Metrika (2018) 81:85–103
https://doi.org/10.1007/s00184-017-0635-y

Three-stage confidence intervals for a linear combination of locations of two negative exponential distributions

Eiichi Isogai[1] · **Chikara Uno**[2]

Received: 24 December 2016 / Published online: 24 November 2017

Abstract Mukhopadhyay and Padmanabhan (Metrika 40:121–128, 1993) considered the construction of fixed-width confidence intervals for the difference of location parameters of two negative exponential distributions via triple sampling when the scale parameters are unknown and unequal. Under the same setting, this paper deals with the problem of fixed-width confidence interval estimation for a linear combination of location parameters, using the above mentioned three-stage procedure.

Keywords Fixed-width interval · Location parameter · Two negative exponentials · Three-stage procedure · Behrens–Fisher situation · Second-order expansions

Mathematics Subject Classification 62L10

1 Introduction

Let $\{X_{i1}, X_{i2}, \ldots\}$ $(i = 1, 2)$ be two independent sequences of random variables where $X_{i1}, X_{i2}, \ldots$ are independent and identically distributed (i.i.d.) random variables with the probability density function (pdf)

$$f(t; \mu_i, \sigma_i) = \frac{1}{\sigma_i} \exp\left(-\frac{t - \mu_i}{\sigma_i}\right) I(t \geq \mu_i).$$

✉ Chikara Uno
uno@math.akita-u.ac.jp

1 Department of Mathematics, Niigata University, Niigata 950-2181, Japan

2 Department of Mathematics, Akita University, Akita 010-8502, Japan

Here $I(\cdot)$ denotes the indicator function of the set $(\cdot)$ and the four parameters $\mu_1,\ \mu_2 \in (-\infty, \infty)$, $\sigma_1,\ \sigma_2 \in (0, \infty)$ are all unknown. This distribution is known as a two-parameter negative exponential distribution (written as $\text{Exp}(\mu_i, \sigma_i)$) and has been widely used in many reliability and life-testing experiments to describe the failure times of complex equipment and some small electrical components. In the paper we consider a linear combination of locations, including the difference of two location parameters. For any given numbers $b_1,\ b_2\ (b_1 b_2 \neq 0)$ and any preassigned numbers $d\,(>0)$ and $0 < \alpha < 1$ we would like to find appropriate sample sizes to construct a confidence interval J for a linear combination $\delta = b_1\mu_1 + b_2\mu_2$ of two location parameters based on the random samples $\{X_{11}, \ldots, X_{1\,n_1}\}$ and $\{X_{21}, \ldots, X_{2\,n_2}\}$ such that $P\{\delta \in J\} \geq 1 - \alpha$ for all fixed values of $\mu_1,\ \mu_2,\ \sigma_1,\ \sigma_2,\ \alpha$ and d and that the length of J is fixed at $2d$. Mukhopadhyay and Padmanabhan (1993) designed three-stage sampling procedures for $\delta = \mu_1 - \mu_2$ and provided the asymptotic second-order expansion of the coverage probability $P\{\delta \in J\} = (1-\alpha) + Ad + o(d)$ as d tends to zero where A is a certain constant. They also gave $P\{\delta \in J\} = (1-\alpha) + o(d)$ with choosing the "fine-tuning" factors. The theory of a three-stage procedure was first established by Hall (1981). Many authors have investigated the sequential estimation problems for the difference of two negative exponential distributions by using purely and/or two-stage procedures, for instance Mukhopadhyay and Hamdy (1984), Mukhopadhyay and Mauromoustakos (1987), Hamdy et al. (1989) and Singh and Chaturvedi (1991). Mukhopadhyay and Zack (2007) dealt with bounded risk estimation of linear combinations of the location and scale parameters. Isogai and Futschik (2010) proposed a purely sequential procedure for a linear combination of locations. Honda (1992) and Yousef et al. (2013) considered the estimation of the mean by a three-stage procedure when the distribution is unspecified.

In the present paper we construct fixed-width confidence intervals for $\delta = b_1\mu_1 + b_2\mu_2$ via the three-stage procedure proposed by Mukhopadhyay and Padmanabhan (1993) when σ_1, σ_2 are unknown and may be unequal, and derive the asymptotic second-order expansion of the coverage probability.

In Sect. 2 we give some preliminaries and design the three-stage procedure. Section 3 provides the main results concerning the asymptotic second-order expansion of the coverage probability. In Sect. 4 we show some simulation results. Section 5 gives all the proofs of the results in Sect. 3.

2 Preliminaries and a three-stage procedure

Having observed $\{X_{i1}, \ldots, X_{in_i}\}$ from the population $\Pi_i\ :\ \text{Exp}(\mu_i, \sigma_i)$, we define for $n_i \geq 2$

$$X_{i\,n_i(1)} = \min\{X_{i1}, \ldots, X_{in_i}\}, \quad U_{i\,n_i} = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - X_{i\,n_i(1)})$$

for $i = 1, 2$ and $X_{i\,n_i(1)}$ and $U_{i\,n_i}$ are the estimators of μ_i and σ_i. Let $\underline{n} = (n_1, n_2)$, b_1 and b_2 $(b_1 b_2 \neq 0)$ be given numbers and $d(>0)$ be a preassigned number. We propose the fixed-width confidence interval of the parameter $\delta = b_1\mu_1 + b_2\mu_2$ with length $2d$

$$J(\underline{n}) = [\,\hat{\delta}(\underline{n}) - d,\ \hat{\delta}(\underline{n}) + d\,], \quad \text{where } \hat{\delta}(\underline{n}) = b_1 X_{1\,n_1(1)} + b_2 X_{2\,n_2(1)}.$$

For a preassigned number $\alpha \in (0, 1)$ we wish to conclude that $P\{\delta \in J(\underline{n})\} \geq 1 - \alpha$.

First of all, we want to find an appropriate sample size C_i which satisfies

$$P\{\delta \in J(\underline{n})\} \geq 1 - \alpha \quad \text{for all } n_i \geq C_i \ (i = 1, 2) \tag{1}$$

for all fixed $\mu_1, \mu_2, \sigma_1, \sigma_2, d$ and α. We will calculate the probability $P\{\delta \in J(\underline{n})\}$. For $i = 1, 2$, let $V_i = |b_i|(X_{i\,n_i(1)} - \mu_i)$ and

$$\beta_i = \frac{n_i}{|b_i|\sigma_i} \ (> 0). \tag{2}$$

V_1 and V_2 are independent and V_i is distributed as $\text{Exp}(0, \beta_i^{-1})$ with pdf $g_i(s) = f(s; 0, \beta_i^{-1})$. First, let us treat the case $b_1 b_2 > 0$. We can easily see

$$\begin{aligned} P\{\delta \in J(\underline{n})\} &= P\{-V_1 - d \leqq V_2 \leqq -V_1 + d\} \\ &= \int_0^d P\{0 < V_2 < -s + d\} g_1(s) ds = 1 - e^{-\beta_1 d} - \beta_1 e^{-\beta_2 d} \int_0^d e^{-(\beta_1 - \beta_2)s} ds, \end{aligned}$$

which provides

$$P\{\delta \in J(\underline{n})\} = \begin{cases} 1 - e^{-\beta_1 d} - (\beta_1 d) e^{-\beta_1 d} & \text{if } \beta_1 = \beta_2, \\ 1 - (\beta_1 d - \beta_2 d)^{-1}\{(\beta_1 d) e^{-\beta_2 d} - (\beta_2 d) e^{-\beta_1 d}\} & \text{if } \beta_1 \neq \beta_2. \end{cases}$$

Let $b_1 b_2 < 0$. By the argument similar to above we get

$$\begin{aligned} P\{\delta \in J(\underline{n})\} &= P\{V_1 - d \leqq V_2 \leqq V_1 + d\} \\ &= 1 - \left\{\beta_2(\beta_1 + \beta_2)^{-1} e^{-\beta_1 d} + \beta_1(\beta_1 + \beta_2)^{-1} e^{-\beta_2 d}\right\}. \end{aligned}$$

Thus, utilizing the indicator function $I(\cdot)$, we have the following lemma.

