g^{(2)}(X_2,X_3)\}$$

$$= \int\left\{\int g^{(1)}(x_1)g^{(2)}(x_1,x_3)dF(x_1)\right\}\left\{\int g^{(1)}(x_2)g^{(2)}(x_2,x_3)dF(x_2)\right\}dF(x_3)$$

$$= \int g_o^2(x_1)dF(x_1)$$

$$= \mathrm{Var}\{g_o(X_1)\}.$$

Then, putting (3.8) into (3.5) gives us

$$(3.11)\qquad \Delta_1 = \beta_1 + r(r-1)\beta_2 + 4(r-1)^2\beta_3.$$

The variance of Q_{1n} is then given by (3.3) with Δ_1 given by (3.11).

Q_{1n} inherits all the good properties of a U-statistic. It is an unbiased estimate of ρ_1. Under the assumption that $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, Q_{1n} converges almost surely to ρ_1 as $n \to \infty$. One drawback is that Q_{1n} may possibly take on negative values. Also, it appears that Q_{1n} would require much time for computation.

3.3 Estimation of ρ_c. In section 2.3 we defined the functionals ρ_c for $c = 1,2,\cdots,r$. Now ρ_c is a regular functional of degree $2r$ and has a kernel given by $q^{(c)}(x_1,\cdots,x_{2r-c}) - q^{(0)}(x_1,\cdots,x_{2r})$ for $c = 1,2,\cdots,r$. The U-statistic estimate of ρ_c is given by $Q_{cn} = U_n^{(c)} - U_n^{(0)}$ for $c = 1,2,\cdots,r$. Then, from (2.1) and Theorem 2.2,

$$\mathrm{Var}\{Q_{cn}\} = \mathrm{Var}\{U_n^{(c)}\} - 2\,\mathrm{Cov}\{U_n^{(c)},U_n^{(0)}\} + \mathrm{Var}\{U_n^{(0)}\}$$

$$= \Delta_c n^{-1} + 0(n^{-2})$$

where

$$\Delta_c = 4r^2\theta^2\rho_1 - 4r(2r-c)\xi_1^{(0,c)} + (2r-c)^2\rho_1^{(c)}$$

for $c = 1,2,\cdots,r$. It is also possible to write

$$\Delta_c = 4r^2\theta^2\mathrm{Var}\{g^{(1)}(X_1)\} - 4r(2r-c)\mathrm{Cov}\{q_1^{(0)}(X_1),q_1^{(c)}(X_1)\}$$
$$+ (2r-c)^2\mathrm{Var}\{q_1^{(c)}(X_1)\}$$
$$= \mathrm{Var}\{(2r-c)q_1^{(c)}(X_1) - 2r\theta g^{(1)}(X_1)\}$$

for $c = 1,2,\cdots,r$. Clearly, for $c = 1$ we have the situation considered in section 3.2.

Suppose we define

$$Q_n = \binom{n}{r}^{-1}\Sigma_{c=1}^{r}\binom{r}{c}\binom{n-r}{r-c}Q_{cn}.$$

Then Q_n is an unbiased estimate of $\mathrm{Var}\{U_n\}$. Also

$$\mathrm{Var}\{Q_n\} = \binom{n}{r}^{-2}\Sigma_{c=1}^{r}\binom{r}{c}^2\binom{n-r}{r-c}^2\mathrm{Var}\{Q_{cn}\}$$
$$+ \binom{n}{r}^{-2}\Sigma_{c=1}^{r}\underset{c\neq d}{\Sigma_{d=1}^{r}}\binom{r}{c}\binom{r}{d}\binom{n-r}{r-c}\binom{n-r}{r-d}\mathrm{Cov}\{Q_{cn},Q_{dn}\}$$
$$= r^4\Delta_1 n^{-3} + 0(n^{-4})$$

where Δ_1 is given by (3.11).

For small values of n, where the higher order terms of $\mathrm{Var}\{U_n\}$ may not be negligible, Q_n might be a satisfactory estimate of $\mathrm{Var}\{U_n\}$. However, it is a very tedious estimate to compute, and we are concerned mainly with large values of n, so that $r^2n^{-1}Q_{1n}$ is to be preferred, so far, as an estimate of $\mathrm{Var}\{U_n\}$.

3.4 Sen's estimate of $\sigma^2 = r^2\rho_1$. In section 2.5 we introduced the W-statistic. Now define

$$s_{wn}^2 = (n-1)^{-1}\Sigma_{i=1}^n (W_{in}-\bar{W}_n)^2. \tag{3.12}$$

Sen [13] introduced (see section 2.6)

$$s_{vn}^2 = (n-1)^{-1}\Sigma_{i=1}^n (V_{in}-U_n)^2$$

as an estimate of ρ_1. Recall that $W_{in} = rV_{in}$ for $i = 1,2,\cdots,n$ and $\bar{W}_n = rU_n$. It then follows that $s_{wn}^2 = r^2 s_{vn}^2$, so that s_{wn}^2 can be considered as an estimate of $\sigma^2 = r^2\rho_1$. Sen [13] showed that s_{wn}^2 is an asymptotically unbiased estimate of $\sigma^2 = r^2\rho_1$, and further, that s_{wn}^2 converges to $\sigma^2 = r^2\rho_1$ in probability as $n \to \infty$. In the next two theorems we derive first order expressions for the bias and the variance of s_{wn}^2, as well as prove that s_{wn}^2 converges almost surely to $\sigma^2 = r^2\rho_1$ as $n \to \infty$.

THEOREM 3.1.

(i) If $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, then

$$\mathcal{E}\{s_{wn}^2\} = r^2\rho_1 + r^2(r-1)[(r-1)\delta_2-2\rho_1]n^{-1} + O(n^{-2}). \tag{3.13}$$

(ii) If $\mathcal{E}\{f(X_1,\cdots,X_r)\}^4 < \infty$, then

$$\mathrm{Var}\{s_{wn}^2\} = r^4\Delta_1 n^{-1} + O(n^{-2}) \tag{3.14}$$

where Δ_1 is given by (3.11).

PROOF. To prove (i) notice that

$$s_{wn}^2 = (n-1)^{-1}\Sigma_{i=1}^n W_{in}^2 - (n-1)^{-1}n\bar{W}_n^2.$$

Then

$$\mathcal{E}\{s_{wn}^2\} = (n-1)^{-1}\Sigma_{i=1}^n \mathcal{E}\{W_{in}\}^2 - (n-1)^{-1}n\mathcal{E}\{\overline{W}_n\}^2$$

$$= (n-1)^{-1}n[\mathrm{Var}\{W_{1n}\} - \mathrm{Var}\{\overline{W}_n\}]$$

which, using Lemma 2.9 and (2.7), reduces to (3.13).

To prove (ii) we express s_{wn}^2 as a linear combination of U-statistics. From the proof of Lemma 2.3

$$U_n^2 = \binom{n}{r}^{-1}\Sigma_{c=o}^r \binom{r}{c}\binom{n-r}{r-c} U_n^{(c)}. \tag{3.15}$$

In a similar fashion we may write

$$(n-1)^{-1}\Sigma_{i=1}^n W_{in}^2 \tag{3.16}$$

$$= (n-1)^{-1}\binom{n-1}{r-1}^{-1} r^2 n \Sigma_{c=1}^r \binom{r-1}{c-1}\binom{n-r}{r-c} U_n^{(c)}.$$

Then, from (3.15) and (3.16), after some rearranging,

$$s_{wn}^2 = (n-1)^{-1}n\binom{n}{r}^{-1}\Sigma_{c=o}^r \binom{r}{c}\binom{n-r}{r-c}[cn-r^2]U_n^{(c)}. \tag{3.17}$$

For $c = 0,1,\cdots,r$ let

$$a_n(c) = (n-1)^{-1}n\binom{n}{r}^{-1}\binom{r}{c}\binom{n-r}{r-c}[cn-r^2].$$

In particular, $a_n(0) = -r^2 + 0(n^{-1})$ and $a_n(1) = r^2 + 0(n^{-1})$. For $c = 2,3,\cdots,r$ notice that $a_n(c) = 0(n^{-1})$. Thus

$$s_{wn}^2 = r^2(U_n^{(1)}-U_n^{(0)}) + \Sigma_{c=o}^r \alpha_n(c)U_n^{(c)} \tag{3.18}$$

where $\alpha_n(c) = 0(n^{-1})$ for $c = 0,1,\cdots,r$. Note that we do not require an explicit expression for $\alpha_n(c)$. For convenience let $T_n = \Sigma_{c=o}^r \alpha_n(c)U_n^{(c)}$, and recall from section 3.2 that $Q_{1n} = U_n^{(1)}-U_n^{(0)}$.

Then

$$(3.19)\quad \mathrm{Var}\{s_{wn}^2\} = r^4\mathrm{Var}\{Q_{1n}\} + 2r^2\mathrm{Cov}\{Q_{1n},T_n\} + \mathrm{Var}\{T_n\}.$$

From (2.1), Theorem 2.2 and the fact that $\alpha_n(c) = 0(n^{-1})$ for $c = 0,1,\cdots,r$, we have that $\mathrm{Var}\{T_n\} = 0(n^{-3})$ and $\mathrm{Cov}\{Q_{1n},T_n\} = 0(n^{-2})$. Therefore, from (3.19) and (3.3)

$$\mathrm{Var}\{s_{wn}^2\} = r^4\Delta_1 n^{-1} + 0(n^{-2})$$

where $\Delta_1 = \beta_1 + 4(r-1)\beta_2 + 4(r-1)^2\beta_3$. This completes the proof.

<u>THEOREM 3.2</u>. <u>If</u> $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, <u>then</u> s_{wn}^2 <u>converges almost surely to</u> $\sigma^2 = r^2\rho_1$ <u>as</u> $n \to \infty$.

<u>PROOF</u>. The theorem follows almost immediately from expression (3.18) in the proof of Theorem 3.1. First, by Hoeffding [9] (or Berk [2]), the U-statistic $Q_{1n} = U_n^{(1)} - U_n^{(0)}$ converges almost surely to its expectation ρ_1 as $n \to \infty$. Secondly, since $\alpha_n(c) = 0(n^{-1})$ for $c = 0,1,\cdots,r$, and $U_n^{(c)}$ is a U-statistic with finite expectation, $c = 0,1,\cdots,r$, then $T_n = \Sigma_{c=o}^r\alpha_n(c)U_n^{(c)}$ converges almost surely to 0 as $n \to \infty$. This completes the proof.

<u>REMARKS</u>. There is not much difference between s_{wn}^2 and r^2Q_{1n} as estimates of $\sigma^2 = r^2\rho_1$. From expression (3.18), $s_{wn}^2 = r^2Q_{1n} + T_n$. The difference T_n converges to 0 almost surely as $n \to \infty$, has expectation of order n^{-1} and variance of order n^{-3}. Both estimates have good properties, including the same first order variance. Both, of course, converge to $\sigma^2 = r^2\rho_1$ almost surely as $n \to \infty$. They only differ on the question of unbiasedness. The estimate r^2Q_{1n} is unbiased whereas s_{wn}^2 is asymptotically unbiased. From the point of view of small sample theory s_{wn}^2 has the good property that it is

always non-negative, whereas it is possible for r^2Q_{1n} to take on negative values. For the problem of finding a fixed-width sequential confidence interval for θ, discussed in Chapter V, the estimate s^2_{wn} serves our purposes better than does r^2Q_{1n}.

3.5 *Estimate of* $\sigma^2 = r^2\rho_1$ *based on the Z-statistic*. Refer to section 2.7 for an introduction to the Z-statistic. For $n > r$ define

$$(3.20)\quad s^2_{zn} = (n-1)^{-1}\Sigma^n_{i=r+1}(Z_i-U_n)^2 + r(n-1)^{-1}(U_r-U_n)^2.$$

Notice that if $r = 1$ then $\sigma^2 = \mathrm{Var}\{f(X_1)\}$ and s^2_{zn} reduces to a sample variance. We now consider the merits of s^2_{zn} as an estimate of $\sigma^2 = r^2\rho_1$.

THEOREM 3.3.

(i) *If* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$, *then*

$$(3.21)\quad \mathcal{E}\{s^2_{zn}\} = r^2\rho_1 + r^2(r-1)^2\delta_2 n^{-1}\log n + An^{-1} + o(n^{-1})$$

where A *is defined by*

$$A = r(r-1)\Sigma^r_{h=1}\binom{r}{h}A_h\delta_h,$$

with $A_1 = -1$, $A_2 = 2\ (\gamma - \Sigma^{r-2}_{i=1}i^{-1})$, $A_h = h(h-2)^{-1}$ *for* $h = 3,4,\cdots,r$ and $\gamma = 0.5772\ldots$ (*Euler's constant*).

