数理解析研究所講究録
第 1860 巻 2013 年 33-41

A note on asymptotic higher-order properties of a two-stage estimation procedure

(二段階推定法の高次漸近特性に関する一考察)[1]

Chikara Uno

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

宇野 力(秋田大学 教育文化学部)

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables from a normal population $N(\mu, \sigma^2)$ where the mean $\mu \in (-\infty, \infty)$ and the variance $\sigma^2 \in (0, \infty)$ are both unknown. Having recorded $X_1, \ldots, X_n$, we define $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $S_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ for $n \geq 2$. Let $d \in (0, \infty)$ and $\alpha \in (0,1)$ be any preassigned numbers. On the basis of the random sample of size n, we consider a confidence interval $I_n = [\overline{X}_n - d, \overline{X}_n + d]$ for μ with confidence coefficient $1 - \alpha$. If we take the sample of size n such that

$$n \geq a^2\sigma^2/d^2 \equiv n_0,$$

where a is the upper $100 \times \alpha/2\%$ point of the standard normal distribution, then it holds that $P(\mu \in I_n) \geq 1 - \alpha$ for all fixed μ, σ^2, α and d. Unfortunately, σ^2 is unknown, so we cannot use the optimal fixed sample size n_0.

Stein's two-stage procedure does not have the asymptotic second-order efficiency. Mukhopadhyay and Duggan (1997) proposed the following two-stage procedure, provided that $\sigma^2 > \sigma_L^2$ where σ_L^2 is positive and known to the experimenter. Let

$$m = m(d) = \max\left\{m_0,\ [a^2\sigma_L^2/d^2]^* + 1\right\},$$

where $m_0\ (\geq 2)$ is a preassigned integer and $[x]^*$ denotes the largest integer less than x. By using the pilot observations $X_1, \ldots, X_m$, calculate S_m^2 and define

$$N = N(d) = \max\left\{m,\ [b_m^2 S_m^2/d^2]^* + 1\right\},$$

where b_m is the upper $100 \times \alpha/2\,\%$ point of the Student's t distribution with $m-1$

[1]本研究は, 科学研究費助成事業(学術研究助成基金助成金)基盤研究 (C) 課題番号 24540107(研究代表者 宇野力)から研究助成を受けています。

degrees of freedom. If $N > m$, then take the second sample $X_{m+1}, \ldots, X_N$. Based on the total observations $X_1, \ldots, X_N$, consider the fixed-width confidence interval $I_N = [\overline{X}_N - d,\, \overline{X}_N + d]$ for μ, where $\overline{X}_N = (X_1 + \cdots + X_N)/N$. Then, it is possible to show the exact consistency, that is, $P(\mu \in I_N) \geq 1 - \alpha$ for all fixed μ, σ^2, d and α. Mukhopadhyay and Duggan (1997) showed that as $d \to 0$

$$\eta + o(n_0^{-1/2}) \leq E(N - n_0) \leq \eta + 1 + o(n_0^{-1/2}),$$

where $\eta = (1/2)(a^2+1)\sigma^2\sigma_L^{-2}$, and so the above two-stage procedure has the asymptotic second-order efficiency. Aoshima and Takada (2000) gave a second-order approximation to the average sample number: $E(N - n_0) = \eta + (1/2) + O(n_0^{-1/2})$ as $d \to 0$, and further Isogai et al. (2012) showed that $E(N - n_0) = \eta + (1/2) + O(n_0^{-1})$ as $d \to 0$. As for the coverage probability, Mukhopadhyay and Duggan (1997) showed that as $d \to 0$

$$1 - \alpha + o(n_0^{-1}) \leq P(\mu \in I_N) \leq 1 - \alpha + 2A\, n_0^{-1} + o(n_0^{-1}),$$

where $A = (1/2)a\phi(a)$ and $\phi(x)$ is the probability density function (p.d.f.) of the standard normal distribution. Aoshima and Takada (2000) gave a second-order approximation to the coverage probability:

$$P(\mu \in I_N) = 1 - \alpha + A\, n_0^{-1} + o(n_0^{-1}) \quad \text{as } d \to 0.$$

