On the Serfling-Wackerly Alternate Approach

by

Raymond J. Carroll*

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina 27514

Institute of Statistics Mimeo Series No. 1052

February, 1976

---

* This research was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2796.

On the Serfling-Wackerly Alternate Approach

by

Raymond J. Carroll*

Department of Statistics
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27514

ABSTRACT

For the problem of fixed-length interval estimation of the mean and median, the limiting distributions of the stopping rules recently introduced by Serfling and Wackerly [8] are compared to those of Chow and Robbins [4] as both the non-coverage probability α and interval length d converge to zero. The procedures based on the sample mean differ in the centering constants, while those based on the sample median also differ in the scaling constants.

Institute of Statistics Mimeo Series #1052.

*This research was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2796.

1. Introduction

The primary purpose of this note is to illustrate the second order differences (where they exist) between two competing approaches to the problem of fixed length confidence interval estimation of the center of symmetry: one due to Chow and Robbins (C-R) [4], the other recently introduced by Serfling and Wackerly (S-W) [8]. The intervals of prescribed coverage probability $1 - 2\alpha$ are of the form $(\overline{X}_n-d, \overline{X}_n+d)$ but the stopping rules differ; C-R fix α and let $d \to 0$, while S-W fix d and let $\alpha \to 0$. S-W show that the C-R approach with stopping time $N(\alpha,d)$ is "inappropriate" for their problem in the sense that if $m(\alpha,d)$ is the required number of observations with F known, $N(\alpha,d)/m(\alpha,d) \to c(d) \neq 1$ as $\alpha \to 0$, although $c(d) = 1$ if $F = \Phi$, the normal distribution function. They also show that $c(d) \to 1$ as $d \to 0$, which suggests that the two approaches are equivalent as both $\alpha,d \to 0$. If the S-W stopping rule is denoted by $M(\alpha,d)$,

$$N(\alpha,d)/M(\alpha,d) \to 1 \quad \text{(a.s.) as } \alpha \to 0, \text{ then } d \to 0. \tag{1}$$

The consistency property (1) says that the rules approximate the same quantity asymptotically, but they still may be different in terms of a limiting distribution. The asymptotic distribution of $N(\alpha,d)$ is easily found (see [6]), and in Section 3 and the Appendix the asymptotic distribution of $M(\alpha,d)$ is obtained, a result of some interest in itself. Contrary to (1), this paper finds that

(2) (a) $N(\alpha,d)$ and $M(\alpha,d)$ have limiting distributions with the same scaling constants as $\alpha,d \to 0$; for large classes of α,d the C-R(S-W) approach has smaller centering constants if the population kurtosis exceeds (is less than) 3.

(b) In the normal case with d fixed, the S-W approach is preferable as $\alpha \to 0$ if $d/\sigma < 2.33$.

The asymptotic distributions of analogous rules based on the sample median ([5], [8]) are also obtained. The differences are more pronounced here; both the scaling and centering constants are different. As an example of the results, Serfling and Wackerly's modification of Geertsema's rule [5] is found to be infinitely more variable than their own as $\alpha \to 0$.

2. The Two Methods

Let $X_1, X_2, \ldots$ be i.i.d. observations from a continuous symmetric distribution $F(x-\theta)$ with unique median θ and with variance σ^2. Let $\Phi(\lambda_\alpha) = 1-\alpha$. The two approaches are:

Chow and Robbins (C-R) [4]. Let $s_n^2 = n^{-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$ and define

$$N(\alpha,d) = \inf\{n: n \ge (\lambda_\alpha s_n/d)^2\}$$
$$n(\alpha,d) = \inf\{n: n \ge (\lambda_\alpha \sigma/d)^2\}.$$

Serfling and Wackerly (S-W) [8]. Let $H_n(x) = n^{-1} \sum_{i=1}^{n} I\{X_i - \overline{X}_n \le x\}$, where $I(A)$ denotes the indicator of the event A. Further, define $m_n(-d) = \inf_{t \le 0} \exp(dt) \int_{-\infty}^{\infty} \exp(tx)\,dH_n(x)$. Then

$$M(\alpha,d) = \inf\{n: n \ge \ln \alpha/\ln m_n(-d) \text{ and } X_i - \overline{X}_n + d < 0 \text{ for some } i \le n\}$$
$$m(\alpha,d) = \inf\left\{n: \Pr\{|\overline{X}_n - \theta| \ge d\} \le \alpha\right\}.$$