(ii) *If* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^4 < \infty$, *then*

$$(3.22)\quad \mathrm{Var}\{s^2_{zn}\} = r^4\Delta n^{-1} + O(n^{-2}(\log n)^2)$$

where $\Delta = \beta_1 + 2(r-1)\beta_2 + 2(r-1)^2\beta_3$.

PROOF. To prove (i) let

$$s_{zn}^{*2} = (n-1)^{-1}\Sigma_{i=r+1}^{n}(Z_i-U_n)^2$$

so that $s_{zn}^2 = s_{zn}^{*2} + r(n-1)^{-1}(U_r-U_n)^2$. Now

$$r(n-1)^{-1}\mathcal{E}\{U_r-U_n\}^2 = rn^{-1}\Sigma_{h=1}^{r}\binom{r}{h}\delta_h + 0(n^{-2}). \tag{3.23}$$

Next

$$\mathcal{E}\{s_{zn}^{*2}\} = (n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Var}\{Z_i-U_n\}^2$$

$$= (n-1)^{-1}[\Sigma_{i=r+1}^{n}\mathrm{Var}\{Z_i\} - 2\Sigma_{i=r+1}^{n}\mathrm{Cov}\{Z_i,U_n\} + (n-r)\mathrm{Var}\{U_n\}].$$

From (2.16) and the corollary to Theorem 2.2, $\mathrm{Cov}\{Z_i,U_n\} = \mathrm{Var}\{U_n\}$ for $i = r+1,\cdots,n$, so that using Lemma 2.11

$$\mathcal{E}\{s_{zn}^{*2}\} = (n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Var}\{Z_i\} - (n-1)^{-1}(n-r)\mathrm{Var}\{U_n\} \tag{3.24}$$

$$= \Sigma_{h=1}^{r}\binom{r}{h}^2\delta_h K_n(h,r) - n^{-1}r^2\rho_1 + 0(n^{-2})$$

where, for $n > r$ and $h = 1,2,\cdots,r$,

$$K_n(h,r) = (n-1)^{-1}\Sigma_{i=r+1}^{n}\binom{i-1}{h}^{-1}[(i-2)h+1].$$

Note that $K_n(1,r) = 1 - n^{-1}(r-1) + 0(n^{-2})$ and that

$$K_n(2,r) = 4(n-1)^{-1}\Sigma_{i=r+1}^{n}(i-2)^{-1} - 2(n-1)^{-1}\Sigma_{i=r+1}^{n}(i-1)^{-1}(i-2)^{-1}.$$

Now, let $\gamma_n = \Sigma_{i=1}^{n}i^{-1} - \log n$ for $n > 2$; then $\lim_{n\to\infty}\gamma_n = \gamma$, where $\gamma = 0.5772...$ is Euler's constant. Also, notice that

$$\Sigma_{i=r+1}^{n}(i-1)^{-1}(i-2)^{-1} = \Sigma_{i=r+1}^{n}[(i-2)^{-1}-(i-1)^{-1}] = (r-1)^{-1} - (n-1)^{-1}.$$

Then $K_n(2,r)$ becomes

$$(3.25)\quad K_n(2,r) = 4n^{-1}[\gamma+\log n-\Sigma_{i=1}^{r-2} i^{-1}] - 2n^{-1}(r-1)^{-1}$$

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

$$= 4n^{-1}\log n + n^{-1}[4\gamma-4\Sigma_{i=1}^{r-2} i^{-1}-2(r-1)^{-1}] + o(n^{-1}).$$

From the theory of infinite series

$$\Sigma_{i=m+1}^{\infty} i^{-1}(i+1)^{-1}\cdots(i+k)^{-1} = k^{-1}(k!)^{-1}\binom{m+k}{k}^{-1}$$

for $k = 1,2,\cdots$ and $m = 0,1,\cdots$. Notice that the above infinite series is of order m^{-k} for $k = 1,2,\cdots$. Then, making use of this series result, for $h = 3,4,\cdots,r$, we have that

$$K_n(h,r) = n^{-1}h(h!)\Sigma_{i=r+1}^{n}(i-2)^{-1}\cdots(i-h)^{-1}$$

$$- n^{-1}(h-1)(h!)\Sigma_{i=r+1}^{n}(i-1)^{-1}\cdots(i-h)^{-1} + O(n^{-2})$$

$$= n^{-1}h(h!)\Sigma_{i=r+1-h}^{\infty} i^{-1}(i+1)^{-1}\cdots(i+h-2)^{-1}$$

$$- n^{-1}(h-1)(h!)\Sigma_{i=r+1-h}^{\infty} i^{-1}(i+1)^{-1}\cdots(i+h-1)^{-1} + O(n^{-2})$$

$$= n^{-1}h^2(h-1)(h-2)^{-1}\binom{r-2}{h-2}^{-1} - n^{-1}h\binom{r-1}{h-1}^{-1} + O(n^{-2})$$

and so

$$(3.26)\quad \Sigma_{h=3}^{r}\binom{r}{h}^2\delta_h K_n(h,r)$$

$$= n^{-1}r\Sigma_{h=3}^{r}\binom{r}{h}\delta_h(h-2)^{-1}[(r-2)h+2] + O(n^{-2}).$$

Combining (3.20), (3.23), (3.24), (3.25) and (3.26) gives us (3.21).

To prove (ii) first set $\theta = 0$, without loss of generality. Then

$$s_{zn}^{*2} = (n-1)^{-1}\Sigma_{i=r+1}^{n} Z_i^2 - (n-1)^{-1}(n+r)U_n^2 + (n-1)^{-1}2rU_rU_n$$

$$= A_n + B_n$$

where we have set

(3.27) $$A_n = (n-1)^{-1}\Sigma_{i=r+1}^{n} Z_i^2$$

and

(3.28) $$B_n = -(n-1)^{-1}(n+r)U_n^2 + (n-1)^{-1}2rU_rU_n.$$

Therefore

(3.29) $$\mathrm{Var}\{s_{zn}^{*2}\} = \mathrm{Var}\{A_n\} + 2\ \mathrm{Cov}\{A_n,B_n\} + \mathrm{Var}\{B_n\}.$$

We now divide the proof of (ii) into two parts. In Part (a) we show that $\mathrm{Var}\{A_n\}$ is given by the expression in (3.22). In Part (b) we show that $\mathrm{Var}\{B_n\}$ and $\mathrm{Cov}\{A_n,B_n\}$ are each of order $n^{-2}\log n$ or higher.

PART (a). From (3.27)

(3.30) $$\mathrm{Var}\{A_n\} = (n-1)^{-2}\Sigma_{i=r+1}^{n}\mathrm{Var}\{Z_i^2\}$$
$$+ 2(n-1)^{-2}\Sigma_{i=r+1}^{n}\Sigma_{j=r+1}^{i-1}\mathrm{Cov}\{Z_i^2,Z_j^2\}.$$

By Lemma 2.12

(3.31) $$(n-1)^{-2}\Sigma_{i=r+1}^{n}\mathrm{Var}\{Z_i^2\} = (n-1)^{-2}\Sigma_{i=r+1}^{n}[r^4\beta_1 + O(i^{-1})]$$
$$= r^4\beta_1 n^{-1} + O(n^{-2}\log n).$$

By Lemma 2.13

$$(3.32)\quad (n-1)^{-2}\Sigma_{i=r+1}^{n}\Sigma_{j=r+1}^{i-1}\mathrm{Cov}\{Z_i^2,Z_j^2\}$$

$$= (n-1)^{-2}\Sigma_{i=r+1}^{n}\Sigma_{j=r+1}^{i-1}\{r^4[(r-1)\beta_2+(r-1)^2\beta_3]i^{-1} + O(i^{-1}j^{-1})\}$$

$$= r^4[(r-1)\beta_2+(r-1)^2\beta_3]n^{-1} + O(n^{-2}(\log n)^2).$$

Putting (3.31) and (3.32) into (3.30) yields

$$(3.33)\qquad \mathrm{Var}\{A_n\} = r^4\Delta n^{-1} + O(n^{-2}(\log n)^2).$$

PART (b). From Lemma 2.3 (with $m = n$), since $\Theta = 0$, we have that $\mathrm{Var}\{U_n^2\} = O(n^{-2})$. It is then a simple matter to show from (3.28) that

$$(3.34)\qquad \mathrm{Var}\{B_n\} = O(n^{-2}).$$

In order to prove

$$(3.35)\qquad \mathrm{Cov}\{A_n,B_n\} = O(n^{-2}\log n)$$

it is sufficient, by (3,28), to show that

$$(3.36)\qquad \mathrm{Cov}\{A_n,U_rU_n\} = O(n^{-1})$$

and

$$(3.37)\qquad \mathrm{Cov}\{A_n,U_n^2\} = O(n^{-2}\log n).$$

We now prove (3.36). First

$$(3.38)\quad \mathrm{Var}\{U_rU_n\} = \mathcal{E}\{U_r^2U_n^2\} - [\mathcal{E}\{U_rU_n\}]^2$$

$$= \mathrm{Cov}\{U_r^2,U_n^2\} + \mathrm{Var}\{U_r\}\mathrm{Var}\{U_n\} - [\mathrm{Cov}\{U_r,U_n\}]^2$$

$$= 0(n^{-1})$$

by Lemma 2.3 (since $\theta = 0$), (2.1) and the corollary to Theorem 2.2. Then (3.36) follows from (3.38), the fact that $\mathrm{Var}\{A_n\} = 0(n^{-1})$ and the Schwarz inequality.

We now prove (3.37). From (3.27) and (2.24)

$$(3.39)\quad \mathrm{Cov}\{A_n,U_n^2\} = (n-1)^{-1}r^2\Sigma_{i=r+1}^n \mathrm{Cov}\{z_i^2,U_n^2\}$$

$$+ (n-1)^{-1}2r\Sigma_{i=r+1}^n \mathrm{Cov}\{z_iP_i,U_n^2\} + (n-1)^{-1}\Sigma_{i=r+1}^n \mathrm{Cov}\{P_i^2,U_n^2\}$$

where $P_i = \Sigma_{h=2}^r \binom{r}{h} z_i^{(h)}$ for $i = r+1,r+2,\cdots,n$. Now, using symmetry and (3.15),

$$(3.40)\quad (n-1)^{-1}\Sigma_{i=r+1}^n \mathrm{Cov}\{z_i^2,U_n^2\}$$

$$= (n-1)^{-1}(n-r)\mathrm{Cov}\{z_1^2,U_n^2\}$$

$$= (n-1)^{-1}(n-r)\mathrm{Cov}\{n^{-1}\Sigma_{i=1}^n z_i^2,U_n^2\}$$

$$= (n-1)^{-1}(n-r)\binom{n}{r}^{-1}\Sigma_{c=o}^r \binom{r}{c}\binom{n-r}{r-c} \mathrm{Cov}\{n^{-1}\Sigma_{i=1}^n z_i^2,U_n^{(c)}\}.$$

Notice that $n^{-1}\Sigma_{i=1}^n z_i^2$ is a U-statistic, so that, applying Theorem 2.2, we obtain $\mathrm{Cov}\{n^{-1}\Sigma_{i=1}^n z_i^2,U_n^{(0)}\} = 0(n^{-2})$. (In this case the ξ_1 of Theorem 2.2 equals zero because $q_1^{(0)}(x_1) = 0$.) For $c = 1,2,\cdots,r$ it is clear, by the Schwarz inequality and the fact that the variance of a U-statistic is of order n^{-1}, that $\mathrm{Cov}\{n^{-1}\Sigma_{i=1}^{n} z_i^2, U_n^{(c)}\} = 0(n^{-1})$. Thus

$$(3.41) \qquad (n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Cov}\{z_i^2, U_n^2\} = 0(n^{-2}).$$

In order to prove that

$$(3.42) \qquad (n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Cov}\{P_i^2, U_n^2\} = 0(n^{-2}\log n)$$

recall from (2.26) that $\mathrm{Var}\{P_i^2\} = 0(i^{-2})$. The Schwarz inequality then implies that $\mathrm{Cov}\{P_i^2, U_n^2\} = 0(i^{-1}n^{-1})$ and (3.42) follows.

