Ann. Inst. Statist. Math.
Vol. 50, No. 1, 119–145 (1998)

SEQUENTIAL FIXED-WIDTH CONFIDENCE INTERVAL FOR THE PRODUCT OF TWO MEANS

SHEN ZHENG*, T. N. SRIRAM** AND ANDREW F. SEILA***

Department of Statistics, Franklin College of Arts and Sciences, University of Georgia, 204 Statistics Building, Athens, GA 30602, U.S.A.

(Received October 17, 1996; revised March 5, 1997)

Abstract. For the product of two population means, the problem of constructing a fixed-width confidence interval with preassigned coverage probability is considered. It is shown that the optimal sample sizes which minimize the total sample size and at the same time guarantee a fixed-width confidence interval of desired coverage depend on the unknown parameters. In order to overcome this, a fully sequential procedure consisting of a sampling scheme and a stopping rule are proposed. It is then shown that the sequential confidence interval is asymptotically consistent and the stopping rule is asymptotically efficient, as the width goes to zero. Furthermore, a second order result for the difference between the expected stopping time and the (total) optimal fixed sample size is established. The theoretical results are supported by appropriate simulations.

Key words and phrases: Coverage probability, fully sequential procedure, sampling scheme, asymptotic consistency, asymptotic efficiency, uniform integrability.

1. Introduction

A sequential procedure for the fixed-width interval estimation of the mean of a single population was investigated in Chow and Robbins (1965) and Starr (1966). In a seminal paper, Robbins *et al.* (1967) considered an analogous procedure for estimating the difference of the means of two populations. Then, Srivastava (1970), Mukhopadhyay (1976), Ghosh and Mukhopadhyay (1980) and, more recently, Mukhopadhyay and Liberman (1989) (also see references therein) and Mukhopadhyay and Sriram (1992) have considered sequential estimation of linear

---

* Now at Equifax J. V. White Technology Center, Alpharetta, GA 30202, Athens, GA 30602, U.S.A.

** Research partially supported by NSF Grant DMS-95-02824.

*** Now at Management Science Department, Terry College of Business, University of Georgia, Athens, GA 30602, U.S.A.

combination of means of multiple populations. While there is substantial literature on the interval estimation of linear combination of means, there is very little known about the sequential fixed-width interval estimation of non-linear functions of means of multiple populations. One simple and interesting non-linear function is the product of means. The purpose of this paper is to construct a sequential fixed-width confidence interval for the product of means of two populations.

The literature on estimation of product of means using Bayesian and frequentist methods is somewhat extensive. See, for instance, Harris (1971), Berry (1977), Page (1985, 1987, 1990, 1995), Rekab (1989), Berger and Bernardo (1989), Sun and Ye (1995) and Hardwick and Stout (1992). The estimation of product of means arises, most obviously, in situations of determining area based on measurements of length and width. It also arises in other practical contexts, however. For instance, in gypsy moth studies, the hatching rate of larvae per unit area can be estimated as the product of the mean of egg masses per unit area times the mean number of larvae hatching per egg mass; see Southwood (1978). Also, in many environmental applications, such as exposure assessment and risk modelling, the estimation of product of means is desired. An example discussed in Sun and Ye (1995) and Yfantis and Flatman (1991) deals with the assessment of risk due to exposure to radiation or various pollutants. Here, it is assumed that the dose per unit time, the units of time per day, and the number of days during which an individual is exposed are three independent normal random variables. The total exposure is the product of the three means. Applications of product mean estimation also arise in the area of reliability, economics and quality control.

Recently, Noble (1992) considered the problem of point estimation of product of two means under a total sample size constraint. He proposed a two-stage procedure (different from a Stein's two-stage procedure) and studied its asymptotic properties. Zheng *et al.* (1996a) independently considered the problem of sequential point estimation of product of two means under a total budget constraint. They proposed a Stein's two-stage procedure and established its asymptotic properties. Furthermore, Zheng *et al.* (1996b) have also extended their sequential point estimation results to the product of two or more population means.

The problem considered here can be described as follows. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be two independent sequences of *i.i.d.* random variables with unknown mean and variance (μ_1, σ_1^2) and (μ_2, σ_2^2), respectively. We want to construct a confidence interval I of width $2d$ and with coverage probability $\approx 1-\beta$ for the parameter $\mu_1\mu_2 (\neq 0)$, where $0 < d < \infty$ and $0 < \beta < 1$ are preassigned constants.

To this end, we could proceed as follows. Take m observations on X and n observations on Y, and let

$$(1.1) \qquad \bar{X}_m = m^{-1}\sum_{i=1}^{m} X_i, \quad \bar{Y}_n = n^{-1}\sum_{i=1}^{n} Y_i$$

and

$$(1.2) \qquad S_{1,m}^2 = m^{-1}\sum_{i=1}^{m}(X_i - \bar{X}_m)^2, \quad S_{2,n}^2 = n^{-1}\sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2$$

be the respective sample means and variances. Then, by the multivariate central limit theorem

$$\begin{pmatrix} \sqrt{m}(\bar{X}_m - \mu_1) \\ \sqrt{m}(S^2_{1,m} - \sigma_1^2) \\ \sqrt{n}(\bar{Y}_n - \mu_2) \\ \sqrt{n}(S^2_{2,n} - \sigma_2^2) \end{pmatrix} \xrightarrow{\mathcal{D}} N(\underline{0}, \Sigma), \quad \text{as} \quad m, n \to \infty, \tag{1.3}$$

provided $EX^4 < \infty$, $EY^4 < \infty$, where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \mu_{3,X} & 0 & 0 \\ \mu_{3,X} & \mu_{4,X} - \sigma_1^4 & 0 & 0 \\ 0 & 0 & \sigma_2^2 & \mu_{3,Y} \\ 0 & 0 & \mu_{3,Y} & (\mu_{4,Y} - \sigma_2^4) \end{pmatrix} \tag{1.4}$$

with $\mu_{3,X} = E(X - \mu_1)^3$, $\mu_{3,Y} = E(Y - \mu_2)^3$, $\mu_{4,X} = E(X - \mu_1)^4$, $\mu_{4,Y} = E(Y - \mu_2)^4$.

If

$$I_d = [\bar{X}_m\bar{Y}_n - d, \bar{X}_m\bar{Y}_n + d] \tag{1.5}$$

is the interval of width $2d$ centered at X_mY_n, then

$$\begin{aligned} \mathrm{P}(\mu_1\mu_2 \in I_d) &= \mathrm{P}(|\bar{X}_m\bar{Y}_n - \mu_1\mu_2| \le d) \\ &= \mathrm{P}\left(\frac{|(\bar{X}_m - \mu_1)\mu_2 + (\bar{Y}_n - \mu_2)\mu_1 + b_{m,n}|}{\sqrt{\sigma_1^2\mu_2^2/m + \sigma_2^2\mu_1^2/n}} \le d_1 \right) \end{aligned} \tag{1.6}$$

where $b_{m,n} = (\bar{X}_m - \mu_1)(\bar{Y}_n - \mu_2)$ and $d_1 = d/\sqrt{\sigma_1^2\mu_2^2/m + \sigma_2^2\mu_1^2/n}$. Note that

$$\frac{|(\bar{X}_m - \mu_1)(\bar{Y}_n - \mu_2)|}{\sqrt{\sigma_1^2\mu_2^2/m + \sigma_2^2\mu_1^2/n}} \le \frac{\sqrt{m}|X_m - \mu_1|}{\sigma_1|\mu_2|}|\bar{Y}_n - \mu_2| \xrightarrow{\mathrm{P}} 0,$$

as $m, n \to \infty$. Assume for large m and n that $m/n = \sigma_1\mu_2/\sigma_2\mu_1$. Then,

$$\frac{(\bar{X}_m - \mu_1)\mu_2 + (\bar{Y}_n - \mu_2)\mu_1}{\sqrt{\sigma_1^2\mu_2^2/m + \sigma_2^2\mu_1^2/n}} \xrightarrow{\mathcal{D}} N(0, 1), \quad \text{as} \quad m, n \to \infty.$$

Therefore, from (1.6) we have

$$\mathrm{P}(\mu_1\mu_2 \in I_d) - 2\Phi\left(\frac{d}{\sqrt{\sigma_1^2\mu_2^2/m + \sigma_2^2\mu_1^2/n}} \right) - 1 + o(1), \tag{1.7}$$

where Φ is the standard normal distribution function. Let a be a number satisfying $\Phi(a) = 1 - \beta/2$. Now, minimize $t = m + n$ subject to

$$\frac{\sigma_1^2\mu_2^2}{m} + \frac{\sigma_2^2\mu_1^2}{n} = \frac{d^2}{a^2} \quad \text{and} \quad \frac{m}{n} = \frac{\sigma_1\mu_2}{\sigma_2\mu_1}. \tag{1.8}$$

It is easily seen that the constraint $m/n = \sigma_1\mu_2/\sigma_2\mu_1$ is redundant. The optimal pair $(m^\star, n^\star)$ which satisfies (1.8) and for which $t = m + n$ is minimum is given by

$$m^\star = \frac{a^2\sigma_1\mu_2}{d^2}\Delta, \qquad n^\star = \frac{a^2\sigma_2\mu_1}{d^2}\Delta, \tag{1.9}$$

where $\Delta = \sigma_1\mu_2 + \sigma_2\mu_1$. For this pair

$$\frac{m^\star}{n^\star} = \frac{\sigma_1\mu_2}{\sigma_2\mu_1},$$

and the total sample size is

$$t^\star = m^\star + n^\star = \frac{a^2}{d^2}\Delta^2. \tag{1.10}$$

When $\sigma_1\mu_2$ and $\sigma_2\mu_1$ are unknown, the optimal fixed sample sizes $(m^\star, n^\star)$ cannot be used in practice. For this case, we shall now give a sequential procedure for determining m and n as random variables in such a way that (1.9) (and hence (1.10)) will hold approximately with probability one. The procedure consists of (i) a sampling scheme which tells us at each stage whether to take the next observation on X or Y, and (ii) a stopping rule which determines m and n and therefore I_d by (1.5).

The rest of the paper is organized as follows. In Section 2 we describe the fully sequential procedure and state the main theorems. In Section 3 a simulation study is presented. The proofs are given in Section 4.

2. Sequential procedure and its asymptotic optimality

Now, we use the notations in (1.1) and (1.2) to define a sampling scheme and a stopping rule. For the rest of the paper, we assume without loss of generality that $\mu_1 > 0$ and $\mu_2 > 0$.

(i) Sampling Scheme: We take an initial sample size $n_0 = ca^{2\alpha}/d^{2\alpha}$, for some suitable $c > 0$ and $0 < \alpha < 1$, on X and Y. Then if at any stage we have taken i observations on X and j on Y with $t = i + j \geq 2n_0$, we take the next observation on X or on Y according as

$$\frac{i}{j} \leq \frac{S_{1,i}\bar{Y}_j}{S_{2,j}\bar{X}_i} \quad \text{or} \quad \frac{i}{j} > \frac{S_{1,i}\bar{Y}_j}{S_{2,j}\bar{X}_i}. \tag{2.1}$$

This procedure generates an infinite sequence of observations and does not depend on the value of β or d.

(ii) Stopping Rule: The stopping time $T = M + N$ is the first integer $T(\geq 2n_0)$ such that, if M observations on X and N observations on Y have been taken, with $M + N = T$ and

$$M \geq \frac{a^2}{d^2}S_{1,M}\bar{Y}_N\hat{\Delta}_{M,N}, \quad \text{and} \quad N \geq \frac{a^2}{d^2}S_{2,N}\bar{X}_M\hat{\Delta}_{M,N}, \tag{2.2}$$

where

$$\hat{\Delta}_{M,N} = S_{1,M}\bar{Y}_N + S_{2,N}\bar{X}_M.$$

When we stop sampling, we compute $\bar{X}_M$ and $\bar{Y}_N$ and propose the confidence interval I_d by (1.5) with (m,n) replaced by (M,N).