As further notation, let $M_0(t) = \int \exp(tx)\,dF(x)$ and let $t_0(d)$ be the point at which $W_0(t) = \exp(dt)M_0(t)$ achieves its infimum. Then, since F has a unique median at zero,

$$\frac{d}{dt} W_o(t)|_{t_o(d)} = 0.$$

Define $W_n(t) = n^{-1} \sum_{i=1}^{n} \exp\left(t(X_i-\bar{X}_n+d)\right)$, and let $t_n(d)$ be the solution to $W_n'(t) = 0$, so that for n large, $m_n(-d) = W_n\left(t_n(d)\right)$. Finally, let

$$m_o(z) = \inf_{t<0} \exp(-zt) \int \exp(tx)dF(x).$$

3. Sample Means

The asymptotic distributions of the rules $N(\alpha,d)$ and $M(\alpha,d)$ are obtained here. The scaling constants are the same but the centering constants are not. For many sequences $\alpha,d \to 0$, the value of the population kurtosis determines the smaller centering constants. In this section, assume $\int xdF(x) = 0$, $\int x^2dF(x) = 1$. Let $\ell n\ x$ and $\log x$ both denote the natural logarithm. Let $A(\alpha,d)$ be the variance of $\psi(X)$, where

$$\psi(y) = \exp\{t_o(d)(y+d)\} - yt_o(d)E\exp\{t_o(d)(X+d)\}.$$

The proof of the following Lemma is given in the appendix. The basic idea is to represent $\log m_n(-d)/m_o(-d)$ as a sum of i.i.d. random variables with remainder term and use the techniques of [6]. Note the result includes the case d fixed.

Lemma 1. Let $\Delta = -\ell n\ m_o(-d)$. Then, as $\alpha \to 0$,

$$m_o(-d)\left[\Delta^3/\left(-A(\alpha,d)\ell n\alpha\right)\right]^{\frac{1}{2}}\left(M(\alpha,d) - (-\ell n\alpha/\Delta)\right) \xrightarrow{L} \Phi.$$

If, for some $\varepsilon > 0$, as $\alpha,d \to 0$

(4a) $$d^{-2}(-\ell n\alpha)^{-1+\varepsilon} \to 0, \quad d^{12}(-\ell n\alpha) \to 0,$$

then

$$\text{(4b)} \qquad \left[d^2/\left(\lambda_\alpha^2(EX^4-1)\right)\right]^{\frac{1}{2}}\left(M(\alpha,d) - (-\ln\alpha/\Delta)\right) \xrightarrow{L} \Phi.$$

Lemma 2. ([6]). As $\alpha,d \to 0$ or as $\alpha \to 0$,

$$\text{(5)} \qquad \left(\frac{d^2}{\lambda_\alpha^2(EX^4-1)}\right)^{\frac{1}{2}} \left(N(\alpha,d) - \lambda_\alpha^2/d^2\right) \xrightarrow{L} \Phi.$$

Now, $\Delta = d^2/2 + d^4(1 - EX^4/3)/8 + o(d^4)$ as $d \to 0$ (see Proposition 1), so that the centering constants $-\ln\alpha/\Delta$ in (4) may be replaced by λ_α^2/d^2 (and the conclusion that the asymptotic distributions are equivalent can be made) if, as $\alpha,d \to 0$,

$$\text{(6)} \qquad H(\alpha,d) = \{\left(1 + d^2(1-EX^4/3)/4 + o(d^2)\right)\lambda_\alpha^2 - (-2\ln\alpha)\}/d\lambda_\alpha \to 0.$$

Note that if $d\lambda_\alpha \sim \lambda_\alpha^\varepsilon$ for some $\varepsilon > 0$, then since $\lambda_\alpha^2/(-2\ln\alpha) \to 1$, (4a) holds but because of Proposition 5,

$$\begin{aligned} H(\alpha,d) &\to -\infty \quad \text{if } EX^4 > 3 \\ &\to 0 \quad \text{if } EX^4 = 3 \\ &\to \infty \quad \text{if } EX^4 < 3, \end{aligned}$$

the result being that the C-R approach has smaller centering constants (and is thus in a sense "preferable" but "inappropriate") if $EX^4>3$, giving (2a).

Lemma 3. Suppose $X_1,X_2,\ldots$ are normally distributed with variance σ^2 and let $\eta = d/\sigma$ be fixed. As $\alpha \to 0$,

$$\text{(7a)} \qquad \lambda_\alpha^{-1}\left(M(\alpha,d) - \lambda_\alpha^2/\eta^2\right) \xrightarrow{L} N(0,\sigma_{SW}^2)$$

$$\text{(7b)} \qquad \lambda_\alpha^{-1}\left(N(\alpha,d) - \lambda_\alpha^2/\eta^2\right) \xrightarrow{L} N(0,\sigma_{CR}^2),$$

where

$$\sigma_{CR}^2/\sigma_{SW}^2 = (\eta^4/2)e^{-\eta^2/2}(1 - e^{-\eta^2} - \eta^2 e^{-\eta^2}).$$

Note that $\sigma_{CR}^2/\sigma_{SW}^2 \to 1$ as $\eta \to 0$ as is predicted by Lemmas 1 and 2, but that $\sigma_{CR}^2/\sigma_{SW}^2 \to 0$ as $\eta \to \infty$. In Table 1, the values of this ratio are presented for various values of η. It appears that for $0 < \eta < 2.33$, $\sigma_{SW}^2 < \sigma_{CR}^2$, while the opposite is true if $2.34 < \eta < \infty$.