To prove

$$(3.43) \qquad (n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Cov}\{z_iP_i, U_n^2\} = 0(n^{-2}\log n)$$

notice from (3.15) that

$$(n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Cov}\{z_iP_i, U_n^2\}$$

$$= \binom{n}{r}^{-1}\Sigma_{c=o}^{r}\binom{r}{c}\binom{n-r}{r-c}(n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Cov}\{z_iP_i, U_n^{(c)}\}.$$

We can therefore establish (3.43) by showing that

$$(3.44) \qquad \mathrm{Cov}\{(n-1)^{-1}\Sigma_{i=r+1}^{n} z_iP_i, U_n^{(1)}\} = 0(n^{-1})$$

and

$$(3.45) \qquad (n-1)^{-1}\Sigma_{i=r+1}^{n}\mathrm{Cov}\{z_iP_i, U_n^{(0)}\} = 0(n^{-2}\log n).$$

Now

$$(3.46) \quad \mathrm{Var}\{(n-1)^{-1}\Sigma_{i=r+1}^{n} z_iP_i\} = (n-1)^{-2}\Sigma_{i=r+1}^{n}\mathrm{Var}\{z_iP_i\}$$

$$+ (n-1)^{-2}2\Sigma_{i=r+1}^{n}\Sigma_{j=r+1}^{i-1}\mathrm{Cov}\{z_iP_i, z_jP_j\}.$$

Then, from (2.38) and (2.52) in the proof of Lemma 2.13

$$\mathrm{Cov}\{z_i P_i, z_j P_j\} = O(i^{-1})$$

for $r < j \leq i$. Thus, from (3.46),

$$(3.47) \qquad \mathrm{Var}\{(n-1)^{-1}\Sigma_{i=r+1}^{n} z_i P_i\} = O(n^{-1}).$$

Therefore (3.44) follows from (3.47), the fact that $\mathrm{Var}\{U_n^{(1)}\} = O(n^{-1})$ and the Schwarz inequality.

To prove (3.45) it is sufficient to prove (by (2.28)) that

$$(3.48) \qquad \mathrm{Cov}\{z_i V_{i-1}^{(h)}, U_n^{(0)}\} = O(i^{-1}n^{-1})$$

and

$$(3.49) \qquad \mathrm{Cov}\{z_i W_{ii}^{*(h)}, U_n^{(0)}\} = O(n^{-2})$$

for $h = 2,3,\cdots,r$ and $i = r+1,r+2,\cdots,n$. Since z_i is independent of $V_{i-1}^{(h)}$, Lemma 2.5 implies that

$$\mathrm{Var}\{z_i V_{i-1}^{(h)}\} = \mathrm{Var}\{z_i\}\mathrm{Var}\{V_{i-1}^{(h)}\} = O(i^{-2})$$

for $h = 2,3,\cdots,r$ and $i = r+1,r+2,\cdots,n$. Also, $\mathrm{Var}\{U_n^{(0)}\} = O(n^{-2})$, so that (3.48) follows from the Schwarz inequality.

To prove (3.49) notice that for $h = 2$

$$\mathrm{Cov}\{z_i W_{ii}^{*(2)}, U_n^{(0)}\} = (i-1)^{-1}\binom{n}{r}^{-1}\binom{n-r}{r}^{-1}$$

$$\cdot\ \Sigma_{j=1}^{i-1}\Sigma^{(0)}\mathcal{E}\{z_i g^{(2)}(X_i,X_j)f(X_{\alpha_1},\cdots,X_{\alpha_r})f(X_{\beta_1},\cdots,X_{\beta_r})\}$$

for $i = r+1,r+2,\cdots,n$. But

$\mathcal{E}\{z_i g^{(2)}(X_i,X_j)f(X_{\alpha_1},\cdots,X_{\alpha_r})f(X_{\beta_1},\cdots,X_{\beta_r})\} = 0$ except possibly when i appears among $(\alpha_1,\cdots,\alpha_r)$ and j appears among $(\beta_1,\cdots,\beta_r)$, or vice versa. The number of possibly non-zero terms is $2\binom{n-2}{r-1}\binom{n-r-1}{r-1}$. Thus, (3.49) holds for h = 2. The cases where h = 3,4,$\cdots$,r follow in analogous fashion. This completes the proof of (3.49).

We have therefore shown that $\mathrm{Var}\{s_{zn}^{*2}\} = r^4\Delta n^{-1} + O(n^{-2}(\log n)^2)$. It is a simple matter to see that $\mathrm{Var}\{(n-1)^{-1}(U_r-U_n)^2\}$ and $\mathrm{Cov}\{s_{zn}^{*2},(n-1)^{-1}(U_r-U_n)^2\}$ are each of order n^{-2}, and so, (3.22) is finally verified.

<u>THEOREM 3.4</u>. <u>Assume</u> $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$. <u>Then</u> s_{zn}^2 <u>converges</u> <u>to</u> $\sigma^2 = r^2\rho_1$ <u>almost surely as</u> $n \to \infty$.

<u>PROOF</u>. Assume that $\Theta = 0$, without loss of generality, and recall that

$$s_{zn}^2 = s_{zn}^{*2} + r(n-1)^{-1}(U_r-U_n)^2 \tag{3.50}$$

where

$$s_{zn}^{*2} = (n-1)^{-1}\Sigma_{i=r+1}^n Z_i^2 - (n-1)^{-1}(n+r)U_n^2 + (n-1)^{-1}2rU_rU_n. \tag{3.51}$$

Clearly $r(n-1)^{-1}(U_r-U_n)^2$ converges to 0 almost surely as $n \to \infty$. By Hoeffding [9] (or Berk [2]), the second and third terms in (3.51) each converge to 0 almost surely as $n \to \infty$. From (3.27) and (2.24)

$$\begin{aligned} A_n &= (n-1)^{-1}\Sigma_{i=r+1}^n[r^2z_i^2 + 2rz_iP_i + P_i^2] \\ &= (n-1)^{-1}r^2\Sigma_{i=r+1}^n z_i^2 + 2r(n-1)^{-1}\Sigma_{i=r+1}^n z_iP_i \\ &\quad + (n-1)^{-1}\Sigma_{i=r+1}^n P_i^2. \end{aligned} \tag{3.52}$$

Now $\mathcal{E}\{z_i^2\} = \mathcal{E}\{g^{(1)2}(X_i)\} = \rho_1$, so that, by the strong law of large numbers, the first term in (3.52) converges to $\sigma^2 = r^2\rho_1$ almost surely as $n \to \infty$. Thus, to complete the proof we need only show that

$$\lim_{n\to\infty} (n-1)^{-1}\Sigma_{i=r+1}^{n} P_i^2 = 0 \qquad \text{(a.s.)} \tag{3.53}$$

and

$$\lim_{n\to\infty} (n-1)^{-1}\Sigma_{i=r+1}^{n} z_i P_i = 0 \qquad \text{(a.s.)}. \tag{3.54}$$

From (2.28)

$$P_i = \Sigma_{h=2}^{r}\binom{r}{h}W_{ii}^{*(h)} - \Sigma_{h=2}^{r}\binom{r}{h}(h-1)V_{i-1}^{(h)} \tag{3.55}$$

for $i = r+1, r+2, \cdots, n$ where $W_{ii}^{*(h)}$ is given by (2.29). From Hoeffding [9], $V_i^{(h)}$ converges almost surely to zero as $i \to \infty$ for $h = 2,3,\cdots,r$. Also, it can be shown, in a proof almost identical to that on pages 108-110 of Wilks [17] (due to Feller [5]), that $W_{ii}^{*(h)}$ converges almost surely to zero as $i \to \infty$ for $h = 2,3,\cdots,r$. Then, (3.55) implies that P_i, and therefore P_i^2, converges almost surely to zero and hence in Césaro-mean a.s. - that is, (3.53) holds. Next, by the Schwarz inequality,

$$\begin{aligned} &|(n-1)^{-1}\Sigma_{i=r+1}^{n} z_i P_i| \\ &\leq [(n-1)^{-1}\Sigma_{i=r+1}^{n} z_i^2]^{1/2}[(n-1)^{-1}\Sigma_{i=r+1}^{n} P_i^2]^{1/2} \end{aligned} \tag{3.56}$$

so that (3.54) holds. This completes the proof.

<u>REMARK 1</u>. If $r = 1$, then $W_{in} = x_i$, $Z_i = x_i$ and $U_n = \bar{x}_n$, the sample mean. Also, both s_{wn}^2 and s_{wn}^2 reduce to the sample variance $(n-1)^{-1}\Sigma_{i=1}^{n}(x_i-\bar{x}_n)^2$. (For notational convenience, when $r = 1$, we

assume that $f(x) = x$.)

REMARK 2. A few comments about the relative merits of s_{wn}^2 and s_{zn}^2 as estimates of $\sigma^2 = r^2\rho_1$ are in order. Both estimates are asymptotically unbiased. However, the bias in the case of s_{zn}^2 is of lower order than that of s_{wn}^2, that is

$$\mathrm{BIAS}(s_{zn}^2) = r^2(r-1)^2\delta_2 n^{-1}\log n + An^{-1} + o(n^{-1})$$

whereas the bias of s_{wn}^2 is of order n^{-1}. Notice that s_{zn}^2 tends to overestimate $\sigma^2 = r^2\rho_1$. The variances of the estimates each have leading terms of order n^{-1}. See (3.14) and (3.22). It is not possible to compare Δ_1 and Δ, as $\beta_2 + (r-1)\beta_3$ may be either negative or non-negative depending upon $f(x_1,\cdots,x_r)$ and the c.d.f. F. The second order term of $\mathrm{Var}\{s_{zn}^2\}$ is of order $n^{-2}(\log n)^2$, which compares with n^{-2}, the order of the second order term of $\mathrm{Var}\{s_{wn}^2\}$. The one advantage, and it is an important one, that s_{zn}^2 has over s_{wn}^2 is that it is by nature more suited for sequential calculation. In s_{wn}^2, each of the W_{in} terms depend upon $x_1,x_2,\cdots,x_n$ and must therefore (in general) be calculated at each stage of the sequential procedure, whereas, in s_{zn}^2, Z_i depends only on the first i observations and need only be calculated once. More will be said about the computation of s_{wn}^2 and s_{zn}^2 at the end of Chapter V. Theoretically, s_{wn}^2 appears a little superior to s_{zn}^2 as an estimate of $\sigma^2 = r^2\rho_1$. However, in many cases, s_{zn}^2 can be calculated with relative ease, and in such cases, might be preferred over s_{wn}^2. Thus, the final choice of estimate depends upon the function $f(x_1,\cdots,x_r)$. In Chapter V the sequential procedure is examined with respect to both s_{wn}^2 and s_{zn}^2.

3.6 Example. Define $\mu = \mathcal{E}\{X_1\}$ and $\mu_j = \mathcal{E}\{(X_1-\mu)^j\}$ for $j = 2,3,\cdots$ (when existent). Assume that $\mu_4 < \infty$ and $\mu_2 > 0$. Let $f(x_1,x_2) = (x_1-x_2)^2/2$, so that $\theta = \mathcal{E}\{(X_1-X_2)^2/2\} = \mu_2$. The corresponding U-statistic is

$$(3.57)\qquad U_n = (n-1)^{-1}\Sigma_{i=1}^n(x_i-\bar{x}_n)^2 = n(n-1)^{-1}(m_2-m_1^2)$$

where $\bar{x}_n$ is the sample mean and $m_j = n^{-1}\Sigma_{i=1}^n x_i^j$ for $j = 1,2,\cdots$. Next, $f_1(x_1) = \mathcal{E}\{(x_1-X_2)^2/2\} = (x_1-\mu)^2/2 + \mu_2/2$ and $\rho_1 = \delta_1 = (\mu_4-\mu_2^2)/4$, so that $\sigma^2 = r^2\rho_1 = \mu_4 - \mu_2^2$.

From (2.4) we obtain $g^{(1)}(x_1) = (x_1-\mu)^2/2 - \mu_2/2$ and $g^{(2)}(x_1,x_2) = -(x_1-\mu)(x_2-\mu)$, so that $\delta_2 = \mu_2^2$. Therefore, it follows from (2.7) that

$$(3.58)\qquad \mathrm{Var}\{U_n\} = (\mu_4-\mu_2^2)n^{-1} + 2\mu_2^2 n^{-1}(n-1)^{-1},$$

as is well known.