The problem considered in this paper can be extended to three or more populations for which a similar sampling scheme and stopping rule can be defined. This will be done elsewhere. Next, we state the main theorems of this paper, the first of which establishes the asymptotic consistency of the sequential confidence interval and the asymptotic efficiency of the stopping rule.

THEOREM 2.1. *For the fully sequential procedure defined in* (2.1) *and* (2.2), *as* $d \to 0$, *the following hold*:

$$T/t^\star \to 1, \qquad \textit{almost surely (a.s.)}. \tag{2.3}$$

If $EX^2 < \infty$ *and* $EY^2 < \infty$, *then*

$$\lim_{d\to 0} \mathrm{P}(\mu_1\mu_2 \in I_d) = 1 - \beta \qquad \textit{(asymptotic consistency)}. \tag{2.4}$$

If $E|X|^{2+\epsilon} < \infty$ *and* $E|Y|^{2+\epsilon} < \infty$ *for some* $\epsilon > 0$, *then*

$$\lim_{d\to 0} ET/t^\star = 1 \qquad \textit{(asymptotic efficiency)}, \tag{2.5}$$

where $t^\star$ *is the optimal total sample size defined in* (1.10).

The next two theorems concern the first and second order properties of the stopping rule T defined in (2.2).

THEOREM 2.2. *Suppose* $EX^4 < \infty$ *and* $EY^4 < \infty$. *Then, for the stopping time* T *defined in* (2.2) *the following hold*:

$$\frac{T - t^\star}{\sqrt{t^\star}} \xrightarrow{\mathcal{D}} N(0,\gamma^2), \tag{2.6}$$

where

$$\gamma^2 = \frac{4\sigma_1^4\sigma_2^2 + 4\sigma_1\sigma_2\mu_2\mu_{3,X} + \mu_2^2(\mu_{4,X} - \sigma_1^4)}{\sigma_1^3\mu_2\Delta} + \frac{4\sigma_1^2\sigma_2^4 + 4\sigma_1\sigma_2\mu_1\mu_{3,Y} + \mu_1^2(\mu_{4,Y} - \sigma_2^4)}{\sigma_2^3\mu_1\Delta}$$

with Δ *defined as in* (1.9), *and* $\mu_{3,X}$, $\mu_{4,X}$, $\mu_{3,Y}$ *and* $\mu_{4,Y}$ *defined as in* (1.4). *Furthermore, if* $EX^{16} < \infty$ *and* $EY^{16} < \infty$, *and* α *in* (2.1) *is such that* $3/4 < \alpha < 1$, *then for some* $d_0 > 0$

$$\left\{ \left(\frac{T - t^\star}{\sqrt{t^\star}} \right)^2, 0 < d < d_0 \right\} \textit{ is uniformly integrable.} \tag{2.7}$$

Table 1. $CV = 1.0$.

α	n_0	λ	d	t^*	$\widehat{ET}$	$\hat{P}$	$\widehat{ET} - t^*$
4/5	5	3.38	0.8	175.03	168.83 (0.735)	0.947 (0.0071)	−6.20
	6	4.50	0.6	311.17	306.17 (0.883)	0.959 (0.0063)	−5.00
	8	5.40	0.5	448.08	444.90 (1.028)	0.947 (0.0071)	3.09
	12	6.75	0.4	700.13	698.65 (1.299)	0.952 (0.0068)	−1.48
8/9	5	3.38	0.8	175.03	170.80 (0.688)	0.949 (0.0070)	−4.23
	8	4.50	0.6	311.17	307.22 (0.866)	0.955 (0.0066)	−3.95
	11	5.40	0.5	448.08	444.75 (1.002)	0.964 (0.0059)	−3.33
	16	6.75	0.4	700.13	695.76 (1.307)	0.950 (0.0069)	−4.37
16/17	5	3.38	0.8	175.03	169.36 (0.691)	0.943 (0.0073)	−5.67
	9	4.50	0.6	311.17	305.64 (0.884)	0.938 (0.0076)	−5.53
	13	5.40	0.5	448.08	443.53 (1.068)	0.955 (0.0066)	−4.55
	19	6.75	0.4	700.13	696.99 (1.330)	0.958 (0.0064)	−3.14
32/33	5	3.38	0.8	175.03	170.22 (0.731)	0.954 (0.0066)	−4.81
	9	4.50	0.6	311.17	307.02 (0.877)	0.953 (0.0067)	−4.15
	14	5.40	0.5	448.08	442.56 (1.054)	0.951 (0.0068)	−5.52
	21	6.75	0.4	700.13	697.07 (1.254)	0.939 (0.0076)	−3.06
64/65	5	3.38	0.8	175.03	169.91 (0.721)	0.934 (0.0079)	−5.12
	10	4.50	0.6	311.17	308.00 (0.875)	0.945 (0.0072)	−3.17
	14	5.40	0.5	448.08	444.16 (1.055)	0.952 (0.0068)	−3.92
	22	6.75	0.4	700.13	694.68 (1.363)	0.953 (0.0067)	−5.45

Consequently,

$$E\left(\frac{T-t^*}{\sqrt{t^*}}\right)^2 = \gamma^2 + o(1). \tag{2.8}$$

THEOREM 2.3. *Suppose* $EX^{16} < \infty$ *and* $EY^{16} < \infty$ *and* α *is as in Theorem 2.2. Then, for* T *defined in* (2.2)

$$\begin{aligned} ET = t^* &- \frac{\sigma_2(\Delta + \sigma_1\mu_2)}{\sigma_1^3\mu_2\Delta}\mu_{3,X} - \frac{\sigma_1(\Delta + \sigma_2\mu_1)}{\sigma_2^3\mu_1\Delta}\mu_{3,Y} \\ &- \frac{\mu_{4,X}}{\sigma_1^4} - \frac{\mu_{4,Y}}{\sigma_2^4} + 2 - \frac{\sigma_1\sigma_2}{\mu_1\mu_2} + C_0 + o(1) \end{aligned} \tag{2.9}$$

as $d \to 0$, *where* $0 < C_0 < 2$ *is a constant.*

Remark. Assertions (2.7) and (2.8) in Theorem 2.2 are of interest mainly because they are used in the proof of Theorem 2.3.

The proofs of the above three theorems are given in Section 4. Next, we support the theoretical results stated above through simulations for normal populations.

Table 2. $CV = 2/3$.

α	n_0	λ	d	t^*	$\widehat{ET}$	$\hat{P}$	$\widehat{ET} - t^*$
4/5	5	4.50	0.8	216.09	212.48 (0.788)	0.948 (0.0070)	−3.61
	6	6.00	0.6	384.16	380.47 (0.987)	0.944 (0.0073)	−3.69
	8	7.20	0.5	553.19	547.52 (1.165)	0.954 (0.0066)	5.67
	12	9.00	0.4	864.36	860.92 (1.475)	0.947 (0.0071)	−3.44
8/9	5	4.50	0.8	216.09	211.37 (0.785)	0.948 (0.0070)	−4.72
	8	6.00	0.6	384.16	379.46 (0.993)	0.949 (0.0070)	−4.70
	11	7.20	0.5	553.19	548.51 (1.185)	0.951 (0.0068)	−4.68
	16	9.00	0.4	864.36	860.25 (1.483)	0.942 (0.0074)	−4.11
16/17	5	4.50	0.8	216.09	212.44 (0.796)	0.944 (0.0073)	−3.65
	9	6.00	0.6	384.16	379.54 (0.995)	0.961 (0.0061)	−4.62
	13	7.20	0.5	553.19	549.84 (1.169)	0.949 (0.0070)	−3.35
	19	9.00	0.4	864.36	861.00 (1.454)	0.954 (0.0066)	−3.36
32/33	5	4.50	0.8	216.09	210.87 (0.764)	0.952 (0.0068)	−5.22
	9	6.00	0.6	384.16	380.14 (0.974)	0.934 (0.0079)	−4.02
	14	7.20	0.5	553.19	551.02 (1.203)	0.948 (0.0070)	−2.17
	21	9.00	0.4	864.36	862.18 (1.432)	0.964 (0.0059)	−2.18
64/65	5	4.50	0.8	216.09	211.38 (0.784)	0.952 (0.0068)	−4.71
	10	6.00	0.6	384.16	379.81 (0.969)	0.957 (0.0064)	−4.35
	14	7.20	0.5	553.19	551.21 (1.224)	0.942 (0.0074)	−1.98
	22	9.00	0.4	864.36	860.39 (1.449)	0.954 (0.0066)	−3.97

3. Simulation study

It is difficult to try to find exact values of $P(\mu_1\mu_2 \in I_d)$ by analytic methods. Instead, we present the results of an experiment using pseudo-random normal deviates (that is, the populations are assumed to be normal) for the values (see displays (1.8) and (2.1))

$$1-\beta = 0.95, \qquad a = 1.96, \qquad c = 1.0 \qquad n_0 = [\max\{5, ca^{2\alpha}/d^{2\alpha}\}]$$

and

$$CV = \frac{\sigma_1\mu_2}{\sigma_2\mu_1}, \qquad \lambda = \frac{\sigma_2\mu_1}{d}$$

for which (1.10) becomes

$$t^* = [1.96\lambda(CV+1)]^2.$$

Values $\alpha = \frac{4}{5}, \frac{8}{9}, \frac{16}{17}, \frac{32}{33}, \frac{64}{65}$; $d = 0.8, 0.6, 0.5, 0.4$; $CV = 1, \frac{2}{3}, \frac{1}{2}, \frac{1}{4}$ ($\mu_1 = \mu_2 = 3$, $(\sigma_1, \sigma_2) = (0.9, 0.9), (0.8, 1.2), (1, 2), (0.6, 2.4)$) were used, and 1,000 sequences of X and Y were generated for each combination. We denote by $\hat{P}$ the coverage frequency of I_d using the stopping rule T in (2.2) and by $\widehat{ET}$ the

Table 3. $CV = 1/2$.

α	n_0	λ	d	t^*	$\widehat{ET}$	$\hat{P}$	$\widehat{ET} - t^*$
4/5	5	7.50	0.8	486.20	482.69 (1.232)	0.954 (0.0066)	−3.51
	6	10.00	0.6	864.36	861.07 (1.652)	0.954 (0.0066)	−3.29
	8	12.00	0.5	1244.68	1241.96 (1.896)	0.952 (0.0068)	−2.72
	12	15.00	0.4	1944.81	1942.48 (2.345)	0.959 (0.0063)	−2.33
8/9	5	7.50	0.8	486.20	482.81 (1.197)	0.943 (0.0073)	−3.39
	8	10.00	0.6	864.36	863.81 (1.562)	0.954 (0.0066)	−0.55
	11	12.00	0.5	1244.68	1241.59 (1.996)	0.957 (0.0064)	−3.09
	16	15.00	0.4	1944.81	1939.00 (2.353)	0.955 (0.0066)	−5.81
16/17	5	7.50	0.8	486.20	481.46 (1.238)	0.946 (0.0072)	−4.74
	9	10.00	0.6	864.36	861.40 (1.626)	0.959 (0.0063)	−2.96
	13	12.00	0.5	1244.68	1239.97 (1.812)	0.952 (0.0068)	−4.71
	19	15.00	0.4	1944.81	1939.65 (2.446)	0.945 (0.0072)	−5.16
32/33	5	7.50	0.8	486.20	482.37 (1.167)	0.951 (0.0068)	−3.83
	9	10.00	0.6	864.36	859.55 (1.604)	0.947 (0.0071)	−4.81
	14	12.00	0.5	1244.68	1242.56 (1.943)	0.954 (0.0066)	−2.12
	21	15.00	0.4	1944.81	1942.89 (2.406)	0.940 (0.0075)	−1.92
64/65	5	7.50	0.8	486.20	480.99 (1.225)	0.951 (0.0068)	−5.21
	10	10.00	0.6	864.36	860.80 (1.569)	0.957 (0.0064)	−3.56
	14	12.00	0.5	1244.68	1242.64 (1.852)	0.947 (0.0071)	−2.04
	22	15.00	0.4	1944.81	1938.52 (2.363)	0.947 (0.0071)	−6.29

average value of T. The following tables give the coverage frequency and average value of T (with standard error), and the difference $\widehat{ET} - t^*$ for various values of α, n_0, λ and d when $CV = 1$, $2/3$, $1/2$ and $1/4$, respectively.