Define $T_d = b_m^2 S_m^2/d^2$, $t_d^* = n_0^{-1/2}(T_d - n_0)$ and $U_d = [T_d]^* + 1 - T_d$. Isogai et al. (2012) showed that as $d \to 0$

$$P(\mu \in I_N) = 1 - \alpha + A\, n_0^{-1} + \varepsilon_d\, n_0^{-3/2} + o(n_0^{-3/2}),$$

where $\varepsilon_d = -A(a^2 + 1)E(t_d^* U_d)$ and $|\varepsilon_d| \leq A(a^2 + 1)\sqrt{\sigma^2/(6\sigma_L^2)} + O(n_0^{-1/2})$. Uno (2013) established the asymptotic independence of t_d^* and U_d, and obtained that

$$P(\mu \in I_N) = 1 - \alpha + A\, n_0^{-1} + o(n_0^{-3/2}) \quad \text{as } d \to 0. \tag{1}$$

In this article, we shall apply the result of Uno (2013) to the slight general case of Mukhopadhyay and Duggan (1999) in Section 2 and give some examples in Section 3.

2. Asymptotic theory

We consider the case of Mukhopadhyay and Duggan (1999) with $\tau = 1$. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables from a population. Several optimal fixed sample sizes which arise from problems in sequential point and interval estimation may be written in the form

$$n_0 = q\theta/h,$$

where q and h are known positive numbers, but θ is the unknown and positive nuisance parameter. We assume that

$$\theta > \theta_L,$$

where $\theta_L (> 0)$ is known to the experimenter. Mukhopadhyay and Duggan (1999) proposed the following two-stage procedure. The initial sample size is defined by

$$m \equiv m(h) = \max\{m_0,\ [q\theta_L/h]^* + 1\},$$

where $m_0\ (\geq 2)$ is a preassigned positive integer. By the pilot sample $X_1, \ldots, X_m$ of size m, we consider an unbiased estimator $V(m)$ of θ satisfying $P\{V(m) > 0\} = 1$. Further, suppose that

$$Y_m = p_m V(m)/\theta \text{ is distributed as } \chi^2_{p_m} \quad \text{with } p_m = c_1 m + c_2,$$

where p_m is a positive integer with a positive integer c_1 and an integer c_2, and $\chi^2_{p_m}$ stands for a chi-square distribution with p_m degrees of freedom. We consider asymptotic theory as $h \to 0$, namely, $n_0 \to \infty$. Then,

$$m \to \infty \quad \text{and} \quad V(m) \xrightarrow{P} \theta \quad \text{as } h \to 0,$$

where "$\xrightarrow{P}$" stands for convergence in probability. Let q^*_m be positive where

$$q^*_m = q + c_3 m^{-1} + O(m^{-2}) \quad \text{as } h \to 0$$

with some real number c_3. Define

$$N \equiv N(h) = \max\{m,\ [q^*_m V(m)/h]^* + 1\}.$$

If $N > m$, then one takes the second sample $X_{m+1}, \ldots, X_N$. The total observations are $X_1, \ldots, X_N$. Throughout the remainder of this article, let

$$T_h = q^*_m V(m)/h, \quad t^*_h = n_0^{-1/2}(T_h - n_0) \quad \text{and} \quad U_h = [T_h]^* + 1 - T_h.$$

Then we obtain the following theorem.

Theorem 1. *U_h and t^*_h are asymptotically independent as $h \to 0$. The asymptotic distribution of U_h is uniform on $(0, 1)$; and the asymptotic distribution of t^*_h is normal with mean 0 and variance $2\theta/(c_1\theta_L)$.*

The proof of Theorem 1 is similar to that of Theorem (i) of Uno (2013). So we omit the details.