We now present s_{wn}^2 and s_{zn}^2, the estimates of $\sigma^2 = \mu_4 - \mu_2^2$. From (2.11)

$$W_{in} = n(n-1)^{-1}[x_i^2 - 2m_1x_i + m_2]$$

for $i = 1,2,\cdots,n$, and so, after some manipulation,

$$(3.59)\qquad s_{wn}^2 = n^3(n-1)^{-3}[m_4 + 8m_1^2m_2 - 4m_1m_3 - m_2^2 - 4m_1^4].$$

The factor $n^3(n-1)^{-3}$ in (3.59) may be omitted without affecting the properties of s_{wn}^2 to any appreciable extent. From (3.20)

$$(3.60)\qquad s_{zn}^2 = (n-1)^{-1}[\Sigma_{i=3}^n Z_i^2 - nU_n^2 + 2U_2^2]$$

where Z_i for $i = 3,4,\cdots$ is given by

$$Z_i = iU_i - (i-1)U_{i-1} \tag{3.61}$$

and U_i is given by (3.57). In this example s_{wn}^2 is just as easy to calculate sequentially as s_{zn}^2, and so, is to be preferred over s_{zn}^2 as an estimate of $\sigma^2 = \mu_4 - \mu_2^2$.

From (3.13) of Theorem 3.1

$$\mathcal{E}\{s_{wn}^2\} = (\mu_4 - \mu_2^2) + (6\mu_2^2 - 2\mu_4)n^{-1} + 0(n^{-2}) \tag{3.62}$$

which indicates that the first order bias of s_{wn}^2 is $(6\mu_2^2 - 2\mu_4)n^{-1}$. From (3.21) of Theorem 3.3

$$\begin{aligned} \mathcal{E}\{s_{zn}^2\} = (\mu_4 - \mu_2^2) &+ 4\mu_2^2 n^{-1} \log n \\ &+ [(4\gamma+1)\mu_2^2 - \mu_4]n^{-1} \\ &+ o(n^{-1}) \end{aligned} \tag{3.63}$$

where $\gamma = 0.5772\ldots$. The dominant term in the bias of s_{zn}^2 is $4\mu_2^2 n^{-1}\log n$, which is non-negative, and so, as we have already noticed, s_{zn}^2 tends to overestimate $\sigma^2 = \mu_4 - \mu_2^2$. The term of order n^{-1} in (3.63) may be negative or non-negative depending upon the c.d.f. F, and therefore, may either help to decrease or increase the bias.

To determine the variances of s_{wn}^2 and s_{zn}^2 assume that $\mu_8 < \infty$ (as well as $\mu_2 > 0$). From the statement of Lemma 2.12

$$\beta_1 = \mathrm{Var}\{g^{(1)2}(X_1)\} = 16^{-1}\mathrm{Var}\{(X_1-\mu)^4 - 2\mu_2(X_1-\mu)^2\}$$

$$= 16^{-1}(\mu_8 + 8\mu_2^2\mu_4 - 4\mu_2\mu_6 - \mu_4^2 - 4\mu_2^4).$$

Next, from section 3.2

$$g_o(x_1) = \mathcal{E}\{g^{(1)}(X_2)g^{(2)}(x_1,X_2)\} = -\mu_3(x_1-\mu)/2.$$

Therefore, by (3.9)

$$\beta_2 = \mathrm{Cov}\{g^{(1)2}(X_1),g_o(X_1)\} = -\mu_3(\mu_5-2\mu_2\mu_3)/8$$

and, by (3.10)

$$\beta_3 = \mathrm{Var}\{g_o(X_1)\} = \mu_3^2\mu_2/4.$$

The variance of s_{wn}^2 is then

(3.64) $$\mathrm{Var}\{s_{wn}^2\} = 16\Delta_1 n^{-1} + 0(n^{-2})$$

where $\Delta_1 = \beta_1 + 4\beta_2 + 4\beta_3$. The variance of s_{zn}^2 is given by

(3.65) $$\mathrm{Var}\{s_{zn}^2\} = 16\Delta n^{-1} + 0(n^{-2}(\log n)^2)$$

where $\Delta = \beta_1 + 2\beta_2 + 2\beta_3$. Notice that $8(\beta_2+\beta_3) = 4\mu_2\mu_3^2 - \mu_3\mu_5$ may be either negative or non-negative depending upon the c.d.f. F, and so, it is not possible to compare the first order terms of (3.64) and (3.65).

For this particular example we might consider

$$s^2 = n^2(n-2)^{-1}(n-3)^{-1}[s_4 - (n^2-3)(n-1)^{-2}s_2^2]$$

as an estimate of $n\,\mathrm{Var}\{U_n\}$ where $s_j = \Sigma_{i=1}^{n}(x_i - \bar{x}_n)^j$ for $j = 2,3,\cdots$. The estimate s^2 is motivated by the classical theory of sampling distributions of sample moments. See, for example, Chapter 12 of Kendall and Stuart [10]. The chief merit of s^2 is that it is an unbiased estimate of $n\,\mathrm{Var}\{U_n\}$. As is evident from (3.59), which may be written as

$$s_{wn}^2 = n^3(n-1)^{-3}[s_4 - s_2^2],$$

s^2 and s_{wn}^2 differ very little (especially for large values of n) with s_{wn}^2 being slightly easier to compute. Since we are mainly concerned with large values of n, we favor s_{wn}^2 over s^2 as an estimate of $n\,\mathrm{Var}\{U_n\} = (\mu_4 - \mu_2^2) + 0(n^{-1})$.

For further examples of U-statistics see Hoeffding [7], [8] and Fraser [6].

CHAPTER IV

CONTRIBUTIONS TO THE ASYMPTOTIC

THEORY OF U-STATISTICS

4.1 Introduction. The results of this chapter were developed with the idea of solving the fixed-width sequential confidence interval problem of Chapter V. We utilize the H-decomposition introduced in Theorem 2.6, namely

$$U_n = \theta + \Sigma_{h=1}^{r} \binom{r}{h} V_n^{(h)},$$

to establish a Kolmogorov-like inequality for U-statistics (Theorem 4.2), and then, to show that, for $\gamma < 1/2$, $n^{\gamma}(U_n - \theta)$ converges almost surely to 0 as $n \to \infty$ (Theorem 4.4). In addition, if for each positive integer s, N_s is a positive integer-valued random variable, then Theorem 4.5 states that the U-statistic U_{N_s} based on $x_1, x_2, \cdots, x_{N_s}$ is asymptotically normal as $s \to \infty$, under certain conditions. Theorem 4.5 is used in Chapter V to establish the asymptotic consistency of the sequential procedure (Theorem 5.2).

4.2 Kolmogorov inequalities. Theorem 2.6 states that for each $h = 1,2,\cdots,r$, $S_n^{(h)} = \binom{n}{h} V_n^{(h)}$ forms a martingale sequence. This fact is used to prove

LEMMA 4.1. *Assume that* $0 < \delta_h < \infty$ *for some* $h = 1,2,\cdots,r$. *Then the following Kolmogorov-like inequality holds: for* $\lambda > 0$

$$(4.1) \qquad P\left\{\max_{h\le\alpha\le n} |S_\alpha^{(h)}| \ge \lambda\delta_h^{1/2}\binom{n}{h}^{1/2}\right\} \le \lambda^{-2}.$$

PROOF. By Lemma 2.5, $\mathcal{E}\{S_n^{(h)2}\} = \binom{n}{h}\delta_h$. Thus, by the Kolmogorov inequality for martingales (see page 399 of Loève [11]), for any $\epsilon > 0$,

$$P\left\{\max_{h\le\alpha\le n} |S_\alpha^{(h)}| \ge \epsilon\right\} \le \epsilon^{-2}\binom{n}{h}\delta_h.$$

Putting $\epsilon = \lambda\delta_h^{1/2}\binom{n}{h}^{1/2}$ completes the proof of (4.1).

REMARKS. The Kolmogorov-like inequality (4.1) can be applied directly to show that for any $\gamma < h/2$, $n^\gamma V_n^{(h)}$ converges almost surely to 0 as $n \to \infty$. See pages 109-110 of Wilks [17]. However, a different approach leads to stronger convergence results, as we shall see in Theorem 4.3. In a somewhat classical approach a proof similar to that on pages 107-108 of Wilks [17], utilizing Lemmas 2.4 and 2.5, can be used to establish (4.1).

We now use Lemma 4.1 to derive a Kolmogorov-like inequality for a U-statistic. From Theorem 2.6

$$S_n = \binom{n}{r}\theta + \binom{n}{r}\Sigma_{h=1}^{r}\binom{r}{h}\binom{n}{h}^{-1}S_n^{(h)}$$

where we have set $S_n = \binom{n}{r}U_n$ for $n \ge r$.

THEOREM 4.2. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and* $\delta_1 > 0$, *and let* $\delta = \Sigma_{h=1}^{r}\binom{r}{h}\delta_h^{1/2}$. *Then*

$$P\left\{\max_{r\le\alpha\le n} |S_\alpha - \binom{\alpha}{r}\theta| \ge \lambda\delta n^{-1/2}\binom{n}{r}\right\} \le r\lambda^{-2}$$

for $\lambda > 0$.

PROOF. First, note that $\delta_h < \infty$ for $h = 1,2,\cdots,r$ as a consequence of our assumption, Lemma 2.1(i), Lemma 2.7 and the Schwarz inequality. Define the events

$$E = \left\{ \max_{r \leq \alpha \leq n} \left| S_\alpha - \binom{\alpha}{r} \theta \right| \geq \lambda \delta n^{-1/2} \binom{n}{r} \right\}$$

and

$$E_h = \left\{ \max_{r \leq \alpha \leq n} \left| S_\alpha^{(h)} \right| \geq \lambda \delta_h^{1/2} \binom{n}{h}^{1/2} \right\}$$

for $h = 1,2,\cdots,r$. If each of $\bar{E}_1, \bar{E}_2, \cdots, \bar{E}_r$ occurs, then for $\alpha \geq r$

$$\begin{aligned} \left| S_\alpha - \binom{\alpha}{r} \theta \right| &\leq \binom{\alpha}{r} \Sigma_{h=1}^r \binom{r}{h} \binom{\alpha}{h}^{-1} \left| S_\alpha^{(h)} \right| \\ &< \binom{n}{r} \Sigma_{h=1}^r \binom{r}{h} \binom{n}{h}^{-1} \lambda \delta_h^{1/2} \binom{n}{h}^{1/2} \\ &\leq \lambda \delta n^{-1/2} \binom{n}{r} . \end{aligned}$$

This implies that $E \subseteq \cup_{h=1}^r E_h$, so that, by Lemma 4.1

$$P(E) \leq P(\cup_{h=1}^r E_h) \leq \Sigma_{h=1}^r P(E_h) \leq r\lambda^{-2},$$

which completes the proof.

4.3 Strong convergence results. The main theorem is

THEOREM 4.3. *Let $\{b_n\}_2^\infty$ be a positive increasing sequence of real numbers with* $\lim_{n\to\infty} b_n = \infty$. *If, for some* $h = 1,2,\cdots,r$, $0 < \delta_h < \infty$ *and*

$$\Sigma_{j=1}^\infty 2^{hj} b_{2^j}^{-2} < \infty, \tag{4.3}$$

then $b_n^{-1} S_n^{(h)}$ *converges almost surely to* 0 *as* $n \to \infty$.