Remark 1.

(i) From the simulation results it seems that the sequential procedure is successful in keeping the coverage frequency close to 0.95 for small values of d.

(ii) The tables also confirm that the difference $ET - t^*$ is bounded. Also, it is interesting to note that the difference $\widehat{ET} - t^*$ is always negative. This is surprising since, after all, the stopping rule T in (2.2) is obtained by mimicking the optimal fixed sample sizes (see (1.9) and (1.10)). Despite this, the simulated expected value of T could be lower than the total fixed sample size. Note that there are other instances in the literature where it has been observed that negative regrets using stopping rules are possible. See, for instance, Martinsek (1983, 1990), Sriram (1991, 1992) and Takada (1992).

Here, for the normal populations (chosen in the simulation), the theoretical bound for $ET - t^*$ is

$$-\frac{\sigma_2(\Delta + \sigma_1\mu_2)}{\sigma_1^3\mu_2\Delta}\mu_{3,X} - \frac{\sigma_1(\Delta + \sigma_2\mu_1)}{\sigma_2^3\mu_1\Delta}\mu_{3,Y} \tag{3.1}$$

Table 4. $CV = 1/4$.

α	n_0	λ	d	t^*	$\widehat{ET}$	$\hat{P}$	$\widehat{ET} - t^*$
4/5	5	9.00	0.8	486.20	480.73 (1.142)	0.952 (0.0068)	−5.47
	6	12.00	0.6	864.36	860.05 (1.574)	0.948 (0.0070)	−4.31
	8	14.40	0.5	1244.68	1239.07 (1.792)	0.948 (0.0070)	−5.61
	12	18.00	0.4	1944.81	1943.82 (2.280)	0.942 (0.0074)	−0.99
8/9	5	9.00	0.8	486.20	482.68 (1.195)	0.939 (0.0076)	−3.52
	8	12.00	0.6	864.36	857.87 (1.555)	0.939 (0.0076)	−6.49
	11	14.40	0.5	1244.68	1240.64 (1.779)	0.951 (0.0068)	−4.04
	16	18.00	0.4	1944.81	1940.99 (2.271)	0.953 (0.0067)	−3.82
16/17	5	9.00	0.8	486.20	480.09 (1.162)	0.941 (0.0075)	−6.11
	9	12.00	0.6	864.36	861.51 (1.517)	0.956 (0.0065)	−2.85
	13	14.40	0.5	1244.68	1240.36 (1.850)	0.954 (0.0066)	−4.32
	19	18.00	0.4	1944.81	1940.12 (2.259)	0.945 (0.0072)	−4.69
32/33	5	9.00	0.8	486.20	484.16 (1.164)	0.948 (0.0070)	−2.04
	9	12.00	0.6	864.36	860.32 (1.543)	0.958 (0.0064)	−4.04
	14	14.40	0.5	1244.68	1243.83 (1.871)	0.951 (0.0068)	−0.85
	21	18.00	0.4	1944.81	1937.30 (2.317)	0.960 (0.0062)	−7.51
64/65	5	9.00	0.8	486.20	481.65 (1.161)	0.963 (0.0060)	−4.55
	10	12.00	0.6	864.36	861.24 (1.603)	0.943 (0.0073)	−3.12
	14	14.40	0.5	1244.68	1237.77 (1.842)	0.959 (0.0063)	−6.91
	22	18.00	0.4	1944.81	1941.87 (2.248)	0.943 (0.0073)	−2.94

$$- \frac{\mu_{4,X}}{\sigma_1^4} - \frac{\mu_{4,Y}}{\sigma_2^4} + 2 - \frac{\sigma_1\sigma_2}{\mu_1\mu_2} + C_0 + o(1)$$
$$= -4 - \frac{\sigma_1\sigma_2}{\mu_1\mu_2} + C_0 + o(1)$$

where $0 < C_0 < 2$. Clearly, the right side of (3.1) is negative for any value of μ_1, μ_2, σ_1 and σ_2.

4. Lemmas and proofs

The following inequalities will be used throughout the paper. Suppose $M > n_0$ and that just before the M-th observation on X there were $(M-1)$ observations on X and J observations on Y. Then, by the sampling scheme (2.1),

$$\frac{M-1}{J} \leq \frac{S_{1,M-1}\bar{Y}_J}{S_{2,J}\bar{X}_{M-1}}. \tag{4.1}$$

This implies that on $[M > n_0]$

$$M < \frac{a^2}{d^2} S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J} + 1 \quad \text{and} \quad J \leq N, \tag{4.2}$$

for otherwise we would have (from (4.1))

$$J \geq (M-1)\frac{S_{2,J}\bar{X}_{M-1}}{S_{1,M-1}\bar{Y}_J}$$
$$\geq \frac{a^2}{d^2} S_{2,J}\bar{X}_{M-1}\hat{\Delta}_{M-1,J}$$

and hence (by (2.2)) sampling would have stopped at $(M-1, J)$ stage. Similarly, let I be the number of observations in X just before the N-th observation is taken from Y. Then, once again by the sampling scheme (2.1)

$$\frac{I}{N-1} > \frac{S_{1,I}\bar{Y}_{N-1}}{S_{2,N-1}\bar{X}_I} \tag{4.3}$$

and on $[N > n_0]$,

$$N < \frac{a^2}{d^2} S_{2,N-1}\bar{X}_I\hat{\Delta}_{I,N-1} + 1 \quad \text{and} \quad I \leq M. \tag{4.4}$$

The positive integer-valued random variables I, J considered above play a crucial role in the analysis given below. Incidentally, the idea of introducing I and J is due to Robbins *et al.* (1967).

PROOF OF THEOREM 2.1. For the assertion (2.3), it is easy to see that as $d \to 0$, $M \to \infty$ and, hence, from (4.1), we have $J \to \infty$. Therefore, from (4.2)

$$\limsup_{d\to 0} \frac{M}{m^\star} \leq 1 \quad \text{a.s.},$$

where $m^\star$ is defined as in (1.9). The reverse inequality for the lim inf is obvious from (2.2). Hence, as $d \to 0$

$$M/m^\star \to 1 \quad \text{a.s.}. \tag{4.5}$$

Similarly, as $d \to 0$

$$N/n^\star \to 1 \quad \text{a.s.}, \tag{4.6}$$

which, together with (4.5) implies (2.3). Assertion (2.4) can be proved using (1.3), arguments leading to (1.7) and the Anscombe's theorem. Finally, for (2.5), note from (4.2) that

$$\frac{M}{m^\star} \leq K_0[S_{1,M-1}^2\bar{Y}_J^2 + S_{1,M-1}\bar{Y}_J S_{2,J}\bar{X}_{M-1}] + K_1$$
$$\leq K_0[s_1^2 y^2 + s_1 y s_2 x] + K_1, \tag{4.7}$$

where

$$s_1^2 = \sup_{m\geq 1} m^{-1}\sum_{i=1}^{m}(X_i-\mu_1)^2, \qquad x = \sup_{m\geq 1}\left|m^{-1}\sum_{i=1}^{m}X_i\right|$$
$$s_2^2 = \sup_{n\geq 1} n^{-1}\sum_{i=1}^{n}(Y_i-\mu_2)^2, \qquad y = \sup_{n\geq 1}\left|n^{-1}\sum_{i=1}^{n}Y_i\right|,$$

and K_0 and K_1 are constants not depending on d. By independence and Theorem 10.3.3 of Chow and Teicher (1978)

$$Es_1^2y^2 = Es_1^2Ey^2 < \infty, \tag{4.8}$$

provided $E|X|^{2+\epsilon} < \infty$ and $E|Y|^2 < \infty$. Similar argument with the Cauchy-Schwarz inequality shows that

$$Es_1ys_2x < \infty, \tag{4.9}$$

provided $E|X|^{2+\epsilon} < \infty$ and $E|Y|^{2+\epsilon} < \infty$. Therefore, from (4.5), (4.7), (4.8), (4.9) and the dominated convergence theorem

$$EM/m^\star \to 1 \quad \text{as} \quad d\to 0.$$

Similarly,

$$EN/n^\star \to 1 \quad \text{as} \quad d\to 0.$$

From these we have (2.5). Hence the theorem. □

In order to prove Theorems 2.2 and 2.3, we need some lemmas, the first of which concerns the lower tail probability rate of N and M defined in (2.2).

LEMMA 4.1. *Assume $E|X|^{2p} < \infty$, $E|Y|^{2p} < \infty$, for $p \geq 2$. Then, for the fully sequential procedure defined in* (2.1) *and* (2.2), α *in* (2.1), *and* $0<\epsilon<1$

$$\mathrm{P}(N \leq (1-\epsilon)n^\star) = O(d^{p\alpha}), \tag{4.10}$$
$$\mathrm{P}(M \leq (1-\epsilon)m^\star) = O(d^{p\alpha}), \tag{4.11}$$

as $d\to 0$, where $m^\star$ and $n^\star$ are defined as in (1.9).

PROOF. For $0<\epsilon<1$ and $\delta>0$, let $D_i(\delta) = [\sigma_i^2-\delta, \sigma_i^2+\delta]$ and $H_i(\delta) = [\mu_i-\delta, \mu_i+\delta]$, $i=1,2$. Then, by (2.2), for some $\delta_0>0$

$$\begin{aligned}\mathrm{P}(N \leq (1-\epsilon)n^\star) &\leq \mathrm{P}(S_{2,N}\bar{X}_M\hat{\Delta}_{M,N} \leq (1-\epsilon)\sigma_2\mu_1\Delta)\\ &\leq \mathrm{P}(|S_{2,N}\bar{X}_M\hat{\Delta}_{M,N} - \sigma_2\mu_1\Delta| > \epsilon\sigma_2\mu_1\Delta)\\ &\leq \mathrm{P}(S_{1,M}^2 \notin D_1(\delta_0)) + \mathrm{P}(S_{2,N}^2 \notin D_2(\delta_0))\\ &\quad + \mathrm{P}(\bar{X}_M \notin H_1(\delta_0)) + \mathrm{P}(\bar{Y}_N \notin H_2(\delta_0)).\end{aligned} \tag{4.12}$$

Consider the first term on the right side of (4.12). Note that

$$\begin{aligned}
\mathrm{P}(S_{1,M}^2 \notin D_1(\delta_0)) &= \mathrm{P}\left(\left|\frac{\sum_{i=1}^{M}(X_i - \bar{X}_M)^2}{M} - \sigma_1^2\right| > \delta_0\right) \\
&\leq \mathrm{P}\left(\left|\frac{\sum_{i=1}^{M}(X_i - \mu_1)^2}{M} - \sigma_1^2\right| > \frac{\delta_0}{2}\right) \\
&\quad + \mathrm{P}(|(\bar{X}_M - \mu_1)^2| > \delta_0/2) \\
&= (i) + (ii).
\end{aligned}$$