Let $\mathbb{R}^+ = (0, \infty)$ and suppose that g: $\mathbb{R}^+ \to \mathbb{R}^+$ is a three-times differentiable function and the third derivative $g^{(3)}(x)$ is continuous at $x = 1$. By Taylor's theorem, we have

$$g(N/n_0) = g(1) + g'(1)n_0^{-1}(N - n_0) + (1/2)g''(1)n_0^{-2}(N - n_0)^2 \\ +(1/6)g^{(3)}(W)n_0^{-3}(N - n_0)^3,$$

where W is a random variable such that $|W - 1| < |(N/n_0) - 1|$. Uno and Isogai (2012) showed that if $\{g^{(3)}(W)n_0^{-3/2}(N - n_0)^3\,;\, 0 < h < h_0\}$ is uniformly integrable for some sufficiently small $h_0 > 0$, then as $h \to 0$

$$E\{g(N/n_0)\} = g(1) + B_0 n_0^{-1} + \epsilon_h n_0^{-3/2} + o(n_0^{-3/2}), \tag{2}$$

where

$$B_0 = (1/2)g'(1) + \Delta(\theta/\theta_L), \quad \Delta = c_3 q^{-1} g'(1) + c_1^{-1} g''(1),$$

$$\epsilon_h = g''(1)E(t_h^* U_h) \quad \text{and} \quad |\epsilon_h| \le |g''(1)|\sqrt{\theta/(6c_1\theta_L)} + O(n_0^{-1/2}).$$

We obtain the next theorem.

Theorem 2. *If $\{g^{(3)}(W)n_0^{-3/2}(N - n_0)^3\,;\, 0 < h < h_0\}$ is uniformly integrable for some sufficiently small $h_0 > 0$, then as $h \to 0$*

$$E\{g(N/n_0)\} = g(1) + B_0 n_0^{-1} + o(n_0^{-3/2}).$$

Proof. It is easily seen from Lemma 2.2 of Mukhopadhyay and Duggan (1999) that $\{|t_h^* U_h|\,;\, 0 < h < h_0\}$ is uniformly integrable for some sufficiently small $h_0 > 0$. Therefore, we have from Theorem 1 that $E(t_h^* U_h) = o(1)$ as $h \to 0$, which yields $\epsilon_h = o(1)$ in (2). □

Remark. If $\Delta = 0$, then the approximation of Theorem 2 does not depend on θ_L up to the order term.

Recall the fixed-width interval estimation of μ of $N(\mu, \sigma^2)$ described in Section 1. We take $q = a^2$, $h = d^2$, $\theta = \sigma^2$, $\theta_L = \sigma_L^2$, $V(m) = S_m^2$ and $q_m^* = b_m^2$. Then we have $p_m = m - 1$ ($c_1 = 1$, $c_2 = -1$) and $q_m^* = b_m^2 = a^2 + c_3 m^{-1} + O(m^{-2})$ with $c_3 = (1/2)a^2(a^2 + 1)$. Taking $g(x) = 2\Phi(a\sqrt{x}) - 1$, where Φ is the cumulative distribution function of $N(0, 1)$, we have $g(1) = 1 - \alpha$, $g'(1) = a\phi(a)$ and $g''(1) = -(1/2)a(a^2 + 1)\phi(a)$. Thus, from Lemma 4.1 of Isogai et al. (2012) and Theorem

2, we obtain $P(\mu \in I_N) = E\{g(N/n_0)\} = 1 - \alpha + (1/2)a\phi(a)n_0^{-1} + o(n_0^{-3/2})$, which becomes the approximation (1). Note that $\Delta \equiv c_3 q^{-1} g'(1) + c_1^{-1} g''(1) = 0$, and so $B_0 = (1/2)a\phi(a)$ does not depend on σ_L^2.

3. Examples

We shall apply our theorem to three problems.

3.1. Bounded risk estimation of the normal mean

We consider a sequence of i.i.d. random variables $X_1, X_2, \ldots$ from a normal population $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R} = (-\infty, \infty)$ and $\sigma^2 \in \mathbb{R}^+$ are both unknown. We assume that there exists a known and positive lower bound σ_L^2 for σ^2 such that $\sigma^2 > \sigma_L^2$. Having recorded $X_1, \ldots, X_n$, we define $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $V(n) = (n-1)^{-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ for $n \geq 2$. On the basis of the random sample $X_1, \ldots, X_n$ of size n, we want to estimate μ by $\overline{X}_n$ under the loss function

$$L_n = (\overline{X}_n - \mu)^2.$$

Then, the risk is given by $R_n = E(L_n) = \sigma^2/n$. For any preassigned $w > 0$, we hope that $R_n = \sigma^2/n \leq w$, which is equivalent to

$$n \geq \sigma^2/w \equiv n_0 .$$

Unfortunately σ is unknown, so we can not use the optimal fixed sample size n_0. Thus we define a two-stage procedure. Let