TABLE 1

Values of the quantity $\sigma_{CR}^2/\sigma_{SW}^2$.

η	$\sigma_{CR}^2/\sigma_{SW}^2$
0.00	1.00
0.05	1.00
0.25	1.01
0.50	1.04
1.00	1.15
1.50	1.25
2.00	1.19
2.25	1.06
2.33	1.00
2.50	.87
3.00	.45
3.50	.16

4. Sample Median.

Serfling and Wackerly [8] also provide a rule $M(\alpha,d)$ based on the median. Geertsema [5] has defined an analogue $N^*(\alpha,d)$ to the C-R approach, while S-W provide a modification of this, call it $N(\alpha,d)$. The results of this section are not as satisfactory technically as those of the previous one, but they do shed insight into the behavior of the three rules.

Letting $X_{n,i}$ denote the ith order statistic in a sample of size n, define

$$b_n = \max\{1, [n/2 - cn^{\frac{1}{2}}/2]\}, \quad a_n = n-b_n+1$$
$$Z_n(c) = n^{\frac{1}{2}}(X_{n,a_n} - X_{n,b_n})/c$$
$$N(\alpha,d) = \text{first time } n \text{ that } Z_n^2(c)/n \le 4d^2/\lambda_\alpha^2 .$$
$$N^*(\alpha,d) = \text{first time } n \text{ that } Z_n^2(\lambda_\alpha)/n \le 4d^2/\lambda_\alpha^2 .$$

Define $M(\alpha,d)$ as in [8]. If T_n is the sample median, set

$$\Delta = -\tfrac{1}{2} \log\left[4\left(F(d) - F^2(d)\right)\right]$$
$$\hat{\Delta}_n = -\tfrac{1}{2} \log\left[4\left(F_n(T_n+d) - F_n^2(T_n+d)\right)\right].$$

<u>Lemma 4</u>. Assume that Z_n is uniformly continuous in probability (see [1]). For the S-W adaptation of Geertsema's rule, as $\alpha,d \to 0$,

(8a) $$\left[2c\left(df(0)/\lambda_\alpha\right)^3\right]^{\frac{1}{2}}\left[N(\alpha,d) - \left(\lambda_\alpha/2df(0)\right)^2\right] \xrightarrow{L} \Phi.$$

For Geertsema's rule as $d \to 0$,

(8b) $$\left[2\left(df(0)\right)^3/\lambda_\alpha^2\right]^{\frac{1}{2}}\left[N^*(\alpha,d) - \left(\lambda_\alpha/2df(0)\right)^2\right] \xrightarrow{L} \Phi.$$

<u>Lemma 5</u>. Define

$$B(F,d) = \left\{1-2F(d)\right\}^2\left[2F(d)\left(1-F(d)\right)\right]^{-2}\left\{F(d)\left(1-F(d)\right)-f(d)\left(1-F(d)\right)/f(0) + \left(f(d)/2f(0)\right)^2\right\}.$$

Then, as $\alpha \to 0$,

(9) $$(2\Delta^3)^{\frac{1}{2}}\lambda_\alpha^{-1}\left(M(\alpha,d) - (-\ell n\alpha/\Delta)\right)/B(F,d)^{\frac{1}{2}} \xrightarrow{L} \Phi.$$

Lemma 6. If σ^2_{CR} denotes the asymptotic variance in (3b) and σ^2_{SW} denotes the asymptotic variance in (9), then as $\alpha, d \to 0$, $\sigma^2_{CR}/\sigma^2_{SW} \to \frac{1}{2}$.

A few points need to be emphasized. First, we do not know if Z_n is uniformly continuous in probability. Second, (8b) holds only as $d \to 0$, while (9) holds only as $\alpha \to 0$. The difficulty with (8b) is that, as $\alpha \to 0$, the necessary appeal to the representation theorem [2] cannot be justified. The difficulty with (9) is that, as $d \to 0$, $n \to \infty$, the asymptotic distribution of the process

$$\left(1 - 2F(d)\right)\left(Y_n(t+d) - f(d)Y_n(0)/f(0)\right)$$

$\left(Y_n(t) = n^{\frac{1}{2}}\left(F_n(t) - F(t)\right)\right)$ is unknown. These are interesting technical problems and more work is needed.