PROOF. From Lemma 4.1 for any $\epsilon > 0$

$$P\{\max_{h\leq\alpha\leq n}|S_\alpha^{(h)}| \geq \epsilon b_n\} \leq \epsilon^{-2}b_n^{-2}\delta_h\binom{n}{h}. \tag{4.4}$$

Then (4.3), (4.4) and the Borel-Cantelli lemma imply that

$$\lim_{j\to\infty} b_{2^j}^{-1}S_{2^j}^{(h)} = 0 \quad \text{(a.s.)}. \tag{4.5}$$

Next, define

$$T_j = \max_{2^j\leq n<2^{j+1}} |S_n^{(h)} - S_{2^j}^{(h)}|$$

for $j = 1,2,\cdots$ and $Y_n = S_{2^j+n}^{(h)} - S_{2^j}^{(h)}$ for $n = 1,2,\cdots$. Then $\{Y_n\}_1^\infty$ is a martingale sequence, so that, by the Kolmogorov inequality for martingales (page 399 of Loève [11])

$$P\{T_j \geq \epsilon b_{2^j}\} \leq \epsilon^{-2}b_{2^j}^{-2}\mathcal{E}\{Y_{2^j}\}^2. \tag{4.6}$$

Now, since $\mathcal{E}\{S_{2^{j+1}}^{(h)}S_{2^j}^{(h)}\} = \mathcal{E}\{S_{2^j}^{(h)}\}^2$, then

$$\begin{aligned} \mathcal{E}\{Y_{2^j}\}^2 &= \mathcal{E}\{S_{2^{j+1}}^{(h)}\}^2 - \mathcal{E}\{S_{2^j}^{(h)}\}^2 \\ &= \delta_h[\binom{2^{j+1}}{h} - \binom{2^j}{h}]. \end{aligned} \tag{4.7}$$

A little computation shows that $\binom{2^{j+1}}{h} - \binom{2^j}{h} \leq K2^{hj}$ for some constant $0 < K < \infty$. Thus (4.3), (4.6), (4.7) and the Borel-Cantelli lemma imply that

$$\lim_{j\to\infty} b_{2^j}^{-1}T_j = 0 \quad \text{(a.s.)}. \tag{4.8}$$

Now, for each n, let j be the positive integer such that

$2^j \leq n < 2^{j+1}$. Then, since $\{b_n\}_2^\infty$ is positive increasing,

$$(4.9) \qquad b_n^{-1}|S_n^{(h)}| \leq b_{2^j}^{-1}|S_{2^j}^{(h)}| + b_{2^j}^{-1}T_j$$

for $n = h, h+1, \cdots$. Combining (4.5), (4.8) and (4.9) completes the proof of the theorem.

COROLLARY. *Assume* $0 < \delta_h < \infty$ *for some* $h = 1,2,\cdots,r$.

(i) *If* $\epsilon > 1/2$, *then* $n^{-h/2}(\log n)^{-\epsilon}S_n^{(h)}$ *converges almost surely to* 0 *as* $n \to \infty$.

(ii) *If* $\gamma < h/2$, *then* $n^\gamma V_n^{(h)}$ *converges almost surely to* 0 *as* $n \to \infty$.

(iii) *If* $\gamma < 1$, *then* $n^\gamma R_n$ *converges almost surely to* 0 *as* $n \to \infty$, *where* R_n *is defined by* (2.6).

PROOF. To prove (i) let $b_n = n^{h/2}(\log n)^\epsilon$ so that, since $2\epsilon > 1$, (4.3) becomes

$$(\log 2)^{-2\epsilon}\Sigma_{j=1}^\infty j^{-2\epsilon} < \infty.$$

To prove (ii) let $b_n = n^{h-\gamma}$. Then, since $h - 2\gamma > 0$, (4.3) becomes

$$\Sigma_{j=1}^\infty 2^{-j(h-2\gamma)} < \infty.$$

Thus $n^{\gamma-h}S_n^{(h)}$ converges almost surely to 0 as $n \to \infty$, which is equivalent to (ii).

Part (iii) follows directly from (ii).

REMARK. Theorem 4.3 may be easily generalized to the following result. Let $\{b_n\}_2^\infty$ be a positive increasing sequence of real numbers with $\lim_{n\to\infty} b_n = \infty$. If, for $n \geq r$, S_n forms a zero-mean martingale sequence such that $\mathcal{E}|S_n|^s\} = 0(n^h)$ for some $s \geq 1$ and some $h > 0$ and

$$\Sigma_{j=1}^{\infty} 2^{hj} b_{2^j}^{-s} < \infty,$$

then $b_n^{-1} S_n$ converges almost surely to 0 as $n \to \infty$.

THEOREM 4.4. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and that* $\delta_1 > 0$. *If* $\gamma < 1/2$, *then* $n^{\gamma}(U_n - \theta)$ *converges almost surely to* 0 *as* $n \to \infty$.

PROOF. The theorem follows directly from the H-decomposition (2.6) and Corollary (ii) of Theorem 4.3.

4.4 *The asymptotic normality of* U_N. In this section we extend to U-statistics Anscombe's theorem on the asymptotic normality of averages of a random number of I.I.D. random variables. From Theorem 2.6 recall that the H-decomposition is

$$U_n = \theta + rV_n^{(1)} + R_n$$

where $R_n = \Sigma_{h=2}^{r} \binom{r}{h} V_n^{(h)}$ and $V_n^{(1)} = n^{-1}\Sigma_{i=1}^{n}(f_1(x_i) - \theta)$.

THEOREM 4.5. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and* $\rho_1 > 0$. *Denote the standard normal c.d.f. by* $\Phi(x)$. *Let* $\{n_s\}$ *be an increasing sequence of positive integers tending to* ∞ *as* $s \to \infty$, *and* $\{N_s\}$ *be a sequence of proper random variables taking on positive integer values such that*

$$p - \lim_{s\to\infty} n_s^{-1} N_s = 1.$$

Then, with $\sigma^2 = r^2 \rho_1$,

$$\lim_{s\to\infty} P\{(U_{N_s} - \theta) \leq n_s^{-1/2} x\sigma\} = \Phi(x).$$

PROOF. Anscombe [1] introduced the following situation. Let

$\{Y_n\}$ be a sequence of random variables. Assume that there exists a real number Θ, a sequence of positive numbers $\{w_n\}$ and a c.d.f. $F(x)$ such that:

C1. For any x such that $F(x)$ is continuous

$$\lim_{n\to\infty} P\{Y_n - \Theta \leq x w_n\} = F(x).$$

C2. Given $\epsilon > 0$ and $\eta > 0$ there exists a large $\nu_{\epsilon,\eta}$ and a small $c > 0$ such that for any $n > \nu_{\epsilon,\eta}$

$$P\{|Y_{n'} - Y_n| < \epsilon w_n \text{ for all } n' \text{ such that } |n'-n| < cn\} \geq 1 - \eta.$$

Theorem 1 of Anscombe [1] states that if $\{Y_n\}$ satisfies C1 and C2, then

$$\lim_{s\to\infty} P\{Y_{N_s} - \Theta \leq x w_{n_s}\} = F(x)$$

at all continuity points of $F(x)$. Let C3 be the condition that $\{w_n\}$ is decreasing, tending to 0 as $n \to \infty$ and $\lim_{n\to\infty} w_{n+1}^{-1} w_n = 1$. Theorem 3 of Anscombe [1] states that C2 is satisfied if Y_n is the average of n I.I.D. random variables, if C1 and C3 hold, and if $F(x)$ is continuous.

We now apply these results to our situation. Here $w_n = n^{-1/2}\sigma$ so that C3 is satisfied. Hoeffding [7] has shown that U_n satisfies C1 with $F(x) = \Phi(x)$. (See Remark 2 following Theorem 2.6.) We now show that U_n satisfies C2. First, $rV_n^{(1)}$ satisfies C2 by Theorem 3 of Anscombe [1]. Thus, from the H-decomposition, given $\epsilon > 0$ and $\eta > 0$ there exists a $\nu_{\epsilon,\eta}$ and a $c > 0$ such that

$$P\{|U_{n'} - U_n - R_{n'} + R_n| < \epsilon\sigma n^{-1/2} \text{ for all } n' \text{ such that } |n'-n| < cn\} \geq 1 - \eta$$

for all $n > \nu_{\epsilon,\eta}$, and hence

(4.10) $$P\{|U_{n'}-U_n|-|R_{n'}-R_n| < \epsilon\sigma n^{-1/2} \text{ for all } n' \text{ such that } |n'-n| < cn\} \geq 1 - \eta$$

for all $n > \nu_{\epsilon,\eta}$. By Corollary (iii) of Theorem 4.3, $\lim_{n\to\infty} n^{1/2}R_n = 0$ (a.s.). Thus given $\epsilon > 0$ and $\eta > 0$ there exists an $N'_{\epsilon,\eta}$ such that

$$P\{|R_{n'}-R_n| < \epsilon\sigma n^{-1/2} \text{ for } n' = n,n+1,\cdots,n+k\} \geq 1 - \eta$$

for all $n > N'_{\epsilon,\eta}$ and for $k = 0,1,\cdots$. Let $N_{\epsilon,\eta} = (1-c)^{-1}N'_{\epsilon,\eta}$. Then $n > N_{\epsilon,\eta}$ implies that $n(1-c) > N'_{\epsilon,\eta}$. Therefore

(4.11) $$P\{|R_{n'}-R_n| < \epsilon\sigma n^{-1/2} \text{ for all } n' \text{ such that } |n'-n| < cn\} \geq 1 - \eta$$

for all $n > N_{\epsilon,\eta}$. Let $\nu = \max(\nu_{\epsilon,\eta},N_{\epsilon,\eta})$. Define the events

$$A = \{|U_{n'}-U_n|-|R_{n'}-R_n| < \epsilon\sigma n^{-1/2} \text{ for all } n' \text{ such that } |n'-n| < cn\},$$

$$B = \{|R_{n'}-R_n| < \epsilon\sigma n^{-1/2} \text{ for all } n' \text{ such that } |n'-n| < cn\}$$

and

$$C = \{|U_{n'}-U_n| < 2\epsilon\sigma n^{-1/2} \text{ for all } n' \text{ such that } |n'-n| < cn\}.$$

Therefore $A \cap B \subseteq C$, and so, from (4.10) and (4.11)

$$P(C) \geq P(A \cap B) = P(A) - P(A \cap \bar{B}) \geq 1 - 2\eta$$

for all $n > \nu$. Thus U_n satisfies C2, and the asymptotic normality of

U_N follows from Theorem 1 of Anscombe [1].

The following corollary is a consequence of Theorems 3.2 and 4.5.

COROLLARY. *Under the assumptions of Theorem 4.5*

$$\lim_{s\to\infty} P\left\{n_s^{1/2} s_{wN_s}^{-1} (U_{N_s} - \theta) \leq x\right\} = \Phi(x) \tag{4.12}$$

where s_{wN_s} *is given by* (3.12).

REMARKS. As a result of Theorem 3.4, s_{wN_s} in (4.12) may be replaced by s_{zN_s}, which is given by (3.20). As a special case of (4.12) it follows that $n^{1/2} s_{wn}^{-1}(U_n - \theta)$ is asymptotically normal with mean 0 and variance 1 as $n \to \infty$. (We may, of course, again substitute s_{zn} for s_{wn}.)

CHAPTER V

SEQUENTIAL FIXED-WIDTH CONFIDENCE INTERVALS FOR REGULAR FUNCTIONALS

5.1 Introduction. Assume that $X_1, X_2, \cdots$ are I.I.D. random variables. Let $f(x_1, \cdots, x_r)$ be the symmetric kernel of a U-statistic U_n whose expectation is θ. The problem is to find a sequential confidence interval for θ of fixed-width 2d, where $d > 0$, and such that the coverage probability either equals, or approaches in some way, a specified α, where $0 < \alpha < 1$. The problem was solved by Chow and Robbins [3] for a special U-statistic, the sample mean. To adapt their procedure to deal with a general U-statistic is the "raison d'être" of Chapter V. Chow and Robbins [3] use $n^{-1}s_n^2$, where s_n^2 is the sample variance, to estimate the unknown variance of the sample mean. In section 5.2, $n^{-1}s_{wn}^2$ is used as an estimate of the unknown variance of U_n, and in section 5.3, $n^{-1}s_{zn}^2$ is used.

The sequential procedure may be simply described as follows: at each stage of sampling the U-statistic U_n and an estimate of its variance are calculated, and sampling is terminated as soon as the approximate coverage probability for the interval $[U_n-d, U_n+d]$, based on a normal approximation, is at least α. It is shown that the coverage probability is, in a certain sense, asymptotically α; that is, the sequential procedures are consistent (Theorem 5.2). It is also shown that the expected sample size of the procedures is asymptotically equal to the sample size of the corresponding non-sequential scheme used when the variance of the U-statistic is known (Theorems 5.3 and 5.8); that is, the sequential procedures are efficient.

In section 5.4 the procedures are illustrated with the estimation of (1) the variance of X_1, and (2) the probability of concordance for bivariate X_1.

5.2 The sequential procedure using s_{wn}^2. For $0 < \alpha < 1$ define "a" (> 0) by

$$(2\pi)^{-1/2} \int_{-a}^{+a} \exp(-u^2/2)du = \alpha.$$

Let $\{a_n\}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} a_n = a$. For $d > 0$ define the stopping variable

(5.1) $N(d) =$ smallest integer $k \geq r$ such that $s_{wk}^2 \leq kd^2a_k^{-2}$.

Define a closed confidence interval $I_N = [U_N-d, U_N+d]$ of width 2d. Notice that N and I_N have properties similar to those stated in the theorem appearing in Chow and Robbins [3].

