Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 1584–1605
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Exact solutions to the Behrens–Fisher Problem: Asymptotically optimal and finite sample efficient choice among☆

Edward J. Dudewicz[a,*], Yan Ma[b], Enping (Shirley) Mai[c], Haiyan Su[b]

[a]Department of Mathematics, Syracuse University, 215 Carnegie Hall, Syracuse, NY 13244-1150, USA
[b]Department of Biostatistics and Computational Biology, University of Rochester, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, USA
[c]Whitman School of Management, Syracuse University, 721 University Avenue, Syracuse, NY 13244-2450, USA

Available online 10 October 2006

Abstract

The problem of testing the equality of two normal means when variances are not known is called the Behrens–Fisher Problem. This problem has three known exact solutions, due, respectively, to Chapman, to Prokof'yev and Shishkin, and to Dudewicz and Ahmed. Each procedure has level alpha and power beta when the means differ by a given amount delta, both set by the experimenter. No single-sample statistical procedures can make this guarantee. The most recent of the three procedures, that of Dudewicz and Ahmed, is asymptotically optimal. We review the procedures, and then compare them with respect to both asymptotic efficiency and also (using simulation) in finite samples. Of these exact procedures, based on finite-sample comparisons the Dudewicz–Ahmed procedure is recommended for practical use.

Keywords: Behrens–Fisher Problem; Exact solutions; Heteroscedasticity; Tests of hypotheses; Testing equality of means; Exact level tests; Tests with specified power; Asymptotically optimal tests; Finite-sample efficiency; Prokof'yev–Shishkin procedure; Chapman procedure; Dudewicz–Ahmed procedure; Comparisons; Recommendation for practical use

1. Introduction

The problem of *testing the* (*null*) *hypothesis* $\mathrm{H}_0 : \mu_1 = \mu_2$, *for two normal distributions when the means* μ_1, μ_2 *and the variances* σ_1^2, σ_2^2 *are all unknown* and hence possibly unequal, is called the *Behrens–Fisher Problem*. The Behrens–Fisher Problem dates back to early 20th century work of the astronomer Behrens in 1929 and the statistician Fisher in 1935. Formally, let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be two independent streams of independent random variables, where $X_1, X_2, \ldots$ are each $\mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_1, Y_2 \ldots$ are each $\mathrm{N}(\mu_2, \sigma_2^2)$. For the Behrens–Fisher Problem of testing the hypothesis $\mathrm{H}_0 : \mu_1 = \mu_2$ when the variances, σ_1^2, σ_2^2 are not known (and thus possibly unequal), for many years no one could find an exact solution, and indeed Linnik proved that "this problem has no solution" (see Bather, 1996, p. 337). Thus, *many different approximate solutions were developed* such as those of Welch, Neyman and Bartlett, Scheffé, Hsu

☆ This paper was the subject of an Invited Plenary Talk by the first author at the International Conference on Statistics, Combinatorics and Related Areas SCRA (10th International Conference of the Forum for Interdisciplinary Mathematics) held at the University of Southern Maine, Portland, Maine, USA.

* Corresponding author.
E-mail addresses: dudewicz@syr.edu (E.J. Dudewicz), yan_ma@urmc.rochester.edu (Y. Ma), emai@syr.edu (E. Mai), haiyan_su@urmc.rochester.edu (H. Su).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2006.09.007

and others. One of the best of these is the solution due to Hsu and Scheffé which acts as if, in random samples of sizes n and m, the statistic

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{s_1^2}{n}+\dfrac{s_2^2}{m}}}$$

has Student's-t distribution with $\min(n, m) - 1$ degrees of freedom and is conservative. However, that solution did not become available until a 1970 paper by Scheffé. Even though it was featured in a junior–senior-graduate level text shortly afterwards (Dudewicz, 1976), and later in an expanded comprehensive form (Dudewicz and Mishra, 1998), that conservative solution has only recently become widely available at the most elementary level (Dudewicz et al., 1989, Moore and McCabe, 1999). Until recently, it seems it was not widely realized that *there are in fact exact solutions. Those are described in Section* 2.

2. Exact solutions of the Behrens–Fisher Problem

Suppose we face the Behrens–Fisher Problem, i.e., for two independent normal distributions $\mathrm{N}(\mu_1, \sigma_1^2)$ and $\mathrm{N}(\mu_2, \sigma_2^2)$ we wish to perform a level α test of the null-hypothesis $\mathrm{H}_0 : \mu_1 = \mu_2$ against the alternative $\mathrm{H}_1 : \mu_1 \neq \mu_2$ (where σ_1^2, σ_2^2 are unknown) such that when $\Delta = \Delta_1$ then $\beta = \beta_1$, where Δ is the absolute difference between the two means, β is the power function of the test, and $\Delta_1 > 0$ and $0 < \beta_1 < 1$ are specified numbers. We now describe *the three procedures that are exact solutions*. In all three procedures *we use the following notation*: $[x]$ denotes the smallest integer not smaller than x; $F_{n_0}(\cdot)$ and $f_{n_0}(\cdot)$ are, respectively, the cumulative distribution function and probability density function of a Student's-t random variable with $n_0 - 1$ degrees of freedom.

2.1. *Dudewicz and Ahmed's (1998) procedure*

The Dudewicz and Ahmed (l998) procedure D–A is both exact *and asymptotically optimal*.

Take initial samples of size n_0, $X_1, X_2, \ldots, X_{n_0}$ and $Y_1, Y_2, \ldots, Y_{n_0}$ with $n_0 \geqslant 2$ from each population, and define

$$\overline{X}_0 = (X_1 + X_2 + \cdots + X_{n_0})/n_0, \quad \overline{Y}_0 = (Y_1 + Y_2 + \cdots + Y_{n_0})/n_0, \tag{1}$$

$$S_1^2 = \sum_{i=1}^{n_0} (X_i - \overline{X}_0)^2/v, \quad S_2^2 = \sum_{i=1}^{n_0} (Y_i - \overline{Y}_0)^2/v, \quad v = n_0 - 1. \tag{2}$$

Calculate

$$N_i = \max\{n_0 + 1, [c^2 S_i(S_1 + S_2)]\}, \quad i = 1, 2, \tag{3}$$

where c is a positive constant to be specified later. In a second stage, observe the random variables $X_{n_0+1}, X_{n_0+2}, \ldots, X_{N_1}$ and $Y_{n_0+1}, \ldots, Y_{N_2}$. Next, calculate

$$\tilde{\tilde{X}} = a_1 X_1 + \cdots + a_{N_1} X_{N_1}, \quad \tilde{\tilde{Y}} = b_1 Y_1 + \cdots + b_{N_2} Y_{N_2}, \tag{4}$$

where the coefficients are chosen to satisfy the constraints:

$$\text{(i)}\ a_1 = \cdots = a_{n_0}, \quad b_1 = \cdots = b_{n_0}. \tag{5}$$

$$\text{(ii)}\ a_1 + \cdots + a_{N_1} = 1, \quad b_1 + \cdots + b_{N_2} = 1. \tag{6}$$

$$\text{(iii)}\ S_1(S_1 + S_2) \sum_{i=1}^{N_1} a_i^2 = \frac{1}{c^2} = S_2(S_1 + S_2) \sum_{i=1}^{N_2} b_i^2. \tag{7}$$

(That such coefficients exist is shown in Aoshima et al. (1996); in particular, a specific solution is demonstrated in their Remark 3.1, p. 65.)

The following is known about the distribution of $\tilde{\tilde{X}} - \tilde{\tilde{Y}}$.

Theorem. *Let*

$$U = \frac{(\tilde{\tilde{X}} - \tilde{\tilde{Y}}) - (\mu_1 - \mu_2)}{1/c},$$

and denote $F_U(z) = P(U \leqslant z)$. *Then* $F_U(z) = g_{n_0}(z|\sigma_2/\sigma_1)$ *is a function only of* n_0 *and* σ_2/σ_1 (*it is not a function of the constant* c).

The D–A solution of the Behrens–Fisher Problem then *rejects* H_0 *if and only if* $|\tilde{\tilde{X}} - \tilde{\tilde{Y}}| > K$. The *level* of the D–A procedure is

$$\begin{aligned} P_{\mathrm{H}_0}(|\tilde{\tilde{X}} - \tilde{\tilde{Y}}| > K) &= 1 - P(-K \leqslant \tilde{\tilde{X}} - \tilde{\tilde{Y}} \leqslant K) = 1 - \{2g_{n_0}(cK|\sigma_2/\sigma_1) - 1\} \\ &= 2\{1 - g_{n_0}(cK|\sigma_2/\sigma_1)\} = \alpha \end{aligned} \tag{8}$$

if (for a given σ_2/σ_1), $g_{n_0}(cK|\sigma_2/\sigma_1) = 1 - \alpha/2$.

Now define the quantity $h_{n_0}(\alpha)$ as the solution of the equation

$$\inf\{g_{n_0}(h|\sigma_2/\sigma_1) : 0 < \sigma_2/\sigma_1 < \infty\} = 1 - \alpha/2. \tag{9}$$

Then the D–A test has level α if we choose K so that

$$cK = h, \tag{10}$$

i.e., so that

$$K = h/c. \tag{11}$$

Note that c is chosen to control the *power* of the test. Also note that $N_i \geqslant n_0 + 1$, which means at least one sample will be taken in stage 2.

2.2. *Chapman's (1950) procedure*

Take initial samples $X_1, X_2, \ldots, X_{n_0}$ and $Y_1, Y_2, \ldots, Y_{n_0}$ (both of size $n_0 \geqslant 2$) from the normal populations $\mathrm{N}(\mu_1, \sigma_1^2)$ and $\mathrm{N}(\mu_2, \sigma_2^2)$ respectively, and calculate

$$\overline{X}(n_0) = \frac{1}{n_0}\sum_{i=1}^{n_0} X_i, \quad \overline{Y}(n_0) = \frac{1}{n_0}\sum_{i=1}^{n_0} Y_i, \tag{12}$$

$$S_1^2 = \frac{1}{n_0 - 1}\sum_{i=1}^{n_0}(X_i - \overline{X}_{n_0})^2, \quad S_2^2 = \frac{1}{n_0 - 1}\sum_{i=1}^{n_0}(Y_i - \overline{Y}_{n_0})^2, \tag{13}$$

$$n_1 = \max(n_0 + 1, [S_1^2/h^2]), \tag{14}$$

$$n_2 = \max(n_0 + 1, [S_2^2/h^2]), \tag{15}$$

where $h > 0$ is to be chosen in such a way that, when $\varDelta = \varDelta_1$, we have $\beta = \beta_1$.

Take $n_1 - n_0$ additional observations $X_{n_0+1}, X_{n_0+2}, \ldots, X_{n_1}$ from $\mathrm{N}(\mu_1, \sigma_1^2)$, and take $n_2 - n_0$ additional observations $Y_{n_0+1}, Y_{n_0+2}, \ldots, Y_{n_2}$ from $\mathrm{N}(\mu_2, \sigma_2^2)$, and then calculate

$$\overline{X}(n_1 - n_0) = \frac{1}{n_1 - n_0}\sum_{i=n_0+1}^{n_1} X_i, \tag{16}$$

$$\overline{Y}(n_2 - n_0) = \frac{1}{n_2 - n_0}\sum_{i=n_0+1}^{n_2} Y_i, \tag{17}$$

$$\tilde{\tilde{X}} = b_1\overline{X}(n_0) + b_2\overline{X}(n_1 - n_0), \tag{18}$$

$$\tilde{\tilde{Y}} = c_1\overline{Y}(n_0) + c_2\overline{Y}(n_2 - n_0), \tag{19}$$

where

$$b_1 = \frac{n_0}{n_1}\left(1 + \left(1 - \frac{n_1}{n_0}\left(1 - \frac{n_1 - n_0}{S_1^2/h^2}\right)\right)^{1/2}\right), \tag{20}$$

$$b_2 = 1 - b_1, \tag{21}$$

$$c_1 = \frac{n_0}{n_2}\left(1 + \left(1 - \frac{n_2}{n_0}\left(1 - \frac{n_2 - n_0}{S_2^2/h^2}\right)\right)^{1/2}\right), \tag{22}$$

$$c_2 = 1 - c_1. \tag{23}$$

The C solution of the Behrens–Fisher Problem then *rejects* $H_0 : \mu_1 = \mu_2$ *in favor of* $H_1 : \mu_1 \neq \mu_2$ *if and only if* $|\tilde{\tilde{X}} - \tilde{\tilde{Y}}| > c_{1-\alpha/2}(n_0)/1/h$ where $c_{1-\gamma}(n_0)$ is the value of c such that

$$\int_{-\infty}^{\infty} F_{n_0}^{k-1}(z + c) f_{n_0}(z)\,dz = 1 - \gamma, \quad k = 2, \tag{24}$$

and $h > 0$ is chosen such that when $\Delta = \Delta_1$ we have $\beta = \beta_1$. Note that the total number of observations here (let us denote it by N) is $n_1 + n_2$, which is also a random variable.