In summary, Lemmas 4 - 6 make it likely that $N(\alpha,d)$ is infinitely more variable and $N^*(\alpha,d)$ is less variable than $M(\alpha,d)$, although more technical work is needed to confirm these impressions. Note again the difference in the centering constants.

It is worthwhile to mention that these results extend readily to the ranking problem ([9]).

ERRATA SHEET

page 8

a) The statement of Proposition 1 should be

Proposition 1. Uniformly as $d \to 0$, $n \to \infty$,

$$t_n(d) - t_o(d) = O\left(n^{-\frac{1}{2}}(\log_2 n)^{\frac{1}{2}}\right) \quad \text{(a.s.)}.$$

page 9

b) The three lines immediately above Proposition 2 should be deleted.

APPENDIX

The proof of Lemma 1 is long and tedious, but the basic idea is as stated in Section 3. We first begin with a number of propositions. Let O, o be the customary "big oh, little oh".

Proposition 1. Uniformly as $d \to 0$, $n \to \infty$,

$$t_n(d) - t_o(d) = n^{-1} \sum_{i=1}^{n} (X_i+d)\exp\left(t_o(d)X_i\right) - \bar{X}_n + O(n^{-1} \log_2 n) \quad \text{(a.s.)}.$$

Proof: It is a simple calculation to show that $t_n(d) \to 0$ (a.s.) as $d \to 0$, $n \to \infty$. As in Lemma 3.2 of [8], we see that for n sufficiently large

$$(10) \qquad 0 = \frac{d}{dt} W_n(t) = A_n^* + \left(t_n(d) - t_o(d)\right)B_n^* \text{, where}$$

$$A_n^* = n^{-1} \sum_{i=1}^{n} (X_i-\bar{X}_n+d)\exp\{t_o(d)(X_i-\bar{X}_n+d)\}$$

$$B_n^* = n^{-1} \sum_{i=1}^{n} (X_i-\bar{X}_n+d)^2\exp\{\gamma_n(d)(X_i-\bar{X}_n+d)\}.$$

Now, $A_n^* = A_n^{(1)} + A_n^{(2)} + A_n^{(3)}$, where

$$\begin{aligned} A_n^{(1)} &= n^{-1} \sum_{i=1}^{n} X_i \exp\{t_o(d)(X_i-\bar{X}_n+d)\} \\ &= e^{t_o(d)(d-\bar{X}_n)} n^{-1} \sum_{i=1}^{n} \{X_i \exp\left(t_o(d)X_i\right) - EX\exp\left(t_o(d)X\right)\} \\ &\qquad + EX\exp\left(t_o(d)(X+d)\right)\{\exp\left(-t_o(d)\bar{X}_n\right)-1\} + EX\exp\left(t_o(d)(X+d)\right). \end{aligned}$$

$$\begin{aligned} A_n^{(2)} &= -n^{-1} \bar{X}_n \sum_{i=1}^{n} \exp\{t_o(d)(X_i-\bar{X}_n+d)\} \\ &= -\bar{X}_n e^{t_o(d)(d-\bar{X}_n)} n^{-1} \sum_{i=1}^{n} \{\exp\left(t_o(d)X_i\right) - E\exp\left(t_o(d)X\right)\} \\ &\qquad - \bar{X}_n E\exp\{t_o(d)(X+d)\}\{\exp\left(-t_o(d)\bar{X}_n\right)-1\} - \bar{X}_n E\exp\{t_o(d)(X+d)\}. \end{aligned}$$

$$A_n^{(3)} = dn^{-1} \sum_{i=1}^{n} \exp\{t_o(d)(X_i - \overline{X}_n + d)\}$$

$$= de^{t_o(d)(d-\overline{X}_n)} n^{-1} \sum_{i=1}^{n} \{\exp\big(t_o(d)X_i\big) - E\exp\big(t_o(d)X\big)\}$$

$$+ dE\exp\{t_o(d)(X+d)\}\Big[\exp\big(-t_o(d)\overline{X}_n\big)-1\Big] + dE\exp\{t_o(d)(X+d)\}.$$

Since $EX\exp\big(t_o(d)(X+d)\big) + dE\exp\big(t_o(d)(X+d)\big) = 0$ and

$$-\overline{X}_n E\exp\{t_o(d)(X+d-\overline{X}_n)\} = -\overline{X}_n + O\big(dn^{-\frac{1}{2}}(\log_2 n)^{\frac{1}{2}}\big) \quad \text{(a.s.)},$$

this means that

$$A_n^* = n^{-1} \sum_{i=1}^{n} (X_i+d)\exp\big(t_o(d)X_i\big) - \overline{X}_n + O\big(dn^{-\frac{1}{2}}(\log_2 n)^{\frac{1}{2}}\big) \quad \text{(a.s.)},$$

and hence that

$$t_n(d) - t_o(d) = O\big(n^{-\frac{1}{2}}(\log_2 n)^{\frac{1}{2}}\big) \quad \text{(a.s.)}.$$

Now, by carrying out the Taylor expansion (10) one more place as in Lemma 3.2 of [8] and going through the above steps once more, one completes the proof.