LEMMA 5.1. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and* $\rho_1 > 0$. *Then*

(i) $N(d)$ *is well-defined and is a non-increasing function of* d,

(ii) $\lim_{d\to o} N(d) = \infty$ (*a.s.*),

(iii) $\lim_{d\to o} \mathcal{E}\{N(d)\} = \infty$, *and*

(iv) $\lim_{d\to o} a^{-2}\sigma^{-2}d^{-2}N(d) = 1$ (*a.s.*).

PROOF. Recall from Theorem 3.2 that $\lim_{n\to\infty} s_{wn}^2 = \sigma^2$ (a.s.). Let $y_n = \sigma^{-2}s_{wn}^2$, $f(n) = a_n^{-2}na^2$ and $t = d^{-2}a^2\sigma^2$. Then all parts of the lemma follow from Lemma 1 of Chow and Robbins [3].

THEOREM 5.2. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and* $\rho_1 > 0$. *Then*

$$\lim_{d\to o} P\{\theta \in I_N\} = \alpha.$$

Hence, for sufficiently short intervals, the coverage probability is approximately equal to α.

PROOF. Let $t = d^{-2}a^2\sigma^2$ and let N_t be defined by (5.1) with d^2 replaced by $t^{-1}a^2\sigma^2$. (Note that $N_t = N(d)$.) Refer to Theorem 4.5 and identify N_t with N_s and t with n_s. Then

$$P\{\theta \in I_N\} = P\{|U_{N_t} - \theta| \leq d\} = P\{|U_{N_t} - \theta| \leq t^{-1/2}a\sigma\}$$

and so the theorem follows.

THEOREM 5.3. Assume $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ and $\rho_1 > 0$. Then

$$\lim_{d\to o} d^2a^{-2}\sigma^{-2}\mathcal{E}\{N(d)\} = 1. \tag{5.2}$$

Before we tackle the proof of Theorem 5.3 we establish a series of four lemmas, which are required in the proof.

LEMMA 5.4. Let $X_1,X_2,\cdots$ be I.I.D. random variables and Y_n a function of $X_1,X_2,\cdots,X_n$. For each $t > 0$ let N_t be a positive integer-valued random variable depending on $(X_1,X_2,\cdots)$ such that the event $\{N_t = n\}$ is in $\mathcal{B}_n$, the σ-field generated by $\{X_1,X_2,\cdots,X_n\}$ for $n = 1,2,\cdots$ (i.e., N_t is a stopping variable). If $\lim_{n\to\infty} Y_n = \theta$ (a.s.) and $\lim_{t\to\infty} N_t = \infty$ (a.s.), then $\lim_{t\to\infty} Y_{N_t} = \theta$ (a.s.).

PROOF. Define the events

$$A = \{(X_1,X_2,\cdots)|\lim_{t\to\infty} N_t = \infty\},\ B = \{(X_1,X_2,\cdots)|\lim_{n\to\infty} Y_n = \theta\}$$

and

$$C = \{(X_1, X_2, \cdots) \,|\, \lim_{t\to\infty} Y_{N_t} = \theta\}.$$

Then $P(A) = P(B) = 1$, which implies that $P(A \cap B) = 1$. But, it can easily be shown that $A \cap B \subseteq C$. Thus $P(C) = 1$.

LEMMA 5.5. *If* $\mathcal{E}\{|f(X_1,\cdots,X_r)|\} < \infty$, *then* $\{U_n\}_r^\infty$ *is a reverse martingale*.

PROOF. The proof appears in Berk [2], but, because of its simple nature and the fact that it is referred to several times in the ensuing pages, is repeated here. Let $n > m \geq r$ and $(\alpha_1,\cdots,\alpha_r)$ be any r-combination from $\{1,2,\cdots,n\}$. Then

$$\mathcal{E}\{f(X_{\alpha_1},\cdots,X_{\alpha_r})|U_n,U_{n+1},\cdots\}$$

$$= \mathcal{E}\{f(X_1,\cdots,X_r)|U_n,U_{n+1},\cdots\} = q \quad \text{(say)}.$$

Sum over all $\binom{n}{r}$ combinations and obtain

$$\Sigma^{(n,r)}\mathcal{E}\{f(X_{\alpha_1},\cdots,X_{\alpha_r})|U_n,U_{n+1},\cdots\} = \binom{n}{r} q$$

so that $\mathcal{E}\{U_n|U_n,U_{n+1},\cdots\} = q$. That is $U_n = q$ (a.s.). Next, sum over all $\binom{m}{r}$ combinations from $\{1,2,\cdots,m\}$ and obtain $\mathcal{E}\{U_m|U_n,U_{n+1},\cdots\} = q$. Thus, $\mathcal{E}\{U_m|U_n,U_{n+1},\cdots\} = U_n$ (a.s.), and $\{U_n\}_r^\infty$ is a reverse martingale.

LEMMA 5.6. *If* $\mathcal{E}\{|f(X_1,\cdots,X_r)|\} < \infty$, *then for any* $\epsilon > 0$

$$\mathcal{E}\{\sup_n n^{-\epsilon}|U_n|\} < \infty. \tag{5.3}$$

PROOF. This method of proof by truncation is similar to that in Hoeffding [9] and Sen [13]. Also see Siegmund [14]. Define

$$f'(x_{\alpha_1},\cdots,x_{\alpha_r}) = \begin{cases} f(x_{\alpha_1},\cdots,x_{\alpha_r}) & \text{if } |f(\cdots)| \le (\max_j \alpha_j)^{\varepsilon/2} \\ 0 & \text{otherwise} \end{cases}$$

and $f''(x_{\alpha_1},\cdots,x_{\alpha_r}) = f(x_{\alpha_1},\cdots,x_{\alpha_r}) - f'(x_{\alpha_1},\cdots,x_{\alpha_r})$. Then set

$$S_n = \Sigma^{(n,r)} f(x_{\alpha_1},\cdots,x_{\alpha_r}),$$

$$S_n' = \Sigma^{(n,r)} f'(x_{\alpha_1},\cdots,x_{\alpha_r})$$

and

$$S_n'' = \Sigma^{(n,r)} f''(x_{\alpha_1},\cdots,x_{\alpha_r}).$$

(a) To prove $\mathcal{E}\{\sup_n n^{-(r+\varepsilon)}|S_n'|\} < \infty$, note that

$$\begin{aligned} \sup_n n^{-(r+\varepsilon)}|S_n'| &\le \sup_n n^{-(r+\varepsilon)} \Sigma^{(n,r)} |f'(x_{\alpha_1},\cdots,x_{\alpha_r})| \\ &\le \sup_n n^{-(r+\varepsilon)} \Sigma^{(n,r)} (\max_j \alpha_j)^{\varepsilon/2} \\ &< \sup_n \Sigma^{(n,r)} (\max_j \alpha_j)^{-r-\varepsilon/2} \\ &\le \sup_n \Sigma_{j=r}^{n} \binom{j-1}{r-1} j^{-r-\varepsilon/2} \\ &< \Sigma_{j=r}^{\infty} j^{-1-\varepsilon/2} < \infty. \end{aligned}$$

(b) To prove $\mathcal{E}\{\sup_n n^{-(r+\varepsilon)}|S_n''|\} < \infty$, note that

$$\mathcal{E}\{\sup_n n^{-(r+\epsilon)}|S_n''|\} \leq \mathcal{E}\{\sup_n n^{-(r+\epsilon)}\Sigma^{(n,r)}|f''(X_{\alpha_1},\cdots,X_{\alpha_r})|\}$$

$$= \mathcal{E}\{\sup_n n^{-(r+\epsilon)}\Sigma_{j=r}^{n}\Sigma_{\alpha's}^{(j-1,r-1)}|f''(X_j,X_{\alpha_2},\cdots,X_{\alpha_r})|\}$$

$$\leq \mathcal{E}\{\Sigma_{j=r}^{\infty} j^{-(r+\epsilon)}\Sigma_{\alpha's}^{(j-1,r-1)}|f''(X_j,X_{\alpha_2},\cdots,X_{\alpha_r})|\}$$

$$\leq \Sigma_{j=r}^{\infty} j^{-(r+\epsilon)}\Sigma_{\alpha's}^{(j-1,r-1)}\mathcal{E}\{|f''(X_j,X_{\alpha_2},\cdots,X_{\alpha_r})|\}$$

$$= \Sigma_{j=r}^{\infty} j^{-(r+\epsilon)}\binom{j-1}{r-1}\int_{[|f(\cdots)|>j^{\epsilon/2}]}|f(x_1,\cdots,x_r)|\Pi_{i=1}^{r}dF(x_i)$$

$$\leq \Sigma_{j=r}^{\infty} j^{-1-\epsilon}b_j$$

$$\leq b_r\Sigma_{j=r}^{\infty} j^{-1-\epsilon} < \infty$$

where we have set

$$b_j = \int_{[|f(x_1,\cdots,x_r)|>j^{\epsilon/2}]}|f(x_1,\cdots,x_r)|\Pi_{i=1}^{r}dF(x_i)$$

for $j = r,r+1,\cdots$, so that $b_j \geq 0$ and $b_j \geq b_{j+1}$.

(c) Finally, we have $S_n = S_n' + S_n''$ and

$$\mathcal{E}\{\sup_n n^{-(r+\epsilon)}|S_n|\} \leq \mathcal{E}\{\sup_n n^{-(r+\epsilon)}|S_n'|\} + \mathcal{E}\{\sup_n n^{-(r+\epsilon)}|S_n''|\}$$

which, along with (a) and (b), proves the lemma.

A positive integer-valued random variable M depending on $(X_1,X_2,\cdots)$ such that, for $n = 1,2,\cdots$, the event $\{M = n\}$ is in $\mathcal{B}_n'$, the σ-field generated by $\{X_n,X_{n+1},\cdots\}$, is called a "reverse stopping variable". The following lemma appears in Simons [15] and follows from Theorem 2.2 on page 302 of Doob [4].

LEMMA 5.7. *Let* $Z_{-m_2}, \cdots, Z_{-m_1}$ *be a martingale where* $-\infty < m_1 < m_2 \leq \infty$ *and let* M *be a reverse stopping variable with* $P\{m_1 \leq M \leq m_2\} = 1$. *Then* $\mathcal{E}\{Z_{-M}\} = \mathcal{E}\{Z_{-m_1}\}$.

PROOF OF THEOREM 5.3.

(a) As in Simons [15] define a reverse stopping variable for $d > 0$ by

$$(5.4)\quad M = \begin{cases} \text{last integer } n \geq n_o \text{ such that } s_{wn}^2 > nd^2 a_n^{-2} & \text{if there is such an } n \\ n_o - 1 & \text{if } s_{wn}^2 \leq nd^2 a_n^{-2} \text{ for all } n \geq n_o \\ \infty & \text{if } s_{wn}^2 > nd^2 a_n^{-2} \text{ infinitely often} \end{cases}$$

where $n_o \geq r + 1$. Let I represent the indicator function and define t and N_t as in the proof of Theorem 5.2. Then for every $t > 0$

$$\begin{aligned} N_t &\leq n_o I_{[M=n_o-1]} + (M+1) I_{[M\geq n_o]} \\ &= MI_{[M\geq n_o]} + n_o I_{[M=n_o-1]} + I_{[M\geq n_o]} \\ &\leq d^{-2} a_M^2 s_{wM}^2 + n_o I_{[M\geq n_o-1]} \\ &\leq ta^{-2}\sigma^{-2} a_M^2 s_{wM}^2 + n_o . \end{aligned}$$

Thus, for every $t > 0$

$$(5.5)\qquad t^{-1}\mathcal{E}\{N_t\} \leq a^{-2}\sigma^{-2}\mathcal{E}\{a_M^2 s_{wM}^2\} + t^{-1} n_o .$$

(b) We now prove $\lim_{t\to\infty} \mathcal{E}\{s_{wM}^2\} = \sigma^2$. From expression (3.18) in the proof of Theorem 3.1

$$\mathcal{E}\{s_{wM}^2\} = r^2\mathcal{E}\{U_M^{(1)}-U_M^{(0)}\} + \Sigma_{c=o}^r \mathcal{E}\{\alpha_M(c)U_M^{(c)}\}. \tag{5.6}$$