$$m = m(w) = \max\left\{m_0,\ \left[\sigma_L^2/w\right]^* + 1\right\},$$

where $m_0 \geq 4$. By using the pilot observations $X_1, \cdots, X_m$, we calculate $V(m)$ and

$$N = N(w) = \max\{m,\ [b_m V(m)/w]^* + 1\},$$

where $b_m = (m-1)/(m-3)$. The risk is given by $R_N = E(\overline{X}_N - \mu)^2$. It follows from (7c.6.2) and (7c.6.7) with $c^2 = w$ and $b^2 = b_m$ in section 7c.6 of Rao (1973) that $R_N \leq w$ for all fixed μ, σ and w. Therefore our requirement is fulfilled. In the notations of Section 2, note that $h = w$, $\theta = \sigma^2$, $\theta_L = \sigma_L^2$, $q = 1$, $p_m = m - 1$ ($c_1 = 1$, $c_2 = -1$) and $q_m^* = b_m = 1 + 2m^{-1} + O(m^{-2})$ with $c_3 = 2$. Taking $g(x) = x^{-1}$ for $x > 0$, we have $R_N = E(\sigma^2/N) = wE\{g(N/n_0)\}$ and $\Delta = 0$. From Proposition 1 of Uno and Isogai (2012) and Theorem 2, we obtain

$$R_N/w = 1 - (1/2)n_0^{-1} + o(n_0^{-3/2}) \quad \text{as } w \to 0.$$

3.2. Fixed-width interval estimation of the negative exponential location

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables from a population having the following p.d.f.:

$$f(x) = \sigma^{-1} \exp\{-(x-\mu)/\sigma\}, \quad x > \mu,$$

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ are both unknown. We assume that there exists a known and positive lower bound σ_L for σ such that $\sigma > \sigma_L$. For any preassigned numbers $d > 0$ and $\alpha \in (0,1)$, we want to construct a confidence interval I_n for the location parameter μ based on the random sample $X_1, \ldots, X_n$ of size n such that the length of I_n is fixed at d and $P\{\mu \in I_n\} \geq 1-\alpha$ for all fixed μ and σ. Having recorded $X_1, \ldots, X_n$, we define $X_{n(1)} = \min\{X_1, \ldots, X_n\}$ and $V(n) = (n-1)^{-1}\sum_{i=1}^{n}(X_i - X_{n(1)})$ for $n \geq 2$, and consider a confidence interval $I_n = [X_{n(1)} - d,\ X_{n(1)}]$ for the location μ. Then $P\{\mu \in I_n\} \geq 1-\alpha$ for all fixed μ, σ, α and d, provided

$$n \geq a\sigma/d \equiv n_0 \quad \text{with } a = \ln(1/\alpha)\,(>0)\,.$$

Mukhopadhyay and Duggan (1999) proposed the following two-stage procedure. Let

$$m = m(d) = \max\{m_0,\ [a\sigma_L/d]^* + 1\}\,,$$

where $m_0 \geq 2$. By using the pilot observations $X_1, \ldots, X_m$, we calculate $V(m)$ and

$$N = N(d) = \max\{m,\ [b_m V(m)/d]^* + 1\}\,,$$

where b_m is the upper $100\alpha\%$ point of the F-distribution with 2 and $2(m-1)$ degrees of freedom. Then the interval $I_N = [X_{N(1)} - d,\ X_{N(1)}]$ is proposed for μ. It follows from (3.3) of Mukhopadhyay and Duggan (1999) that $P\{\mu \in I_N\} \geq 1-\alpha$ for all fixed μ, σ, d and α. Then, let $h = d$, $\theta = \sigma$, $\theta_L = \sigma_L$, $q = a$, $p_m = 2m-2$ ($c_1 = 2$, $c_2 = -2$) and $q_m^* = b_m = a + (1/2)a^2 m^{-1} + O(m^{-2})$ with $c_3 = (1/2)a^2$ in the notations of Section 2. Taking $g(x) = 1 - e^{-ax}$ for $x > 0$, we have $P\{\mu \in I_N\} = E\{1 - \exp(-Nd/\sigma)\} = E\{g(N/n_0)\}$ and $\Delta = 0$. From Proposition 2 of Uno and Isogai (2012) and Theorem 2, we obtain

$$P\{\mu \in I_N\} = 1 - \alpha + (1/2)a\alpha n_0^{-1} + o(n_0^{-3/2}) \quad \text{as } d \to 0.$$