<u>Proposition 2.</u> As $d \to 0$,

$$t_o(d) = -d - d^3(\tfrac{1}{2} - EX^4/6) + o(d^3).$$

$$\Delta = -\ell n\, m_o(-d) = d^2/2 + d^4(1 - EX^4/3)/8 + o(d^4).$$

<u>Proof:</u> The key to the expansion of $t_o(d)$ lies in the relation

$$0 = E(X+d)\exp\big(t_o(d)(X+d)\big).$$

By using the facts $EX = EX^3 = 0$, $EX^2 = 1$, one will obtain $t_o(d)$ by progressively adding terms to the Taylor expansion of $\exp\big(t_o(d)(X+d)\big)$. Now, $m_o(-d) = M_o\big(t_o(d)\big)\exp\big(dt_o(d)\big)$, where

$$M_o\big(t_o(d)\big) = 1 + t_o^2(d)EX^2/2 + t_o^4(d)EX^4/24 + o(d^4).$$

Thus,

$$\Delta = -\ln m_o(-d) = -t_o(d)d - \big(t_o^2(d)/2 + t_o^4(d)EX^4/24 - t_o^4(d)/8 + o(d^4)\big),$$

so that using the expansion for $t_o(d)$ yields the result.

Proposition 3. Uniformly as $d \to 0$, $n \to \infty$,

$$m_n(-d) - m_o(-d) = n^{-1}\sum_{i=1}^{n}\{\psi(X_i) - E\psi(X)\} + O(n^{-1}\log_2 n) \quad \text{(a.s.)}$$

and

$$\log\frac{m_n(-d)}{m_o(-d)} = \big(d^2/m_o(-d)\big)n^{-1}\sum_{i=1}^{n}\{(X_i^2-1)\big(1+d^2(1-EX^4/3)\big)/4 + d(X_i+X_i^3/3)/2$$
$$+ d^2(X_i^4-EX^4)/24\} + o(d^4) + O(n^{-1}\log_2 n) \quad \text{(a.s.)}.$$

As $\alpha \to 0$,

$$\log\big(m_n(-d)/m_o(-d)\big) = n^{-1}\sum_{i=1}^{n}\{\psi(X_i)-E\psi(X)\}/m_o(-d) + O(n^{-1}\log_2 n) \quad \text{(a.s.)}.$$

Proof: By a Taylor expansion,

$$\exp\big(t_n(d)X_i\big) = \exp\big(t_o(d)X_i\big) + \big(t_n(d)-t_o(d)\big)X_i\exp\big(t_o(d)X_i\big)$$
$$+ \big(t_n(d)-t_o(d)\big)^2X_i^2\exp\big(Z_i(d)\big)/2,$$

so that, since $m_n(-d) = W_n\big(t_n(d)\big)$,

$$(11)\quad \exp\{-t_n(d)(d-\bar{X}_n)\}m_n(-d)$$

$$= \Big\{n^{-1}\sum_{i=1}^{n}\{\exp\big(t_o(d)X_i\big) - E\exp\big(t_o(d)X\big)\} + E\exp\big(t_o(d)X\big)$$

$$+ \big(t_n(d)-t_o(d)\big)n^{-1}\sum_{i=1}^{n}\{X_i\exp\big(t_o(d)X_i\big) - EX\exp\big(t_o(d)X\big)\}$$

$$+ \big(t_n(d)-t_o(d)\ EX\exp\big(t_o(d)X\big) + O\big(n^{-1}(\log_2 n)\big)\Big\}$$

$$= n^{-1}\sum_{i=1}^{n}\{\exp\big(t_o(d)X_i\big) - E\exp\big(t_o(d)X\big)\} + E\exp\big(t_o(d)X\big)$$

$$- d\big(t_n(d)-t_o(d)\big)E\exp\big(t_o(d)X\big) + O(n^{-1}\log_2 n)\quad \text{(a.s.)}.$$

Now,

$$\exp\{t_n(d)(d-\bar{X}_n)\}$$

$$= \exp\{dt_o(d)\} + \{d\big(t_n(d)-t_o(d)\big)-t_n(d)\bar{X}_n\}\exp\big(dt_o(d)\big) + O(n^{-1}\log_2 n)\quad \text{(a.s.)}.$$