Observe that

$$\left\{\sum_{i=1}^{m}[(X_i - \mu_1)^2 - \sigma_1^2]/m, m \geq n_0\right\}$$

is a reverse martingale with respect to $\mathcal{G}_m = \sigma\{\bar{Z}_m, \bar{Z}_{m+1}, \ldots\}$ where $\bar{Z}_k = \sum_{i=1}^{k}[(X_i - \mu_1)^2 - \sigma_1^2]/k$. Therefore, by the maximal inequality for reverse submartingales (see Chow *et al.* (1971), display (4.39) and Stout (1974), (second) Lemma 3.3.1)

$$\begin{aligned}
(i) &\leq \mathrm{P}\left(\sup_{m \geq n_0}\left|\frac{\sum_{i=1}^{m}[(X_i - \mu_1)^2 - \sigma_1^2]}{m}\right| > \frac{\delta_0}{2}\right) \\
&\leq K_0 E\left|\frac{\sum_{i=1}^{n_0}[(X_i - \mu_1)^2 - \sigma_1^2]}{n_0}\right|^p \\
&= O(n_0^{-p/2}),
\end{aligned}$$

where the last equality follows from Marcinkiewicz-Zygmund Inequality (see Chow and Teicher (1978), Corollary 10.3.2) and the assumption that $E|X_1|^{2p} < \infty$ for $p \geq 2$. Similarly, $(ii) = O(n_0^{-p/2})$. Since $n_0 = ca^{2\alpha}/d^{2\alpha}$ (see (2.1)), the required result in (4.10) follows from the above arguments. Argue as for (4.10) to show (4.11). □

The rest of the lemmas also concern the integer-valued random variables I and J introduced in (4.1) and (4.4).

LEMMA 4.2. *Under the assumptions of Lemma* 4.1, *the following hold for I and J defined in* (4.3) *and* (4.1), *respectively, as $d \to 0$:*

$$\mathrm{P}(I \leq (1-\epsilon)m^*) = O(d^{p\alpha}) \tag{4.13}$$

$$\mathrm{P}(J \leq (1-\epsilon)n^*) = O(d^{p\alpha}) \tag{4.14}$$

for $0 < \epsilon < 1$ and α as defined in (2.1). *Furthermore, as $d \to 0$*

$$\frac{I}{m^*} \to 1 \quad a.s. \qquad and \qquad \frac{J}{n^*} \to 1 \quad a.s. \tag{4.15}$$

PROOF. First we will show (4.14). Then, (4.13) will follow in a similar way. Observe from (4.1) that

$$\begin{aligned}[J \le (1-\epsilon)n^\star] &= \left[\frac{M-1}{J} \le \frac{S_{1,M-1}\bar{Y}_J}{S_{2,J}\bar{X}_{M-1}}, J \le (1-\epsilon)n^\star\right] \\ &\subset \left[M-1 \le (1-\epsilon)m^\star \frac{\sigma_2\mu_1}{\sigma_1\mu_2}\frac{S_{1,M-1}\bar{Y}_J}{S_{2,J}\bar{X}_{M-1}}\right] \\ &\subset [M \le (1-\epsilon')m^\star] \cup [S^2_{1,M-1} \notin D_1(\delta_0)] \cup [\bar{X}_{M-1} \notin H_1(\delta_0)] \\ &\quad \cup [S^2_{2,J} \notin D_2(\delta_0)] \cup [\bar{Y}_J \notin H_2(\delta_0)].\end{aligned}$$

for some $\epsilon' \in (0,1)$ and $\delta_0 > 0$. Now argue as for the right side of (4.12) and use (4.11) to get the desired result in (4.14). As for the first assertion in (4.15), it follows from (4.3) and (4.6) that

$$\liminf_{d\to 0} \frac{I}{m^\star} \ge \lim_{d\to 0} \frac{S_{1,I}\bar{Y}_{N-1}}{S_{2,N-1}\bar{X}_I}\frac{N-1}{m^\star} = 1 \quad \text{a.s.}$$

Since $I \le M$ (see (4.4)), by (4.5), $\limsup_{d\to 0} I/m^\star \le 1$ a.s. Hence the first assertion in (4.15). The second assertion in (4.15) follows similarly. Hence the lemma. □

For the rest of the lemmas let

(4.16) $$L = M \quad \text{or} \quad I \qquad \text{and} \qquad H = N \quad \text{or} \quad J,$$

where M, I, N, and J are random variables defined as in (2.2), (4.1) and (4.3).

LEMMA 4.3. *Assume that $EX^{8p} < \infty$, $EY^{8p} < \infty$ for $p \ge 1$. Then, for some $d_0 \in (0,1)$*

(4.17) $$\{(L/m^\star)^{4p}; 0 < d < d_0\} \quad \text{and} \quad \{(H/n^\star)^{4p}; 0 < d < d_0\}$$

are uniformly integrable (u.i.). Furthermore, if $E|X|^{2p} < \infty$, $E|Y|^{2p} < \infty$ for $p \ge 2$ and α in (2.1) is such that $\alpha > 3/4$, then for some $d_0 \in (0,1)$

(4.18) $$\{(m^\star/L)^{3p/2}; 0 < d < d_0\} \quad \text{and} \quad \{(n^\star/H)^{3p/2}; 0 < d < d_0\}$$

are u.i.

PROOF. Since by definition $I \le M$ and $J \le N$, it suffices to show that assertion (4.17) holds for $L = M$ and $H = N$. We only show (4.17) for $L = M$. The second assertion in (4.17) with $H = N$ follows similarly. Now, from (4.7) for $p \ge 1$

(4.19) $$(M/m^\star)^{4p} \le K_2[s_1^{8p}y^{8p} + s_1^{4p}y^{4p}s_2^{4p}x^{4p}] + K_3,$$

where s_1^2, x^2, s_2^2 and y^2 are as defined in (4.7) and K_2, K_3 are positive constants not depending on d. Once again use independence, Theorem 10.3.3 of Chow and Teicher (1978), the Cauchy-Schwarz Inequality and the moment assumption to show that the right hand side of (4.19) is integrable. Hence, $\{(M/m^\star)^{4p}; 0 < d < d_0\}$ is u.i. Hence, the assertion (4.17).

Next for (4.18), we show $\{(n^\star/H)^{3p/2}, 0 < d < d_0\}$ is uniformly integrable. Once again, since $n^\star/N \leq n^\star/J$, it suffices to show that $\{(n^\star/J)^{3p/2}, 0 < d < d_0\}$ is uniformly integrable. Observe that for $0 < \epsilon < 1$

$$\left(\frac{n^\star}{J}\right)^{3p/2} = \left(\frac{n^\star}{J}\right)^{3p/2} I_{[n_0 \leq J \leq (1-\epsilon)n^\star]} + \left(\frac{n^\star}{J}\right)^{3p/2} I_{[J > (1-\epsilon)n^\star]}. \tag{4.20}$$

The second term on the right side of (4.20) is less than $(1-\epsilon)^{-3p/2}$. For the first term on the right side of (4.20), recall that $n_0 = ca^{2\alpha}/d^{2\alpha}$ and use (4.14) to get

$$\begin{aligned} E\left(\frac{n^\star}{J}\right)^{3p/2} I_{\{n_0 \leq J \leq (1-\epsilon)n^\star\}} &\leq \left(\frac{n^\star}{n_0}\right)^{3p/2} \mathrm{P}(J \leq (1-\epsilon)n^\star) \\ &= O(d^{-3p(1-\alpha)})O(d^{p\alpha}) \\ &= o(1), \end{aligned}$$

as $d \to 0$, since by assumption $\alpha > 3/4$. That $\{(n^\star/J)^{3p/2}, 0 < d < d_0\}$ is uniformly integrable now follows easily from (4.15) and the above arguments. Similarly, $\{(m^\star/L)^{3p/2}; 0 < d < d_0\}$ is uniformly integrable. Hence the lemma. □

LEMMA 4.4. *The following hold for some $d_0 > 0$:*

(i) *If $E|X|^{4r} < \infty$ and $E|Y|^{4r} < \infty$ for $r \geq 1$, then*

$$\left\{\left(\sum_{i=1}^{M}(X_i - \mu_1)\right)^{4r} / m^{\star 2r}; 0 < d < d_0\right\}$$

and

$$\left\{\left(\sum_{i=1}^{N}(Y_i - \mu_2)\right)^{4r} / n^{\star 2r}; 0 < d < d_0\right\}$$

are uniformly integrable.

(ii) *If $E|X|^{8r} < \infty$ and $E|Y|^{8r} < \infty$ for $r \geq 1$, then*

$$\left\{\left(\sum_{i=1}^{M}[(X_i - \mu_1)^2 - \sigma_1^2]\right)^{4r} / m^{\star 2r}, 0 < d < d_0\right\}$$

and

$$\left\{\left(\sum_{i=1}^{N}[(Y_i - \mu_2)^2 - \sigma_2^2]\right)^{4r} / n^{\star 2r}, 0 < d < d_0\right\}$$

are all uniformly integrable. Furthermore, if $E|X|^{8r} < \infty$ *and* $E|Y|^{8r} < \infty$ *for* $r \geq 1$, *then assertion* (i) *above holds with* M *and* N *replaced by the positive integer-valued random variables* I *and* J *defined in* (4.3) *and* (4.1), *respectively. Also, if* $E|X|^{8r+2} < \infty$ *and* $E|Y|^{8r+2} < \infty$ *for* $r \geq 1$, *then assertion* (ii) *above holds with* M *and* N *replaced by* I *and* J, *respectively.*

PROOF. First we prove the uniform integrability of

$$\left\{ \left(\sum_{i=1}^{M} (X_i - \mu_1) \right)^{4r} / m^{\star 2r}, 0 < d < d_0 \right\}.$$

This follows from Chow and Yu (1981) once we make the following observations. Define σ-algebras

$$\mathcal{D}_n = \sigma\{Y_1, Y_2, \ldots, Y_n, X_i, i \geq 1\}, \qquad n \geq 1$$

and

$$\mathcal{F}_m = \sigma\{X_1, X_2, \ldots, X_m, Y_i \geq 1, i \geq 1\}, \qquad m \geq 1.$$

Observe that the event $\{M = m\}$ is measurable with respect to $\mathcal{F}_m$ and $\{\sum_{i=1}^{m}(X_i - \mu_1), \mathcal{F}_m, m \geq 1\}$ is a martingale. Furthermore, it is possible to show using (4.7) and arguments similar to (4.19) that $\{(M/m^\star)^{2r}; 0 < d < d_0\}$ is uniformly integrable, provided $E|X|^{4r} < \infty$ and $E|Y|^{4r} < \infty$ for $r \geq 1$. Therefore, the first assertion in (i) follows from Lemma 5 of Chow and Yu (1981). The second assertion in (i) and the assertions in (ii) follow similarly.

As for assertions (i) and (ii) for M and N replaced by the positive integer-valued random variables I and J, respectively, first note that I and J are not stopping times. This means that it is not possible to use Lemma 5 of Chow and Yu (1981) for I and J. But, from Theorem 2 of Chow *et al.* (1979) and (4.17) of Lemma 4.3 above, it follows that assertation (i) holds with M and N replaced by I and J, under the assumption that $E|X|^{8r} < \infty$ and $E|Y|^{8r} < \infty$ for $r \geq 1$. Assertion (ii) for I and J follows similarly. Hence the lemma. □

Remark 2. In order to prove assertions (i) and (ii) in Lemma 4.4 for I and J one needs finiteness of $8r$ and $(8r+1)$, $r \geq 1$, moments, whereas for M and N one only needs finiteness of $4r$ and $8r$ moments. As indicated in the proof, the increase in moment assumptions (for I and J) is due to the fact that (a) I and J are not necessarily stopping times, and (b) for positive integer-valued random variables, Theorem 2 of Chow *et al.* (1979) requires $\{(I/m^\star)^{4r}; 0 < d < d_0\}$ and $\{(J/n^\star)^{4r}; 0 < d < d_0\}$ to be uniformly integrable, which hold (by Lemma 4.3) when $E|X|^{8r} < \infty$ and $E|Y|^{8r} < \infty$. Consequently, the next lemma and Theorems 2.2 and 2.3 require more moments.

