REVSTAT – Statistical Journal
Volume 14, Number 3, June 2016, 297–309

SEQUENTIAL ESTIMATION OF A COMMON LOCATION PARAMETER OF TWO POPULATIONS

Authors: Agnieszka Stępień-Baran
– Institute of Applied Mathematics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland
A.Stepien-Baran@mimuw.edu.pl

Received: June 2014 Revised: January 2015 Accepted: January 2015

Abstract:

- The problem of sequentially estimating a common location parameter of two independent populations from the same distribution with an unknown location parameter and known but different scale parameters is considered in the case when the observations become available at random times. Certain classes of sequential estimation procedures are derived under a location invariant loss function and with the observation cost determined by convex functions of the stopping time and the number of observations up to that time.

Key-Words:

- *location parameter; location invariant loss function; minimum risk equivariant estimator; optimal stopping time; risk function.*

AMS Subject Classification:

- 62L12, 62L15, 62F10.

1. INTRODUCTION

The paper concerns the invariant problem of sequentially estimating a common location parameter of two independent populations from the same distribution with an unknown location parameter and known but different scale parameters, in the special case when the observations arrive at random times. For example, in studying the effectiveness of experimental safety devices of mobile constructions relevant data may become available only as a result of accidents. Medical data (such as data on drug abuse or an asymptomatic disease) can sometimes only be obtained when patients voluntarily seek help or are somehow otherwise identified and examined, at random times. Other examples are data resulting from an undersea survey of containerized radioactive waste, from archeological discoveries, from market research or from planning the assortment of production (when the orders come forward at random times).

The estimation problem of a common location of two independent populations has been extensively discussed in the literature. Rao and Reddy (1988) studied the estimation of the unknown common location parameter of two symmetric distributions with different scale parameters. They derived asymptotic distributions and the asymptotic relative efficiencies of proposed estimators: the mean, the median, the average of the mean and the median and the Hodges–Lehmann estimator. Baklizi (2004) considered estimation of the common location parameter of several exponentials. It is found that the proposed estimators are effective in taking advantage of the available prior information. Farsipour and Asgharzadeh (2002) investigated the problem of estimating the common mean of two normal distributions. They derived a class of risk unbiased estimators which linearly combines the means of the two samples from both distributions. Mitra and Sinha (2007) studied some aspects of the problem of estimation of a common mean of two normal populations from an asymptotic point of view. They also considered the Bayes estimate of the common mean under Jeffrey's prior. Chang *et al.* (2012) considered the problem of estimating the common mean of two normal distributions with unknown ordered variances. They gave a broad class of estimators which includes the estimators proposed by Nair (1982) and Elfessi *et al.* (1992) and showed that the estimators stochastically dominate the estimators which do not take into account the order restriction on variances, including the one given by Graybill and Deal (1959). Then they proposed a broad class of individual estimators of two ordered means when unknown variances are ordered.

The problem of estimating with delayed observations was investigated by Starr *et al.* (1976), who considered the case of Bayes estimation of a mean of normally distributed observations with known variance. Some of their results were generalized by Magiera (1996). He dealt with estimation of the mean value parameter of the exponential family of distributions. Jokiel-Rokita and Stępień (2009) studied the model with delayed observations for estimating a location parameter.

We consider the following model. Let the samples $(X_1, ..., X_{n_1})$ and $(Y_1, ..., Y_{n_2})$ be independent and have a joint distribution P_θ with a Lebesgue p.d.f.

$$f\left(\frac{x_1-\theta}{\sigma_1}, ..., \frac{x_{n_1}-\theta}{\sigma_1}\right),$$

and

$$f\left(\frac{y_1-\theta}{\sigma_2}, ..., \frac{y_{n_2}-\theta}{\sigma_2}\right),$$

respectively, where f is known, σ_1, $\sigma_2 > 0$ are known and different scale parameters, and $\theta \in \mathbb{R}$ is an unknown location parameter.

The set of observations is bounded, i.e., the statistician can receive at most $N = n_1 + n_2$ observations. It is assumed that X_i is observed at time t_i, $i = 1, ..., n_1$, where $t_1, ..., t_{n_1}$ are the values of the order statistics of positive i.i.d. random variables $U_1, ..., U_{n_1}$ which are obtained before the conducted observations $X_1, ..., X_{n_1}$ and independent of $X_1, ..., X_{n_1}$. Similarly: Y_i is observed at time s_i, $i = 1, ..., n_2$, where $s_1, ..., s_{n_2}$ are the values of the order statistics of positive i.i.d. random variables $V_1, ..., V_{n_2}$ which are obtained before the conducted observations $Y_1, ..., Y_{n_2}$ and independent of $Y_1, ..., Y_{n_2}$. Furthermore, it is assumed that the samples $(U_1, ..., U_{n_1})$ and $(V_1, ..., V_{n_2})$ are independent.

Let

$$k_1(t) = \sum_{i=1}^{n_1} \mathbf{1}_{[0,t]}(U_i) \tag{1.1}$$

and

$$k_2(t) = \sum_{i=1}^{n_2} \mathbf{1}_{[0,t]}(V_i) \tag{1.2}$$

denote the number of observations which have been made by time $t \geq 0$ for the sample $(X_1, ..., X_{n_1})$ and $(Y_1, ..., Y_{n_2})$, respectively, and let $\mathcal{F}_{1,t} = \sigma\{k_1(r),\ r \leq t,\ X_1, ..., X_{k_1(t)}\}$ and $\mathcal{F}_{2,t} = \sigma\left\{k_2(r),\ r \leq t,\ Y_1, ..., Y_{k_2(t)}\right\}$ be the informations which is available at time t.