Multiplying the two sides of (11) by this last expression yields the expansion for $m_n(-d)$. Thus,

$$\big(m_n(-d)-m_o(-d)\big)\exp\{-dt_o(d)\}$$

$$= n^{-1}\sum_{i=1}^{n}\{\exp\big(t_o(d)X_i\big) - E\exp\big(t_o(d)X\big) - t_o(d)X_i\,E\exp\big(t_o(d)X\big)\}$$

$$+ O(n^{-1}\log_2 n).$$

We have

$$\exp\big(t_o(d)X_i\big) = \sum_{K=0}^{4}\big(t_o(d)X_i\big)^K/K! + o(d^4)$$

$$M_o(t_o) = E\exp\big(t_o(d)X\big) = 1 + d^2/2 + d^4(\tfrac{1}{2} - EX^4/8) + o(d^4)$$

$$t_o(d)M_o(t_o) = -d - d^3(1 - EX^4/6) + o(d^4).$$

This leads to

$$\big(m_n(-d)-m_o(-d)\big)\exp\{-dt_o(d)\}$$

$$= d^2n^{-1}\sum_{i=1}^{n}\{(X_i^2-1)\big(1+(1-EX^4/3)d^2\big)/4 + d(X_i+X_i^3/3)/2 + d^2(X_i^4-EX^4)/24\}$$

$$+ O(n^{-1}\log_2 n) + o(d^4)$$

and completes the proof.

Proposition 4. If $A(\alpha,d) = \int \big(\psi(x) - \int \psi(y)dF(y)\big)^2 dF(x)$, then

$$A(\alpha,d) = d^4(EX^4-1)/4 + o(d^4) \quad \text{as} \quad d \to 0.$$

Proof: $\psi(X) = \exp\big(t_o(d)(X+d)\big) - X\,M_o\big(t_o(d)\big)t_o(d)\exp\big(dt_o(d)\big)$, so that

$$A(\alpha,d) = \mathrm{Var}\big(\psi(X)\big)$$

$$= \exp\big(2t_o(d)d\big)\Big[M_o\big(2t_o(d)\big) - M_o^2\big(t_o(d)\big) + t_o^2(d)M_o^2\big(t_o(d)\big)$$

$$+ 2dt_o(d)M_o^2\big(t_o(d)\big)\Big].$$

Now, simple Taylor expansions show that

$$M_o\big(t_o(d)\big) = 1 + d^2/2 + d^4(1-EX^4/4)/2 + o(d^4)$$

$$M_o^2\big(t_o(d)\big) = 1 + d^2 + d^4(5-EX^4)/4 + o(d^4)$$

$$M_o\big(2t_o(d)\big) = 1 + 2d^2 + 2d^4 + o(d^4),$$

which, with a few computations, yields the proposition.

Proposition 5. For all $\varepsilon > 0$,

$$\big(\lambda_\alpha^2 - (-2\ell n\alpha)\big)/\lambda_\alpha^\varepsilon \to 0 \quad \text{as} \quad \alpha \to 0.$$

Proof: It is well known that $\lambda_\alpha^2/(-2\ell n\alpha) \to 1$ as $\alpha \to 0$. Since $\lambda_\alpha = -\Phi^{-1}(\alpha)$, by L'Hospital's rule applied to $\lambda_\alpha^2/(-2\ell n\alpha)$,

$$H(\alpha) = \lambda_\alpha \exp\{\ell n\alpha + \tfrac{1}{2}\lambda_\alpha^2\} \to (2\pi)^{-\frac{1}{2}} .$$

Thus, $\left(H(\alpha)\right)^{\lambda_\alpha^{-\varepsilon}} \to 1$ if $\varepsilon > 0$, so that

$$\exp\{(\ell n\alpha + \tfrac{1}{2}\lambda_\alpha^2)/\lambda_\alpha^\varepsilon\} \to 1.$$

Proof of Lemma 1. By the nature of the stopping rule, $M(\alpha,d)$ stops the first time n that

$$-\log\left(m_n(-d)/m_o(-d)\right) \geq -\ell n\alpha/n - \Delta.$$

It is possible, using techniques in [6], page 298, to neglect the excess terms so that as $\alpha \to 0$, from Proposition 3,

$$m_o(-d)\Delta\left(M(\alpha,d)A(\alpha,d)\right)^{-\frac{1}{2}}\left(M(\alpha,d) - (-\ell n\alpha/\Delta)\right) \xrightarrow{L} \Phi.$$

Since $M(\alpha,d)/(-\ell n\alpha/\Delta) \to 1$ (a.s.), this completes the first part of Lemma 1. Letting $M = M(\alpha,d)$, note that

$$\frac{M(\alpha,d)}{(-\ell n\alpha/\Delta)} \geq \left[1 - \Delta^{-1}\log\left(m_M(-d)/m_o(-d)\right)\right]^{-1}$$

$$\frac{M(\alpha,d)-1}{(-\ell n\alpha/\Delta)} \leq \left[1 - \Delta^{-1}\log\left(m_{M-1}(-d)/m_o(-d)\right)\right]^{-1} ,$$

so that as $\alpha,d \to 0$, $M(\alpha,d)(-\ell n\alpha/\Delta)^{-1} \to 1$ (a.s.) if