Lemma 1 *For any fixed* $\underline{n} = (n_1, n_2)$ *with* $n_i \geq 2$ $(i = 1, 2)$ *we have*

$$P\{\delta \in J(\underline{n})\} = \begin{cases} 1 - \left\{\beta_2(\beta_1 + \beta_2)^{-1} e^{-\beta_1 d} + \beta_1(\beta_1 + \beta_2)^{-1} e^{-\beta_2 d}\right\} \\ \quad \textit{for } b_1 b_2 < 0, \\ 1 - e^{-\beta_1 d} - (\beta_1 d) e^{-\beta_1 d} \\ \quad + (\beta_1 d) e^{-\beta_1 d}\left\{1 + ((\beta_2 - \beta_1) d)^{-1}(e^{-(\beta_2 - \beta_1) d} - 1)\right\} I(\beta_1 \neq \beta_2) \\ \quad \textit{for } b_1 b_2 > 0. \end{cases}$$

Let any $\alpha \in (0, 1)$ be fixed. For $b_1 b_2 < 0$ we get

$$\beta_2(\beta_1 + \beta_2)^{-1} e^{-\beta_1 d} + \beta_1(\beta_1 + \beta_2)^{-1} e^{-\beta_2 d} \leq \alpha$$

if $e^{-\beta_i d} \leq \alpha$ $(i = 1, 2)$ which is equivalent to $n_i \geq a|b_i|\sigma_i/d \equiv C_i$ with $a = \ln \alpha^{-1}$. Hence from Lemma 1 we get $P\{\delta \in J(\underline{n})\} \geq 1 - \alpha$ for all $n_i \geq C_i$ $(i = 1, 2)$ which gives (1). Next we consider the case $b_1 b_2 > 0$. Let $u(x) = (1 + x)e^{-x}$ for $x > 0$. We can easily show that the function $u(x)$ is strictly decreasing on $(0, \infty)$ with $u(0) = 1$ and $u(+\infty) = 0$ and hence there exists a unique solution $a_0 (> 0)$ satisfying that $u(a_0) = \alpha$. Let us define the function $h(x, y)$ on $\mathbb{R}_+^2$ as

$$h(x, y) = \begin{cases} (x - y)^{-1}(xe^{-y} - ye^{-x}) & \text{when } x \neq y, \\ u(x) & \text{when } x = y, \end{cases}$$

where $\mathbb{R}_+ = (0, \infty)$. After some calculations we have

$$h(x, y) \leq h(a_0, y) \leq u(a_0) = \alpha \quad \text{for all } x \geq a_0 \text{ and } \; y \geq a_0.$$

It follows from Lemma 1 that $P\{\delta \in J(\underline{n})\} = 1 - h(\beta_1 d, \beta_2 d)$, which, together with the above inequality, yields that $P\{\delta \in J(\underline{n})\} \geq 1 - \alpha$ if $\beta_i d \geq a_0$ for $i = 1, 2$. Let $C_i = a_0|b_i|\sigma_i/d$. From (2) we get that $\beta_i d \geq a_0$ if $n_i \geq C_i$ for $i = 1, 2$. Therefore we have that $P\{\delta \in J(\underline{n})\} \geq 1 - \alpha$ if $n_i \geq C_i$ for $i = 1, 2$, which gives (1). We call C_i $(i = 1, 2)$ the optimal fixed sample size. From the above results, we obtain

Proposition 1 *Let*

$$C_i = \frac{a_*|b_i|\sigma_i}{d} \; (i = 1, 2) \quad \textit{and} \tag{3}$$

$$a_* = \begin{cases} a = \ln \alpha^{-1} & \textit{for } \; b_1 b_2 < 0, \\ a_0 \textit{ with } (1 + a_0)e^{-a_0} = \alpha & \textit{for } \; b_1 b_2 > 0. \end{cases} \tag{4}$$

Then for all $n_i \geq C_i$ $(i = 1, 2)$ *we have*

$$P\{\delta \in J(\underline{n})\} \geq 1 - \alpha \quad \textit{for all fixed } \mu_1, \mu_2, \sigma_1, \sigma_2, d \textit{ and } \alpha,$$

where $\underline{n} = (n_1, n_2)$ *with* $n_i \geq 2$ $(i = 1, 2)$.

Since the optimal fixed sample size C_i of (3) is unknown, we will define a three-stage procedure which is similar to that designed by Mukhopadhyay and Padmanabhan (1993). First we take the pilot sample $X_{i1}, \ldots, X_{im}$ and calculate $X_{i\,m(1)}$ and $U_{i\,m}$ for $i = 1, 2$, where the starting sample size $m (\geq 2)$ satisfies $m = O(d^{-1/r})$ for some $r > 1$ as $d \to 0$. We also choose and fix any two numbers $\rho_i \in (0, 1)$ $(i = 1, 2)$. Let any $d (> 0)$ be fixed and define

$$T_i = T_i(d) = \max\left\{m, \; \langle \rho_i a_* |b_i| d^{-1} U_{i\,m} \rangle + 1\right\} \quad \text{for } i = 1, 2. \tag{5}$$

If $T_i > m$, then we take the second sample $X_{i\,m+1}, \ldots, X_{iT_i}$ for $i = 1, 2$. Using the combined sample $X_{i1}, \ldots, X_{iT_i}$, we calculate $X_{i\,T_i(1)}$ and $U_{i\,T_i}$ and define

$$N_i = N_i(d) = \max\left\{T_i, \; \langle a_* |b_i| d^{-1} U_{i\,T_i} \rangle + 1\right\} \quad \text{for } i = 1, 2, \tag{6}$$

where $\langle x \rangle$ stands for the largest integer less than x. If $N_i > T_i$, then we take the third sample $X_{i\,T_i+1}, \ldots, X_{i\,N_i}$ for $i = 1, 2$. Using all the combined sample $X_{i1}, \ldots, X_{i\,N_i}$ $(i = 1, 2)$, we construct a confidence interval of $\delta = b_1\mu_1 + b_2\mu_2$ as

$$J(\underline{N}) = [\,\hat{\delta}(\underline{N}) - d,\ \hat{\delta}(\underline{N}) + d\,], \tag{7}$$

where $\underline{N} = (N_1, N_2)$ and $\hat{\delta}(\underline{N}) = b_1 X_{1\,N_1(1)} + b_2 X_{2\,N_2(1)}$.

3 Main results

In this section we will derive the asymptotic second-order expansions of the expected sample size $E(N_i)$ for (6) and coverage probability $P\{\delta \in J(\underline{N})\}$ for (7). Theorem 1 gives the asymptotic second-order expansion of $E(N_i)$ for $i = 1, 2$.

Theorem 1 *We have*

$$E(N_i) = C_i + \eta_i + o(1) \quad \text{as } d \to 0,$$

where $\eta_i = \frac{1}{2} - \rho_i^{-1} \in (-\infty, -\frac{1}{2})$.

The following theorem shows the asymptotic second-order expansion of the coverage probability $P\{\delta \in J(\underline{N})\}$.

Theorem 2 *As* $d \to 0$ *we have*

$$P\{\delta \in J(\underline{N})\} = 1 - \alpha + A_\alpha d + o(d),$$

where

$$(0 >)\ A_\alpha = \begin{cases} \dfrac{1}{4}\alpha \left\{ \displaystyle\sum_{i=1}^{2} \left(1 - (a+3)\rho_i^{-1}\right) (|b_i|\sigma_i)^{-1} \right\} & \text{for } b_1 b_2 < 0, \\[2ex] a_0 e^{-a_0} \left\{ \displaystyle\sum_{i=1}^{2} \left(\dfrac{1}{4} - \dfrac{1}{6}(a_0+3)\rho_i^{-1}\right) (|b_i|\sigma_i)^{-1} \right\} & \text{for } b_1 b_2 > 0. \end{cases}$$

Remark 1 Theorems 1 and 2 generalize the results of Mukhopadhyay and Padmanabhan (1993) for estimating the difference $\delta = \mu_1 - \mu_2$ $(b_1 = 1, b_2 = -1)$.

Remark 2 The approximation to $P\{\delta \in J(\underline{N})\}$ becomes better as ρ_i increases, since the absolute value of A_α gets smaller as ρ_i increases.