Choice of h in Chapman's procedure: Denoting the integral $\int_{-\infty}^{\infty} F_{n_0}^{k-1}(z+h) f_{n_0}(z)\,dz$ by $P_{n_0}(h)$ the power function of the test is

$$\beta\left(\frac{\Delta}{h}\right) = P_{n_0}\left(-c_{1-\alpha/2}(n_0) - \frac{\Delta}{h}\right) + P_{n_0}\left(-c_{1-\alpha/2}(n_0) + \frac{\Delta}{h}\right). \tag{25}$$

For level α, $c_{1-\alpha/2}(n_0)$ is obtained from the integral equation

$$\int_{-\infty}^{\infty} F_{n_0}^{k-1}(z + c) f_{n_0}(z)\,dz = 1 - \frac{\alpha}{2}. \tag{26}$$

Then for a desired value β_1 for $\beta(\Delta/h)$, the equation

$$P_{n_0}\left(-c_{1-\alpha/2}(n_0) - \frac{\Delta}{h}\right) + P_{n_0}\left(-c_{1-\alpha/2}(n_0) + \frac{\Delta}{h}\right) = \beta_1 \tag{27}$$

is solved for Δ/h. Finally, by using $\Delta_1$for Δ in Δ/h, we obtain h.

For $x < 0$, we use the relation $P_{n_0}(-x) = 1 - P_{n_0}(x)$ to find $P_{n_0}(x)$.

2.3. Prokof'yev and Shishkin (1974) procedure

Stage 1: Take $Z_1, \ldots, Z_{n_0}$ of size $n_0 \geqslant 2$, where $Z_i = X_i - Y_i$, and compute

$$\overline{Z}(n_0) = \frac{\sum_{i=1}^{n_0} Z_i}{n_0}, \tag{28}$$

$$S^2 = \frac{\sum (Z_i - \overline{Z}(n_0))^2}{n_0 - 1}, \tag{29}$$

$$n = \max(n_0, [S^2/d]). \tag{30}$$

Note: $d > 0$ is chosen such that $\beta(\Delta/\sqrt{d}) = \int_{\infty}^{-L-\Delta/\sqrt{d}} f_{n_0}(x)\,dx + \int_{L-\Delta/\sqrt{d}}^{\infty} f_{n_0}(x)\,dx$, where $\Delta = |\mu_1 - \mu_2|$, and L is found by solving $\alpha/2 = \int_L^{\infty} f_{n_0}(x)\,dx$ for given β (e.g., $\beta = 0.99$).

Stage 2: Take additional samples $Z_{n_0+1}, \ldots, Z_n$ and calculate

$$\overline{Z} = \frac{\sum_{i=1}^{n} Z_i}{n}, \tag{31}$$

$$T = \frac{\sqrt{n}\overline{Z}}{S}. \tag{32}$$

Note that n could be equal to n_0; in this case, no additional samples will be needed in stage 2.

The P–S solution of the Behrens–Fisher Problem then *rejects* $\mathrm{H}_0 : \mu_1 = \mu_2$ *if and only if* $|T| > L$, where L is such that $\alpha/2 = \int_L^\infty f_{n_0}(x)\,\mathrm{d}x$.

For example, if $n_0=10$, in stage 1: $Z_1=X_1-Y_1, \ldots, Z_{10}=X_{10}-Y_{10}$. Determine $n=\max(n_0, [S^2/d])=(10, [S^2/d])$. Note that n increases as d decreases, if $[S^2/d]$ is greater than 10. To get the power of test when $|\mu_1-\mu_2|=\Delta$, we need $\beta(\Delta/\sqrt{d}) = P_{\text{when } |\mu_1-\mu_2|=\Delta}$ (test rejects $\mathrm{H}_0 : \mu_1 = \mu_2$) (see Taneja and Dudewicz, 1993, p. 450).

In Section 3 we asymptotically compare the three exact solutions given above; this leads to a definite preference for the D–A procedure. Then in Section 4 we compare the procedures for non-asymptotic sample sizes, giving the comprehensive performance comparison needed to make definitive preference recommendations; this also leads to a definite preference for the D–A procedure. As simulation methods are used in Section 4, the random number generators (rng's) used are also carefully considered and documented there.

3. Asymptotic comparisons of the three exact solutions of the Behrens–Fisher Problem

3.1. *Comparison of Dudewicz–Ahmed procedure to the optimal when σ_1, σ_2 known*

In previous work, Dudewicz and Ahmed (1999, p. 173) compared their procedure to the optimal in terms of asymptotic sample size needed and found that

$$\text{As } \Delta_1 \to 0, \quad \frac{N_1+N_2}{n_1^*+n_2^*} \to 1 \text{ almost surely} \tag{33}$$

and

$$\text{As } \Delta_1 \to 0, \quad \frac{E(N_1+N_2)}{n_1^*+n_2^*} \to 1, \tag{34}$$

where N_1, N_2 are the Dudewicz–Ahmed procedure's sample sizes, and n_1^*, n_2^* are the optimal fixed sample sizes. From (33) and (34) *we see that the Dudewicz–Ahmed procedure is asymptotically optimal*.

3.2. *The optimal when σ_1, σ_2 known*

A brief *synopsis of the* "*optimal*" (needed for our further work below) is as follows.

Suppose $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are two independent streams of independent random variables, and we observe $X_1, \ldots, X_{n_1}$ each $\mathrm{N}(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_{n_2}$ each $\mathrm{N}(\mu_2, \sigma_2^2)$ with σ_1^2 and σ_2^2 both known, then

$$\text{reject } \mathrm{H}_0 : \mu_1 = \mu_2 \text{ if and only if } |\overline{X} - \overline{Y}| > K_1. \tag{35}$$

The level of the test with critical region (35) is α if

$$1-\alpha = P_{\mathrm{H}_0}(-K_1 \leqslant \overline{X} - \overline{Y} \leqslant K_1) = 2\Phi(K_1/(\sigma_1^2/n_1+\sigma_2^2/n_2)^{0.5}) - 1.$$

The power of the test with critical region (35) is, when $|\mu_1-\mu_2| = \Delta_1$,

$$\begin{aligned} &P_{\Delta_1}(|\overline{X} - \overline{Y}| > K_1) \\ &\quad = 2 - \Phi(\Phi^{-1}(1-\alpha/2) + \Delta_1/(\sigma_1^2/n_1+\sigma_2^2/n_2)^{0.5}) - \Phi(\Phi^{-1}(1-\alpha/2) - \Delta_1/(\sigma_1^2/n_1+\sigma_2^2/n_2)^{0.5}). \end{aligned} \tag{36}$$

If the total sample size is fixed at $n_1 + n_2 = n$, the *best* allocation of the sample size is that which maximizes the power (36):

$$n_1 = \frac{\sigma_1}{\sigma_1+\sigma_2} n, \quad n_2 = \frac{\sigma_2}{\sigma_1+\sigma_2} n. \tag{37}$$

The optimal power in fact equals

$$\Phi\left(\Phi^{-1}(1-\alpha/2) + \frac{\sqrt{n}\Delta_1}{\sigma_1+\sigma_2}\right) - \Phi\left(\Phi^{-1}(1-\alpha/2) - \frac{\sqrt{n}\Delta_1}{\sigma_1+\sigma_2}\right).$$

Let

$$\frac{\sqrt{n}\Delta_1}{\sigma_1+\sigma_2} = Q_1. \tag{38}$$

Then

$$\beta_1 = 2 - \Phi(\Phi^{-1}(1-\alpha/2) + Q_1) - \Phi(\Phi^{-1}(1-\alpha/2) - Q_1). \tag{39}$$

Denote the optimal fixed sample sizes by n_1^*, n_2^*; from (37) and (38) we have

$$n_i^* = \frac{Q_1^2}{\Delta_1^2}\sigma_i(\sigma_1+\sigma_2), \quad i = 1, 2, \tag{40}$$

where Q_1 solves (39).

3.3. Comparison of P–S procedure to the optimal when σ_1, σ_2 known

The P–S procedure (see Section 2) rejects $H_0 : \mu_1 = \mu_2$if and only if $|T| > L$ where L solves $\alpha/2 = \int_L^\infty f_{n_0}(x)\,dx = P(X > L) = 1 - P(X \leqslant L) = 1 - \int_{-\infty}^L f_{n_0}(x)\,dx = 1 - F_{n_0}(L)$, so $F_{n_0} = 1 - \alpha/2$ and hence $L = F_{n_0}^{-1}(1 - \alpha/2)$.

When $\Delta = \Delta_1$, $\beta = \beta_1$, the power function of the P–S test is (Taneja and Dudewicz, 1993, p. 450)

$$\begin{aligned}\beta_1\left(\frac{\Delta_1}{\sqrt{d}}\right) &= \int_{-\infty}^{-L-\Delta_1/\sqrt{d}} f_{n_0}(x)\,dx + \int_{L-\Delta_1/\sqrt{d}}^{\infty} f_{n_0}(x)\,dx \\ &= F_{n_0}(-L-\Delta_1/\sqrt{d}) + (1 - F_{n_0}(L-\Delta_1/\sqrt{d})) \\ &= 2 - F_{n_0}(L+\Delta_1/\sqrt{d}) - F_{n_0}(L-\Delta_1/\sqrt{d}) \\ &= 2 - F_{n_0}(F_{n_0}^{-1}(1-\alpha/2)+\Delta_1/\sqrt{d}) - F_{n_0}(F_{n_0}^{-1}(1-\alpha/2)-\Delta_1/\sqrt{d}).\end{aligned} \tag{41}$$

Now suppose the initial sample size $n_0 = n_0(\Delta_1)$ is looked at when chosen in a sequence such that $n_0(\Delta_1) \to \infty$ with $\Delta_1^2 n_0(\Delta_1) \to 0$, as $\Delta_1 \to 0$. Let $\Delta_1/\sqrt{d} = Q$; then since $F_{n_0}(x) \to \Phi(x)$ asymptotically as $n_0 \to \infty$ where $\Phi(x)$ is the c.d.f. of the standard normal distribution, (41) in the limit as $n_0 \to \infty$ becomes $\beta_1 = 2 - \Phi(\Phi^{-1}(1-\alpha/2)+Q) - \Phi(\Phi^{-1}(1-\alpha/2)-Q)$, which is the same as Eq. (39) with Q_1 replaced by Q. Hence we must have $Q = Q(n_0) \to Q_1$ as $n_0 \to \infty$. It follows from (40) that

$$n_1^* + n_2^* = \frac{Q_1^2}{\Delta_1^2}(\sigma_1+\sigma_2)^2, \tag{42}$$

where n_1^*, n_2^* are the optimal fixed sample sizes.

Now to compare the Prokof'yev–Shishkin procedure to the optimal. Note that from (30) we have $2S^2/d \leqslant 2n \leqslant 2n_0 + 2S^2/d$, so

$$\frac{2S^2\Delta_1^2}{dQ_1^2(\sigma_1+\sigma_2)^2} \leqslant \frac{2n}{n_1^*+n_2^*} \leqslant o(n_0\Delta_1^2) + \frac{2S^2\Delta_1^2}{dQ_1^2(\sigma_1+\sigma_2)^2}. \tag{43}$$

Noting that $\Delta_1^2/d = Q^2 \to Q_1^2$ as $\Delta_1 \to 0$ $(n_0 \to \infty)$, and $S^2 \to \sigma_1^2 + \sigma_2^2$, we have

$$\frac{2n}{n_1^* + n_2^*} \to \frac{2(\sigma_1^2 + \sigma_2^2)}{(\sigma_1 + \sigma_2)^2} = 1 + \left(\frac{1 - \sigma_2/\sigma_1}{1 + \sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0. \tag{44}$$

From Eqs. (43) and (44), we see that $2n/(n_1^* + n_2^*)$ is bounded for all n and converges to

$$1 + \left(\frac{1 - \sigma_2/\sigma_1}{1 + \sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0.$$

Hence, by the Dominated Convergence Theorem (e.g., Loève, 1963, p. 125), we then obtain

$$\frac{E(2n)}{n_1^* + n_2^*} \to 1 + \left(\frac{1 - \sigma_2/\sigma_1}{1 + \sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0. \tag{45}$$

3.4. Asymptotic comparison of Prokof'yev–Shishkin procedure to Dudewicz–Ahmed procedure

To compare the Prokof'yev–Shishkin procedure to the Dudewicz–Ahmed procedure note that (44) and (33) imply

$$\frac{2n}{n_1^* + n_2^*} \Big/ \frac{(N_1 + N_2)}{(n_1^* + n_2^*)} = \frac{2n}{(N_1 + N_2)} \to 1 + \left(\frac{1 - \sigma_2/\sigma_1}{1 + \sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0, \tag{46}$$

while (45) and (34) imply

$$\frac{E(2n)}{n_1^* + n_2^*} \Big/ \frac{E(N_1 + N_2)}{(n_1^* + n_2^*)} = \frac{E(2n)}{E(N_1 + N_2)} \to 1 + \left(\frac{1 - \sigma_2/\sigma_1}{1 + \sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0. \tag{47}$$

Denote

$$1 + \left(\frac{1 - \sigma_2/\sigma_1}{1 + \sigma_2/\sigma_1}\right)^2 \equiv R.$$

Then clearly: $1 \leqslant R \leqslant 2$; $R = 1$, when $\sigma_2/\sigma_1 = 1$; and R increases to 2 as σ_2/σ_1 decreases to 0 or increases to ∞. Hence the Dudewicz–Ahmed procedure is twice as efficient as the Prokof'yev–Shishkin procedure when the variances are very different, and $\Delta_1 \to 0$; and D–A is more efficient as long as $\sigma_1 \neq \sigma_2$. (The procedures are equally efficient when $\sigma_1 = \sigma_2$.)