The next lemma is somewhat similar to the Theorem in Ghosh and Mukhopadyay (1980). However, our proof for it is a bit more complicated than theirs for the following reasons. Our stopping times M and N defined in (2.2) are very different from the ones considered in Ghosh and Mukhopadhyay (1980) in that the stopping boundaries (see (2.2)) depend on both populations, whereas for the stopping rules defined in Ghosh and Mukhopadhyay, the stopping boundaries depend only on one population. Due to this reason, a sampling scheme is not necessary in their analysis, whereas a sampling scheme (see (2.2) and also, (4.2) and (4.4)) is absolutely necessary for our analysis. See proof below.

LEMMA 4.5. *For L and H defined in* (4.16) *the following hold: If $EX^{16} < \infty$ and $EY^{16} < \infty$ and α in* (2.1) *is such that $1 > \alpha > 3/4$ then for some $d_0 > 0$*

$$\{(L - m^\star)^2/m^\star, 0 < d < d_0\}$$

and

$$\{(H - n^\star)^2/n^\star, 0 < d < d_0\}$$

are uniformly integrable.

PROOF. We only prove that $(N - n^\star)^2/n^\star$ and $(J - n^\star)^2/n^\star$ are uniformly integrable. The rest of the assertions can be proved similarly. Note from (2.2) and (4.4) that

$$\frac{a^2}{d^2}S_{2,N}\bar{X}_M\hat{\Delta}_{M,N} - n^\star \le N - n^\star \le \frac{a^2}{d^2}S_{2,N-1}\bar{X}_I\hat{\Delta}_{I,N-1} - n^\star + 1 + n_0 I_{[N=n_0]}$$

which implies for some generic constant K (not depending on d)

$$\begin{aligned}
(4.21)\quad \frac{(N-n^\star)^2}{n^\star}
&\le \frac{K}{d^2}(S_{2,N}\bar{X}_M\hat{\Delta}_{M,N} - \sigma_2\mu_1\Delta)^2 \\
&\quad + \frac{K}{d^2}(S_{2,N-1}\bar{X}_I\hat{\Delta}_{I,N-1} - \sigma_2\mu_1\Delta)^2 \\
&\quad + K\frac{n_0^2}{n^\star}I_{[N=n_0]} + K \\
&\le Km^\star(S_{1,M}^2 - \sigma_1^2)^4 + Km^\star(S_{1,I}^2 - \sigma_1^2)^4 + Kn^\star(S_{2,N}^2 - \sigma_2^2)^4 \\
&\quad + Kn^\star(S_{2,N-1}^2 - \sigma_2^2)^4 + Km^\star(S_{1,M}^2 - \sigma_1^2)^2 + Km^\star(S_{1,I}^2 - \sigma_1^2)^2 \\
&\quad + Kn^\star(S_{2,N}^2 - \sigma_2^2)^2 + Kn^\star(S_{2,N-1}^2 - \sigma_2^2)^2 + Km^\star(\bar{X}_M - \mu_1)^8 \\
&\quad + Km^\star(\bar{X}_M - \mu_1)^4 + Km^\star(\bar{X}_M - \mu_1)^2 + Kn^\star(\bar{Y}_N - \mu_2)^8 \\
&\quad + Kn^\star(\bar{Y}_N - \mu_2)^4 + Km^\star(\bar{Y}_N - \mu_2)^2 + Km^\star(\bar{X}_I - \mu_1)^8 \\
&\quad + Km^\star(\bar{X}_I - \mu_1)^4 + Km^\star(\bar{X}_I - \mu_1)^2 + Kn^\star(\bar{Y}_{N-1} - \mu_2)^8 \\
&\quad + Kn^\star(\bar{Y}_{N-1} - \mu_2)^4 + Kn^\star(\bar{Y}_{N-1} - \mu_2)^2 + Kd^{(-4\alpha+2)}I_{[N=n_0]} \\
&\quad + K + Km^\star B,
\end{aligned}$$

where

$$(4.22)\qquad B = (S_{2,N} - \sigma_2)^2(\bar{X}_M - \mu_1)^2(S_{1,M} - \sigma_1)^2$$

$$
\begin{aligned}
&+ (S_{2,N} - \sigma_2)^2(\bar{X}_M - \mu_1)^2(\bar{Y}_N - \mu_2)^2 \\
&+ (S_{1,M} - \sigma_1)^2(\bar{X}_M - \mu_1)^2(\bar{Y}_N - \mu_2)^2 \\
&+ (S_{2,N} - \sigma_2)^2(\bar{Y}_N - \mu_2)^2(S_{1,M} - \sigma_1)^2 \\
&+ (S_{2,N-1} - \sigma_2)^2(\bar{X}_I - \mu_1)^2(S_{1,I} - \sigma_1)^2 \\
&+ (S_{2,N-1} - \sigma_2)^2(\bar{X}_I - \mu_1)^2(\bar{Y}_{N-1} - \mu_2)^2 \\
&+ (S_{1,I} - \sigma_1)^2(\bar{X}_I - \mu_1)^2(\bar{Y}_{N-1} - \mu_2)^2 \\
&+ (S_{2,N-1} - \sigma_2)^2(\bar{Y}_{N-1} - \mu_2)^2(S_{1,I} - \sigma_1)^2 \\
&+ (S_{2,N} - \sigma_2)^2(\bar{X}_M - \mu_1)^2(S_{1,M} - \sigma_1)^2(\bar{Y}_N - \mu_2)^2 \\
&+ (S_{2,N-1} - \sigma_2)^2(\bar{X}_I - \mu_1)^2(S_{1,I} - \sigma_1)^2(\bar{Y}_{N-1} - \mu_2)^2.
\end{aligned}
$$

Observe that

$$m^\star(\bar{X}_M - \mu_1)^2 = \frac{m^{\star 2}}{M^2}\frac{[\sum_{i=1}^M (X_i - \mu_1)]^2}{m^\star}.$$

By the Cauchy-Schwarz Inequality and Lemma 4.4 we have that $\{m^\star(\bar{X}_M - \mu_1)^2, 0 < d < d_0\}$ is uniformly integrable. Also observe that

$$
\begin{aligned}
m^\star(\bar{X}_M - \mu_1)^4 &= \frac{m^{\star 3}}{M^4}\frac{[\sum_{i=1}^M (X_i - \mu_1)]^4}{m^{\star 2}} \\
&\le \frac{m^{\star 3}}{n_0^4}\frac{[\sum_{i=1}^M (X_i - \mu_1)]^4}{m^{\star 2}} \\
&= O(d^{(8\alpha-6)})\frac{[\sum_{i=1}^M (X_i - \mu_1)]^4}{m^{\star 2}}.
\end{aligned}
$$

Since $3/4 < \alpha < 1$, by Lemma 4.4 we have that

$$\{m^\star(\bar{X}_M - \mu_1)^4, 0 < d < d_0\}$$

is uniformly integrable. Also observe that from (4.10) of Lemma 4.1

$$
\begin{aligned}
Ed^{(-4\alpha+2)}I_{[N=n_0]} &= d^{(-4\alpha+2)}O(d^{4\alpha}) \\
&= o(1)
\end{aligned}
$$

as $d \to 0$. Therefore,

$$\{d^{-4\alpha+2}I_{[N=n_0]}; 0 < d < d_0\}$$

is uniformly integrable. Next, consider the first term of $m^\star B$. By (4.22)

$$
\begin{aligned}
m^\star(S_{2,N} - \sigma_2)^2(\bar{X}_M - \mu_1)^2(S_{1,M} - \sigma_1)^2 &\le Km^\star(\bar{X}_M - \mu_1)^4(S_{2,N} - \sigma_2)^4 \\
&\quad + Km^\star(S_{1,M} - \sigma_1)^4 \\
&\le Km^\star(\bar{X}_M - \mu_1)^4(S_{2,N}^4 + \sigma_2^4) \\
&\quad + Km^\star(S_{1,M}^2 - \sigma_1^2)^4.
\end{aligned}
$$

Using the arguments above and Lemma 4.4 (ii), it can be shown that

$$\{m^\star(\bar{X}_M - \mu_1)^4(S_{2,N}^4 + \sigma_2^4); 0 < d < d_0\}$$

and

$$\{m^\star(S_{1,M}^2-\sigma_1^2)^4; 0<d<d_0\}$$

are uniformly integrable. Similarly, we can show that all the remaining terms in (4.21) and (4.22) are uniformly integrable. Hence, we have the uniform integrability of $\{(N-n^\star)^2/n^\star, 0<d<d_0\}$.

To show $(J-n^\star)^2/n^\star$ is u.i., observe that

$$\frac{(J-n^\star)^2}{n^\star}\le\frac{(N-n^\star)^2}{n^\star}+\frac{(J-n^\star)^2}{n^\star}I_{[J<n^\star]}.$$

It suffices to show that

$$\{(J-n^\star)^2I_{[J<n^\star]}/n^\star; 0<d<d_0\}$$

is u.i. Now,

$$E\frac{(J-n^\star)^2}{n^\star}I_{[J\le n^\star-B\sqrt{n^\star},J<n^\star]}=2\int_B^\infty t\mathrm{P}[J\le n^\star-t\sqrt{n^\star},J<n^\star]dt + B^2\mathrm{P}[J\le n^\star-B\sqrt{n^\star},J<n^\star]. \tag{4.23}$$