Define $Z_{-n}^{(c)} = U_n^{(c)}$ and $Z_{-\infty}^{(c)} = \lim_{n\to\infty} Z_{-n}^{(c)}$ for $c = 0,1,\cdots,r$. Then $Z_{-\infty}^{(c)} = \lim_{n\to\infty} U_n^{(c)} = \rho_c + \theta^2$ (a.s.) for $c = 0,1,\cdots,r$. (Recall that $\rho_o = 0$.) Then, by Lemma 5.5, $\{Z_{-\infty}^{(c)},\cdots,Z_{-r}^{(c)}\}$ is a martingale. Therefore, from Lemma 5.7 with $m_1 = n_o - 1$ and $m_2 = \infty$, we obtain

$$\mathcal{E}\{U_M^{(c)}\} = \mathcal{E}\{U_{n_o-1}^{(c)}\} = \rho_c + \theta^2 \tag{5.7}$$

for $c = 0,1,\cdots,r$. In particular, $\mathcal{E}\{U_M^{(1)}\} = \rho_1 + \theta^2$ and $\mathcal{E}\{U_M^{(0)}\} = \theta^2$. From (5.1) and (5.4) note that, for every $t > 0$, $N_t \leq M + 1$, so that, as a consequence of Lemma 5.1(ii), $\lim_{t\to\infty} M = \infty$ (a.s.). Also, Lemma 5.4 implies that $\lim_{t\to\infty} U_M^{(c)} = \rho_c + \theta^2$ (a.s.) for $c = 0,1,\cdots,r$. Now $\alpha_n(c) = 0(n^{-1})$, so that $\lim_{t\to\infty} \alpha_M(c)U_M^{(c)} = 0$ (a.s.) for $c = 0,1,\cdots,r$. Furthermore, by Lemma 5.6, $\mathcal{E}\{\sup_n\alpha_n(c)|U_n^{(c)}|\} < \infty$ for $c = 0,1,\cdots,r$. We then use the Lebesgue dominated convergence theorem to obtain

$$\lim_{t\to\infty} \Sigma_{c=o}^r \mathcal{E}\{\alpha_M(c)U_M^{(c)}\} = 0. \tag{5.8}$$

Finally, from (5.6), (5.7) and (5.8) we conclude that

$$\lim_{t\to\infty} \mathcal{E}\{s_{wM}^2\} = \sigma^2. \tag{5.9}$$

(c) Guided by Pratt [12], we next show that

$$\lim_{t\to\infty} \mathcal{E}\{a_M^2 s_{wM}^2\} = a^2\sigma^2. \tag{5.10}$$

From Lemma 5.4 it follows that $\lim_{t\to\infty} s_{wM}^2 = \sigma^2$ (a.s.) and

$\lim_{t\to\infty} a_M^2 s_{wM}^2 = a^2\sigma^2$ (a.s.). Now, let $A = \inf_n a_n^2$ and $B = \sup_n a_n^2$. Then, for every $t > 0$, $As_{wM}^2 \leq a_M^2 s_{wM}^2 \leq Bs_{wM}^2$. Thus

$$0 \leq a^2\sigma^2 - A\sigma^2 = \mathcal{E}\{\lim_{t\to\infty} (a_M^2 s_{wM}^2 - As_{wM}^2)\}$$

and, by Fatou's lemma,

$$(5.11) \qquad 0 \leq a^2\sigma^2 - A\sigma^2 \leq \liminf_{t\to\infty} \mathcal{E}\{a_M^2 s_{wM}^2 - As_{wM}^2\}$$

$$= \liminf_{t\to\infty} \mathcal{E}\{a_M^2 s_{wM}^2\} - A\sigma^2.$$

Also

$$0 \leq B\sigma^2 - a^2\sigma^2 = \mathcal{E}\{\lim_{t\to\infty} (Bs_{wM}^2 - a_M^2 s_{wM}^2)\}$$

and, by invoking Fatou's lemma once more,

$$(5.12) \qquad 0 \leq B\sigma^2 - a^2\sigma^2 \leq \liminf_{t\to\infty} \mathcal{E}\{Bs_{wM}^2 - a_M^2 s_{wM}^2\}$$

$$= B\sigma^2 - \limsup_{t\to\infty} \mathcal{E}\{a_M^2 s_{wM}^2\}.$$

Then (5.10) follows from (5.11) and (5.12).

(d) We conclude from (5.5) and (5.10) that $\limsup_{t\to\infty} t^{-1}\mathcal{E}\{N_t\} \leq 1$. However, Fatou's lemma implies that $\liminf_{t\to\infty} t^{-1}\mathcal{E}\{N_t\} \geq 1$. This completes the proof of Theorem 5.3.

5.3 *The sequential procedure using s_{zn}^2.* The results of section 5.2 also hold if s_{zn}^2 is used as an estimate of $\sigma^2 = r^2\rho_1$. Throughout this section N(d) is defined by (5.1) with s_{zk}^2 substituted for s_{wk}^2. Results analogous to Lemma 5.1 and Theorem 5.2 follow immediately. We now consider the analog of Theorem 5.3.

THEOREM 5.8. *Assume* $\mathcal{E}\{f(X_1,\cdots,X_r)\}^2 < \infty$ *and* $\rho_1 > 0$. *Then*

(5.13) $$\lim_{d\to o} d^2 a^{-2}\sigma^{-2}\mathcal{E}\{N(d)\} = 1.$$

PROOF.

(a) Examine the proof of Theorem 5.3. It is clear that, in analogy to (5.5), for every $t > 0$

(5.14) $$t^{-1}\mathcal{E}\{N_t\} \leq a^{-2}\sigma^{-2}\mathcal{E}\{a_M^2 s_{zM}^2\} + t^{-1}n_o$$

where $t = d^{-2}a^2\sigma^2$. In order to establish (5.13), it is sufficient to prove

(5.15) $$\lim_{t\to\infty} \mathcal{E}\{s_{zM}^2\} = \sigma^2.$$

For, assume for the moment that (5.15) is true. Then it is easy to see that Part (c) of the proof of Theorem 5.3 can be applied, as it stands, to prove that

(5.16) $$\lim_{t\to\infty} \mathcal{E}\{a_M^2 s_{zM}^2\} = a^2\sigma^2.$$

As a result of (5.14), (5.16) and Part (d) of the proof of Theorem 5.3, it follows that (5.13) is true.

We therefore begin the proof of (5.15). First, set $\theta = 0$, without loss of generality. Then from the proof of Theorem 3.3

$$s_{zn}^2 = s_{zn}^{*2} + r(n-1)^{-1}(U_r - U_n)^2$$

and

$$s_{zn}^{*2} = (n-1)^{-1}\Sigma_{i=r+1}^n Z_i^2 - (n-1)^{-1}(n+r)U_n^2 + (n-1)^{-1}2rU_rU_n.$$

We establish (5.15) by proving each of the following four statements:

$$(5.17) \qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} Z_i^2\} = \sigma^2,$$

$$(5.18) \qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}(M+r)U_M^2\} = 0,$$

$$(5.19) \qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}U_r U_M\} = 0 \quad \text{and}$$

$$(5.20) \qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}(U_r-U_M)^2 = 0.$$

(b) Proof of (5.17). From the proof of Lemma 2.12 recall that

$$(5.21) \qquad Z_i^2 = r^2 z_i^2 + 2rz_i P_i + P_i^2$$

where $P_i = \Sigma_{h=2}^{r}\binom{r}{h} Z_i^{(h)}$ and

$$(5.22)\quad P_i^2 = \Sigma_{h=2}^{r}\binom{r}{h}^2 Z_i^{(h)2} + \Sigma_{\substack{h_1=2\\h_1\neq h_2}}^{r}\Sigma_{h_2=2}^{r}\binom{r}{h_1}\binom{r}{h_2} Z_i^{(h_1)} Z_i^{(h_2)}$$

for $i = r+1,\cdots,n$. Now write

$$(5.23)\quad \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} z_i^2\} = \mathcal{E}\{(M-r)^{-1}\Sigma_{i=r+1}^{M} z_i^2\} + \mathcal{E}\{b(M)\Sigma_{i=r+1}^{M} z_i^2\}$$

where $b(M) = 0(M^{-2})$. Clearly, $(n-r)^{-1}\Sigma_{i=r+1}^{n} z_i^2$ is a reverse martingale, so that, by Lemma 5.7

$$(5.24) \qquad r^2\mathcal{E}\{(M-r)^{-1}\Sigma_{i=r+1}^{M} z_i^2\} = r^2\mathcal{E}\{z_{r+1}^2\} = r^2\rho_1 = \sigma^2.$$

Recall that $\lim_{t\to\infty} M = \infty$ (a.s.). Therefore, from Lemmas 5.4 and 5.6 and the Lebesgue dominated convergence theorem, we obtain

$$(5.25) \qquad \lim_{t\to\infty} \mathcal{E}\{b(M)\Sigma_{i=r+1}^{M} z_i^2\} = 0.$$

Putting (5.24) and (5.25) into (5.23) yields

$$(5.26) \qquad \lim_{t\to\infty} r^2 \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} z_i^2\} = \sigma^2.$$

Because of (5.21) and (5.26), in order to prove (5.17), we need only prove

$$(5.27) \qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} P_i^2\} = 0$$

and

$$(5.28) \qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} z_i P_i\} = 0.$$

But, by the Schwarz inequality (for both the summation and the expectation),

$$|\mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} z_i P_i\}| \leq \mathcal{E}\{((M-1)^{-1}\Sigma_{i=r+1}^{M} z_i^2)^{1/2}((M-1)^{-1}\Sigma_{i=r+1}^{M} P_i^2)^{1/2}\}$$

$$(5.29) \qquad \leq [\mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} z_i^2\}\,\mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} P_i^2\}]^{1/2}.$$

From (5.29) and (5.26), notice that (5.27) implies (5.28).

We now prove (5.27). From the proof of Lemma 2.12 recall that

$$(5.30) \qquad z_i^{(h)} = W_{ii}^{*(h)} - (h-1)V_{i-1}^{(h)}$$

so that

$$(5.31) \qquad z_i^{(h)2} = W_{ii}^{*(h)2} - 2(h-1)W_{ii}^{*(h)}V_{i-1}^{(h)} + (h-1)^2 V_{i-1}^{(h)2}$$

for $h = 2,3,\cdots,r$ and $i = r+1,\cdots,n$. Then, because of (5.22), (5.31) and an argument similar to (5.29), in order to prove (5.27) it is sufficient to prove

$$(5.32)\qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} W_{ii}^{*(h)2}\} = 0$$

and

$$(5.33)\qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} V_{i-1}^{(h)2}\} = 0$$

for $h = 2,3,\cdots,r$.

To prove (5.32), for $h = 2,3,\cdots,r$; $c = 0,1,\cdots,h-1$ and $i = r+1,\cdots,n$, define

$$(5.34)\quad V_i^{(h,c)} = \left[\binom{i-1}{h-1}\binom{h-1}{c}\binom{i-h}{h-1-c}\right]^{-1} \cdot \Sigma^{(c)} g^{(h)}(x_j, x_{\alpha_2}, \cdots, x_{\alpha_h}) g^{(h)}(x_j, x_{\beta_2}, \cdots, x_{\beta_h})$$

where the summation $\Sigma^{(c)}$ is over all combinations $(\alpha_2,\cdots,\alpha_h)$ and $(\beta_2,\cdots,\beta_h)$ each chosen from $\{1,2,\cdots,i-1\}$ and such that there are exactly c integers in common. (Compare (5.34) with (2.3) which defines the $U_n^{(c)}$'s.) Then, in analogy to (3.15),

$$(5.35)\qquad W_{ii}^{*(h)2} = \binom{i-1}{h-1}^{-1}\Sigma_{c=o}^{h-1}\binom{h-1}{c}\binom{i-h}{h-1-c} V_i^{(h,c)}$$

for $h = 2,3,\cdots,r$ and $i = r+1,\cdots,n$. Because of (5.35), (5.32) becomes

$$(5.36)\quad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M}\binom{i-1}{h-1}^{-1}\binom{i-h}{h-1} V_i^{(h,0)}\}$$
$$+ \lim_{t\to\infty} \Sigma_{c=1}^{h-1}\binom{h-1}{c}\mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M}\binom{i-1}{h-1}^{-1}\binom{i-h}{h-1-c} V_i^{(h,c)}\} = 0$$

for h = 2,3,$\cdots$,r, which is to be proved.