3.5. Comparison of the Chapman procedure to the optimal when σ_1, σ_2 known

Denote the integral $\int_{-\infty}^{\infty} F_{n_0}(z + h) f_{n_0}(z)\,\mathrm{d}z$ by $P_{n_0}(h)$. Then

$$\begin{aligned} P_{n_0}(h) &= \int_{-\infty}^{\infty} P(T_1 \leqslant z + h) f_{n_0}(z)\,\mathrm{d}z = \int_{-\infty}^{\infty} P(T_1 \leqslant z + h | T_2 = z) f_{T_2}(z)\,\mathrm{d}z \\ &= P(T_1 \leqslant T_2 + h) = P(T_1 - T_2 \leqslant h), \end{aligned} \tag{48}$$

where T_1, T_2 are independent $t_{n_0 - 1}$ random variables.

Since the Student's-t random variables T_2, $-T_2$ are identically distributed,

$$P_{n_0}(h) = P(T_1 - T_2 \leqslant h) = P(T_1 + T_2 \leqslant h), \tag{49}$$

i.e., $P_{n_0}(h) = \int_{-\infty}^{\infty} F_{n_0}(z + h) f_{n_0}(z)\,\mathrm{d}z$ is the c.d.f. of the sum of two i.i.d. Student's-t variables with $n_0 - 1$ degrees of freedom.

Since T_1 and T_2 are independent and each asymptotically N(0, 1) as $\Delta_1 \to 0$ $(n_0 \to \infty)$, $T_1 + T_2$ is asymptotically N(0, 2), as $\Delta_1 \to 0$ $(n_0 \to \infty)$. Hence $(T_1 + T_2)/\sqrt{2}$ is asymptotically N(0, 1), as $\Delta_1 \to 0$ $(n_0 \to \infty)$ and we have

$$1 - \frac{\alpha}{2} = P_{n_0}(c) = P\left(\frac{T_1 + T_2}{\sqrt{2}} \leqslant \frac{c}{\sqrt{2}}\right) \to \Phi\left(\frac{c}{\sqrt{2}}\right) = 1 - \frac{\alpha}{2} \quad \text{as } \Delta_1 \to 0 \ (n_0 \to \infty) \tag{50}$$

so that

$$c_{1-\alpha/2} = \sqrt{2}\Phi^{-1}\left(1-\frac{\alpha}{2}\right). \tag{51}$$

When $\Delta = \Delta_1$, $\beta = \beta_1$, the power function of Chapman's test is (see Taneja and Dudewicz, 1993, p. 452)

$$\begin{aligned}\beta_1\left(\frac{\Delta_1}{h}\right) &= P_{n_0}\left(-c_{1-\alpha/2}(n_0) - \frac{\Delta_1}{h}\right) + P_{n_0}\left(-c_{1-\alpha/2} + \frac{\Delta_1}{h}\right)\\ &= 2 - P_{n_0}\left(c_{1-\alpha/2}(n_0) + \frac{\Delta_1}{h}\right) - P_{n_0}\left(c_{1-\alpha/2} - \frac{\Delta_1}{h}\right);\end{aligned} \tag{52}$$

letting $\Delta_1 \to 0$ and $n_0 \to \infty$, and using (50) and (51), (52) becomes

$$\begin{aligned}\beta_1 &= 2 - \Phi\left(\frac{\sqrt{2}\Phi^{-1}(1-\alpha/2) + \Delta_1/h}{\sqrt{2}}\right) - \Phi\left(\frac{\sqrt{2}\Phi^{-1}(1-\alpha/2) - \Delta_1/h}{\sqrt{2}}\right)\\ &= 2 - \Phi(\Phi^{-1}(1-\alpha/2) + \Delta_1/(\sqrt{2}h)) - \Phi(\Phi^{-1}(1-\alpha/2) - \Delta_1/(\sqrt{2}h)).\end{aligned} \tag{53}$$

Letting $\Delta_1/\sqrt{2}h = Q$, (53) becomes $\beta_1 = 2 - \Phi(\Phi^{-1}(1-\alpha/2) + Q) - \Phi(\Phi^{-1}(1-\alpha/2) - Q)$, which is the same as Eq. (39) that Q_1 solves. Hence $Q \to Q_1$ as $n_0 \to \infty$.

Now to compare the Chapman procedure to the optimal, note that from (14) and (15) we have

$$\frac{s_1^2+s_2^2}{h^2} \leqslant n_1 + n_2 \leqslant (2n_0+2) + \frac{s_1^2+s_2^2}{h^2}. \tag{54}$$

It follows from (54) and (42) that

$$\frac{(s_1^2+s_2^2)/h^2}{Q_1^2(\sigma_1+\sigma_2)^2/\Delta_1^2} \leqslant \frac{n_1+n_2}{n_1^*+n_2^*} \leqslant \frac{2n_0+2}{Q_1^2(\sigma_1+\sigma_2)^2/\Delta_1^2} + \frac{(s_1^2+s_2^2)/h^2}{Q_1^2(\sigma_1+\sigma_2)^2/\Delta_1^2}.$$

Noting that we choose n_0 such that $\Delta_1^2 n_0 \to 0$ and $n_0 \to \infty$ as $\Delta_1 \to 0$, we have

$$\frac{2n_0+2}{Q_1^2(\sigma_1+\sigma_2)^2/\Delta_1^2} \to 0 \quad \text{as } \Delta_1 \to 0.$$

Finally, since $(\Delta_1/\sqrt{2}h)^2 = Q^2 \to Q_1^2$ as $\Delta_1 \to 0$, $n_0 \to \infty$, and $s_1^2 \to \sigma_1^2$, $s_2^2 \to \sigma_2^2$,

$$\frac{n_1+n_2}{n_1^*+n_2^*} \to \frac{2(\sigma_1^2+\sigma_2^2)}{(\sigma_1+\sigma_2)^2} = 1 + \left(\frac{1-\sigma_2/\sigma_1}{1+\sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0. \tag{55}$$

By the Dominated Convergence Theorem, we then obtain

$$\frac{E(n_1+n_2)}{n_1^*+n_2^*} \to 1 + \left(\frac{1-\sigma_2/\sigma_1}{1+\sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0. \tag{56}$$

3.6. Asymptotic comparison of the Chapman and Dudewicz–Ahmed solutions

Next, comparing Chapman procedure to the Dudewicz–Ahmed procedure using (55) and (33),

$$\frac{(n_1+n_2)}{(n_1^*+n_2^*)} \Bigg/ \frac{(N_1+N_2)}{(n_1^*+n_2^*)} = \frac{n_1+n_2}{N_1+N_2} \to 1 + \left(\frac{1-\sigma_2/\sigma_1}{1+\sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0 \tag{57}$$

while (56) and (34) imply

$$\frac{E(n_1+n_2)}{(n_1^*+n_2^*)} \Bigg/ \frac{E(N_1+N_2)}{(n_1^*+n_2^*)} = \frac{E(n_1+n_2)}{E(N_1+N_2)} \to 1 + \left(\frac{1-\sigma_2/\sigma_1}{1+\sigma_2/\sigma_1}\right)^2 \quad \text{as } \Delta_1 \to 0. \tag{58}$$

This is equal to the same value R as the efficiency of the P–S procedure to the D–A procedure found at (47) so the Dudewicz–Ahmed procedure is twice as efficient as the Chapman procedure when the variances are very different, and $\Delta_1 \to 0$; and is more efficient as long as $\sigma_1 \neq \sigma_2$ (while it is equally efficient when $\sigma_1 = \sigma_2$).

3.7. Asymptotic comparison of the Prokof'yev–Shishkin and Chapman procedures

We finally compare the Prokof'yev–Shishkin and Chapman procedures: (46) and (57) imply

$$\frac{(2n)}{(N_1+N_2)} \Bigg/ \frac{(n_1+n_2)}{(N_1+N_2)} = \frac{2n}{n_1+n_2} \to 1 \quad \text{as } \Delta_1 \to 0, \tag{59}$$

$$\frac{E(2n)}{E(N_1+N_2)} \Bigg/ \frac{E(n_1+n_2)}{E(N_1+N_2)} = \frac{E(2n)}{E(n_1+n_2)} \to 1 \quad \text{as } \Delta_1 \to 0. \tag{60}$$

Thus, the Chapman and Prokov'yev–Shishkin procedures have asymptotically the same efficiency.

3.8. Summary of asymptotic comparisons of exact solutions D–A, P–S, and C

Eqs. (47) and (58) show that *when we compare Dudewicz–Ahmed's procedure to the Prokof'yev–Shishkin or Chapman procedure*, the asymptotic efficiency is the same quantity R. This R depends on the ratio of the two population standard deviations, as $\Delta_1 \to 0$, $n_0 \to \infty$, $\Delta_1^2 n_0(\Delta_1) \to 0$. Except for the $\sigma_1/\sigma_2 = 1$ case where all three procedures have the same asymptotic efficiency, *the Dudewicz–Ahmed procedure always has better asymptotic efficiency*.

4. Finite-sample comparisons of the three exact solutions of the Behrens–Fisher Problem

4.1. Summary of results thus far

For the Behrens–Fisher Problem of testing H_0: $\mu_1 = \mu_2$ when the variances of independent sources of observations are not known (σ_1^2, σ_2^2 unknown, hence possibly unequal), there are no single-sample procedures which have level α and power β_1 when $|\mu_1 - \mu_2| = \Delta_1$ ($\alpha < \frac{1}{2}$, $\alpha < \beta_1 < 1$). There are three exact procedures, all of which are two-stage procedures. The first was given by Chapman (1950), and has sample size $n_1 + n_2$. The second exact solution was given by Prokof'yev and Shishkin (1974), and has sample size $2n$. The latest exact solution was given by Dudewicz and Ahmed (1998), and has sample size $N_1 + N_2$.

These procedures were detailed in Section 2. If one knew the variances, the sample size needed to meet the same level and power at the same difference Δ_1 of means would be $n_1^* + n_2^*$ (see (40)).

Dudewicz and Ahmed (1999) showed that, if the initial sample size n_0 is made a function of Δ_1, say $n_0(\Delta_1)$ such that

$$n_0(\Delta_1) \to \infty \quad \text{and} \quad \Delta_1^2 n_0(\Delta_1) \to 0 \ \text{ as } \Delta_1 \to 0, \tag{61}$$

then

$$\frac{E(N_1+N_2)}{(n_1^*+n_2^*)} \to 1. \tag{62}$$

This means *the Dudewicz–Ahmed procedure is asymptotically optimal*. In Section 3 we showed that (under the same limits as above)

$$\frac{E(n_1+n_2)}{E(N_1+N_2)} \to 1 + \left(\frac{1-\sigma_2/\sigma_1}{1+\sigma_2/\sigma_1}\right)^2 = R, \qquad \frac{E(2n)}{E(N_1+N_2)} \to R \quad \text{and} \quad \frac{E(2n)}{E(N_1+N_2)} \to 1, \tag{63}$$

where R is in the interval $[1, 2]$ and equals 1 if $\sigma_2 = \sigma_1$ (and approaches 2 as $\sigma_2/\sigma_1 \to 0$ or ∞). *That is, the Dudewicz–Ahmed procedure is more efficient asymptotically than either the Chapman procedure or the Prokof'yev–Shishkin procedure*, and may require only 50% as much sampling; the Chapman and Prokof'yev–Shishkin procedures are equally efficient asymptotically.