3.3. Selecting the best normal population

Suppose there exist $k\,(\geq 2)$ independent populations π_i, $i = 1, \ldots, k$ and each π_i has a normal distribution $N(\mu_i, \sigma^2)$, where the mean μ_i and the common variance σ^2 are unknown. Let us denote $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)'$ and write $\mu_{[1]} \leq \cdots \leq \mu_{[k-1]} \leq \mu_{[k]}$ for the ordered μ values. Along the line of Bechhofer (1954), we consider the problem of selecting the population associated with the largest $\mu_{[k]}$, referred to as the best population, while guaranteeing

$$P\{CS\} \geq P^* \quad \text{whenever } \boldsymbol{\mu} \in \Omega(\delta) \tag{3}$$

for given $\delta\,(> 0)$ and $P^* \in (k^{-1}, 1)$, where $\Omega(\delta) = \{\boldsymbol{\mu} : \ \mu_{[k]} - \mu_{[k-1]} \geq \delta\}$ and the complementary subspace $\Omega^c(\delta)$ is called the indifference zone. Here and elsewhere, "CS" stands for "Correct Selection". Let $X_{i1}, X_{i2}, \dots$ be i.i.d. random variables from π_i for $i = 1, \dots, k$. Having recorded $X_{i1}, \dots, X_{in}$ with fixed $n\,(\geq 2)$ from each π_i, we compute $\overline{X}_{in} = n^{-1}\sum_{j=1}^{n} X_{ij}$ and $\overline{X}_{[kn]} = \max_{1\leq i\leq k} \overline{X}_{in}$. If σ^2 were known, one implements the following selection rule (SR) for fixed n:

SR_n : Select the population which gives rise to the largest sample mean $\overline{X}_{[kn]}$ as the best population. (4)

Then, it follows from the equation (2.2) of Aoshima and Aoki (2000) that

$$\inf_{\boldsymbol{\mu}\in\Omega(\delta)} P\left\{CS_{(SR_n)}\right\} = \int_{-\infty}^{\infty} \Phi^{k-1}(y + \sqrt{n\delta^2/\sigma^2})\phi(y)dy,$$

where $CS_{(SR_n)}$ stands for "Correct Selection" under the selection rule SR_n. The infimum is attained when $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta$, which is known as the least favorable configuration. Let

$$H(x) = \int_{-\infty}^{\infty} \Phi^{k-1}(y + \sqrt{x})\phi(y)dy, \quad x > 0$$

and $z = z(k, P^*)$ is a positive constant which satisfies the integral equation $H(z^2) = P^*$. The requirement (3) is satisfied if

$$n \geq z^2\sigma^2/\delta^2 \equiv n_0 \,.$$

Since σ^2 is unknown, we can not use the optimal fixed sample size n_0. The two-stage procedure proposed by Bechhofer et al. (1954) satisfies (3) and hence it has the exact consistency.

Let us assume that $\sigma^2 > \sigma_L^2$ where $\sigma_L^2(> 0)$ is known, and define

$$m = m(\delta) = \max\left\{m_0,\ \left[z^2\sigma_L^2/\delta^2\right]^* + 1\right\}, \tag{5}$$

where $m_0 \geq 2$. Take the initial sample $X_{i1}, \dots, X_{im}$ from each π_i and compute $\overline{X}_{im}$, $i = 1, \dots, k$ and $V(m) = k^{-1}\sum_{i=1}^{k} V_{im}$ where $V_{im} = (m-1)^{-1}\sum_{j=1}^{m}(X_{ij} - \overline{X}_{im})^2$. Aoshima and Aoki (2000) proposed

$$N = N(\delta) = \max\left\{m,\ \left[t^2V(m)/\delta^2\right]^* + 1\right\}, \tag{6}$$