We now examine the first term of (5.36), which may be written as

$$(5.37)\qquad \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} V_i^{(h,0)}\}$$
$$+ \lim_{t\to\infty} \mathcal{E}\{(M-1)^{-1}\Sigma_{i=r+1}^{M} a(h,i)V_i^{(h,0)}\}$$

where $a(h,i) = 0(i^{-1})$ for h = 2,3,$\cdots$,r. Notice that the first term of (5.37) is equal to

$$(5.38)\qquad \lim_{t\to\infty} \mathcal{E}\{(M-r)^{-1}\Sigma_{i=r+1}^{M} V_i^{(h,0)}\}$$
$$+ \lim_{t\to\infty} \mathcal{E}\{b(M)\Sigma_{i=r+1}^{M} V_i^{(h,0)}\}$$

for h = 2,3,$\cdots$,r and where $b(M) = 0(M^{-2})$. By Lemma 2.4, $\mathcal{E}\{V_i^{(h,0)}\} = 0$ for h = 2,3,$\cdots$,r and i = r+1,$\cdots$,n. Also, $(n-r)^{-1}\Sigma_{i=r+1}^{n} V_i^{(h,0)}$ is a reverse martingale (the proof is similar to that of Lemma 5.5) so that, by Lemma 5.7,

$$(5.39)\qquad \mathcal{E}\{(M-r)^{-1}\Sigma_{i=r+1}^{M} V_i^{(h,0)}\} = 0$$

for h = 2,3,$\cdots$,r. Next, Theorem 4.3 can be adapted to show that $\lim_{n\to\infty} b(n)\Sigma_{i=r+1}^{n} V_i^{(h,0)} = 0$ (a.s.), and also, Lemma 5.6 can be adapted to show that $\mathcal{E}\{\sup_n b(n)\Sigma_{i=r+1}^{n} |V_i^{(h,0)}|\} < \infty$, for h = 2,3,$\cdots$,r. Thus, from the Lebesgue dominated convergence theorem

$$(5.40)\qquad \lim_{t\to\infty} \mathcal{E}\{b(M)\Sigma_{i=r+1}^{M} V_i^{(h,0)}\} = 0$$

for h = 2,3,$\cdots$,r. From (5.38), (5.39) and (5.40) we find that the first term of (5.37) equals zero. In a similar fashion it is possible to show that the second term of (5.37), and hence, the first term of (5.36), equals zero.

We now examine the second term of (5.36). Again, Theorem 4.3 can be adapted to show that

$$(5.41)\qquad \lim_{n\to\infty}(n-1)^{-1}\Sigma_{i=r+1}^{n} i^{-1} V_i^{(h,c)} = 0 \qquad \text{(a.s.)}$$

and a proof similar to that of Lemma 5.6 demonstrates that

$$(5.42)\qquad \mathcal{E}\{\sup_n (n-1)^{-1}\Sigma_{i=r+1}^{n} i^{-1} |V_i^{(h,c)}|\} < \infty$$

for $h = 2,3,\cdots,r$ and $c = 0,1,\cdots,h-1$. Then (5.41), (5.42) and the Lebesgue dominated convergence theorem combine to show that the second term of (5.36) equals zero. This completes the proof of (5.32).

The proof of (5.33) is similar to the proof of (5.32) and is therefore omitted. We have thus established (5.27), and hence, (5.17).

(c) Proof of (5.18), (5.19) and (5.20). From the proof of Theorem 3.1 recall that

$$(5.43)\qquad U_n^2 = \binom{n}{r}^{-1}\Sigma_{c=o}^{r}\binom{r}{c}\binom{n-r}{r-c} U_n^{(c)}$$

with $U_n^{(c)}$ given by (2.3). Notice that $\binom{n}{r}^{-1}\binom{r}{c}\binom{n-r}{r-c} = 0(n^{-c})$ for $c = 0,1,\cdots,r$. For $c = 0$, by Lemma 5.7, $\mathcal{E}\{U_M^{(0)}\} = 0$. For $c = 1,2,\cdots,r$, by Lemma 5.6,

$$(5.44)\qquad \mathcal{E}\{\sup_n \binom{n}{r}^{-1}\binom{r}{c}\binom{n-r}{r-c} |U_n^{(c)}|\} < \infty.$$

Also, for $c = 1,2,\cdots,r$

$$(5.45)\qquad \lim_{n\to\infty}\binom{n}{r}^{-1}\binom{r}{c}\binom{n-r}{r-c} U_n^{(c)} = 0 \qquad \text{(a.s.)}.$$

Thus, by (5.43), (5.44), (5.45) and the Lebesgue dominated convergence theorem, (5.18) holds. Both (5.19) and (5.20) can be easily proved using the Schwarz inequality. This completes the proof of (5.15) and the theorem.

5.4 Examples.

EXAMPLE 1. We continue our discussion of the example of section 3.6 in which $\Theta = \text{Var}\{X_1\}$. To be specific let $a_k = a$ for $k = 2,3,\cdots$ although any positive sequence $\{a_k\}$ such that $\lim_{k\to\infty} a_k = a$ would do since we are only investigating asymptotic behavior. From (5.1) define for $d > 0$

$$(5.46)\quad N(d) = \text{smallest integer } k \geq 2 \text{ such that } s_{wk}^2 \leq kd^2a^{-2}$$

where s_{wk}^2 is given by (3.59). Then $I_N = [U_N - d, U_N + d]$ is a sequential confidence interval for $\Theta = \mu_2$ having width equal to 2d and coverage probability approximately equal to α, for small values of d. The sequential procedure is asymptotically efficient in the sense that (5.2) holds. In this case s_{wn}^2 is not difficult to calculate sequentially as it depends only on the first four sample moments.

We could also define N(d) by (5.46) with s_{wk}^2 replaced by s_{zk}^2 where s_{zk}^2 is computed using (3.57), (3.61) and (3.60), in that order. However, for this example, the procedure using s_{wk}^2 is to be preferred.

Note, incidentally, that the sequential procedure is invariant under a location shift.

EXAMPLE 2. Suppose that $x_1 = (x_1^{(1)}, x_1^{(2)}), \cdots, x_n = (x_n^{(1)}, x_n^{(2)})$ is a bivariate random sample of a random variable $X = (X^{(1)}, X^{(2)})$ with continuous marginal distribution functions. Let

$$s(u) = \begin{cases} -1 & u < 0 \\ 0 & u = 0 \\ +1 & u > 0 \end{cases}$$

and $f(x_1,x_2) = s(x_1^{(1)}-x_2^{(1)})s(x_1^{(2)}-x_2^{(2)})$. The corresponding U-statistic is

$$(5.47) \qquad U_n = n^{-1}(n-1)^{-1}\Sigma_{\alpha_1\neq\alpha_2} s(x_{\alpha_1}^{(1)}-x_{\alpha_2}^{(1)})s(x_{\alpha_1}^{(2)}-x_{\alpha_2}^{(2)})$$

and is referred to as the difference sign covariance of the sample. See Hoeffding [7]. Two points x_1 and x_2 are said to be concordant if $s(x_1^{(1)}-x_2^{(1)})s(x_1^{(2)}-x_2^{(2)}) = +1$ and are discordant if $s(x_1^{(1)}-x_2^{(1)})s(x_1^{(2)}-x_2^{(2)}) = -1$. Let

$$(5.48) \quad \pi = P\{X_1 \text{ and } X_2 \text{ are concordant}\} = P\{(X_1^{(1)}-X_2^{(1)})(X_1^{(2)}-X_2^{(2)}) > 0\}.$$

Then the expectation of the U-statistic is $\theta = \mathcal{E}\{s(X_1^{(1)}-X_2^{(1)})s(X_1^{(2)}-X_2^{(2)})\} = 2\pi - 1$. Now, let C_n equal the number of concordant pairs among $x_1,x_2,\cdots,x_n$. Then (5.47) becomes

$$(5.49) \qquad U_n = 4n^{-1}(n-1)^{-1}C_n - 1.$$

Next, for $i = 1,2,\cdots,n$, let T_{in} equal the number of points among $\{x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n\}$ concordant with x_i. Then $\Sigma_{i=1}^n T_{in} = 2C_n$ and $C_n = \Sigma_{i=2}^n T_{ii}$, so that $C_{n+1} = C_n + T_{n+1,n+1}$.

To determine s_{wn}^2 notice that $W_{in} = 4(n-1)^{-1}T_{in} - 2$ and so

$$W_{in} - \overline{W}_n = 4(n-1)^{-1}T_{in} - 8n^{-1}(n-1)^{-1}C_n$$

for $i = 1,2,\cdots,n$. Thus, after some arrangement

$$(5.50) \qquad s_{wn}^2 = 16(n-1)^{-3}[\Sigma_{i=1}^n T_{in}^2 - 4n^{-1}C_n^2].$$

Define

$$a_i(n+1) = \begin{cases} 1 & \text{if } x_i \text{ and } x_{n+1} \text{ are concordant} \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1,2,\cdots,n$. Then $T_{i,n+1} = T_{in} + a_i(n+1)$ for $i = 1,2,\cdots,n$. Notice that the T_{in}'s may be arranged in a triangle as follows:

$$\begin{array}{llll} T_{12} & T_{22} & & \\ T_{13} & T_{23} & T_{33} & \\ \vdots & & & \\ T_{1n} & T_{2n} & \cdots & T_{nn} \end{array}$$

Suppose that the observations $x_1,x_2,\cdots,x_n$ have been taken and that $C_n,T_{1n},\cdots,T_{nn}$ have been determined numerically. Now, if a further observation x_{n+1} is taken, then $T_{n+1,n+1}$ can be determined either by plotting x_{n+1} and comparing with $x_1,x_2,\cdots,x_n$, or otherwise. Then compute C_{n+1} from $C_{n+1} = C_n + T_{n+1,n+1}$. The $T_{i,n+1}$'s are determined, for $i = 1,2,\cdots,n$, from the last row of the above triangle, and finally, $s^2_{w,n+1}$ is given by (5.50).

The estimate s^2_{zn} is given by

$$(5.51) \qquad s^2_{zn} = (n-1)^{-1}[\Sigma^n_{i=3} Z^2_i - nU^2_n + 2U^2_2]$$

where, for $n > 2$,

$$(5.52) \qquad Z_n = nU_n - (n-1)U_{n-1}.$$

Suppose that C_n, U_n and $\Sigma^n_{i=3}Z^2_i$ are known numerically and a further observation x_{n+1} is taken. Determine $T_{n+1,n+1}$ and C_{n+1} as before.

Compute U_{n+1} from (5.49) and Z_{n+1} from (5.52). Then, finally $s^2_{z,n+1}$ can be computed from (5.51).

Define N(d) using either s^2_{wn} or s^2_{zn} as an estimate of σ^2. Then $I_N = [U_N-d, U_N+d]$ is a sequential confidence interval for $\theta = 2\pi - 1$ having fixed-width equal to 2d and coverage probability approximately equal to α, for small values of d.

BIBLIOGRAPHY

[1] Anscombe, F. J. (1952). Large-sample theory of sequential estimation. J. Roy. Statist. Soc. Ser. B. **48** 600-607.

[2] Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. Ann. Math. Statist. **37** 51-58.

[3] 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.

[4] Doob, J. L. (1953). Stochastic Processes. Wiley, New York.

[5] Feller, W. (1957). An Introduction to Probability Theory and its Applications, **1**. Second edition. Wiley, New York.

[6] Fraser, D. A. S. (1957). Nonparametric Methods in Statistics. Wiley, New York.

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

[8] Hoeffding, W. (1960). An upper bound for the variance of Kendall's tau and of related statistics. In Contributions to Probability and Statistics, edited by I. Olkin and others. Stanford University Press, Stanford.

[9] Hoeffding, W. (1961). The strong law of large numbers for U-statistics. Institute of Statistics Mimeo Series No. 302, University of North Carolina, Chapel Hill.

[10] Kendall, M. G. and Stuart, A. (1958). The Advanced Theory of Statistics, **1**. Hafner, New York.

[11] Loeve, M. (1960). Probability Theory. Second edition. Van Nostrand, New York.

[12] Pratt, J. W. (1960). On interchanging limits and integrals. Ann. Math. Statist. **31** 74-77.

[13] Sen, P. K. (1960). On some convergence properties of U-statistics. Calcutta Statist. Assn. Bull. **10** 1-18.

[14] Siegmund, D. (1969). On moments of the maximum of normed partial sums. Ann. Math. Statist. **40** 527-531.

[15] Simons, G. (1968). On the cost of not knowing the variance when making a fixed-width confidence interval for the mean. Ann. Math. Statist. **39** 1946-1952.

[16] Starr, N. (1966). The performance of a sequential procedure for the fixed-width interval estimation of the mean. Ann. Math. Statist. **37** 36-50.

[17] Wilks, S. S. (1962). Mathematical Statistics. Wiley, New York.