Since the results thus far are asymptotic under a particular limiting situation (see (61)), and we also do not know how large the sample sizes need to be to approach these limits, it is not yet definitive that the D–A procedure is the best of the three, and the one to be used in practice. For that evaluation the finite-sample (non-limiting) comparisons of this Section 4 are needed to complete the comparison of the three procedures.

4.2. Simulation comparison of C, P–S, and D–A procedures

For given α, β, Δ, σ_1^2, σ_1^2/σ_2^2, n_0 we generate $X_1, X_2, \ldots, X_{n_0}$ and $Y_1, Y_2, \ldots, Y_{n_0}$ as independent random variables from respective $\mathrm{N}(0, \sigma_1^2)$ and $\mathrm{N}(\Delta, \sigma_2^2)$ populations independently by use of the Box–Muller transformation (using only the sine form), so using U_1, U_2, U_3, U_4 from a good random number generator (to be discussed shortly) we calculate

$$V_1 = \sqrt{-2\log(U_1)}\sin(2\pi U_2), \quad V_2 = \sqrt{-2\log(U_3)}\sin(2\pi U_4) \tag{64}$$

and then (with the means and variances taken above)

$$X_1 = \mu_1 + \sigma_1 V_1, \quad Y_1 = \mu_2 + \sigma_2 V_2. \tag{65}$$

We continue in the stream $U_5, \ldots$ to find $X_2, \ldots, X_{n_0}$ and $Y_2, \ldots, Y_{n_0}$. Using the constant h from procedure C in (14) and (15) we can find its needed total sample size $n_1 + n_2$; using the constant d needed for procedure P–S in (30) we can find its total needed sample size $2n$; and using the constant c needed for procedure D–A in (3) we can find its total sample size $N_1 + N_2$.

4.3. Comparison of D–A and P–S procedures

Denote the D–A total sample size just discussed by N_1 ($=N_1 + N_2$ above) and the P–S total sample size by M_1 ($=2n$ above). Replicating this simulation 10,000 times, we obtain $N_1, \ldots, N_{10,000}$ from D–A and $M_1, \ldots, M_{10,000}$ from P–S. The ratio

$$\frac{\overline{N}}{\overline{M}} = \frac{(N_1 + N_2 + \cdots + N_{10,000})/10,000}{(M_1 + M_2 + \cdots + M_{10,000})/10,000} = \frac{N_1 + N_2 + \cdots + N_{10,000}}{M_1 + M_2 + \cdots + M_{10,000}} \tag{66}$$

is then calculated, and procedure D–A is preferable to procedure P–S if the ratio is less than one (preferable in the sense of a smaller expected total sample size at the specified level, power, etc.). The specified parameters used in this study are all combinations of

$$\begin{gathered} n_0 = 5, 10, \quad \alpha = 0.01, 0.05, \quad \beta = 0.90, 0.95, 0.99, \quad \Delta = 0.25, 0.50, 1.0, 1.5, 2.0, \\ \sigma_1^2 = 0.1, 1.0, 10.0 \quad \text{and} \quad \sigma_2^2/\sigma_1^2 = 1, 1.5, 2, 5, 10, 20, 100 \end{gathered} \tag{67}$$

so we find $2 \times 2 \times 3 \times 5 \times 3 \times 7 = 1260$ ratios, which are given in Table 1. The column of Table 1 which has $\sigma_2^2/\sigma_1^2 = \infty$ takes $\overline{N}/\overline{M}$ to equal

$$\frac{E(N_1+N_2)}{E(2n)} \approx \frac{c^2(\sigma_1+\sigma_2)^2}{2\dfrac{\sigma_1^2+\sigma_2^2}{d}} = \frac{dc^2}{2}\,\frac{\dfrac{\sigma_1^2}{\sigma_2^2} + \dfrac{\sigma_2^2}{\sigma_2^2} + \dfrac{2\sigma_1\sigma_2}{\sigma_2^2}}{\dfrac{\sigma_1^2}{\sigma_2^2} + \dfrac{\sigma_2^2}{\sigma_2^2}} \to \frac{dc^2}{2}. \tag{68}$$

The needed constants were obtained as follows: d for procedure P–S from Table 1 of Taneja and Dudewicz (1993, p. 461); h for procedure C from that same table; and c for procedure D–A from Table II of Dudewicz and Ahmed (1999, p. 167). Specifically, suppose $n_0 = 5$, $\alpha = 0.01$, $\beta = 0.99$, $\sigma_1^2 = 0.1$, $\sigma_2^2/\sigma_1^2 = 1$. Then in Table 1 of Taneja and Dudewicz (1993) we find

$$\Delta/\sqrt{d} = 8.3392776,$$

which means we will use

$$d = (\Delta/8.3392776)^2 \tag{69}$$

as the d for procedure P–S. Similarly, in Table 1 of Taneja and Dudewicz (1993), for the same set of parameters, we find

$$\Delta/h = 11.2755190,$$

Table 1
$\overline{N}$ of Dudewicz–Ahmed procedure over $\overline{M}$ of Prokof'yev–Shishkin procedure, for selected n_0, α, β, Δ, σ_1^2, and σ_2^2/σ_1^2 in (a)–(l)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
(a) $n_0=5$, $\sigma_1^2=0.1$, $\alpha=0.01$									
0.25	0.99	0.940803	0.931733	0.915415	0.828636	0.753824	0.688384	0.588293	0.501058
	0.95	0.940433	0.931393	0.915092	0.828331	0.753538	0.688103	0.588035	0.501286
	0.90	0.942863	0.933843	0.917524	0.830535	0.755538	0.689923	0.589575	0.502366
0.50	0.99	0.941147	0.932111	0.915691	0.829037	0.754137	0.688586	0.588348	0.501058
	0.95	0.940481	0.931664	0.915509	0.828900	0.754020	0.688411	0.588120	0.501286
	0.90	0.942938	0.934219	0.917882	0.831201	0.756110	0.690287	0.589678	0.502366
1.00	0.99	0.943216	0.933566	0.917111	0.830716	0.755452	0.689416	0.588567	0.501058
	0.95	0.955528	0.939601	0.921935	0.832991	0.756394	0.689798	0.588460	0.501286
	0.90	0.972190	0.950114	0.929527	0.837472	0.759547	0.691978	0.590094	0.502366
1.50	0.99	0.994150	0.963634	0.939755	0.840495	0.760027	0.691428	0.588977	0.501058
	0.95	1.076480	1.029700	0.993529	0.865032	0.770742	0.695259	0.589159	0.501286
	0.90	1.115040	1.069620	1.030480	0.887615	0.782599	0.700942	0.590982	0.502366
2.00	0.99	1.106820	1.060920	1.022040	0.881220	0.778764	0.698518	0.589639	0.501058
	0.95	1.175190	1.145800	1.115490	0.957415	0.821506	0.716552	0.590845	0.501286
	0.90	1.189090	1.171290	1.149310	1.003480	0.854083	0.733083	0.593737	0.502366
(b) $n_0=5$, $\sigma_1^2=0.1$, $\alpha=0.05$									
0.25	0.99	0.941086	0.932068	0.915749	0.828965	0.754098	0.688625	0.588471	0.501428
	0.95	0.940624	0.931552	0.915338	0.828594	0.753742	0.688260	0.588112	0.501103
	0.90	0.942668	0.933624	0.917222	0.830416	0.755386	0.689748	0.589350	0.502143
0.50	0.99	0.941136	0.932251	0.916286	0.829558	0.754595	0.688961	0.588562	0.501428
	0.95	0.940918	0.931871	0.916208	0.829683	0.754632	0.688853	0.588278	0.501103
	0.90	0.943947	0.934676	0.918286	0.831927	0.756578	0.690505	0.589562	0.502143
1.00	0.99	0.961399	0.943102	0.925002	0.834555	0.757523	0.690449	0.588930	0.501428
	0.95	1.040400	0.997174	0.966599	0.851798	0.764748	0.693104	0.588972	0.501103
	0.90	1.094380	1.047930	1.010080	0.875435	0.776207	0.698209	0.590557	0.502143
1.50	0.99	1.091890	1.045390	1.007430	0.873369	0.774781	0.697105	0.589700	0.501428
	0.95	1.180340	1.156630	1.127070	0.971313	0.831030	0.721065	0.591298	0.501103
	0.90	1.194630	1.183800	1.167720	1.037580	0.881380	0.747627	0.594908	0.502143
2.00	0.99	1.181680	1.158490	1.129820	0.975517	0.833818	0.722222	0.591788	0.501428
	0.95	1.199420	1.196970	1.191690	1.111590	0.958783	0.795497	0.599811	0.501103
	0.90	1.199950	1.199590	1.198240	1.156700	1.030090	0.851659	0.610364	0.502143
(c) $n_0=5$, $\sigma_1^2=1$, $\alpha=0.01$									
0.25	0.99	0.940669	0.931633	0.915307	0.828525	0.753734	0.688324	0.588276	0.501058
	0.95	0.940252	0.931219	0.914898	0.828154	0.753397	0.688014	0.588009	0.501286
	0.90	0.942703	0.933647	0.917291	0.830321	0.755366	0.689812	0.589544	0.502366
0.50	0.99	0.940707	0.931666	0.915350	0.828560	0.753764	0.688345	0.588281	0.501058
	0.95	0.940301	0.931277	0.914957	0.828215	0.753444	0.688046	0.588017	0.501286
	0.90	0.942765	0.933713	0.917364	0.830389	0.755421	0.689848	0.589555	0.502366
1.00	0.99	0.940849	0.931812	0.915498	0.828710	0.753882	0.688428	0.588304	0.501058
	0.95	0.940577	0.931499	0.915165	0.828448	0.753633	0.688172	0.588051	0.501286
	0.90	0.943038	0.933976	0.917583	0.830676	0.755633	0.689996	0.589596	0.502366
1.50	0.99	0.941096	0.931965	0.915679	0.828962	0.754073	0.688562	0.588339	0.501058
	0.95	0.940588	0.931644	0.915377	0.828841	0.753946	0.688387	0.588108	0.501286
	0.90	0.942839	0.934011	0.917966	0.831125	0.756021	0.690238	0.589663	0.502366
2.00	0.99	0.941021	0.932166	0.915892	0.829300	0.754346	0.688749	0.588393	0.501058
	0.95	0.940453	0.931765	0.915824	0.829422	0.754339	0.688682	0.588185	0.501286
	0.90	0.942984	0.934313	0.918090	0.831827	0.756500	0.690553	0.589755	0.502366