Remark 3 When $b_1 b_2 > 0$, one can consider the confidence interval

$$J^*(\underline{n}) = \begin{cases} \left[\hat{\delta}(\underline{n}) - d,\ \hat{\delta}(\underline{n})\right] & \text{for } b_1 > 0,\ b_2 > 0, \\ \left[\hat{\delta}(\underline{n}),\ \hat{\delta}(\underline{n}) + d\right] & \text{for } b_1 < 0,\ b_2 < 0 \end{cases}$$

Table 1 $\delta = \frac{1}{2}(\mu_1 + \mu_2) = 1$ for $\mu_1 = 2, \sigma_1 = 1$ and $\mu_2 = 0, \sigma_2 = 1$

$\alpha = 0.05$		$d = 0.06$	s.e.	$d = 0.03$	s.e.
	C_1	39.532167		79.064333	
	C_2	39.532167		79.064333	
$\rho_1 = 0.4$	$E(T_1)$	16.358697	0.006378	32.121238	0.009982
$\rho_2 = 0.4$	$E(T_2)$	16.362222	0.006379	32.133641	0.009981
$\eta_1 = -2$	$E(N_1)$	37.114968	0.010976	76.670732	0.015255
$\eta_2 = -2$	$E(N_2)$	37.099661	0.010969	76.690959	0.015243
	$E(\hat{\delta}(\underline{N}))$	1.030393	0.000025	1.013693	0.000010
	$P\{\delta \in J(\underline{N})\}$	0.896762	0.000304	0.927890	0.000259
	$A_\alpha d$	−0.029501		−0.014750	
$\rho_1 = 0.6$	$E(T_1)$	24.224945	0.009669	47.947887	0.014990
$\rho_2 = 0.6$	$E(T_2)$	24.217171	0.009659	47.931609	0.014982
$\eta_1 = -1.167$	$E(N_1)$	38.252309	0.009479	77.820806	0.012495
$\eta_2 = -1.167$	$E(N_2)$	38.255471	0.009476	77.805422	0.012483
	$E(\hat{\delta}(\underline{N}))$	1.028458	0.000023	1.013227	0.000010
	$P\{\delta \in J(\underline{N})\}$	0.916636	0.000276	0.937332	0.000242
	$A_\alpha d$	−0.018841		−0.009421	

with fixed-width d. By the same arguments as Lemma 1, we have

$$P\{\delta \in J^*(\underline{n})\} = P\{0 \leqq V_1 + V_2 \leqq d\} = P\{-V_1 - d \leqq V_2 \leqq -V_1 + d\}$$

and hence, it holds for all $n_i \geq C_i$ $(i = 1, 2)$ that $P\{\delta \in J^*(\underline{n})\} \geq 1 - \alpha$ for all fixed $\mu_1, \mu_2, \sigma_1, \sigma_2, d$ and α. Therefore, when $b_1 b_2 > 0$, the length of the confidence interval (7) is indeed considered as a half length.

4 Simulation results

We shall present some simulation results which were carried out by means of Borland C++. We consider two cases when $\delta = \frac{1}{2}(\mu_1 + \mu_2)$ $(b_1 b_2 > 0)$ and $\delta = \mu_1 - \mu_2$ $(b_1 b_2 < 0)$. We choose $\rho_1 = \rho_2 = 0.4,\ 0.6$ in (5) and take $\alpha = 0.05$ $(1 - \alpha = 0.95)$ in Tables 1, 2, 5 and 6 and $\alpha = 0.10,\ 0.01$ in Tables 3 and 4, respectively. About (5) and (6), we have $a_* = a_0 = 4.74386$ with $(1 + a_0)e^{-a_0} = 0.05$ for $b_1 b_2 > 0$ and $a_* = a = \ln(0.05)^{-1} = 2.99573$ for $b_1 b_2 < 0$. From Taylor's expansion and calculus, one can find an approximation $\tilde{a}_0$ to a_0 such as

$$a_0 \risingdotseq \tilde{a}_0 = a + \frac{a}{a-1} \ln a \quad \text{with } a = \ln \alpha^{-1}.$$

For $\alpha = 0.05$, we have $\tilde{a}_0 = 4.64269$ with $(1 + \tilde{a}_0)e^{-\tilde{a}_0} = 0.05435$. For $\alpha = 0.1$, we also have $a_0 = 3.88972$ with $(1 + a_0)e^{-a_0} = 0.1$ and $\tilde{a}_0 = 3.77691$ with $(1 +$

Table 2 $\delta = \frac{1}{2}(\mu_1 + \mu_2) = 1$ for $\mu_1 = 2, \sigma_1 = 2$ and $\mu_2 = 0, \sigma_2 = 1$

$\alpha = 0.05$		$d = 0.06$	s.e.	$d = 0.03$	s.e.
	C_1	79.064333		158.128667	
	C_2	39.532167		79.064333	
$\rho_1 = 0.4$	$E(T_1)$	32.120166	0.012905	63.750527	0.020002
$\rho_2 = 0.4$	$E(T_2)$	16.369592	0.006384	32.136720	0.009987
$\eta_1 = -2$	$E(N_1)$	76.435833	0.015986	155.800940	0.021296
$\eta_2 = -2$	$E(N_2)$	37.129127	0.010973	76.703653	0.015235
	$E(\hat{\delta}(\underline{N}))$	1.029041	0.000024	1.013367	0.000010
	$P\{\delta \in J(\underline{N})\}$	0.910766	0.000285	0.934339	0.000248
	$A_\alpha d$	−0.022126		−0.011063	
$\rho_1 = 0.6$	$E(T_1)$	47.949343	0.019381	95.299930	0.029960
$\rho_2 = 0.6$	$E(T_2)$	24.223336	0.009662	47.925842	0.014983
$\eta_1 = -1.167$	$E(N_1)$	78.222371	0.013653	157.126273	0.017727
$\eta_2 = -1.167$	$E(N_2)$	38.237131	0.009465	77.811653	0.012503
	$E(\hat{\delta}(\underline{N}))$	1.027530	0.000022	1.013062	0.000010
	$P\{\delta \in J(\underline{N})\}$	0.926187	0.000261	0.940754	0.000236
	$A_\alpha d$	−0.014131		−0.007065	

$\tilde{a}_0)e^{-\tilde{a}_0} = 0.10936$. Further, for $\alpha = 0.01$, we have $a_0 = 6.63835$ with $(1+a_0)e^{-a_0} = 0.01$ and $\tilde{a}_0 = 6.55596$ with $(1 + \tilde{a}_0)e^{-\tilde{a}_0} = 0.01074$. In all tables below, the three-stage procedure defined by (5) and (6) was carried out with 1,000,000 independent replications under $d = 0.06$ (moderate) and $d = 0.03$ (sufficiently small). In each table, $E(T_i)$, $E(N_i)$, $E(\hat{\delta}(\underline{N}))$ and $P\{\delta \in J(\underline{N})\}$ stand for the averages of 1,000,000 independent replications and "s.e." stands for each standard error. Let the size of the pilot sample be $m = \langle d^{-2/3} \rangle + 1$ for each population. Thus, $m = 7$ for $d = 0.06$ and $m = 11$ for $d = 0.03$.

For estimating $\delta = \frac{1}{2}(\mu_1 + \mu_2)$, we take $\text{Exp}(2, 1)$ as Π_1 and $\text{Exp}(0, 1)$ as Π_2 in Table 1, where the variances are equal and also take $\text{Exp}(2, 2)$ as Π_1 and $\text{Exp}(0, 1)$ as Π_2 in Table 2, where the variances are unequal. In both Tables 1 and 2, we estimate $\delta = 1$ with $\alpha = 0.05$ and the optimal fixed sample sizes C_1 and C_2 are calculated by (3) with $b_1 = b_2 = 0.5$ and $a_* = a_0 = 4.74386$. We have from Theorem 1

$$E(N_i) - C_i \approx \eta_i, \tag{8}$$

where $\eta_i = -2$ for $\rho_i = 0.4$ and $\eta_i = -1.167$ for $\rho_i = 0.6$. It seems from Tables 1 and 2 that each N_i underestimates C_i as the above approximation. We also have from Theorem 2

$$P\{\delta \in J(\underline{N})\} - (1 - \alpha) \approx A_\alpha d. \tag{9}$$

It also seems from Tables 1 and 2 that the coverage probabilities $P\{\delta \in J(\underline{N})\}$ are less than 0.95, for $A_\alpha d < 0$. However, as d becomes sufficiently small ($d = 0.03$), the coverage probabilities $P\{\delta \in J(\underline{N})\}$ get closer to 0.95 in both tables. In Tables 3

Table 3 $\delta = \frac{1}{2}(\mu_1 + \mu_2) = 1$ for $\mu_1 = 2, \sigma_1 = 1$ and $\mu_2 = 0, \sigma_2 = 1$