where $t = t(k, P^*)$ is a positive constant such that $E\{H(t^2 Y_m/p_m)\} = P^*$. Here, $Y_m = p_m V(m)/\sigma^2$ has the distribution $\chi^2_{p_m}$ with $p_m = k(m-1)$. In the notations of Section 2, note that $q = z^2$, $\theta = \sigma^2$, $\theta_L = \sigma_L^2$, $h = \delta^2$, $c_1 = k$, $c_2 = -k$ and $q_m^* = t^2$. Secondly, one takes the additional sample $X_{i(m+1)}, \ldots, X_{iN}$ of size $N - m$ from each π_i and computes $\overline{X}_{iN} = \sum_{j=1}^N X_{ij}/N$, $i = 1, \ldots, k$. Then, we implement the selection rule SR_N given by (4) associated with $\overline{X}_{[kN]} = \max_{1\le i\le k} \overline{X}_{iN}$. For the two-stage procedure defined by (5) and (6), Aoshima and Aoki (2000) showed the exact consistency, namely, $\inf_{\mu\in\Omega(\delta)} P\{CS_{(SR_N)}\} \ge P^*$ for each fixed δ. It follows from the equation (2.9) of Aoshima and Aoki (2000) that as $\delta \to 0$

$$t^2 = z^2 + c_3 m^{-1} + O(m^{-2}), \quad \text{where } c_3 = -\frac{z^4 H''(z^2)}{kH'(z^2)}.$$

Here, H' and H'' are the first and second derivatives of H, respectively. Taking $g(x) = H(z^2 x)$ for $x > 0$, we have $\inf_{\mu\in\Omega(\delta)} P\{CS_{(SR_N)}\} = E\{g(N/n_0)\}$ and $\Delta = 0$. From Proposition 4 of Uno and Isogai (2012) and Theorem 2, we obtain

$$\inf_{\mu\in\Omega(\delta)} P\{CS_{(SR_N)}\} = P^* + (1/2)z^2 H'(z^2) n_0^{-1} + o(n_0^{-3/2}).$$

Mukhopadhyay and Duggan (1999) proposed

$$N^\dagger = N^\dagger(\delta) = \max\left\{m,\ \left[z^2 V(m)/\delta^2\right]^* + 1\right\}. \tag{7}$$

For the two-stage procedure defined by (5) and (7), the exact consistency does not hold and $\Delta = k^{-1} z^4 H''(z^2)$. Hence, from Proposition 3 of Uno and Isogai (2012) and Theorem 2, we have

$$\inf_{\mu\in\Omega(\delta)} P\{CS_{(SR_{N^\dagger})}\} = P^* + B_0^\dagger n_0^{-1} + o(n_0^{-3/2}),$$

where $B_0^\dagger = (1/2)z^2 H'(z^2) + k^{-1} z^4 H''(z^2)\sigma^2\sigma_L^{-2}$, which depends on σ_L^2.

References

Aoshima, M., Aoki, M. (2000). Two-stage procedure having exact consistency and second order properties for the s best selection. *Sequential Anal.* **19**, 115–131.

Aoshima, M., Takada, Y. (2000). Second-order properties of a two-stage procedure for comparing several treatments with a control. *J. Japan Statist. Soc.* **30**, 27–41.

Bechhofer, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variance. *Ann. Math. Statist.* **25**, 16-39.

Bechhofer, R. E., Dunnett, C. W., Sobel, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika* **41**, 170-176.

Isogai, E., Uno, C., Takeuchi, D. (2012). Two-stage procedure having exact third-order properties for a normal mean. *Commun. Statist.-Theor. Meth.* **41**, 4332–4347.

Mukhopadhyay, N., Duggan, W.T. (1997). Can a two-stage procedure enjoy second-order properties? *Sankhyā, Ser. A* **59**, 435–448.

Mukhopadhyay, N., Duggan, W.T. (1999). On a two-stage procedure having second-order properties with applications. *Ann. Inst. Statist. Math.* **51**, 621–636.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications, Second edition*, Wiley, New York.

Uno, C. (2013). Asymptotic theory for a two-stage procedure in sequential interval estimation of a normal mean. *Statist. Probab. Lett.* **83**, 1420–1423.

Uno, C., Isogai, E. (2012). Asymptotic properties of a two-stage procedure and their applications. *Far East J. Theor. Stat.* **40**, 75–105.