Table 1(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
(d) $n_0=5$, $\sigma_1^2=1$, $\alpha=0.05$									
0.25	0.99	0.940943	0.931906	0.915576	0.828767	0.753954	0.688523	0.588444	0.501428
	0.95	0.940351	0.931316	0.914998	0.828249	0.753477	0.688085	0.588064	0.501103
	0.90	0.942311	0.933256	0.916912	0.829977	0.755050	0.689520	0.589287	0.502143
0.50	0.99	0.941012	0.931973	0.915641	0.828824	0.754004	0.688556	0.588453	0.501428
	0.95	0.940472	0.931440	0.915102	0.828349	0.753563	0.688142	0.588080	0.501103
	0.90	0.942420	0.933412	0.917035	0.830136	0.755159	0.689596	0.589308	0.502143
1.00	0.99	0.941183	0.932144	0.915861	0.829089	0.754219	0.688693	0.588489	0.501428
	0.95	0.940719	0.931707	0.915470	0.828790	0.753930	0.688376	0.588144	0.501103
	0.90	0.942513	0.933708	0.917390	0.830635	0.755606	0.689890	0.589391	0.502143
1.50	0.99	0.941246	0.932326	0.916166	0.829576	0.754534	0.688924	0.588550	0.501428
	0.95	0.940433	0.931755	0.915690	0.829479	0.754499	0.688772	0.588251	0.501103
	0.90	0.942898	0.933742	0.918020	0.831803	0.756405	0.690405	0.589526	0.502143
2.00	0.99	0.941108	0.932437	0.916490	0.830106	0.755021	0.689232	0.588635	0.501428
	0.95	0.945800	0.933964	0.917874	0.830970	0.755515	0.689311	0.588409	0.501103
	0.90	0.956241	0.940925	0.922889	0.834291	0.757959	0.691200	0.589725	0.502143
(e) $n_0=5$, $\sigma_1^2=10$, $\alpha=0.01$									
0.25	0.99	0.940658	0.931621	0.915297	0.828513	0.753725	0.688318	0.588274	0.501058
	0.95	0.940230	0.931198	0.914881	0.828136	0.753382	0.688005	0.588006	0.501286
	0.90	0.942684	0.933629	0.917270	0.830299	0.755349	0.689801	0.589541	0.502366
0.50	0.99	0.940662	0.931625	0.915301	0.828517	0.753728	0.688320	0.588275	0.501058
	0.95	0.940236	0.931204	0.914887	0.828142	0.753387	0.688008	0.588007	0.501286
	0.90	0.942689	0.933636	0.917277	0.830306	0.755355	0.689805	0.589542	0.502366
1.00	0.99	0.940679	0.931636	0.915314	0.828533	0.753741	0.688329	0.588277	0.501058
	0.95	0.940262	0.931226	0.914913	0.828165	0.753406	0.688021	0.588010	0.501286
	0.90	0.942723	0.933666	0.917306	0.830337	0.755377	0.689820	0.589546	0.502366
1.50	0.99	0.940718	0.931659	0.915340	0.828556	0.753760	0.688342	0.588281	0.501058
	0.95	0.940299	0.931262	0.914954	0.828210	0.753436	0.688041	0.588016	0.501286
	0.90	0.942770	0.933711	0.917358	0.830381	0.755416	0.689845	0.589553	0.502366
2.00	0.99	0.940729	0.931710	0.915372	0.828595	0.753789	0.688361	0.588286	0.501058
	0.95	0.940356	0.931326	0.914997	0.828261	0.753478	0.688072	0.588024	0.501286
	0.90	0.942806	0.933763	0.917397	0.830458	0.755464	0.689878	0.589562	0.502366
(f) $n_0=5$, $\sigma_1^2=10$, $\alpha=0.05$									
0.25	0.99	0.940925	0.931886	0.915557	0.828749	0.753939	0.688514	0.588441	0.501428
	0.95	0.940316	0.931283	0.914965	0.828213	0.753451	0.688068	0.588059	0.501103
	0.90	0.942270	0.933218	0.916866	0.829933	0.755017	0.689497	0.589281	0.502143
0.50	0.99	0.940931	0.931892	0.915565	0.828755	0.753944	0.688517	0.588442	0.501428
	0.95	0.940329	0.931294	0.914976	0.828224	0.753461	0.688074	0.588061	0.501103
	0.90	0.942284	0.933233	0.916880	0.829948	0.755028	0.689504	0.589283	0.502143
1.00	0.99	0.940958	0.931920	0.915590	0.828778	0.753965	0.688530	0.588445	0.501428
	0.95	0.940376	0.931342	0.915020	0.828269	0.953497	0.688097	0.588067	0.501103
	0.90	0.942359	0.933290	0.916939	0.830006	0.755072	0.689535	0.589291	0.502143
1.50	0.99	0.941000	0.931954	0.915631	0.828821	0.753997	0.688552	0.588452	0.501428
	0.95	0.940452	0.931430	0.915099	0.828338	0.753552	0.688134	0.588078	0.501103
	0.90	0.942410	0.933387	0.917042	0.830108	0.755149	0.689584	0.589305	0.502143
2.00	0.99	0.941076	0.932007	0.915689	0.828871	0.754050	0.688583	0.588460	0.501428
	0.95	0.940600	0.931527	0.915203	0.828430	0.753641	0.688191	0.588094	0.501103
	0.90	0.942532	0.933539	0.917136	0.830228	0.755255	0.689657	0.589323	0.502143

Table 1(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
(g) $n_0 = 10$, $\sigma_1^2 = 0.1$, $\alpha = 0.01$									
0.25	0.99	0.974079	0.964011	0.946309	0.852698	0.772238	0.701997	0.594664	0.501470
	0.95	0.974027	0.963959	0.946254	0.852688	0.772228	0.701959	0.594605	0.501408
	0.90	0.974268	0.964235	0.946543	0.853004	0.772528	0.702197	0.594796	0.501559
0.50	0.99	0.972813	0.963629	0.946479	0.853209	0.772795	0.702373	0.594769	0.501470
	0.95	0.970782	0.962699	0.945937	0.853603	0.772951	0.702473	0.594752	0.501408
	0.90	0.970606	0.962208	0.946475	0.854171	0.773439	0.702834	0.594968	0.501559
1.00	0.99	1.054480	1.019580	0.992418	0.875032	0.779821	0.704557	0.595179	0.501470
	0.95	1.091350	1.073520	1.052020	0.924772	0.801053	0.710046	0.595345	0.501408
	0.90	1.097640	1.089460	1.076410	0.964191	0.824709	0.718011	0.595786	0.501559
1.50	0.99	1.099740	1.098350	1.094090	1.022600	0.877023	0.740829	0.596469	0.501470
	0.95	1.100000	1.099950	1.099580	1.076500	0.965057	0.797500	0.600283	0.501408
	0.90	1.100000	1.099980	1.099940	1.090070	1.009580	0.837650	0.605381	0.501559
2.00	0.99	1.100000	1.100000	1.099960	1.091410	1.017260	0.846132	0.606592	0.501470
	0.95	1.100000	1.100000	1.100000	1.099230	1.074580	0.936869	0.627066	0.501408
	0.90	1.100000	1.100000	1.100000	1.099920	1.088820	0.985480	0.643398	0.501559
(h) $n_0 = 10$, $\sigma_1^2 = 0.1$, $\alpha = 0.05$									
0.25	0.99	0.978081	0.968000	0.950129	0.856215	0.775440	0.704874	0.597082	0.503495
	0.95	0.978588	0.968576	0.950840	0.856916	0.776047	0.705406	0.597478	0.503807
	0.90	0.979096	0.969411	0.951735	0.857828	0.776903	0.706133	0.598060	0.504284
0.50	0.99	0.974321	0.966607	0.949845	0.857030	0.776165	0.705397	0.597223	0.503495
	0.95	0.979717	0.967720	0.951962	0.858859	0.777205	0.706235	0.597702	0.503807
	0.90	1.000140	0.978436	0.960018	0.862510	0.778880	0.707200	0.598334	0.504284
1.00	0.99	1.091400	1.073800	1.052800	0.927242	0.803796	0.712788	0.597810	0.503495
	0.95	1.099660	1.097990	1.093510	1.019650	0.874670	0.740942	0.599149	0.503807
	0.90	1.099960	1.099720	1.098720	1.060320	0.930835	0.774287	0.601368	0.504284
1.50	0.99	1.100000	1.099950	1.099580	1.077020	0.967021	0.799950	0.602712	0.503495
	0.95	1.100000	1.100000	1.100000	1.097870	1.056500	0.901301	0.619825	0.503807
	0.90	1.100000	1.100000	1.100000	1.099710	1.083020	0.962812	0.637637	0.504284
2.00	0.99	1.100000	1.100000	1.100000	1.099290	1.075270	0.938958	0.629366	0.503495
	0.95	1.100000	1.100000	1.100000	1.099980	1.097600	1.041290	0.673462	0.503807
	0.90	1.100000	1.100000	1.100000	1.100000	1.099570	1.075040	0.706636	0.504284
(i) $n_0 = 10$, $\sigma_1^2 = 1$, $\alpha = 0.01$									
0.25	0.99	0.973934	0.963904	0.946160	0.852503	0.772092	0.701885	0.594635	0.501470
	0.95	0.973819	0.963794	0.946048	0.852406	0.772003	0.701802	0.594562	0.501408
	0.90	0.974118	0.964082	0.946338	0.852668	0.772243	0.702018	0.594743	0.501559
0.50	0.99	0.973983	0.963968	0.946206	0.852559	0.772144	0.701924	0.594645	0.501470
	0.95	0.973884	0.963852	0.946134	0.852497	0.772077	0.701856	0.594577	0.501408
	0.90	0.974175	0.964119	0.946410	0.852798	0.772333	0.702079	0.594760	0.501559
1.00	0.99	0.974167	0.964086	0.946405	0.852828	0.772361	0.702069	0.594687	0.501470
	0.95	0.973856	0.964050	0.946441	0.852802	0.772385	0.702069	0.594635	0.501408
	0.90	0.974024	0.964318	0.946752	0.853241	0.772693	0.702311	0.594830	0.501559
1.50	0.99	0.973168	0.963889	0.946493	0.853303	0.772698	0.702329	0.594755	0.501470
	0.95	0.970962	0.962620	0.946064	0.853294	0.772888	0.702394	0.594733	0.501408
	0.90	0.970693	0.962752	0.946605	0.853890	0.773314	0.702757	0.594941	0.501559
2.00	0.99	0.970361	0.962169	0.946172	0.853735	0.773248	0.702654	0.594843	0.501470
	0.95	0.978798	0.964526	0.948289	0.854890	0.773741	0.702885	0.594863	0.501408
	0.90	0.996465	0.973958	0.955278	0.857779	0.774642	0.703348	0.595100	0.501559

Table 1(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
(j) $n_0 = 10$, $\sigma_1^2 = 1$, $\alpha = 0.05$									
0.25	0.99	0.977870	0.967803	0.949981	0.855948	0.775216	0.704725	0.597038	0.503495
	0.95	0.978477	0.968422	0.950575	0.856495	0.775709	0.705170	0.597409	0.503807
	0.90	0.979408	0.969329	0.951498	0.857315	0.776453	0.705842	0.597977	0.504284
0.50	0.99	0.977927	0.967855	0.950031	0.856034	0.775293	0.704777	0.597052	0.503495
	0.95	0.978507	0.968509	0.950670	0.856616	0.775815	0.705247	0.597432	0.503807
	0.90	0.979433	0.969423	0.951574	0.857490	0.776597	0.705943	0.598005	0.504284
1.00	0.99	0.977898	0.968004	0.950320	0.956417	0.775587	0.704985	0.597110	0.503495
	0.95	0.977676	0.968137	0.950890	0.857230	0.776274	0.705563	0.597517	0.503807
	0.90	0.977535	0.968713	0.951512	0.858042	0.777179	0.706348	0.598114	0.504284
1.50	0.99	0.974815	0.966655	0.949882	0.856832	0.776083	0.705315	0.597208	0.503495
	0.95	0.976489	0.966763	0.950729	0.858204	0.777120	0.706094	0.597668	0.503807
	0.90	0.989725	0.972324	0.955090	0.860613	0.778283	0.707041	0.598303	0.504284
2.00	0.99	0.981529	0.967881	0.951910	0.858390	0.776883	0.705793	0.597344	0.503495
	0.95	1.034670	1.002330	0.977964	0.870024	0.780437	0.707127	0.597874	0.503807
	0.90	1.069490	1.039410	1.011380	0.889391	0.787533	0.709301	0.598568	0.504284
(k) $n_0 = 10$, $\sigma_1^2 = 10$, $\alpha = 0.01$									
0.25	0.99	0.973921	0.963896	0.946145	0.852481	0.772076	0.701874	0.594632	0.501470
	0.95	0.973801	0.963776	0.946027	0.852376	0.771980	0.701787	0.594558	0.501408
	0.90	0.974095	0.964067	0.946313	0.852633	0.772214	0.701999	0.594737	0.501559
0.50	0.99	0.973928	0.963900	0.946149	0.852487	0.772081	0.701877	0.594633	0.501470
	0.95	0.973805	0.963781	0.946035	0.852384	0.771988	0.701792	0.594559	0.501408
	0.90	0.974098	0.964073	0.946320	0.852645	0.772222	0.702005	0.594739	0.501559
1.00	0.99	0.973942	0.963916	0.946167	0.852515	0.772101	0.701891	0.594637	0.501470
	0.95	0.973819	0.963800	0.946061	0.852417	0.772018	0.701814	0.594565	0.501408
	0.90	0.974116	0.964106	0.946362	0.852685	0.772261	0.702029	0.594746	0.501559
1.50	0.99	0.973968	0.963930	0.946183	0.852551	0.772141	0.701917	0.594644	0.501470
	0.95	0.973843	0.963819	0.946091	0.852486	0.772071	0.701848	0.594575	0.501408
	0.90	0.974126	0.964145	0.946399	0.852762	0.772320	0.702072	0.594758	0.501559
2.00	0.99	0.974017	0.963995	0.946241	0.852617	0.772179	0.701949	0.594653	0.501470
	0.95	0.973895	0.963871	0.946212	0.852560	0.772139	0.701894	0.594588	0.501408
	0.90	0.974253	0.964221	0.946506	0.852831	0.772414	0.702131	0.594774	0.501559
(l) $n_0 = 10$, $\sigma_1^2 = 10$, $\alpha = 0.05$									
0.25	0.99	0.977855	0.967789	0.949966	0.855924	0.775194	0.704709	0.597033	0.503495
	0.95	0.978459	0.968389	0.950555	0.856455	0.775675	0.705146	0.597403	0.503807
	0.90	0.979387	0.969306	0.951456	0.857268	0.776411	0.705814	0.597969	0.504284
0.50	0.99	0.977863	0.967793	0.949972	0.855933	0.775202	0.704714	0.597035	0.503495
	0.95	0.978470	0.968394	0.950563	0.856468	0.775686	0.705153	0.597405	0.503807
	0.90	0.979397	0.969316	0.951467	0.857285	0.776424	0.705824	0.597972	0.504284
1.00	0.99	0.977849	0.967808	0.949998	0.855973	0.775233	0.704734	0.597041	0.503495
	0.95	0.978518	0.968415	0.950595	0.856523	0.775732	0.705185	0.597414	0.503807
	0.90	0.979426	0.969352	0.951515	0.857353	0.776479	0.705863	0.597983	0.504284
1.50	0.99	0.977897	0.967862	0.950027	0.856026	0.775279	0.704766	0.597051	0.503495
	0.95	0.978538	0.968482	0.950640	0.856609	0.775805	0.705243	0.597429	0.503807
	0.90	0.979449	0.969415	0.951600	0.857450	0.776572	0.705931	0.598001	0.504284
2.00	0.99	0.977878	0.967836	0.950127	0.856110	0.775356	0.704819	0.597064	0.503495
	0.95	0.978581	0.968522	0.950713	0.856721	0.775906	0.705298	0.597451	0.503807
	0.90	0.979531	0.969424	0.951720	0.857621	0.776714	0.706020	0.598028	0.504284