$\alpha = 0.1$		$d = 0.06$	s.e.	$d = 0.03$	s.e.
	C_1	32.414333		64.828667	
	C_2	32.414333		64.828667	
$\rho_1 = 0.4$	$E(T_1)$	13.572373	0.005135	26.430996	0.008169
$\rho_2 = 0.4$	$E(T_2)$	13.567443	0.005135	26.431895	0.008166
$\eta_1 = -2$	$E(N_1)$	30.219554	0.009540	62.445478	0.013771
$\eta_2 = -2$	$E(N_2)$	30.216410	0.009543	62.459390	0.013772
	$E(\hat{\delta}(\underline{N}))$	1.037604	0.000031	1.016986	0.000013
	$P\{\delta \in J(\underline{N})\}$	0.828111	0.000377	0.865192	0.000342
	$A_\alpha d$	− 0.050034		− 0.025017	
$\rho_1 = 0.6$	$E(T_1)$	19.955724	0.007895	39.400128	0.012280
$\rho_2 = 0.6$	$E(T_2)$	19.969642	0.007912	39.406992	0.012302
$\eta_1 = -1.167$	$E(N_1)$	31.077383	0.008467	63.536216	0.011328
$\eta_2 = -1.167$	$E(N_2)$	31.096707	0.008464	63.547090	0.011330
	$E(\hat{\delta}(\underline{N}))$	1.035697	0.000029	1.016316	0.000012
	$P\{\delta \in J(\underline{N})\}$	0.848593	0.000358	0.879527	0.000326
	$A_\alpha d$	− 0.031765		− 0.015883	

Table 4 $\delta = \frac{1}{2}(\mu_1 + \mu_2) = 1$ for $\mu_1 = 2, \sigma_1 = 1$ and $\mu_2 = 0, \sigma_2 = 1$

$\alpha = 0.01$		$d = 0.06$	s.e.	$d = 0.03$	s.e.
	C_1	55.319583		110.639167	
	C_2	55.319583		110.639167	
$\rho_1 = 0.4$	$E(T_1)$	22.642284	0.009009	44.721256	0.013999
$\rho_2 = 0.4$	$E(T_2)$	22.641949	0.009016	44.752411	0.013993
$\eta_1 = -2$	$E(N_1)$	52.726017	0.013360	108.268277	0.017973
$\eta_2 = -2$	$E(N_2)$	52.733431	0.013352	108.264365	0.017937
	$E(\hat{\delta}(\underline{N}))$	1.020788	0.000017	1.009523	0.000007
	$P\{\delta \in J(\underline{N})\}$	0.969137	0.000173	0.983566	0.000127
	$A_\alpha d$	− 0.007855		− 0.003928	
$\rho_1 = 0.6$	$E(T_1)$	33.685548	0.013554	66.834883	0.020986
$\rho_2 = 0.6$	$E(T_2)$	33.677759	0.013541	66.881905	0.020985
$\eta_1 = -1.167$	$E(N_1)$	54.221485	0.011316	109.515043	0.014776
$\eta_2 = -1.167$	$E(N_2)$	54.195915	0.011316	109.489136	0.014778
	$E(\hat{\delta}(\underline{N}))$	1.019543	0.000015	1.009299	0.000007
	$P\{\delta \in J(\underline{N})\}$	0.978853	0.000144	0.986617	0.000115
	$A_\alpha d$	− 0.005063		− 0.002531	

and 4, we carried out simulations for $\alpha = 0.1$ with $a_0 = 3.88972$ and $\alpha = 0.01$ with $a_0 = 6.63835$, respectively, under the same settings in Table 1. The results in Tables 3 and 4 behave as in Table 1.

Table 5 $\delta = \mu_1 - \mu_2 = 1$ for $\mu_1 = 2, \sigma_1 = 1$ and $\mu_2 = 1, \sigma_2 = 1$

$\alpha = 0.05$		$d = 0.06$	s.e.	$d = 0.03$	s.e.
	C_1	49.928871		99.857742	
	C_2	49.928871		99.857742	
$\rho_1 = 0.4$	$E(T_1)$	20.484368	0.008127	40.432489	0.012608
$\rho_2 = 0.4$	$E(T_2)$	20.468928	0.008115	40.429592	0.012616
$\eta_1 = -2$	$E(N_1)$	47.349820	0.012631	97.522519	0.017090
$\eta_2 = -2$	$E(N_2)$	47.345715	0.012619	97.488300	0.017068
	$E(\hat{\delta}(\underline{N}))$	0.999963	0.000039	1.000016	0.000016
	$P\{\delta \in J(\underline{N})\}$	0.910361	0.000286	0.934870	0.000247
	$A_\alpha d$	− 0.020984		− 0.010492	
$\rho_1 = 0.6$	$E(T_1)$	30.439973	0.012213	60.370934	0.018905
$\rho_2 = 0.6$	$E(T_2)$	30.448254	0.012216	60.402815	0.018933
$\eta_1 = -1.167$	$E(N_1)$	48.734646	0.010734	98.687033	0.014019
$\eta_2 = -1.167$	$E(N_2)$	48.751315	0.010728	98.704855	0.014007
	$E(\hat{\delta}(\underline{N}))$	0.999974	0.000035	1.000000	0.000015
	$P\{\delta \in J(\underline{N})\}$	0.926278	0.000261	0.941096	0.000235
	$A_\alpha d$	− 0.013489		− 0.006745	

In Tables 5 and 6, we consider the estimation of $\delta = \mu_1 - \mu_2$, where our three-stage procedure defined by (5) and (6) concides with the one of Mukhopadhyay and Padmanabhan (1993). We take $\text{E}_{\text{XP}}(2, 1)$ as Π_1 and $\text{E}_{\text{XP}}(1, 1)$ as Π_2 in Table 5 and also take $\text{E}_{\text{XP}}(2, 2)$ as Π_1 and $\text{E}_{\text{XP}}(1, 1)$ as Π_2 in Table 6. In both tables, we estimate $\delta = 1$ and the optimal fixed sample sizes C_1 and C_2 are calculated by (3) with $b_1 = 1$, $b_2 = -1$ and $a_* = a = 2.99573$. The simulation results in Tables 5 and 6 also seem to have the trends as above including the properties (8) and (9). Throughout these tables, we can verify Remark 2 for ρ_i.

Hamdy (1997), Hamdy et al. (2015) and Son et al. (1997) treated theories on the type II errors of sequential procedures and gave simulation results for one-sample case. For the present two-sample case, it is still open.

5 Proofs of Theorems 1 and 2

In this section we will give the proofs of two theorems in Sect. 3. Let $\mu_1' = b_1\mu_1$, $\mu_2' = -b_2\mu_2$, $\sigma_i' = |b_i|\sigma_i$ for $b_1b_2 < 0$ and $\mu_i' = b_i\mu_i$, $\sigma_i' = |b_i|\sigma_i$ $(i = 1, 2)$ for $b_1b_2 > 0$. Then without any loss of generality δ can be written as

$$\delta = \begin{cases} \mu_1 - \mu_2 & \text{when} \quad b_1b_2 < 0, \\ \mu_1 + \mu_2 & \text{when} \quad b_1b_2 > 0. \end{cases}$$

Throughout this section we use this form. Thus, $b_1 = b_2 = 1$ for both cases.

Table 6 $\delta = \mu_1 - \mu_2 = 1$ for $\mu_1 = 2, \sigma_1 = 2$ and $\mu_2 = 1, \sigma_2 = 1$

$\alpha = 0.05$		$d = 0.06$	s.e.	$d = 0.03$	s.e.
	C_1	99.857742		199.715485	
	C_2	49.928871		99.857742	
$\rho_1 = 0.4$	$E(T_1)$	40.428351	0.016293	80.346714	0.025218
$\rho_2 = 0.4$	$E(T_2)$	20.478699	0.008111	40.423224	0.012618
$\eta_1 = -2$	$E(N_1)$	97.236447	0.017891	197.349920	0.023874
$\eta_2 = -2$	$E(N_2)$	47.361480	0.012616	97.472558	0.017082
	$E(\hat{\delta}(\underline{N}))$	0.998088	0.000036	0.999648	0.000015
	$P\{\delta \in J(\underline{N})\}$	0.921438	0.000269	0.939450	0.000239
	$A_\alpha d$	− 0.015738		− 0.007869	
$\rho_1 = 0.6$	$E(T_1)$	60.431824	0.024456	120.278800	0.037800
$\rho_2 = 0.6$	$E(T_2)$	30.458167	0.012218	60.403670	0.018917
$\eta_1 = -1.167$	$E(N_1)$	99.262021	0.015482	198.918548	0.019982
$\eta_2 = -1.167$	$E(N_2)$	48.751595	0.010723	98.686575	0.014023
	$E(\hat{\delta}(\underline{N}))$	0.998742	0.000033	0.999771	0.000015
	$P\{\delta \in J(\underline{N})\}$	0.933273	0.000250	0.943806	0.000230
	$A_\alpha d$	− 0.010117		− 0.005058	