Note that $J\ge n_0$ and $n^\star-t\sqrt{n^\star}>n_0$ implies $0<d<K/t$, for some constant $K>0$. By (4.1) and arguments in Lemmas 4.1 and 4.2, we have that there exist $B_0(=K/d_0)>0$ and $\delta_0>0$ such that for any $t\ge B>B_0$,

$$\begin{aligned}
\mathrm{P}[J\le n^\star-t\sqrt{n^\star},J<n^\star] &= \mathrm{P}[n_0\le J\le n^\star-t\sqrt{n^\star},J<n^\star]\\
&\le \mathrm{P}\left[n_0\le J\le n^\star-t\sqrt{n^\star},\frac{S_{2,J}\bar{X}_{M-1}}{S_{1,M-1}\bar{Y}_J}(M-1)\le J\right]\\
&\le \mathrm{P}\left[tK_0\le\frac{|(M-1)S_{2,J}\bar{X}_{M-1}-n^\star S_{1,M-1}\bar{Y}_J|}{\sqrt{n^\star}}\right]\\
&\quad+\mathrm{P}[\bar{Y}_J\notin H_2(\delta_0)\text{ or }S_{1,M-1}^2\notin D_1(\delta_0)]\\
&\le \mathrm{P}\left[tK_0\le\frac{|(M-1)S_{2,J}\bar{X}_{M-1}-n^\star S_{1,M-1}\bar{Y}_J|}{\sqrt{n^\star}}\right]\\
&\quad+O(d^{4\alpha})\\
&\le \mathrm{P}\left[K_1t\le\frac{|m^\star S_{2,J}\bar{X}_{M-1}-n^\star S_{1,M-1}\bar{Y}_J|}{\sqrt{n^\star}}\right]\\
&\quad+\mathrm{P}\left[K_1t\le\frac{|(M-m^\star-1)S_{2,J}\bar{X}_{M-1}|}{\sqrt{n^\star}}\right]+O(d^{4\alpha})\\
&\le \mathrm{P}\left[K_2t\le\frac{|\sigma_1\mu_2S_{2,J}\bar{X}_{M-1}-\sigma_2\mu_1S_{1,M-1}\bar{Y}_J|}{d}\right]\\
&\quad+\mathrm{P}\left[K_2t\le\frac{|M-m^\star-1|}{\sqrt{m^\star}}\right]+O(d^{4\alpha})\\
&\le \mathrm{P}\left[K_3t\le\frac{|S_{2,J}\bar{X}_{M-1}-\sigma_2\mu_1|}{d}\right]\\
&\quad+\mathrm{P}\left[K_3t\le\frac{|S_{1,M-1}\bar{Y}_J-\sigma_1\mu_2|}{d}\right]\\
&\quad+\mathrm{P}\left[K_3t\le\frac{|M-m^\star|}{\sqrt{m^\star}}\right]+O(d^{4\alpha})
\end{aligned}$$

where K_0, K_1, K_2 and K_3 are positive constants, $D_1(\delta_0)$ and $H_2(\delta_0)$ are defined as in the proof of Lemma 4.1. Using arguments similar to the ones in the proof of Lemmas 4.4 and 4.5, we can show that for $\epsilon > 0$, $\{(S_{2,J}\bar{X}_{M-1} - \sigma_2\mu_1)^{(2+\epsilon)}/d^{(2+\epsilon)}, 0 < d < d_0\}$ and $\{(S_{1,M-1}\bar{Y}_J - \sigma_1\mu_2)^{(2+\epsilon)}/d^{(2+\epsilon)}, 0 < d < d_0\}$ are u.i. Therefore,

$$\text{(4.24)} \qquad \begin{aligned} \mathrm{P}[J \le n^\star - t\sqrt{n^\star}, J < n^\star] \le O(d^{4\alpha}) + t^{-(2+\epsilon)}O(1) \\ + \mathrm{P}\left[K_3 t \le \frac{|M - m^\star|}{\sqrt{m^\star}}\right]. \end{aligned}$$

Since $\sup_{0<d<K/t} d^{4\alpha} = O(t^{-4\alpha})$, by (4.23), (4.24) and the uniform integrability of $\{(M - m^\star)^2/m^\star; 0 < d < d_0\}$ we have that

$$\{(J - n^\star)^2/n^\star; 0 < d < d_0\}$$

is u.i. This completes our proof of the Lemma. □

LEMMA 4.6. *Assume that* $EX^4 < \infty$, $EY^4 < \infty$. *Let* (Z_1, Z_2, Z_3, Z_4) *be a multivariate normal vector with mean* $\underline{0}$ *and* Σ *defined in* (1.4). *Then for* L *and* H *defined in* (4.16), *as* $d \to 0$

$$\begin{aligned} &\frac{a\sqrt{m^\star}}{d}(\sigma_1\mu_2\Delta - S_{1,L}\bar{Y}_H\hat{\Delta}_{L,H})(\bar{X}_L - \mu_1) \\ &\qquad \xrightarrow{\mathcal{D}} -a_0 Z_1(a_1 Z_1 + a_2 Z_2 + a_3 Z_3 + a_4 Z_4), \\ &\frac{a\sqrt{n^\star}}{d}(\sigma_2\mu_1\Delta - S_{2,H}\bar{X}_L\hat{\Delta}_{L,H})(\bar{Y}_H - \mu_2) \\ &\qquad \xrightarrow{\mathcal{D}} -b_0 Z_3(b_1 Z_1 + b_2 Z_2 + b_3 Z_3 + b_4 Z_4), \\ &\frac{a\sqrt{m^\star}}{d}(\sigma_1\mu_2\Delta - S_{1,L}\bar{Y}_H\hat{\Delta}_{L,H})(S_{1,L}^2 - \sigma_1^2) \\ &\qquad \xrightarrow{\mathcal{D}} -a_0 Z_2(a_1 Z_1 + a_2 Z_2 + a_3 Z_3 + a_4 Z_4), \\ &\frac{a\sqrt{n^\star}}{d}(\sigma_2\mu_1\Delta - S_{2,H}\bar{X}_L\hat{\Delta}_{L,H})(S_{2.H}^2 - \sigma_2^2) \\ &\qquad \xrightarrow{\mathcal{D}} -b_0 Z_4(b_1 Z_1 + b_2 Z_2 + b_3 Z_3 + b_4 Z_4), \end{aligned}$$

where

$$\begin{aligned} a_0 &= 1/\sqrt{\sigma_1\mu_2\Delta}, & b_0 &= 1/\sqrt{\sigma_2\mu_1\Delta}, \\ a_1 &= \sigma_1\sigma_2\mu_2, & b_1 &= (2\mu_1\sigma_2^2 + \sigma_1\sigma_2\mu_2)\sqrt{\sigma_2\mu_1}/\sqrt{\sigma_1\mu_2}, \\ a_2 &= \mu_2^2 + \mu_2\mu_1\sigma_2/2\sigma_1, & b_2 &= \sigma_2\mu_1\mu_2\sqrt{\sigma_2\mu_1}/2\sigma_1\sqrt{\sigma_1\mu_2}, \\ a_3 &= (2\sigma_1^2\mu_2 + \sigma_1\sigma_2\mu_1)\sqrt{\sigma_1\mu_2}/\sqrt{\sigma_2\mu_1}, & b_3 &= \sigma_1\sigma_2\mu_1, \\ a_4 &= \mu_1\mu_2\sigma_1\sqrt{\sigma_1\mu_2}/2\sigma_2\sqrt{\sigma_2\mu_1}, & b_4 &= \mu_1^2 + \sigma_1\mu_1\mu_2/2\sigma_2. \end{aligned}$$

PROOF. By (1.3), (4.5), (4.6), (4.15) and the Anscombe's Theorem

$$\begin{pmatrix} \sqrt{m^\star}(\bar{X}_L - \mu_1) \\ \sqrt{m^\star}(S^2_{1,L} - \sigma_1^2) \\ \sqrt{n^\star}(\bar{Y}_H - \mu_2) \\ \sqrt{n^\star}(S^2_{2,H} - \sigma_2^2) \end{pmatrix} \xrightarrow{\mathcal{D}} N(\underline{0}, \Sigma), \qquad \text{as} \quad d \to 0, \tag{4.25}$$

where Σ is defined as in (1.4). The required results now follow from routine calculations. □

LEMMA 4.7. *Assume that* $EX^{16} < \infty$, $EY^{16} < \infty$ *and* α *in* (2.1) *is such that* $3/4 < \alpha < 1$. *Then for* L *and* H *defined in* (4.16), *as* $d \to 0$

$$E\frac{a^2}{d^2}(S^2_{1,L} - \sigma_1^2) = a_0^4(a_1\mu_{3,X} + a_2(\mu_{4,X} - \sigma_1^4)) + o(1), \tag{4.26}$$

$$E\frac{a^2}{d^2}(S^2_{2,H} - \sigma_2^2) = -b_0^4(b_3\mu_{3,Y} + b_4(\mu_{4,Y} - \sigma_2^4)) + o(1), \tag{4.27}$$

$$E\frac{a^2}{d^2}(\bar{X}_L - \mu_1) = -a_0^4(a_1\sigma_1^2 + a_2\mu_{3,X}) + o(1), \tag{4.28}$$

$$E\frac{a^2}{d^2}(\bar{Y}_H - \mu_2) = -b_0^4(b_3\sigma_2^2 + b_4\mu_{3,Y}) + o(1), \tag{4.29}$$

where the constants $a_0, \ldots, a_4$ *and* $b_0, \ldots, b_4$ *are as defined in Lemma* 4.6. *Furthermore*

$$E\frac{a^2}{d^2}(S^2_{1,L} - \sigma_1^2)(S^2_{2,H} - \sigma_2^2) - E\frac{a^2}{d^2}(\bar{X}_L - \mu_1)(\bar{Y}_H - \mu_2) = o(1) \tag{4.30}$$

$$E\frac{a^2}{d^2}(S^2_{1,L} - \sigma_1^2)(\bar{Y}_H - \mu_2) = E\frac{a^2}{d^2}(S^2_{2,H} - \sigma_2^2)(\bar{X}_L - \mu_1) = o(1) \tag{4.31}$$

$$E\frac{a^2}{d^2}(S^2_{1,L} - \sigma_1^2)(\bar{X}_L - \mu_1) = \frac{\mu_{3,X}}{\sigma_1\mu_2\Delta} + o(1) \tag{4.32}$$

$$E\frac{a^2}{d^2}(S^2_{2,H} - \sigma_2^2)(\bar{Y}_H - \mu_2) = \frac{\mu_{3,Y}}{\sigma_2\mu_1\Delta} + o(1) \tag{4.33}$$

and the expectations of all the third and fourth order crossed-product terms are $o(1)$.

PROOF. We only prove (4.28) for $L = I$. The rest of the assertions can be proved similarly. Observe that

$$\frac{a^2}{d^2}(\bar{X}_I - \mu_1) = \frac{1}{\sigma_1\mu_2\Delta}\sum_{i=1}^{I}(X_i - \mu_1) + \frac{1}{\sigma_1\mu_2\Delta}\frac{\sum_{i=1}^{I}(X_i - \mu_1)}{\sqrt{m^\star}}\frac{(m^\star - I)}{\sqrt{m^\star}}\frac{m^\star}{I}.$$

By Wald's Lemma and Lemma 4.4 we have

$$E\frac{a^2}{d^2}(\bar{X}_I - \mu_1) = \frac{1}{\sigma_1\mu_2\Delta}E\frac{\sum_{i=1}^{I}(X_i - \mu_1)}{\sqrt{m^\star}}\frac{(m^\star - I)}{\sqrt{m^\star}}\frac{m^\star}{I} + o(1). \tag{4.34}$$

By the inequality $2|ab| \leq (a^2 + b^2)$, (4.18), Lemma 4.4 and Lemma 4.5

$$\left\{ \frac{\sum_{i=1}^{I}(X_i - \mu_1)}{\sqrt{m^\star}} \frac{(m^\star - I)}{\sqrt{m^\star}} \frac{m^\star}{I}, 0 < d < d_0 \right\} \tag{4.35}$$

is u.i. for some $d_0 > 0$. To compute the expectation in (4.34) it remains to find the limit distribution of the r.v. in (4.35). Since $I \leq M$, by (4.2)

$$\begin{aligned} \frac{(m^\star - I)}{\sqrt{m^\star}} &\geq \frac{(m^\star - M)}{\sqrt{m^\star}} \\ &\geq \frac{(m^\star - a^2 S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J}/d^2 - 1 - n_0 I_{[M=n_0]})}{\sqrt{m^\star}} \\ &= \frac{a}{d\sqrt{\sigma_1\mu_2\Delta}}(\sigma_1\mu_2\Delta - S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J}) \\ &\quad - \left(\frac{1}{\sqrt{m^\star}} + \frac{n_0}{\sqrt{m^\star}} I_{[M=n_0]}\right) \\ &= \frac{a}{d\sqrt{\sigma_1\mu_2\Delta}}(\sigma_1\mu_2\Delta - S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J}) + o_p(1). \end{aligned} \tag{4.36}$$