(12) $$\Delta^{-1}\log\left(m_M(-d)/m_o(-d)\right) \to 0 \quad \text{(a.s.)}.$$

Once (12) is shown, the proof will be completed by once again using the technique in [6] together with the expansion of $\log\big(m_n(-d)/m_o(-d)\big)$ as well as Proposition 4. Now $\Delta = O(d^2)$ and $\log\big(m_M(-d)/m_o(-d)\big) = O(M^{-1}\log_2 M)$, so that it suffices to show

(13) $$d^{-2}M^{-1}\log_2 M \to 0 \quad \text{(a.s.)}.$$

A sufficient condition for (13) to hold is that for some $\varepsilon > 0$

$$d^{-2}M^{-1+\varepsilon} \to 0 \quad \text{(a.s.)}.$$

Now, as $\alpha, d \to 0$, $M(\alpha,d) \geq -\ell n\alpha$, so that

$$d^{-2}M(\alpha,d)^{-1+\varepsilon} \leq d^{-2}(-\ell n\alpha)^{-1+\varepsilon} \to 0.$$

Proof of Lemma 3. By Lemma 1, since $\Delta = \eta^2/2$ and $(-2\ell n\alpha)/\lambda_\alpha^2 \to 1$, we have

$$\lambda_\alpha^{-1}\big(M(\alpha,d) - (-2\ell n\alpha)/\eta^2\big) \xrightarrow{L} N(0,\sigma_{SW}^2),$$

where $\sigma_{SW}^2 = 4A(\alpha,d)/\eta^6$. By Lemma 2,

$$\lambda_\alpha^{-1}\big(N(\alpha,d) - \lambda_\alpha^2/\eta^2\big) \xrightarrow{L} N(0,\sigma_{CR}^2),$$

where $\sigma_{CR}^2 = (EX^4-1)/d^2 = 2/\eta^2$. In the normal case $A(\alpha,d) = 1 - e^{-\eta^2} - \eta^2 e^{-\eta^2}$, completing the proof because, by Proposition 5, $(-2\ell n\alpha - \lambda_\alpha^2)/\lambda_\alpha \to 0$.

Proof of Lemma 4. In the notation of Block and Gastwirth [3] with $m \sim cn^{\frac{1}{2}}/2$, as $n \to \infty$ $Z_n^2 \to \xi^2$ (a.s.) by [2] and

$$(2m)^{\frac{1}{2}}(Z_n^2-\xi^2)/2\xi^2 \xrightarrow{L} \Phi,$$

where $\xi = 1/f(0)$. Now, $N(\alpha,d)/(\xi/2k_\alpha)^2 \to 1$ (a.s.), where $k_\alpha = d/\lambda_\alpha$. Neglecting excess terms, $N(\alpha,d) \approx Z^2_{N(\alpha,d)}/(2k_\alpha)^2$ so that

$$N(\alpha,d) - (\xi/2k_\alpha)^2 \approx (2k_\alpha)^{-2}(Z^2_{N(\alpha,d)} - \xi^2).$$

Since $m \sim cN(\alpha,d)^{\frac{1}{2}}/d \sim c\xi/4k_\alpha$, the uniform continuity in probability guarantees that as $\alpha, d \to 0$,

$$\left(2c(d/\xi\lambda_\alpha)^3\right)^{\frac{1}{2}}\left(N(\alpha,d) - (\xi/2k_\alpha)^2\right) \xrightarrow{L} \Phi.$$

Geertsema's rule basically replaces c by λ_α; by steps similar to the above, as $d \to 0$,

$$(2d^3/\xi^3\lambda_\alpha^2)^{\frac{1}{2}}\left(N^*(\alpha,d) - (\xi/2k_\alpha)^2\right) \xrightarrow{L} \Phi.$$

Proof of Lemma 5. By once again neglecting excess over the boundary,

$$M(\alpha,d) - (-\ln\alpha/\Delta) \approx (-\ln\alpha)(\Delta\hat{\Delta}_M)^{-1}(\Delta-\hat{\Delta}_M).$$