Let $Y_{i\,2}, Y_{i\,3}, \ldots$ be i.i.d. random variables according to $\mathrm{Exp}(0, \sigma_i)$ and $Y_{1\,j}$'s and $Y_{2\,j}$'s be independent. Also let $\{X_{1j}, X_{2j} : j \geq 1\}$ and $\{Y_{1j}, Y_{2j} : j \geq 2\}$ be independent. Set $\overline{Y}_{i\,n} = \sum_{j=2}^{n} Y_{i\,j}/(n-1)$ for $n \geq 2$ $(i = 1, 2)$. From Lemma 6.1 of Lombard and Swanepoel (1978) $\{(n-1)U_{i\,n},\ n \geq 2\}$ and $\{(n-1)\overline{Y}_{i\,n},\ n \geq 2\}$ are identically distributed. Let us define for $i = 1, 2$

$$R_i = \max\left\{m,\ \langle \rho_i a_* d^{-1} \overline{Y}_{i\,m} \rangle + 1\right\} \quad \text{and} \quad S_i = \max\left\{R_i,\ \langle a_* d^{-1} \overline{Y}_{i\,R_i} \rangle + 1\right\}.$$

Then we get the following lemma.

Lemma 2 *For* $i = 1, 2$ (T_i, N_i) *and* (R_i, S_i) *are identically distributed, and* S_1 *and* S_2 *are independent.*

Proof Let any $m \leq k \leq n$ be fixed. Then

$$\begin{aligned} &P\{T_i \leq k,\ N_i \leq n\} \\ &= \sum_{t=m}^{k} P\{t = \max\{m,\ \langle \rho_i a_* d^{-1} U_{i\,m} \rangle + 1\},\ \max\{t,\ \langle a_* d^{-1} U_{i\,t} \rangle + 1\} \leq n\} \\ &= \sum_{t=m}^{k} P\{t = \max\{m,\ \langle \rho_i a_* d^{-1} \overline{Y}_{i\,m} \rangle + 1\} = R_i,\ \max\{t,\ \langle a_* d^{-1} \overline{Y}_{i\,t} \rangle + 1\} \leq n\} \\ &= P\{R_i \leq k,\ S_i \leq n\}, \end{aligned}$$

which shows that (T_i, N_i) and (R_i, S_i) are identically distributed. It is obvious that S_1 and S_2 are independent. This completes the proof. □

Lemma 2 implies that we can use results of Mukhopadhyay (1990) for (10) below, from which we can derive the desired results for (6) and (7).

Let $Y'_{ij} = Y_{ij}/\sigma_i$ and $\lambda_i = a_* \sigma_i d^{-1} = C_i$. Then $Y'_{i2}, Y'_{i3}, \ldots$ are i.i.d. random variables according to $\text{E}_{\text{XP}}(0, 1)$, and R_i and S_i can be rewritten as

$$R_i = \max\left\{m,\ \langle \rho_i \lambda_i \overline{Y'}_{im} \rangle + 1\right\} \quad \text{and} \quad S_i = \max\left\{R_i,\ \langle \lambda_i \overline{Y'}_{iR_i} \rangle + 1\right\}. \tag{10}$$

From Theorems 2 and 3 of Mukhopadhyay (1990) we have

Lemma 3 *Let* $i = 1, 2$.

(i) For $k = 1, 2, 3, \ldots$

$$E(S_i^k) = C_i^k + \frac{1}{2} k C_i^{k-1} \{(k-3) + \rho_i\}/\rho_i + o(C_i^{k-1}) \quad \textit{and}$$
$$E(S_i) = C_i + \eta_i + o(1) \quad \textit{as } d \to 0.$$

(ii) Let $\tilde{S}_i = C_i^{-1/2}(S_i - C_i)$. *Then*

$$\tilde{S}_i \xrightarrow{\mathcal{D}} \sqrt{\rho_i^{-1}}\, Z_i \textit{ as } d \to 0$$

and for each $p \geq 1$ $\{\tilde{S}_i^p,\ 0 < d \leq d_0\}$ *is uniformly integrable for some* $d_0 > 0$, *where* Z_1 *and* Z_2 *are independent and identically distributed random variables according to the standard normal distribution and "*$\xrightarrow{\mathcal{D}}$*" stands for convergence in distribution.*

The uniform integrability of $\{\tilde{S}_i^p,\ 0 < d \leq d_0\}$ for each $p \geq 1$ in Lemma 3 will be shown in "Appendix".

Proof of Theorem 1 For both cases Theorem 1 is an immediate consequence of Lemmas 2 and 3. □

Proof of Theorem 2 From the point of view of Lemma 1 we need to show it separately.

Case 1 $b_1 b_2 < 0$. Thus $\delta = \mu_1 - \mu_2$. In the proof of Theorem 1 of Mukhopadhyay and Padmanabhan (1993) they provided the equation

$$\begin{aligned} P\{\delta \in J(\underline{N})\} &= (1-\alpha) + \frac{1}{2} d e^{-a} \sum_{i=1}^{2} \sigma_i^{-1} E(S_i - C_i) \\ &\quad - \frac{1}{4} d e^{-a} (1+a) \sum_{i=1}^{2} \sigma_i^{-1} E(\tilde{S}_i^2) + \frac{1}{2} d e^{-a} (\sigma_1 \sigma_2)^{-1/2} E(\tilde{S}_1 \tilde{S}_2) + E(K) \\ &\equiv (1-\alpha) + E(K_1) - E(K_2) + E(K_3) + E(K), \quad \text{say,} \end{aligned}$$

where K is similarly given as Mukhopadhyay and Padmanabhan (1993). Lemmas 2 and 3 yield

$$E(K_1) - E(K_2) = \frac{1}{4}\alpha \left\{ \sum_{i=1}^{2} \left(2\eta_i - (a+1)\rho_i^{-1} \right) \sigma_i^{-1} \right\} d + o(d) \text{ and } E(K_3) = o(d).$$

Mukhopadhyay and Padmanabhan (1993) showed that $E(K) = o(d)$. Therefore, recalling $\sigma_i = |b_i|\sigma_i$, the above results give Theorem 2.

Case 2 $b_1 b_2 > 0$. Thus $\delta = \mu_1 + \mu_2$. Lemmas 4, 5 and 7 (which are given later) imply the desired result. Therefore the proof of Theorem 2 is complete. □

Let us give Lemmas 4, 5, 6 and 7. We introduce the following real valued functions of $(x, y) \in \mathbb{R}^2$:

$$g(x) = \begin{cases} 1 + x^{-1}(e^{-x} - 1) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases} \tag{11}$$
$$A(x) = 1 - e^{-a_0 x} - a_0 x e^{-a_0 x} \text{ and } B(x, y) = a_0 x e^{-a_0 x} g(a_0(y - x)).$$

Throughout the rest of this section let $Q_i = S_i / C_i$ for $i = 1, 2$. Lemma 4 shows an expression of the coverage probability.

Lemma 4 *Let* $b_1 b_2 > 0$. *Then we have*

$$P\{\delta \in J(\underline{N})\} = E\left[A(Q_1)\right] + E\left[B(Q_1, Q_2)\right]. \tag{12}$$

Proof Lemma 1 implies

$$P\{\delta \in J(\underline{n})\} = A(n_1/C_1) + B(n_1/C_1, n_2/C_2). \tag{13}$$

Since $\{X_{1n_1(1)}, X_{2n_2(1)}\}$ and $\{U_{i2}, \ldots, U_{in_i}\ i = 1, 2\}$ are independent, two events $\{\delta \in J(\underline{n})\}$ and $\{\underline{N} = \underline{n}\}$ for any fixed $\underline{n}$ are also independent. Hence from (13) and Lemma 2 we get

$$\begin{aligned} P\{\delta \in J(\underline{N})\} &= \sum_{n_1 \geq m} \sum_{n_2 \geq m} P\{\delta \in J(\underline{n})\} P\{S_1 = n_1\} P\{S_2 = n_2\} \\ &= E\left[A(Q_1)\right] + E\left[B(Q_1, Q_2)\right], \end{aligned}$$

which leads to the lemma. Thus the proof is complete. □

We will evaluate each quantity in (12).