On the other hand, by (4.4)

$$\begin{aligned} \frac{(m^\star - I)}{\sqrt{m^\star}} &< \frac{m^\star - (N-1)S_{1,I}\bar{Y}_{N-1}/S_{2,N-1}\bar{X}_I}{\sqrt{m^\star}} \\ &\leq \frac{m^\star - \dfrac{S_{1,I}\bar{Y}_{N-1}}{S_{2,N-1}\bar{X}_I}\dfrac{a^2}{d^2}S_{2,N}\bar{X}_M\hat{\Delta}_{M,N}}{\sqrt{m^\star}} + \frac{S_{1,I}\bar{Y}_{N-1}}{S_{2,N-1}\bar{X}_I}\frac{1}{\sqrt{m^\star}} \\ &= \frac{m^\star - \dfrac{a^2}{d^2}S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}}{\sqrt{m^\star}} \\ &\quad + \frac{S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}}{\sigma_1\mu_2\Delta}\sqrt{m^\star}\left(1 - \frac{S_{2,N}\bar{X}_M}{S_{2,N-1}\bar{X}_I}\right) \\ &\quad + \frac{S_{1,I}\bar{Y}_{N-1}}{S_{2,N-1}\bar{X}_I}\frac{1}{\sqrt{m^\star}} \\ &= \frac{1}{\sigma_1\mu_2\Delta}\sqrt{m^\star}(\sigma_1\mu_2\Delta - S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}) + D + o_p(1). \end{aligned} \tag{4.37}$$

As for D in (4.37), write

$$\begin{aligned} D &= \frac{S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}}{\sigma_1\mu_2\Delta S_{2,N-1}\bar{X}_I}\sqrt{m^\star}(S_{2,N-1}\bar{X}_I - S_{2,N}\bar{X}_M) \\ &= D_1\sqrt{m^\star}[(S_{2,N-1} - \sigma_2)\mu_1 + (\bar{X}_I - \mu_1)\sigma_2 \\ &\qquad - (S_{2,N} - \sigma_2)\mu_1 - (\bar{X}_M - \mu_1)\sigma_2] \\ &\quad + o_p(1) \\ &= D_1\sqrt{m^\star}[(S_{2,N-1} - S_{2,N})\mu_1 + (\bar{X}_I - \bar{X}_M)\sigma_2] + o_p(1) \end{aligned} \tag{4.38}$$

where $D_1 = S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}/S_{2,N-1}\bar{X}_I\sigma_1\mu_2\Delta \overset{\text{a.s.}}{\to} 1/\sigma_2\mu_1$. By (4.5), (4.6), (4.15) and the Anscombe's Theorem, it can be shown that

$$D_1\sqrt{m^\star}[(S_{2,N-1} - S_{2,N})\mu_1 + (\bar{X}_I - \bar{X}_M)\sigma_2] = o_p(1). \tag{4.39}$$

Using similar arguments, one can show that

$$\begin{aligned}\frac{a}{d\sqrt{\sigma_1\mu_2\Delta}}&(\sigma_1\mu_2\Delta - S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J}) \\ &= \frac{1}{\sigma_1\mu_2\Delta}\sqrt{m^\star}(\sigma_1\mu_2\Delta - S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}) \\ &\quad + o_p(1).\end{aligned} \tag{4.40}$$

By (4.36), (4.37), (4.38), (4.39) and (4.40) we have that

$$\frac{m^\star - I}{\sqrt{m^\star}} = \frac{1}{\sigma_1\mu_2\Delta}\sqrt{m^\star}(\sigma_1\mu_2\Delta - S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}) + o_p(1). \tag{4.41}$$

Let

$$A = \frac{1}{\sigma_1\mu_2\Delta}\sqrt{m^\star}(\sigma_1\mu_2\Delta - S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N}). \tag{4.42}$$

From (4.41) and the first assertion in Lemma 4.6 it can be shown that

$$\begin{aligned}\frac{\sum_{i=1}^I(X_i - \mu_1)}{\sqrt{m^\star}}&\frac{(m^\star - I)}{\sqrt{m^\star}}\frac{m^\star}{I} \\ &= \sqrt{m^\star}(\bar{X}_I - \mu_1)A + o_p(1) \\ &= \frac{m^\star}{\sigma_1\mu_2\Delta}(\sigma_1\mu_2\Delta - S_{1,I}\bar{Y}_{N-1}\hat{\Delta}_{M,N})(\bar{X}_I - \mu_1) \\ &\quad + o_p(1) \\ &\overset{\mathcal{D}}{\to} -a_0^2 Z_1\sum_{i=1}^4 a_i Z_i,\end{aligned} \tag{4.43}$$

where $Z_i's$ are as defined in Lemma 4.6. Therefore, by (4.34), (4.35) and (4.43)

$$\begin{aligned}E\frac{a^2}{d^2}(\bar{X}_I - \mu_1) &= -a_0^4 a_1 EZ_1^2 - a_0^4 a_2 EZ_1Z_2 + o(1) \\ &= a_0^4(a_1\sigma_1^2 + a_2\mu_{3,X}) + o(1)\end{aligned}$$

as $d \to 0$. Hence the proof of (4.28). The rest of the assertions follow similarly with a fair amount of work. Hence the lemma. □

PROOF OF THEOREM 2.2. Recall that $T = M + N$. For (2.6), by the sampling scheme (2.1), stopping rule (2.2), (4.2) and (4.4)

$$\begin{aligned}\frac{a^2}{d^2}\hat{\Delta}_{M,N}^2 \le T \le \frac{a^2}{d^2}S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J} &+ \frac{a^2}{d^2}S_{2,N-1}\bar{X}_I\hat{\Delta}_{I,N-1} \\ &+ 2 + n_0(I_{[M=n_0]} + I_{[N=n_0]}).\end{aligned} \tag{4.44}$$

Write the left hand side of (4.44) as

$$\left(\frac{a^2}{d^2}\hat{\Delta}^2_{M,N} - t^\star\right)/\sqrt{t^\star} = (\hat{\Delta}_{M,N}/\Delta + 1)\frac{a}{d}(\hat{\Delta}_{M,N} - \Delta).$$

Note that

$$\hat{\Delta}_{M,N} \overset{\text{a.s.}}{\to} \Delta.$$

Moreover, by (4.25)

$$\begin{aligned}\frac{a}{d}(\hat{\Delta}_{M,N} - \Delta) &= \frac{a}{d}(S_{1,M} - \sigma_1)\mu_2 + \frac{a}{d}(S_{2,N} - \sigma_2)\mu_1 \\ &\quad + \frac{a}{d}(\bar{X}_M - \mu_1)\sigma_2 + \frac{a}{d}(\bar{Y}_N - \mu_2)\sigma_1 + o_p(1) \\ &= \frac{\mu_2\sqrt{m^\star}(S^2_{1,M} - \sigma_1^2)}{2\sigma_1\sqrt{\sigma_1\mu_2\Delta}} + \frac{\mu_1\sqrt{n^\star}(S^2_{2,N} - \sigma_2^2)}{2\sigma_2\sqrt{\sigma_2\mu_1\Delta}} \\ &\quad + \frac{\sigma_2\sqrt{m^\star}(\bar{X}_M - \mu_1)}{\sqrt{\sigma_1\mu_2\Delta}} + \frac{\sigma_1\sqrt{n^\star}(\bar{Y}_N - \mu_2)}{\sqrt{\sigma_2\mu_1\Delta}} + o_p(1) \\ &\overset{\mathcal{D}}{\to} \frac{\mu_2 Z_2}{2\sigma_1\sqrt{\sigma_1\mu_2\Delta}} + \frac{\mu_1 Z_4}{2\sigma_2\sqrt{\sigma_2\mu_1\Delta}} + \frac{\sigma_2 Z_1}{\sqrt{\sigma_1\mu_2\Delta}} + \frac{\sigma_1 Z_3}{\sqrt{\sigma_2\mu_1\Delta}} \\ &\sim N(0, \gamma^2/4),\end{aligned}$$

where γ^2 is as defined in (2.6). Hence

$$\frac{a^2}{d^2}(\hat{\Delta}^2_{M,N} - t^\star)/\sqrt{t^\star} \overset{\mathcal{D}}{\to} N(0, \gamma^2). \tag{4.45}$$

Similarly, we can show that for the right hand side of (4.44)

$$\begin{aligned}&\Bigg[\frac{a^2}{d^2}S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J} + \frac{a^2}{d^2}S_{2,N-1}\bar{X}_I\hat{\Delta}_{I,N-1} \\ &\qquad + 2 + n_0(I_{[M=n_0]} + I_{[N=n_0]}) - t^\star\Bigg]/\sqrt{t^\star} \\ &\overset{\mathcal{D}}{\to} N(0, \gamma^2).\end{aligned} \tag{4.46}$$

Hence, (2.6) follows from (4.44), (4.45) and (4.46).

The uniform integrability of $\{(T - t^\star)^2/t^\star, 0 < d < d_0\}$ follows from Lemma 4.5 since $T = M + N$ and $t^\star = m^\star + n^\star$. From (2.6), (2.7) and Theorem 4 of Billingsley (see Billingsley (1968) p. 32), we get (2.8). This completes the proof for Theorem 2.2. □

PROOF OF THEOREM 2.3. By (2.2) and (1.9)

$$\begin{aligned}E(T - m^\star - n^\star) &\geq \frac{a^2}{d^2}E(S_{1,M}\bar{Y}_N\hat{\Delta}_{M,N} - \sigma_1\mu_2\Delta) \\ &\quad + \frac{a^2}{d^2}E(S_{2,N}\bar{X}_M\hat{\Delta}_{M,N} - \sigma_2\mu_1\Delta) \\ &= (i) + (ii).\end{aligned} \tag{4.47}$$

As for (i) in (4.47), write

$$
\begin{aligned}
(i) &= \frac{a^2}{d^2}E[S_{1,M}^2\bar{Y}_N^2 + S_{1,M}\bar{Y}_N S_{2,N}\bar{X}_M - \sigma_1^2\mu_2^2 - \sigma_1\mu_2\sigma_2\mu_1] \\
&= \frac{a^2}{d^2}E(S_{1,M}^2 - \sigma_1^2)\mu_2^2 + \frac{a^2}{d^2}E(\bar{Y}_N^2 - \mu_2^2)\sigma_1^2 + \frac{a^2}{d^2}E(S_{1,M}^2 - \sigma_1^2)(\bar{Y}_N^2 - \mu_2^2) \\
&\quad + \frac{a^2}{d^2}E(S_{1,M} - \sigma_1)\mu_2\mu_1\sigma_2 + \frac{a^2}{d^2}E(S_{2,N} - \sigma_2)\sigma_1\mu_2\mu_1 \\
&\quad + \frac{a^2}{d^2}E(\bar{Y}_N - \mu_2)\sigma_1\sigma_2\mu_1 \\
&\quad + \frac{a^2}{d^2}E(\bar{X}_M - \mu_1)\sigma_1\sigma_2\mu_2 + \frac{a^2}{d^2}E(S_{1,M} - \sigma_1)(\bar{X}_M - \mu_1)\sigma_2\mu_2 \\
&\quad + \frac{a^2}{d^2}E(S_{2,N} - \sigma_2)(\bar{Y}_N - \mu_2)\sigma_1\mu_1 + o(1),
\end{aligned}
$$

where we used (4.30), (4.31) and the last line of Lemma 4.7. Use (4.25) and (4.29) of Lemma 4.7 to get

$$
\begin{aligned}
\frac{a^2}{d^2}E(\bar{Y}_N^2 - \mu_2^2) &= 2\mu_2\frac{a^2}{d^2}E(\bar{Y}_N - \mu_2) + \frac{a^2}{d^2}E(\bar{Y}_N - \mu_2)^2 \\
&= -2\mu_2 b_0^4(b_3\sigma_2^2 + b_4\mu_{3,Y}) + \frac{\sigma_2^2}{\sigma_2\mu_1\Delta} + o(1).
\end{aligned}
$$