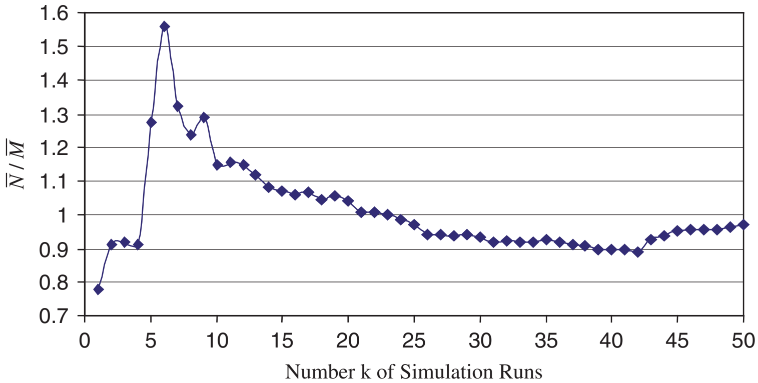

Fig. 1. The ratio $\overline{N}/\overline{M}$ (D–A average sample size $\overline{N}$divided by P–S average sample size $\overline{M}$) in the first k replications, $k \leqslant 50$.

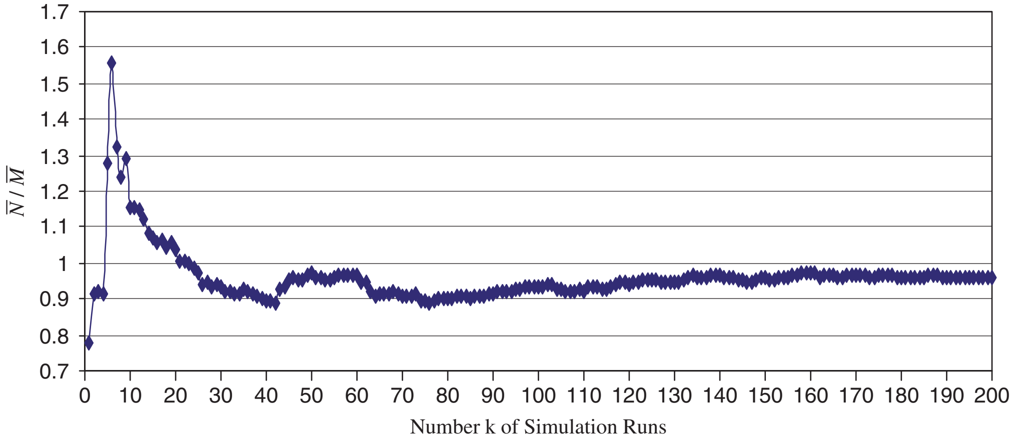

Fig. 2. The ratio $\overline{N}/\overline{M}$ (D–A average sample size $\overline{N}$divided by P–S average sample size $\overline{M}$) in the first k replications, $0 < k \leqslant 200$.

hence we will use

$$h = \Delta/11.2755190 \tag{70}$$

as the h for procedure C. And, in Table II of Dudewicz and Ahmed (1999), for the same set of parameters, we find

$$Q = 8.35$$

hence (see Dudewicz and Ahmed, 1999, Eq. (21), p. 161) we take

$$c = Q/\Delta = 8.35/\Delta \tag{71}$$

as the c for procedure D–A.

Random number generator (*rng*) *for the P–S vs. D–A comparison*: In order to compare solutions due to P–S and D–A, we consider random number generators (rng's) which did well in the tests reported in Karian and Dudewicz (1999), especially the ones called URN03, URN41, and URN37. URN03, while excellent, would have needed more programming consideration for our C++ program because URN03 uses a feedback shift sequence as well as a multiplicative congruential sequence. URN41 is set up for 64-bit computer architecture, but not for the 32-bit architecture we had available. We finally settled on URN37, which was studied in Dudewicz et al. (1985) and is based on the sequence $x_i = 5^{13}x_{i-1} \bmod 2^{31}$, with the x_0 specified in Karian and Dudewicz (1999, p. 492) (where the first $100x_i$ produced are given and were used as a check of our programming of the rng). Figs. 1–3 show the ratio (66) in the first 50, first 200,

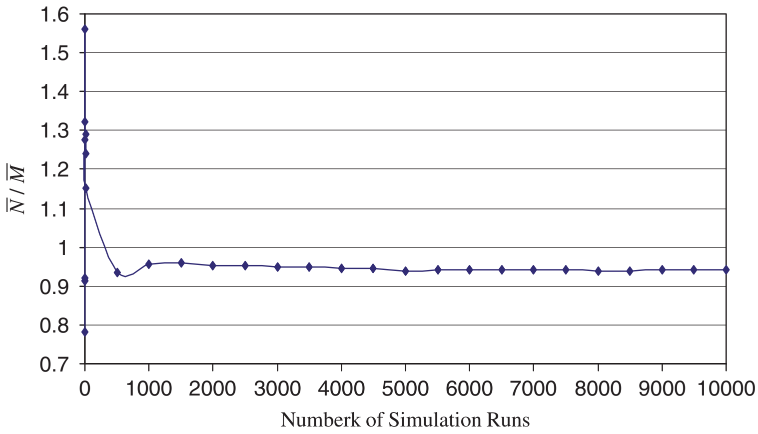

Fig. 3. The ratio $\overline{N}/\overline{M}$ (D–A average sample size $\overline{N}$divided by P–S average sample size $\overline{M}$) in the first k replications, $0 < k \leqslant 10,000$.

and all 10,000 replications to show the approach to the limit is well-attained after many fewer than 10,000 replications (so the results of Table 1 are based on sufficient replications).

Discussion of Table 1 (P–S vs. D–A): Regression analysis of the Table 1 values yielded the model

$$1\hat{/}\text{Ratio} = 1.055 - 0.009n_0 - 0.357\alpha + 0.166\beta - 0.04\Delta + 0.005\sigma_1^2 + 0.006\frac{\sigma_2^2}{\sigma_1^2} \tag{72}$$

with $R^2 = 0.77$ and standard error of estimate 0.1149, and all coefficients significant at level of significance 0.1. As this is a model for the reciprocal of the values in Table 1, values greater than 1 indicate the P–S procedure needs larger sample sizes than the D–A procedure. We note that the model (72) indicates P–S gets worse by comparison as: power increases, variance increases, variance ratio increases, difference to be detected decreases, level of the test decreases. Thus, D–A does better for "harder" problems (problems with higher power requirement, larger variances, smaller level, and smaller difference to be detected with specified power). The regression meta-model only indicates P–S will improve as n_0 increases.

Looking in further detail at the values in Table 1, we focus on those where the ratio is greater than 1, i.e., where it seems P–S does better. Table 2 shows analysis of the ratios greater than 1 in Table 1a. In particular, we see that with $n_0 = 5$ (so that in the first stage 10 total observations are taken by both P–S and D–A), $\overline{N}$ and $\overline{M}$ are in the range of 10.2–14.1; this indicates that the fact that P–S does not always need a second stage (see (30)) will play a large role. In fact, P–S had $n = n_0$ in over 50% of the cases in these ratios, while procedure D–A (see (3)) needs at least sample size $n_0 + 1$ no matter how small S_1 and S_2 are. Since $\overline{N} = 12.7$ and $\overline{M} = 11.6$ in Table 2, this is the main reason for ratios greater than 1. Thus, the P–S procedure is better only when the problem is so "easy" that n_0 samples may be sufficient; and, in these cases, while the ratio may be as large as 1.18, the savings in numbers of observations is 18% of a small sample size, so the benefit of P–S is not large in these cases. On the other hand, as the problem becomes "harder" sample sizes increase, and then there will be a large benefit of using procedure D–A, which in any case is favored over most of Table 1 (ratios less than 1). *Based on finite-sample comparisons, we can conclude that procedure D–A is preferable to procedure P–S.*

4.4. Comparison of D–A and C procedures

This comparison proceeds similarly to the D–A vs. P–S comparison just discussed in Section 4.3. For it, we use another rng, URN35 ("beware the use of one rng" as you "beware the man of one book"). Other details are comparable (except that the column headed $\sigma_2^2/\sigma_1^2 = \infty$ arises from use of a variance ratio of 10^{10}). The results are given in Table 3, from which we see: D–A is preferred when the variance ratio is greater than 2, and it can be better by up to 50% (i.e., D–A may require on average as little as 50% of the sample size procedure C requires). When the variance ratio is less than or equal to 2, sometimes procedure C is preferred, but never by more than 4%. Since the variances are unknown

Table 2
Analysis of comparison ratios greater than 1 from Table 1a

Δ	β	σ_2^2/σ_1^2	Ratio	# ($n = n_0$) in P–S	%	$\overline{N}$	$\overline{M}$
$n_0 = 5$, $\sigma_1^2 = 0.1$, $\alpha = 0.01$							
1.50	0.95	1.0	1.07648	7134	0.7134	12.6762	11.7756
		1.5	1.02970	5906	0.5906	13.5302	13.1400
1.50	0.90	1.0	1.11504	7991	0.7991	12.2978	11.0290
		1.5	1.06962	6878	0.6878	12.8042	11.9708
		2.0	1.03048	5938	0.5938	13.5220	13.1220
2.00	0.99	1.0	1.10682	7807	0.7807	12.3546	11.1622
		1.5	1.06092	6691	0.6691	12.9212	12.1792
		2.0	1.02204	5715	0.5715	13.7037	13.4082
2.00	0.95	1.0	1.17519	9382	0.9382	12.0276	10.2346
		1.5	1.14580	8688	0.8688	12.1288	10.5854
		2.0	1.11549	7933	0.7933	12.3563	11.0770
2.00	0.90	1.0	1.18909	9711	0.9711	12.0053	10.0962
		1.5	1.17129	9272	0.9272	12.0411	10.2802
		2.0	1.14931	8725	0.8725	12.1526	10.5738
		5.0	1.00348	5294	0.5294	14.1187	14.0698