Lemma 5 *We have as* $d \to 0$

$$E\left[A(Q_1)\right] = (1 - \alpha) + a_0 e^{-a_0}\{\eta_1 + (1 - a_0)\rho_1^{-1}/2\}\sigma_1^{-1} d + o(d).$$

Proof Let $h(x) = e^{-a_0 x} + a_0 x e^{-a_0 x}$ for x. Then by using Taylor's expansion around one and $(1 + a_0)e^{-a_0} = \alpha$ we get

$$h(x) = \alpha - a_0^2 e^{-a_0}(x-1) - \frac{1}{2}a_0^2(1 - a_0 w_1)e^{-a_0 w_1}(x-1)^2,$$

where w_1 satisfies that $|w_1 - 1| < |x - 1|$. Using $\tilde{S}_1 = C_1^{-1/2}(S_1 - C_1)$ in Lemma 3, we have

$$\begin{aligned} E[A(Q_1)] &= 1 - E[h(Q_1)] \\ &= (1-\alpha) + a_0^2 e^{-a_0} C_1^{-1} E(S_1 - C_1) + \frac{1}{2} a_0^2 C_1^{-1} E\{(1 - a_0 W_1)e^{-a_0 W_1} \tilde{S}_1^2\} \\ &\equiv (1-\alpha) + K_1 + K_2, \quad \text{say}, \end{aligned} \tag{14}$$

where W_1 is a positive random variable satisfying $|W_1 - 1| < |Q_1 - 1|$. From Lemma 3 we get

$$K_1 = a_0 e^{-a_0} \eta_1 \sigma_1^{-1} d + o(d) \quad \text{as } d \to 0. \tag{15}$$

Since $Q_1 \xrightarrow{P} 1$ by Lemma 3 (ii) where "$\xrightarrow{P}$" means convergence in probability, we have that $(1 - a_0 W_1)e^{-a_0 W_1}\tilde{S}_1^2 \xrightarrow{\mathcal{D}} (1 - a_0)e^{-a_0}\rho_1^{-1} Z_1^2$. Using $a_0 W_1 > 0$, we get that $|(1 - a_0 W_1)e^{-a_0 W_1}\tilde{S}_1^2| \le \tilde{S}_1^2$, which, together with Lemma 3, implies that $\{(1 - a_0 W_1)e^{-a_0 W_1}\tilde{S}_1^2\}$ is uniformly integrable. Thus we get

$$K_2 = \frac{1}{2} a_0 e^{-a_0}(1 - a_0)\rho_1^{-1}\sigma_1^{-1} d + o(d) \quad \text{as } d \to 0. \tag{16}$$

Therefore, combining (14)–(16), we obtain the desired result. Therefore the proof is complete. □

The following lemma is used to evaluate the expectation $E[B(Q_1, Q_2)]$.

Lemma 6 *As* $d \to 0$ *we have the following results:*

(i) $E[(Q_1 - 1)e^{-a_0 Q_1}] = a_0^{-1} e^{-a_0}(\eta_1 - a_0\rho_1^{-1})\sigma_1^{-1} d + o(d)$,
(ii) $E[(Q_1 - 1)^2 e^{-a_0 Q_1}] = a_0^{-1} e^{-a_0}\rho_1^{-1}\sigma_1^{-1} d + o(d)$,
(iii) $E[(Q_2 - 1)e^{-a_0 Q_1}] = a_0^{-1} e^{-a_0}\eta_2\sigma_2^{-1} d + o(d)$,
(iv) $E[(Q_2 - 1)^2 e^{-a_0 Q_1}] = a_0^{-1} e^{-a_0}\rho_2^{-1}\sigma_2^{-1} d + o(d)$,
(v) $E[(Q_1 - 1)e^{-a_0 Q_1}(Q_2 - 1)] = o(d)$,
(vi) $E[(Q_1 - 1)^j e^{-a_0 Q_1}(Q_2 - 1)^{3-j}] = o(d)$ *for* $j = 1, 2, 3$.

Proof First we will prove (i). Let $h(x) = e^{-a_0 x}$. Taylor's expansion and Lemma 3 give

$$\begin{aligned} E[(Q_1 - 1)e^{-a_0 Q_1}] &= C_1^{-1} E[(S_1 - C_1)h(Q_1)] \\ &= C_1^{-1} e^{-a_0} E(S_1 - C_1) - C_1^{-1} a_0 E(e^{-a_0 W_1}\tilde{S}_1^2) \end{aligned}$$

$$= e^{-a_0} C_1^{-1} \eta_1 - a_0 C_1^{-1} E[e^{-a_0 W_1} \tilde{S}_1^2] + o(d),$$

where W_1 is a positive random variable satisfying $|W_1 - 1| < |Q_1 - 1|$. Since $E[e^{-a_0 W_1} \tilde{S}_1^2] = e^{-a_0} \rho_1^{-1} + o(1)$, we obtain (i). (ii) follows from the fact that $Q_1 - 1 = (a_0 \sigma_1)^{-1/2} d^{1/2} \tilde{S}_1$ and $E[e^{-a_0 Q_1} \tilde{S}_1^2] = e^{-a_0} \rho_1^{-1} + o(1)$. Similarly, we can show (iii)–(vi). This completes the proof. □

Lemma 7 *As $d \to 0$ we have*

$$E[B(Q_1, Q_2)] = a_0 e^{-a_0} d \left[\sigma_1^{-1} \left\{ -\frac{1}{2}(\eta_1 + (1 - a_0)\rho_1^{-1}) - \frac{1}{6} a_0 \rho_1^{-1} \right\} + \sigma_2^{-1} \left\{ \frac{1}{2} \eta_2 - \frac{1}{6} a_0 \rho_2^{-1} \right\} \right] + o(d).$$

Proof Let $g(x)$ be defined as in (11). Taylor's expansion for e^{-x} implies

$$g(x) = \frac{1}{2} x - \frac{1}{6} x^2 + \frac{1}{24} e^{-w} x^3 \quad \text{for all } x,$$

where $w = \theta x$ for some $\theta = \theta(x) \in (0, 1)$. Hence from (11) we get

$$\begin{aligned} E[B(Q_1, Q_2)] &= a_0 E[Q_1 e^{-a_0 Q_1} g(a_0(Q_2 - Q_1))] \\ &= \frac{1}{2} a_0^2 E[Q_1 e^{-a_0 Q_1} (Q_2 - Q_1)] - \frac{1}{6} a_0^3 E[Q_1 e^{-a_0 Q_1} (Q_2 - Q_1)^2] \\ &\quad + \frac{1}{24} a_0^4 E[Q_1 e^{-a_0 Q_1} e^{-W} (Q_2 - Q_1)^3] \\ &\equiv \frac{1}{2} a_0^2 K_1 - \frac{1}{6} a_0^3 K_2 + \frac{1}{24} a_0^4 K_3, \quad \text{say}, \end{aligned} \tag{17}$$

where $W = \theta a_0 (Q_2 - Q_1)$ for some $\theta = \theta(Q_1, Q_2) \in (0, 1)$. We will evaluate each term K_i for $i = 1, 2, 3$. It follows from Lemma 6 that

$$\begin{aligned} K_1 &= E[(Q_1 - 1) e^{-a_0 Q_1} (Q_2 - 1)] - E[(Q_1 - 1)^2 e^{-a_0 Q_1}] \\ &\quad + E[e^{-a_0 Q_1} (Q_2 - 1)] - E[(Q_1 - 1) e^{-a_0 Q_1}] \\ &= -a_0^{-1} e^{-a_0} \{\eta_1 + (1 - a_0)\rho_1^{-1}\} \sigma_1^{-1} d + a_0^{-1} e^{-a_0} \eta_2 \sigma_2^{-1} d + o(d). \end{aligned} \tag{18}$$

Similarly,

$$\begin{aligned} K_2 &= E[(Q_1 - 1) e^{-a_0 Q_1} (Q_2 - 1)^2] - 2E[(Q_1 - 1)^2 e^{-a_0 Q_1} (Q_2 - 1)] \\ &\quad + E[(Q_1 - 1)^3 e^{-a_0 Q_1}] + E[e^{-a_0 Q_1} (Q_2 - 1)^2] + E[e^{-a_0 Q_1} (Q_1 - 1)^2] \\ &\quad - 2E[(Q_1 - 1) e^{-a_0 Q_1} (Q_2 - 1)] \\ &= a_0^{-1} e^{-a_0} \left(\sum_{i=1}^{2} \rho_i^{-1} \sigma_i^{-1} \right) d + o(d). \end{aligned} \tag{19}$$