Then, by expansions of $\log(1+x)$ and $F(T_n+d) - F(d)$,

$$n^{\frac{1}{2}}(\Delta-\hat{\Delta}_n) = \tfrac{1}{2} n^{\frac{1}{2}} \log\left(1 + \frac{\{F_n(T_n+d)-F(d)\}\{1-F_n(T_n+d)-F(d)\}}{F(d)-F^2(d)}\right)$$

$$= \left(1-2F(d)\right)\{2F(d)\left(1-F(d)\right)\}^{-1} n^{\frac{1}{2}}\left(F_n(T_n+d)-F(T_n+d)+T_n f(d)\right) + o(1) \text{ (a.s.)}.$$

By the representation theorem [2], if $Y_n(t) = n^{\frac{1}{2}}\left(F_n(t)-F(t)\right)$, then

$$n^{\frac{1}{2}}(\Delta-\hat{\Delta}_n) = \left(1-2F(d)\right)\{2F(d)\left(1-F(d)\right)\}^{-1}\left(Y_n(T_n+d)-f(d)Y_n(0)/f(0)\right) + o(1) \quad \text{(a.s.)}.$$

Now, $M(\alpha,d)/(-\ln\alpha/\Delta) \to 1$ (a.s.); denoting $M = M(\alpha,d)$, since the empirical process Y_n is weakly convergent under random sample sizes [7] and

$T_M \to 0$ (a.s.), this yields

$$\left(M/B(F,d)\right)^{\frac{1}{2}}(\Delta-\hat{\Delta}_M) \longrightarrow \Phi,$$

so that

$$\left(2\Delta^3/B(F,d)\lambda_\alpha^2\right)^{\frac{1}{2}}\left(M(\alpha,d) - (-\ln\alpha/\Delta)\right) \xrightarrow{L} \Phi.$$

Proof of Lemma 6. The ratio of the variances is

$$\frac{\sigma_{CR}^2}{\sigma_{SW}^2} = \frac{\lambda_\alpha^2/\left[2d^3\left(f(0)\right)^3\right]}{\lambda_\alpha^2 B(F,d)/2\Delta^3}\ .$$

Since $F(d) = \frac{1}{2} + df(0) + d^2f'(0) + d^3f''(0) + o(d^3)$, we have

$$\Delta = 2d^2f^2(0) + o(d^2).$$

Also,

$$\left(1-2F(d)\right)^2 = 4d^2f^2(0) + o(d^2)$$

$$F(d)\left(1-F(d)\right) = \tfrac{1}{4} - d^2f^2(0) + o(d^2).$$

Two more Taylor expansions yield

$$f(d)\left(1-F(d)\right)/f(0) = \tfrac{1}{2} + d\left(\frac{f'(0)}{2f(0)} - f(0)\right) + d^2\left(\frac{f''(0)}{2f(0)} - 2f'(0)\right) + o(d^2)$$

$$\left(f(d)/2f(0)\right)^2 = \tfrac{1}{4}\left\{1 + 2d\left(f'(0)/f(0)\right) + d^2\left[\frac{2f''(0)}{f(0)} + \left(\frac{f'(0)}{f(0)}\right)^2\right] + o(d^2)\right\},$$

so that

$$B(F,d) = 16d^2f^2(0)\left(df(0)+o(d)\right) = 16\left(df(0)\right)^3 + o(d^3).$$

Hence, as $d \to 0$, $\sigma_{CR}^2/\sigma_{SW}^2 \to \frac{1}{2}$.

ACKNOWLEDGEMENT

The author wishes to thank Professor Wackerly for sending preprints of [8], [9].

REFERENCES

[1] Anscombe, F.J., Large sample theory of sequential estimation, *Proceedings of the Cambridge Philosophical Society, 48,* (October 1952), 600-7.

[2] Bahadur, R.R., A note on quantiles in large samples, *Annals of Mathematical Statistics, 37,* (June 1966), 577-80.

[3] Block, D.A. and Gastwirth, J.L., On a simple estimate of the reciprocal of the density function, *Annals of Mathematical Statistics, 39,* (June 1968), 1083-85.

[4] Chow, Y.S. and Robbins, H., On the asymptotic theory of fixed-width confidence intervals for the mean, *Annals of Mathematical Statistics, 36,* (April 1965), 457-62.

[5] Geertsema, J.C., Sequential confidence intervals based on rank tests, *Annals of Mathematical Statistics, 41,* (June 1970), 1016-26.

[6] Gut, A., On the moments and limit distributions of some first passage times, *Annals of Probability, 2,* (April 1974), 277-308.

[7] Pyke, R., The weak convergence of the empirical process with random sample size, *Proceedings of the Cambridge Philosophical Society, 64,* (January 1968), 155-60.

[8] Serfling, R.J. and Wackerly, D.D., Asymptotic theory of sequential fixed-width confidence intervals for location parameters, Florida State University Technical Report M316, September, 1974.

[9] Wackerly, D. D., An alternative sequential approach to the problem of selecting the best of k populations, University of Florida Technical Report No. 91 (February, 1975).