Table 3
$\overline{N}$ of Dudewicz–Ahmed procedure over $\overline{M}$ of Chapman procedure, for selected n_0, α, β, Δ, σ_1^2, and σ_2^2/σ_1^2 in (a)–(l)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
(a) $n_0 = 5$, $\sigma_1^2 = 0.1$, $\alpha = 0.01$									
0.25	0.99	1.033900	1.024020	1.006000	0.910072	0.827351	0.755012	0.644441	0.548434
	0.95	1.000120	0.990546	0.973140	0.880345	0.800317	0.730328	0.623352	0.530477
	0.90	1.001210	0.991682	0.974255	0.881388	0.801272	0.731197	0.624081	0.531094
0.50	0.99	1.032920	1.023420	1.005670	0.910171	0.827510	0.755156	0.644489	0.548434
	0.95	0.998381	0.989512	0.972457	0.880469	0.800566	0.730528	0.623421	0.530477
	0.90	0.998327	0.990059	0.973080	0.881362	0.801491	0.731396	0.624155	0.531094
1.0	0.99	1.014050	1.010330	0.995398	0.906378	0.825985	0.754645	0.644477	0.548434
	0.95	0.971194	0.967822	0.954066	0.871360	0.796125	0.728550	0.623164	0.530477
	0.90	0.967443	0.962676	0.949047	0.868211	0.794522	0.728029	0.623645	0.531094
1.5	0.99	0.981947	0.977702	0.965807	0.887076	0.814742	0.748831	0.643450	0.548434
	0.95	0.974540	0.959115	0.938729	0.849636	0.778951	0.717915	0.620873	0.530477
	0.90	0.981479	0.965553	0.943605	0.846032	0.773278	0.712561	0.618242	0.529162
2.0	0.99	0.986139	0.974396	0.956358	0.868352	0.797802	0.737157	0.640606	0.548434
	0.95	0.994896	0.984404	0.967338	0.858424	0.773394	0.707963	0.616456	0.530477
	0.90	0.997964	0.991726	0.978686	0.873515	0.780150	0.709063	0.615479	0.531094
(b) $n_0 = 5$, $\sigma_1^2 = 0.1$, $\alpha = 0.05$									
0.25	0.99	0.971116	0.961780	0.944866	0.854746	0.777039	0.709080	0.605205	0.515031
	0.95	0.914835	0.906005	0.890079	0.805153	0.731912	0.667852	0.569965	0.485021
	0.90	0.904692	0.896122	0.880347	0.796368	0.723896	0.660529	0.563674	0.479654
0.50	0.99	0.969515	0.960966	0.944427	0.854933	0.777306	0.709300	0.605278	0.515031
	0.95	0.911752	0.904112	0.888962	0.805111	0.732154	0.668119	0.570068	0.485021
	0.90	0.900143	0.892771	0.877630	0.795651	0.723871	0.660682	0.563770	0.479654
1.0	0.99	0.948920	0.943020	0.928158	0.846161	0.772755	0.707256	0.605010	0.515031
	0.95	0.934321	0.910011	0.885863	0.792069	0.721866	0.662157	0.568914	0.485021
	0.90	0.954403	0.925628	0.895493	0.785268	0.711391	0.651969	0.561743	0.479654

Table 3(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
1.5	0.99	0.970898	0.951268	0.927898	0.829448	0.757281	0.696920	0.602655	0.515031
	0.95	0.992861	0.978805	0.955117	0.821061	0.724271	0.654775	0.564592	0.485021
	0.90	0.998095	0.991175	0.976263	0.846980	0.734347	0.653465	0.556926	0.479654
2.0	0.99	0.995460	0.985520	0.967253	0.850339	0.758538	0.689936	0.598347	0.515031
	0.95	0.999850	0.998278	0.992971	0.899529	0.774860	0.675963	0.562188	0.485021
	0.90	1.000000	0.999758	0.997921	0.935034	0.807900	0.692664	0.558037	0.479654
(c) $n_0 = 5$, $\sigma_1^2 = 1$, $\alpha = 0.01$									
0.25	0.99	1.034010	1.024060	1.006020	0.910007	0.827279	0.754962	0.644426	0.548434
	0.95	1.000160	0.990540	0.973086	0.880219	0.800199	0.730247	0.623327	0.530477
	0.90	1.001330	0.991692	0.974216	0.881245	0.801131	0.731098	0.624052	0.531094
0.50	0.99	1.033990	1.024040	1.006010	0.910030	0.827303	0.754979	0.644431	0.548434
	0.95	1.000170	0.990550	0.973098	0.880262	0.800237	0.730274	0.623336	0.530477
	0.90	1.001330	0.991707	0.974242	0.881294	0.801179	0.731132	0.624062	0.531094
1.0	0.99	1.033830	1.023950	1.005990	0.910119	0.827388	0.755047	0.644452	0.548434
	0.95	1.000040	0.990501	0.973145	0.880417	0.800394	0.730380	0.623367	0.530477
	0.90	1.001090	0.991576	0.974201	0.881455	0.801346	0.731251	0.624100	0.531094
1.5	0.99	1.033240	1.023560	1.005770	0.910150	0.827503	0.755134	0.644482	0.548434
	0.95	0.998889	0.989671	0.972755	0.880514	0.800558	0.730511	0.623412	0.530477
	0.90	0.995596	0.986968	0.969987	0.878259	0.798570	0.728734	0.621880	0.529162
2.0	0.99	1.031370	1.022340	1.005000	0.910098	0.827573	0.755222	0.644518	0.548434
	0.95	0.994720	0.987323	0.970742	0.879994	0.800465	0.730546	0.623447	0.530477
	0.90	0.993028	0.986183	0.970330	0.880391	0.801167	0.731318	0.624177	0.531094
(d) $n_0 = 5$, $\sigma_1^2 = 1$, $\alpha = 0.05$									
0.25	0.99	0.971054	0.961705	0.944762	0.854595	0.776903	0.708986	0.605178	0.515031
	0.95	0.914519	0.905703	0.889751	0.804822	0.731650	0.667685	0.569918	0.485021
	0.90	0.904422	0.895710	0.879917	0.795931	0.723565	0.660303	0.563613	0.479654
0.50	0.99	0.971087	0.961738	0.944800	0.854646	0.776948	0.709017	0.605187	0.515031
	0.95	0.914649	0.905833	0.889879	0.804940	0.731741	0.667745	0.569934	0.485021
	0.90	0.904636	0.895929	0.880106	0.796100	0.723687	0.660380	0.563634	0.479654
1.0	0.99	0.971075	0.961779	0.944908	0.854837	0.777119	0.709135	0.605221	0.515031
	0.95	0.914637	0.905980	0.890161	0.805321	0.732054	0.667956	0.569992	0.485021
	0.90	0.904444	0.896073	0.880374	0.796544	0.724069	0.660637	0.563708	0.479654
1.5	0.99	0.969951	0.961264	0.944598	0.854978	0.777310	0.709295	0.605271	0.515031
	0.95	0.912327	0.904767	0.889030	0.805265	0.732206	0.668113	0.570063	0.485021
	0.90	0.901235	0.893651	0.878577	0.796113	0.724029	0.660738	0.563768	0.479654
2.0	0.99	0.966278	0.958671	0.942793	0.854463	0.777199	0.709329	0.605310	0.515031
	0.95	0.907903	0.900149	0.885397	0.803249	0.731299	0.667789	0.570041	0.485021
	0.90	0.898492	0.889015	0.873755	0.792255	0.722067	0.659816	0.563658	0.479654
(e) $n_0 = 5$, $\sigma_1^2 = 10$, $\alpha = 0.01$									
0.25	0.99	1.034020	1.024070	1.006020	0.910000	0.827272	0.754957	0.644424	0.548434
	0.95	1.000160	0.990537	0.973081	0.880206	0.800187	0.730239	0.623325	0.530477
	0.90	1.001330	0.991688	0.974212	0.881230	0.801117	0.731088	0.624049	0.531094
0.50	0.99	1.034020	1.024060	1.006020	0.910002	0.827275	0.754958	0.644425	0.548434
	0.95	1.000160	0.990538	0.973082	0.880210	0.800191	0.730242	0.623326	0.530477
	0.90	1.001330	0.991691	0.974215	0.881235	0.801122	0.731091	0.624050	0.531094
1.0	0.99	1.034000	1.024060	1.006020	0.910011	0.827284	0.754965	0.644427	0.548434
	0.95	1.000160	0.990539	0.973090	0.880227	0.800206	0.730252	0.623329	0.530477
	0.90	1.001330	0.991696	0.974222	0.881256	0.801140	0.731105	0.624054	0.531094

Table 3(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
1.5	0.99	1.033990	1.024050	1.006010	0.910026	0.827299	0.754977	0.644430	0.548434
	0.95	1.000160	0.990530	0.973096	0.880258	0.800231	0.730270	0.623334	0.530477
	0.90	0.997687	0.988087	0.970708	0.878085	0.798259	0.728467	0.621791	0.529162
2.0	0.99	1.033960	1.024040	1.006010	0.910041	0.827320	0.754993	0.644435	0.548434
	0.95	1.000140	0.990534	0.973115	0.880297	0.800272	0.730297	0.623342	0.530477
	0.90	1.001320	0.991680	0.974249	0.881340	0.801218	0.731160	0.624070	0.531094
(f) $n_0 = 5$, $\sigma_1^2 = 10$, $\alpha = 0.05$									
0.25	0.99	0.971044	0.961696	0.944748	0.854578	0.776888	0.708977	0.605176	0.515031
	0.95	0.914467	0.905663	0.889703	0.804785	0.731622	0.667667	0.569913	0.485021
	0.90	0.904350	0.895644	0.879859	0.795881	0.723527	0.660279	0.563607	0.479654
0.50	0.99	0.971046	0.961701	0.944754	0.854583	0.776893	0.708980	0.605177	0.515031
	0.95	0.914481	0.905677	0.889717	0.804797	0.731630	0.667673	0.569915	0.485021
	0.90	0.904375	0.895667	0.879879	0.795898	0.723539	0.660287	0.563609	0.479654
1.0	0.99	0.971060	0.961716	0.944772	0.854605	0.776910	0.708992	0.605180	0.515031
	0.95	0.914540	0.905735	0.889775	0.804846	0.731666	0.667696	0.569921	0.485021
	0.90	0.904462	0.895755	0.879951	0.795966	0.723588	0.660318	0.563618	0.479654
1.5	0.99	0.971084	0.961738	0.944806	0.854643	0.776943	0.709013	0.605186	0.515031
	0.95	0.914656	0.905830	0.889859	0.804926	0.731725	0.667735	0.569932	0.485021
	0.90	0.904614	0.895879	0.880080	0.796076	0.723672	0.660369	0.563631	0.479654
2.0	0.99	0.971105	0.961762	0.944836	0.854688	0.776984	0.709041	0.605194	0.515031
	0.95	0.914727	0.905919	0.889995	0.805020	0.731815	0.667788	0.569946	0.485021
	0.90	0.904803	0.896016	0.880247	0.796209	0.723781	0.660437	0.563650	0.479654
(g) $n_0 = 10$, $\sigma_1^2 = 0.1$, $\alpha = 0.01$									
0.25	0.99	1.012300	1.002620	0.985009	0.891156	0.810151	0.739301	0.631000	0.536982
	0.95	0.989976	0.980582	0.963353	0.871608	0.792373	0.723057	0.617110	0.525149
	0.90	0.988225	0.978909	0.961758	0.870166	0.791075	0.721864	0.616073	0.524258
0.50	0.99	1.008970	1.000380	0.983575	0.891032	0.810319	0.739494	0.631073	0.536982
	0.95	0.983816	0.976669	0.960727	0.871056	0.792400	0.723199	0.617192	0.525149
	0.90	0.979311	0.973069	0.957260	0.868890	0.790733	0.721851	0.616138	0.524258
1.0	0.99	0.974187	0.969982	0.956928	0.876268	0.802543	0.735755	0.630488	0.536982
	0.95	0.962215	0.949654	0.932238	0.848538	0.778326	0.715592	0.615774	0.525149
	0.90	0.968732	0.952314	0.932446	0.842964	0.772468	0.711202	0.613991	0.524258
1.5	0.99	0.984675	0.970124	0.950203	0.854906	0.782205	0.721524	0.627013	0.536982
	0.95	0.994232	0.982935	0.964417	0.853081	0.767390	0.701873	0.610582	0.525149
	0.90	0.997649	0.990611	0.976401	0.866735	0.772261	0.701014	0.607927	0.524258
2.0	0.99	0.998395	0.992866	0.981340	0.879863	0.787407	0.715833	0.621793	0.536982
	0.95	0.999867	0.998394	0.993707	0.912450	0.803521	0.714248	0.606309	0.525149
	0.90	0.999975	0.999559	0.997275	0.935063	0.824419	0.725667	0.605380	0.524258
(h) $n_0 = 10$, $\sigma_1^2 = 0.1$, $\alpha = 0.05$									
0.25	0.99	0.981849	0.972528	0.955456	0.864431	0.785856	0.717110	0.612032	0.520827
	0.95	0.948859	0.939898	0.923414	0.835466	0.759474	0.693004	0.591399	0.503246
	0.90	0.941711	0.932889	0.916677	0.829479	0.754065	0.688050	0.587136	0.499604
0.50	0.99	0.976323	0.969038	0.953092	0.863925	0.785885	0.717283	0.612118	0.520827
	0.95	0.939174	0.932933	0.917911	0.833436	0.758765	0.692880	0.591450	0.503246
	0.90	0.929812	0.923459	0.908203	0.825566	0.752397	0.687400	0.587114	0.499604
1.0	0.99	0.958554	0.944742	0.927594	0.842651	0.772569	0.709959	0.610779	0.520827
	0.95	0.973199	0.950551	0.923780	0.815943	0.740776	0.680472	0.588485	0.503246
	0.90	0.985461	0.966481	0.940140	0.818567	0.735580	0.672677	0.582771	0.499604