Finally, we will calculate the term K_3. Since $W = \theta a_0(Q_2 - Q_1)$, it is easy to see that $e^{-a_0 Q_1} e^{-W} = e^{-a_0\{(1-\theta)Q_1+\theta Q_2\}}$, which implies that $0 < e^{-a_0 W} e^{-W} \le 1$, for $a_0\{(1-\theta)Q_1 + \theta Q_2\} \ge 0$. Thus we have

$$\begin{aligned}|K_3| &\le E[|Q_1\{(Q_2-1)-(Q_1-1)\}|^3]\\ &\le E[|Q_1(Q_2-1)|^3] + 3E[|Q_1(Q_2-1)^2(Q_1-1)|]\\ &\quad + 3E[|Q_1(Q_1-1)^2(Q_2-1)|]\\ &\quad + E[|Q_1(Q_1-1)^3|] \equiv K_{31} + 3K_{32} + 3K_{33} + K_{34}, \quad \text{say.}\end{aligned}$$

Let $s_j = (a_0\sigma_j)^{-1/2}$ for $j = 1, 2$. Recall that $Q_j - 1 = s_j d^{1/2}\tilde{S}_j$. The uniform integrability of $\{\tilde{S}_j^p,\ 0 < d \le d_0\}$ for each $p \ge 1$ gives that $\sup_{0<d\le d_0} E(|\tilde{S}_j|^p)$ is bounded from above for each $p \ge 1$. Let us evaluate each term K_{3j} for $j = 1, 2, 3, 4$. Since $\tilde{S}_1$ and $\tilde{S}_2$ are independent, we have for some positive constant M

$$\begin{aligned}|K_{31}| &\le E[|Q_1-1||Q_2-1|^3] + E[|Q_2-1|^3]\\ &\le ME(|\tilde{S}_1|)E(|\tilde{S}_2|^3)d^2 + ME(|\tilde{S}_2|^3)d^{3/2} = O(d^2 + d^{3/2}) = o(d).\end{aligned}$$

In the same way we get that $K_{3j} = o(d)$ for $j = 2, 3, 4$. Therefore we obtain

$$K_3 = o(d). \tag{20}$$

Combining (17)–(20), we obtain the desired result of the lemma. This completes the proof. □

Acknowledgements The authors thank the anonymous referees for their constructive comments and suggestions which helped to improve the paper. The first author was supported by JSPS KAKENHI Grant-Number 26400193.

Appendix

In this appendix we will give the uniform integrability of $\{\tilde{S}_i^p,\ 0 < d \le d_0\}$ for each $p \ge 1$ in Lemma 3. Let $Y_2, Y_3, \ldots$ be a sequence of independent and identically distributed positive continuous random variables having a finite mean $\theta = E(Y_2)$. We consider the following three-stage procedure defined by Mukhopadhyay (1990):

$$R = R(d) = \max\{m,\ N_1\} \quad \text{and} \quad S = S(d) = \max\{R,\ N_2\},$$

where $N_1 = \langle \rho\lambda\overline{Y}_m \rangle + 1$, $N_2 = \langle \lambda\overline{Y}_R \rangle + 1$, $0 < \rho < 1$, $0 < \lambda < \infty$, $\overline{Y}_n = (n-1)^{-1}\sum_{i=2}^n Y_i$ for $n \ge 2$ and $m = m(d)\ (\ge 2)$ is the starting sample size such that $m \to \infty$ as $d \to 0$. Let $n^* = \lambda\theta$ and we suppose the following conditions

$$\lambda = \lambda(m) \to \infty \text{ as } m \to \infty, \quad \limsup_{d\to 0} m/n^* < \rho^2 \tag{21}$$

and for some $r > 1$, as $m \to \infty$

$$n^* = O(m^r). \tag{22}$$

In the following we assume that $E(Y_2^p) < \infty$ for some $p \geq 2$ and let M denote a generic positive constant, not depending on d. Let $V_j = Y_j/\theta$ for $j = 2, 3, \cdots$ and $\overline{V}_n = \sum_{j=2}^n V_j/(n-1)$. Then $N_1 = \langle \rho n^* \overline{V}_m \rangle + 1$ and $N_2 = \langle n^* \overline{V}_R \rangle + 1$. For $\varepsilon \in (0, 1)$, define a set $B_{m,\varepsilon}$ by $B_{m,\varepsilon} = \left\{ \overline{V}_m < 1 - \varepsilon \right\}$.

Lemma 8 *As $d \to 0$, we have $P(B_{m,\varepsilon}) = O(m^{-p/2})$.*

Proof Since $\{\overline{V}_n - 1, \ n \geq m\}$ is a reversed martingale, we have from the submartingale inequality,

$$P(B_{m,\varepsilon}) \leq P\left\{ \sup_{n \geq m} \left| \overline{V}_n - 1 \right| > \varepsilon \right\} \leq \varepsilon^{-p} E \left| \overline{V}_m - 1 \right|^p = O(m^{-p/2}).$$

□

Lemma 9 *As $d \to 0$, we have*

$$P(R \neq N_1) = O(m^{-p/2}) \quad \textit{and} \quad P(S \neq N_2) = O(m^{-p/2}). \tag{23}$$

Proof Fix $\varepsilon_0 \in (0, 1-\rho)$. By (21) and Lemma 8, for sufficiently small d,

$$P(R \neq N_1) \leq P(\overline{V}_m < m/(\rho n^*)) \leq P(\overline{V}_m < 1 - \varepsilon_0) = O(m^{-p/2}),$$

which implies the left side of (23). Next,

$$\begin{aligned} &P(S \neq N_2) \leq P(\langle \rho\lambda \overline{Y}_m \rangle + 1 > \langle \lambda \overline{Y}_R \rangle + 1, \ R = N_1) + P(R \neq N_1) \\ &\leq P(\rho\lambda \overline{Y}_m > \lambda \overline{Y}_R) + O(m^{-p/2}) \end{aligned}$$

from the left side of (23). The first term is evaluated as follows.

$$\begin{aligned} &P(\rho\lambda \overline{Y}_m > \lambda \overline{Y}_R) \\ &= \ P(\rho n^* \overline{V}_m > n^* \overline{V}_R, \ \overline{V}_R < 1 - \varepsilon_0) + P(\rho n^* \overline{V}_m > n^* \overline{V}_R, \ \overline{V}_R \geq 1 - \varepsilon_0) \\ &\leq \ P\left(\left| \overline{V}_R - 1 \right| > \varepsilon_0 \right) + P(\rho \overline{V}_m > 1 - \varepsilon_0). \end{aligned}$$

As in the proof of Lemma 8, we have that $P\left(\left| \overline{V}_R - 1 \right| > \varepsilon_0 \right) = O(m^{-p/2})$ and $P(\rho \overline{V}_m > 1 - \varepsilon_0) = P\left(\left| \overline{V}_m - 1 \right| > (1 - \varepsilon_0 - \rho)/\rho \right) = O(m^{-p/2})$. Hence, the right side of (23) holds. □

Lemma 10 *If $0 < q < p/(2r)$, where r is as in (22), then $\{(n^*/R)^q, 0 < d \leq d_0\}$ and $\{(n^*/S)^q, 0 < d \leq d_0\}$ are uniformly integrable for some $d_0 > 0$.*

Proof Note that $(n^*/S)^q \le (n^*/R)^q$. From Lemma 1 of Chow and Yu (1981), it suffices to show that $P(R < \varepsilon_1 n^*) = o(n^{*-q})$ for some $\varepsilon_1 \in (0, 1)$. By choosing $\varepsilon_1 \in (0, \rho)$, we have from (22)

$$P(R < \varepsilon_1 n^*) \le P(\rho \overline{V}_m < \varepsilon_1) \le P\left(\left|\overline{V}_m - 1\right| > 1 - \varepsilon_1/\rho\right) = o(n^{*-q}).\square$$

Lemma 11 *For* $0 < q \le p, \{(R/n^*)^q,\ 0 < d \le d_0\}$ *and* $\{(S/n^*)^q,\ 0 < d \le d_0\}$ *are uniformly integrable for some* $d_0 > 0$.