Similarly, by (4.25) and (4.26)

$$
\begin{aligned}
\frac{a^2}{d^2}E(S_{1,M} - \sigma_1) &= \frac{1}{2\sigma_1}\frac{a^2}{d^2}E(S_{1,M}^2 - \sigma_1^2) - \frac{1}{4\sigma_1^3}\frac{a^2}{d^2}E(S_{1,M}^2 - \sigma_1^2)^2 + o(1) \\
&= -\frac{a_0^4[a_1\mu_{3,X} + a_2(\mu_{4,X} - \sigma_1^4)]}{2\sigma_1} \\
&\quad - \frac{1}{4\sigma_1^4\mu_2\Delta}E[\sqrt{m^\star}(S_{1,M}^2 - \sigma_1^2)]^2 + o(1) \\
&= -\frac{a_0^4[a_1\mu_{3,X} + a_2(\mu_{4,X} - \sigma_1^4)]}{2\sigma_1} - \frac{(\mu_{4,X} - \sigma_1^4)}{4\sigma_1^4\mu_2\Delta} + o(1),
\end{aligned}
$$

and by (4.32)

$$
\begin{aligned}
\frac{a^2}{d^2}E(S_{1,M} - \sigma_1)(\bar{X}_M - \mu_1) &= \frac{1}{2\sigma_1}\frac{a^2}{d^2}E(S_{1,M}^2 - \sigma_1^2)(\bar{X}_M - \mu_1) + o(1) \\
&= \frac{\mu_{3,X}}{2\sigma_1^2\mu_2\Delta} + o(1).
\end{aligned}
$$

Similarly, by (4.27) and (4.33)

$$
\frac{a^2}{d^2}E(S_{2,N} - \sigma_2) = -\frac{b_0^4[b_3\mu_{3,Y} + b_4(\mu_{4,Y} - \sigma_2^4)]}{2\sigma_2} - \frac{(\mu_{4,Y} - \sigma_2^4)}{4\sigma_2^4\mu_1\Delta} + o(1)
$$

and

$$\frac{a^2}{d^2}E(S_{2,N}-\sigma_2)(\bar{Y}_N-\mu_2)=\frac{\mu_{3,Y}}{2\sigma_2^2\mu_1\Delta}+o(1).$$

Collecting the terms from above

$$\begin{aligned}(i) = &-\mu_2^2a_0^4[a_1\mu_{3,X}+a_2(\mu_{4,X}-\sigma_1^4)]-2\mu_2\sigma_1^2b_0^4(b_3\sigma_2^2+b_4\mu_{3,X})+\frac{\sigma_1^2\sigma_2^2}{\sigma_2\mu_1\Delta}\\ &-\mu_1\mu_2\sigma_2\left[\frac{a_0^4[a_1\mu_{3,X}+a_2(\mu_{4,X}-\sigma_1^4)]}{2\sigma_1}+\frac{(\mu_{4,X}-\sigma_1^4)}{4\sigma_1^4\mu_2\Delta}\right]\\ &-\mu_1\mu_2\sigma_1\left[\frac{b_0^4[b_3\mu_{3,Y}+b_4(\mu_{4,Y}-\sigma_2^4)]}{2\sigma_2}+\frac{(\mu_{4,Y}-\sigma_2^4)}{4\sigma_2^4\mu_1\Delta}\right]\\ &-\mu_1\sigma_1\sigma_2b_0^4(b_3\sigma_2^2+b_4\mu_{3,Y})-\mu_2\sigma_1\sigma_2a_0^4(a_1\sigma_1^2+a_2\mu_{3,X})+\sigma_2\mu_2\frac{\mu_{3,X}}{2\sigma_1^2\mu_2\Delta}\\ &+\sigma_1\mu_1\frac{\mu_{3,Y}}{2\sigma_2^2\mu_1\Delta}+o(1).\end{aligned}$$

Similar arguments yield

$$\begin{aligned}(ii) = &-\mu_1^2b_0^4[b_1\mu_{3,Y}+b_2(\mu_{4,Y}-\sigma_2^4)]-2\mu_1\sigma_2^2a_0^4(a_1\sigma_1^2+a_2\mu_{3,X})+\frac{\sigma_1^2\sigma_2^2}{\sigma_1\mu_2\Delta}\\ &-\mu_1\mu_2\sigma_2\left[\frac{a_0^4[a_1\mu_{3,X}+a_2(\mu_{4,X}-\sigma_1^4)]}{2\sigma_1}+\frac{(\mu_{4,X}-\sigma_1^4)}{4\sigma_1^4\mu_2\Delta}\right]\\ &-\mu_1\mu_2\sigma_1\left[\frac{b_0^4[b_3\mu_{3,Y}+b_4(\mu_{4,Y}-\sigma_2^4)]}{2\sigma_2}+\frac{(\mu_{4,Y}-\sigma_2^4)}{4\sigma_2^4\mu_1\Delta}\right]\\ &-\mu_1\sigma_1\sigma_2b_0^4(b_3\sigma_2^2+b_4\mu_{3,Y})-\mu_2\sigma_1\sigma_2a_0^4(a_1\sigma_1^2+a_2\mu_{3,X})+\sigma_2\mu_2\frac{\mu_{3,X}}{2\sigma_1^2\mu_2\Delta}\\ &+\sigma_1\mu_1\frac{\mu_{3,Y}}{2\sigma_2^2\mu_1\Delta}+o(1).\end{aligned}$$

Hence by (4.47) and calculation

$$\begin{aligned}(4.48)\qquad E(T-m^\star-n^\star)\geq &-\frac{\sigma_2(\Delta+\sigma_1\mu_2)}{\sigma_1^3\mu_2\Delta}\mu_{3,X}-\frac{\sigma_1(\Delta+\sigma_2\mu_1)}{\sigma_2^3\mu_1\Delta}\mu_{3,Y}\\ &-\frac{\mu_{4,X}}{\sigma_1^4}-\frac{\mu_{4,Y}}{\sigma_2^4}+2-\frac{\sigma_1\sigma_2}{\mu_1\mu_2}+o(1).\end{aligned}$$

For the reverse inequality, use (4.44) and similar arguments with $(S_{1,M-1}\bar{Y}_J\hat{\Delta}_{M-1,J}-\sigma_1\mu_2\Delta)$ and $(S_{2,N-1}\bar{X}_I\hat{\Delta}_{I,N-1}-\sigma_2\mu_1\Delta)$ to show that

$$\begin{aligned}(4.49)\qquad E(T-m^\star-n^\star)\leq &-\frac{\sigma_2(\Delta+\sigma_1\mu_2)}{\sigma_1^3\mu_2\Delta}\mu_{3,X}-\frac{\sigma_1(\Delta+\sigma_2\mu_1)}{\sigma_2^3\mu_1\Delta}\mu_{3,Y}\\ &-\frac{\mu_{4,X}}{\sigma_1^4}-\frac{\mu_{4,Y}}{\sigma_2^4}+2-\frac{\sigma_1\sigma_2}{\mu_1\mu_2}+2+o(1).\end{aligned}$$

From (4.48) and (4.49) we have (2.9). Hence the theorem. □

References

Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors, *J. Amer. Statist. Assoc.*, **84**, 200–207.

Berry, D. (1977). Optimal sampling schemes for estimating system reliability by testing components-I: fixed sample size, *J. Amer. Statist. Assoc.*, **69**, 485–491.

Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

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.

Chow, Y. S. and Teicher, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales*, Springer, New York.

Chow, Y. S. and Yu, K. F. (1981). The performance of a sequential procedure for the estimation of the mean, *Ann. Statist.*, **9** 184–189.

Chow, Y. S., Hsiung, C. A. and Lai, T. L. (1979). Extended renewal theory and moment convergence in Anscome's theorem, *Ann. Probab.*, **7**, 304–318.

Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin, Boston.

Ghosh, M. and Mukhopadhyay, N. (1980). Sequential point estimation of the difference of two normal means, *Ann. Statist.*, **8**, 221–225.

Hardwick, J. and Stout, Q. (1992). Optimal allocation for estimating the product of two means, *Proceeding of the 23rd Symposium on the Interface: Computing Science and Statistics*, 592–596, Interface Foundation of North American, Fairfax.

Harris, B. (1971). Hypothesis testing and confidence intervals for products and quotients of Poisson parameters with applications to reliability, *J. Amer. Statist. Assoc.*, **66**, 609–613.

Martinsek, A. T. (1983). Second order approximation to the risk of a sequential procedure, *Ann. Statist.*, **11**, 827–836.

Martinsek, A. T. (1990). Sequential point estimation in regression models with nonnormal errors, *Sequential Anal.*, **9**, 243–268.

Mukhopadhyay, N. (1976). Sequential estimation of a linear function of means of three normal populations, *J. Amer. Statist. Assoc.*, **71**, 149–153.

Mukhopadhyay, N. and Liberman, S. (1989). Sequential estimation of a linear function of mean vectors, *Sequential Anal.*, **8**, 381–394.

Mukhopadhyay, N. and Sriram, T. N. (1992). On sequential comparisons of means of first-order autoregressive models, *Metrika*, **39**, 155–164.

Noble, W. (1992). First order allocation, Doctoral Dissertation, Department of Statistics and Probability, Michigan State University.

Page, C. (1985). Allocation schemes for estimating the product of positive parameters, *J. Amer. Statist. Assoc.*, **80**, 449–454.

Page, C. (1987). Sequential designs for estimating products of parameters, *Sequential Anal.*, **6**, 351–371.

Page, C. (1990). Allocation proportional to coefficients of variation when estimating the product of parameters, *J. Amer. Statist. Assoc.*, **85**, 1134–1139.

Page, C. (1995). Adaptive Allocation for estimation, *IMS Lecture Notes–Monograph Series*, **25**, 213–222.

Rekab, K. (1989). Asymptotic efficiency in sequential designs for estimation, *Sequential Anal.*, **8**, 269–280.

Robbins, H., Simons, G. and Starr, N. (1967). A sequential analogue of the Behrens-Fisher problem. *Ann. Math. Statist.*, **38**, 1384–1391.

Southwood, T. (1978). *Ecological Methods with Particular Reference to the Study of Insect Populations*, Chapman and Hall, London.

Sriram, T. N. (1991). Second order approximation to the risk of a sequential procedure measured under squared relative error loss, *Statistics and Decisions*, **9**, 375–392.

Sriram, T. N. (1992). An improved sequential procedure for estimating the regression parameter in regression models with symmetric errors, *Ann. Statist.*, **20**, 1441–1453.

Srivastava, M. S. (1970). On a sequential analogue of Behrens-Fisher problem, *J. Roy. Statist. Soc. Ser. B*, **32**, 144–148.

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

Stout, W. F. (1974). *Almost Sure Convergence*, Academic Press, New York.

Sun, D. and Ye, K. (1995). Reference prior Bayesian analysis for normal mean products, *J. Amer. Statist. Assoc.*, **90**, 589–597.

Takada, Y. (1992). A sequential procedure with asymptotically negative regret for estimating the normal mean, *Ann. Statist.*, **20**, 562–569.

Yfantis, E. A. and Flatman, G. T. (1991). On confidence interval for the product of three means of three normally distributed populations, *Journal of Chemometrics*, **5**, 309–319.

Zheng, S., Seila, A. F. and Sriram, T. N. (1996*a*). Asymptotically risk efficient two stage procedure for estimating the product of two means, Tech. Report No. 96-22, Department of Statistics, University of Georgia.

Zheng, S., Seila, A. F. and Sriram, T. N. (1996*b*). Asymptotically risk efficient two stage procedure for estimating the product of $k(\geq 2)$ means, *Statistics and Decisions* (under revision).