Table 3(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
1.5	0.99	0.993829	0.982211	0.962677	0.848692	0.762395	0.696745	0.605760	0.520827
	0.95	0.999359	0.995498	0.986072	0.878780	0.767352	0.683575	0.582779	0.503246
	0.90	0.999892	0.998627	0.994040	0.908782	0.789621	0.692184	0.578186	0.499604
2.0	0.99	0.999850	0.998262	0.993321	0.909252	0.799014	0.709423	0.601610	0.520827
	0.95	1.000000	0.999908	0.999176	0.956063	0.843186	0.727162	0.585056	0.503246
	0.90	1.000000	0.999992	0.999858	0.977004	0.879669	0.752197	0.586444	0.499604
(i) $n_0 = 10$, $\sigma_1^2 = 1$, $\alpha = 0.01$									
0.25	0.99	1.012420	1.002680	0.985017	0.891013	0.810012	0.739203	0.630971	0.536982
	0.95	0.990122	0.980597	0.963321	0.871391	0.792172	0.722919	0.617068	0.525149
	0.90	0.988445	0.978934	0.961686	0.869919	0.790834	0.721697	0.616023	0.524258
0.50	0.99	1.012410	1.002680	0.985035	0.891063	0.810063	0.739236	0.630981	0.536982
	0.95	0.990155	0.980632	0.963379	0.871464	0.792243	0.722969	0.617083	0.525149
	0.90	0.988459	0.978964	0.961732	0.870016	0.790918	0.721755	0.616040	0.524258
1.0	0.99	1.012040	1.002450	0.984981	0.891213	0.810223	0.739358	0.631019	0.536982
	0.95	0.989472	0.980325	0.963304	0.871672	0.792459	0.723138	0.617134	0.525149
	0.90	0.987367	0.978398	0.961483	0.870189	0.791148	0.721941	0.616102	0.524258
1.5	0.99	1.009830	1.001110	0.984029	0.891155	0.810347	0.739487	0.631070	0.536982
	0.95	0.985272	0.977605	0.961334	0.871225	0.792431	0.723243	0.617186	0.525149
	0.90	0.980877	0.974219	0.958252	0.869302	0.790912	0.721934	0.616149	0.524258
2.0	0.99	1.002740	0.996298	0.980467	0.889836	0.809897	0.739396	0.631083	0.536982
	0.95	0.974205	0.969193	0.954623	0.868209	0.791097	0.722684	0.617142	0.525149
	0.90	0.967386	0.963430	0.949106	0.864597	0.788605	0.720924	0.616022	0.524258
(j) $n_0 = 10$, $\sigma_1^2 = 1$, $\alpha = 0.05$									
0.25	0.99	0.981971	0.972518	0.955380	0.864195	0.785631	0.716955	0.611986	0.520827
	0.95	0.948827	0.939693	0.923133	0.835026	0.759113	0.692754	0.591328	0.503246
	0.90	0.941961	0.932894	0.916453	0.828984	0.753620	0.687742	0.587048	0.499604
0.50	0.99	0.981975	0.972522	0.955386	0.864204	0.785639	0.716960	0.611987	0.520827
	0.95	0.948843	0.939708	0.923147	0.835042	0.759125	0.692763	0.591330	0.503246
	0.90	0.941982	0.932912	0.916471	0.829004	0.753636	0.687752	0.587051	0.499604
1.0	0.99	0.981998	0.972538	0.955406	0.864240	0.785669	0.716980	0.611993	0.520827
	0.95	0.948891	0.939763	0.923202	0.835102	0.759172	0.692795	0.591339	0.503246
	0.90	0.942054	0.932985	0.916537	0.829082	0.753698	0.687792	0.587062	0.499604
1.5	0.99	0.981995	0.972577	0.955430	0.864291	0.785715	0.717013	0.612002	0.520827
	0.95	0.948966	0.939838	0.923311	0.835209	0.759251	0.692848	0.591354	0.503246
	0.90	0.942131	0.933066	0.916648	0.829203	0.753793	0.687859	0.587081	0.499604
2.0	0.99	0.981981	0.972557	0.955475	0.864370	0.785780	0.717057	0.612016	0.520827
	0.95	0.949019	0.939859	0.923377	0.835329	0.759351	0.692920	0.591374	0.503246
	0.90	0.942072	0.933064	0.916683	0.829356	0.753913	0.687944	0.587105	0.499604
(k) $n_0 = 10$, $\sigma_1^2 = 10$, $\alpha = 0.01$									
0.25	0.99	1.012430	1.002680	0.985012	0.890999	0.809998	0.739193	0.630968	0.536982
	0.95	0.990118	0.980587	0.963307	0.871366	0.792150	0.722904	0.617064	0.525149
	0.90	0.988440	0.978925	0.961674	0.869889	0.790808	0.721679	0.616018	0.524258
0.50	0.99	1.012430	1.002680	0.985012	0.891004	0.810003	0.739196	0.630969	0.536982
	0.95	0.990122	0.980589	0.963313	0.871374	0.792157	0.722909	0.617065	0.525149
	0.90	0.988446	0.978931	0.961679	0.869899	0.790816	0.721685	0.616019	0.524258
1.0	0.99	1.012420	1.002680	0.985018	0.891023	0.810022	0.739209	0.630973	0.536982
	0.95	0.990118	0.980593	0.963323	0.871406	0.792185	0.722928	0.617071	0.525149
	0.90	0.988463	0.978948	0.961707	0.869936	0.790851	0.721709	0.616026	0.524258

Table 3(*Continued*)

Δ	β	σ_2^2/σ_1^2							
		1	1.5	2	5	10	20	100	∞
1.5	0.99	1.012390	1.002680	0.985018	0.891055	0.810054	0.739232	0.630979	0.536982
	0.95	0.990128	0.980628	0.963367	0.871461	0.792234	0.722962	0.617080	0.525149
	0.90	0.988452	0.978972	0.961736	0.870003	0.790914	0.721749	0.616038	0.524258
2.0	0.99	1.012380	1.002660	0.985038	0.891095	0.810095	0.739262	0.630988	0.536982
	0.95	0.990092	0.980631	0.963379	0.871530	0.792292	0.723006	0.617093	0.525149
	0.90	0.988366	0.978957	0.961756	0.870077	0.790981	0.721804	0.616054	0.524258
(l) $n_0=10$, $\sigma_1^2=10$, $\alpha=0.05$									
0.25	0.99	0.981971	0.972518	0.955380	0.864195	0.785631	0.716955	0.611986	0.520827
	0.95	0.948827	0.939693	0.923133	0.835026	0.759113	0.692754	0.591328	0.503246
	0.90	0.941961	0.932894	0.916453	0.828984	0.753620	0.687742	0.587048	0.499604
0.50	0.99	0.981975	0.972522	0.955386	0.864204	0.785639	0.716960	0.611987	0.520827
	0.95	0.948843	0.939708	0.923147	0.835042	0.759125	0.692763	0.591330	0.503246
	0.90	0.941982	0.932912	0.916471	0.829004	0.753636	0.687752	0.587051	0.499604
1.0	0.99	0.981998	0.972538	0.955406	0.864240	0.785669	0.716980	0.611993	0.520827
	0.95	0.948891	0.939763	0.923202	0.835102	0.759172	0.692795	0.591339	0.503246
	0.90	0.942054	0.932985	0.916537	0.829082	0.753698	0.687792	0.587062	0.499604
1.5	0.99	0.981995	0.972577	0.955430	0.864291	0.785715	0.717013	0.612002	0.520827
	0.95	0.948966	0.939838	0.923311	0.835209	0.759251	0.692848	0.591354	0.503246
	0.90	0.942131	0.933066	0.916648	0.829203	0.753793	0.687859	0.587081	0.499604
2.0	0.99	0.981981	0.972557	0.955475	0.864370	0.785780	0.717057	0.612016	0.520827
	0.95	0.951204	0.942116	0.925560	0.837292	0.761150	0.694558	0.592774	0.504437
	0.90	0.942072	0.933064	0.916683	0.829356	0.753913	0.687944	0.587105	0.499604

in practice, and the gain of D–A over C may be large (up to 50%), while the maximal gain of C over D–A is at most 4%, *we conclude*: *procedure D–A is preferred over procedure C*, based on finite-sample comparisons.

4.5. Comparison of P–S and C procedures

Comparison of procedures P–S and C was given in Taneja and Dudewicz (1993), where it was concluded *C was preferred to P–S*, since with "tight" requirements (high power at a close alternative) C needed less sample size in all cases (a savings of up to 9% in sample size), while when P–S looked better it might be a case of low sample size and due to the n_0 and n_0+1difference seen above in D–A vs. P–S. As this comparison was done in detail by Taneja and Dudewicz (1993), here we just give this brief summary of their conclusions. We also note the consistency of our results and theirs. E.g., for the case $n_0=10$, $\sigma_1^2=1$, $\alpha=0.01$, $\Delta=0.25$, $\beta=0.99$, $\sigma_2^2/\sigma_1^2=5$, the results:

$$\frac{\text{avg.}C}{\text{avg.}P\text{–}S}=0.9312 \quad \text{(Table 2i of Taneja and Dudewicz (1993, p. 468)),}$$

$$\frac{\text{avg.}D\text{–}A}{\text{avg.}P\text{–}S}=0.8525 \quad \text{(Table 1 of this paper),}$$

$$\frac{\text{avg.}D\text{–}A}{\text{avg.}C}=0.8910 \quad \text{(Table 3 of this paper)}$$

are consistent since $0.8525/0.8910=0.9568$ (close to 0.9312).

5. Recommendations

Serious statistical work requires procedures which meet the stated requirements. For the Behrens–Fisher Problem, there are three procedures which meet the desired level and power at a given difference of means:

C due to Chapman (1950),
P–S due to Prokof'yev and Shishkin (1974), and
D–A due to Dudewicz and Ahmed (1999).

We have seen in Section 3 that asymptotically D–A requires sample sizes that are never larger and may be up to 50% smaller, in comparison to each of C and P–S. In Section 4 we saw that for finite samples D–A also may save up to 50% of the sample size, is uniformly better with tight requirements (low level, high power, close alternative), and in any case never needs more than a few % more in the worst case (loose requirements). We conclude that the D–A procedure should be used in practice.

References

Aoshima, M., Hyakutake, H., Dudewicz, E.J., 1996. An asymptotically optimal fixed-width confidence interval for the difference of two normal means. Sequantial Anal. 15, 61–70.

Bather, J., 1996. A conversation with Herman Chernoff. Statist. Sci. 11, 335–350.

Chapman, D.G., 1950. Some two sample tests. Ann. Math. Statist. 21, 601–606.

Dudewicz, E.J., 1976. Introduction to Statistics and Probability. Holt, Rinehart and Winston, New York.

Dudewicz, E.J., Ahmed, S.U., 1998. New exact and asymptotically optimal solution to the Behrens–Fisher Problem, with tables. Amer. J. Math. Management Sci. 18, 359–426.

Dudewicz, E.J., Ahmed, S.U., 1999. New exact and asymptotically optimal heteroscedastic statistical procedures and tables, II. Amer. J. Math. Management Sci. 19, 157–180.

Dudewicz, E.J., Mishra, S.N., 1998. Modern Mathematical Statistics. Wiley, New York.

Dudewicz, E.J., Karian, Z.A., Marshall III, R.J., 1985. Random number generation on microcomputers. Modeling and Simulation on Microcomputers: 1985. The Society for Computer Simulation, La Jolla, CA. pp. 9–14.

Dudewicz, E.J., Chen, P., Taneja, B.K., 1989. Modern Elementary Probability and Statistics, with Statistical Programming in SAS, MINITAB, & BMDP (Second Printing). American Sciences Press, Inc., Columbus, OH.

Karian, Z.A., Dudewicz, E.J., 1999. Modern Statistical, Systems, and GPSS Simulation. second ed.. CRC Press LLC, Boca Raton, FL.

Loève, M., 1963. Probability Theory. third ed.. D. Van Nostrand Company, Inc., Princeton, NJ.

Moore, D.S., McCabe, G.P., 1999. Introduction to the Practice of Statistics. third ed.. W.H. Freeman and Company, New York.

Prokof'yev, V.N., Shishkin, A.D., 1974. Successive classification of normal sets with unknown variances. Radio Engng. Electron. Phys. 19 (2), 141–143.

Taneja, B.K., Dudewicz, E.J., 1993. Exact solutions to the Behrens–Fisher Problem. Multiple Comparisons, Selection, and Applications in Biometry. Marcel Dekker, Inc., New York. pp. 447–477.