Proof From Corollary 4.1 of Gut (2005), if $E\left\{\sup_{0<d\le d_0}(R/n^*)^q\right\} < \infty$, then $\{(R/n^*)^q,\ 0 < d \le d_0\}$ is uniformly integrable. By the definition of R, Doob's maximal inequality for the reversed martingale and (21),

$$\begin{aligned} E\left\{\sup_{0<d\le d_0}(R/n^*)^q\right\} &\le M \cdot E\left[\sup_{0<d\le d_0}\{(m/n^*)^q + (\rho\overline{V}_m + (1/n^*))^q\}\right] \\ &\le M + M\rho^q E\left(\sup_{0<d\le d_0}\overline{V}_m^q\right) \le M + ME\left(\sup_{n\ge 2}\overline{V}_n^q\right) \le M \quad \text{for } 1 < q \le p, \end{aligned}$$

which yields the uniform integrability of $\{(R/n^*)^q,\ 0 < d \le d_0\}$ for $1 < q \le p$. When $0 < q \le 1$, we have that $\sup_{0<d\le d_0} E(R/n^*)^{q\zeta} = \sup_{0<d\le d_0} E(R/n^*)^p < \infty$ for $\zeta = p/q > 1$. Therefore, $\{(R/n^*)^q,\ 0 < d \le d_0\}$ is uniformly integrable for $0 < q \le p$. Next, we shall show the uniform integrability of $\{(S/n^*)^q,\ 0 < d \le d_0\}$. Since $S \le N_2 + R$, it suffices to show that $E\left\{\sup_{0<d\le d_0}(N_2/n^*)^q\right\} < \infty$ which can be proved similarly. $\square$

Lemma 12 *For* $0 < q \le p$,

$$\left\{\left|n^{*-\frac{1}{2}}\sum_{j=2}^{R}(V_j - 1)\right|^q,\ 0 < d \le d_0\right\} \text{ and } \left\{\left|n^{*-\frac{1}{2}}\sum_{j=2}^{S}(V_j - 1)\right|^q,\ 0 < d \le d_0\right\}$$

are uniformly integrable for some $d_0 > 0$.

Proof Follows from Lemma 5 of Chow and Yu (1981) and Lemma 11. $\square$

Proposition 2 *We assume that* $E(Y_2^p) < \infty$ *for some* $p \ge 2$. *Let* $\tilde{S} = n^{*-\frac{1}{2}}(S - n^*)$. *Under the conditions* (21) *and* (22), *if* $0 < q < p/(2r + 1)$, *then* $\left\{\tilde{S}^q,\ 0 < d \le d_0\right\}$ *is uniformly integrable for some* $d_0 > 0$.

Proof Now,

$$\begin{aligned} |\tilde{S}^q| &= |n^{*-1/2}(S - n^*)|^q \\ &= |n^{*-1/2}(\langle n^*\overline{V}_R\rangle + 1 - n^*)|^q I(S = N_2) + |n^{*-1/2}(R - n^*)|^q I(S \ne N_2) \\ &\equiv K_1 + K_2, \quad \text{say}. \end{aligned}$$

Since $K_3 \equiv n^{*-1/2}(\langle n^*\overline{V}_R\rangle + 1 - n^*\overline{V}_R) \le n^{*-1/2} \le 1$ and $0 < R/(R-1) \le 2$, we have for some $\zeta > 1$, $u = 2r+1$ and $v = \frac{1}{2r}+1$,

$$\begin{aligned} E(K_1^{\zeta}) \le\ & E|n^{*-1/2}(n^*\overline{V}_R - n^*) + K_3|^{q\zeta} \\ \le\ & ME\left|n^{*-\frac{1}{2}}\sum_{j=2}^{R}(V_j-1)\cdot(R/(R-1))\cdot(n^*/R)\right|^{q\zeta} + M \\ \le\ & M\left\{E\left|n^{*-\frac{1}{2}}\sum_{j=2}^{R}(V_j-1)\right|^{uq\zeta}\right\}^{\frac{1}{u}}\left\{E(n^*/R)^{vq\zeta}\right\}^{\frac{1}{v}} + M = O(1) \end{aligned}$$

by Lemmas 10 and 12. Finally, for some $\zeta > 1$, $u_0 = r+1$ and $v_0 = \frac{1}{r}+1$, we have from (23) and Lemma 11

$$\begin{aligned} E(K_2^{\zeta}) &\le M\, n^{*\frac{1}{2}q\zeta}\{E(R/n^*)^{u_0 q\zeta}+1\}^{\frac{1}{u_0}}\{P(S \ne N_2)\}^{\frac{1}{v_0}} = O(m^{q\zeta r/2 - p/(2v_0)}) \\ &= O(1). \end{aligned}$$

Hence, the proposition is proved. □

Proof of the uniform integrability We will show the uniform integrability of $\{\tilde{S}_i^{\,p},\ 0 < d \le d_0\}$ for each $p \ge 1$. Let $Y'_{ij} = Y_{ij}/\sigma_i$ and $C_i = \lambda_i = a_*\sigma_i d^{-1}$, where Y_{ij} has the exponential distribution $\mathrm{Exp}(0, \sigma_i)$. Then $Y'_{i2}, Y'_{i3}, \ldots$ are i.i.d random variables according to $\mathrm{Exp}(0, 1)$, and R_i and S_i defined by (10) can be written as $R_i = \max\{m,\ N_{1i}\}$ and $S_i = \max\{R_i,\ N_{2i}\}$, where $N_{1i} = \langle\rho_i\lambda_i\overline{Y'}_{im}\rangle + 1$, $N_{2i} = \langle\lambda_i\overline{Y'}_{iR_i}\rangle + 1$ and $0 < \rho_i < 1$. Put $n^* = C_i$, $\lambda = \lambda_i$, $\rho = \rho_i$, $R = R_i$, $S = S_i$, $Y_j = Y_{ij}$ and $V_j = Y'_{ij}$ for $i = 1, 2$. Since $E(Y_{i2}^p) < \infty$ for all $p > 0$ and $m = O(d^{-1/r})$ for some $r > 1$, the conditions (21) and (22) are satisfied. Therefore from Proposition 2, $\{\tilde{S}_i^{\,p},\ 0 < d \le d_0\}$ is uniformly integrable for some $d_0 > 0$. □

References

Chow YS, Yu KF (1981) The performance of a sequential procedure for the estimation of the mean. Ann Stat 9:184–189

Gut A (2005) Probability, graduate course. Springer, Berlin

Hall P (1981) Asymptotic theory and triple sampling for sequential estimation of a mean. Ann Stat 9:1229–1238

Hamdy HI (1997) Performance of fixed width confidence intervals under type II errors: the exponential case. S Afr Stat J 31:259–269

Hamdy HI, Al-Mahmeed M, Nigm A, Son MS (1989) Three-stage estimation procedure for the exponential location parameters. Metron 47:279–294

Hamdy HI, Son MS, Yousef AS (2015) Sensitivity analysis of multistage sampling to departure of an underlying distribution from normality with computer simulations. Seq Anal 34:532–558

Honda T (1992) Estimation of the mean by three stage procedure. Seq Anal 11:73–89

Isogai E, Futschik A (2010) Sequential estimation of a linear function of location parameters of two negative exponential distributions. J Stat Plan Inference 140:2416–2424

Lombard F, Swanepoel JWH (1978) On finite and infinite confidence sequences. S Afr Stat J 12:1–24

Mukhopadhyay N (1990) Some properties of a three-stage procedure with applications in sequential analysis. Sankhyā A52:218–231

Mukhopadhyay N, Hamdy HI (1984) On estimating the difference of location parameters of two negative exponential distributions. Can J Stat 12:67–76

Mukhopadhyay N, Mauromoustakos A (1987) Three-stage estimation procedures for the negative exponential distributions. Metrika 34:83–93

Mukhopadhyay N, Padmanabhan AR (1993) A note on three-stage confidence intervals for the difference of locations: the exponential case. Metrika 40:121–128

Mukhopadhyay N, Zack S (2007) Bounded risk estimation of linear combinations of the location and scale parameters in exponential distributions under two-stage sampling. J Stat Plan Inference 137:3672–3686

Singh RK, Chaturvedi A (1991) A note on sequential estimation of the difference between location parameters of two negative exponential distributions. J Indian Stat Assoc 29:107–114

Son MS, Haugh LD, Hamdy HI, Costanza MC (1997) Controlling type II error while constructing triple sampling fixed precision confidence intervals for the normal mean. Ann Inst Stat Math 49:681–692

Yousef AS, Kimber AC, Hamdy HI (2013) Sensitivity of normal-based triple sampling sequential point estimation to the normality assumption. J Stat Plan Inference 143:1606–1618