The problem is to estimate the parameter θ. If observation is stopped at time t, the loss incurred is defined by

$$\mathcal{L}_t(\theta, d) := \mathcal{L}(\theta, d) + c_A k_1(t) + c_B k_2(t) + c_1(t) + c_2(t), \tag{1.3}$$

where $\mathcal{L}(\theta, d)$ denotes the loss associated with estimation, when θ is the true value of the parameter and d is the chosen estimate. The functions $c_1(t)$ and $c_2(t)$ represents the cost of observing the processes up to time t ($k_1(t)$ and $k_2(t)$, respectively). It is supposed to be a differentiable and increasing convex functions such that $c_1(0) = 0$ and $c_2(0) = 0$. The constants $c_A \geq 0$ and $c_B \geq 0$ are the cost of taking one observation X_i and Y_i, respectively.

The family $\{P_\theta : \theta \in \mathbb{R}\}$ is invariant under the location transformations $x \mapsto x + \alpha$ ($y \mapsto y + \alpha$) with $\alpha \in \mathbb{R}$. Consequently, the decision problem is invariant under location transformations if and only if $\mathcal{L}(\theta, a) = \mathcal{L}(\theta + \alpha, a + \alpha)$ for all $\alpha \in \mathbb{R}$, which is equivalent to

$$\mathcal{L}(\theta, a) = \delta(a - \theta) \tag{1.4}$$

for a Borel function $\delta(\cdot)$ on $\mathbb{R}$. An estimator d of the parameter θ is *location equivariant* if and only if

$$d(X_1 + \alpha, ..., X_{n_1} + \alpha, Y_1 + \alpha, ..., Y_{n_2} + \alpha) = d(X_1, ..., X_{n_1}, Y_1, ..., Y_{n_2}) + \alpha.$$

Suppose that we agree to take at least one observation. If we observe the process for $t \geq t_1$ units of time, then the conditional expected loss, given $k_1(t)$ and $k_2(t)$, associated with an equivariant estimator $d\left(\mathbf{X}_{k_1(t)}, \mathbf{Y}_{k_2(t)}\right)$ based on the random size samples $\mathbf{X}_{k_1(t)} = \left(X_1, ..., X_{k_1(t)}\right)$ and $\mathbf{Y}_{k_2(t)} = \left(Y_1, ..., Y_{k_2(t)}\right)$ is of the form

$$\begin{aligned} \mathcal{R}_t\left(\theta, d\left(\mathbf{X}_{k_1(t)}, \mathbf{Y}_{k_2(t)}\right)\right) &:= E_\theta\left[\mathcal{L}_t\left(\theta, d\left(\mathbf{X}_{k_1(t)}, \mathbf{Y}_{k_2(t)}\right)\right) \Big| k_1(t), k_2(t)\right] \\ &= h_1(k_1(t)) + h_2(k_2(t)) + c_1(t) + c_2(t), \end{aligned} \tag{1.5}$$

where E_θ means the expectation with respect to the conditional distribution given θ. The functions h_1 and h_2 depend only on the loss function δ.

The form of the risk function $\mathcal{R}_t(\theta, d)$, given by (1.5), follows from the fact that the risk of any equivariant estimator of the parameter θ in the invariant problem of estimation is independent of θ (see e.g. Lehmann and Casella 1998, Theorem 3.1.4). Hence, if an equivariant estimator exists which minimizes the constant risk, it is called *the minimum risk equivariant (MRE) estimator.*

In Section 2 we present the method of finding a stopping time which minimizes the expected risk associated with a MRE estimator of the parameter θ over all stopping times. We consider a situation when the common distributions of the random variables $U_1, ..., U_{n_1}$ and $V_1, ..., V_{n_2}$, respectively, which can be interpreted as the lifetimes of n_1 and n_2 objects are known exactly. In Section 3 we apply the results of Section 2 to estimate a common location parameter of two normal distributions under the squared error loss and a LINEX loss function. Additionally, in Section 4 some illustrative simulations are given.

2. THE OPTIMAL STOPPING TIME

Suppose that in the estimation problem of the parameter θ with the loss function $\mathcal{L}(\theta, d)$ there exists an MRE estimator, denoted by d^*. We look for a stopping time τ^* which minimizes the expected risk

$$(2.1)\quad E\left[\mathcal{R}_\tau\left(\theta, d^*\left(\mathbf{X}_{k_1(\tau)}, \mathbf{Y}_{k_2(\tau)}\right)\right)\right] = E[h_1(k_1(\tau)) + h_2(k_2(\tau)) + c_1(\tau) + c_2(\tau)]$$

over all stopping times $\tau \geq t_1$, $\tau \in \mathcal{T}$, where $\mathcal{T}$ denotes the class of $(\mathcal{F}_{1,t}, \mathcal{F}_{2,t})$-measurable functions. Such a stopping time will be called *an optimal stopping time*. Then we construct *an optimal sequential estimation procedure* of the form $\left(\tau^*, d^*\left(\mathbf{X}_{k_1(\tau^*)}, \mathbf{Y}_{k_2(\tau^*)}\right)\right)$.

Let the random variables $U_1, ..., U_{n_1}$ be independent and have a common known distribution function G_1. Suppose that $G_1(0) = 0$, $G_1(t) > 0$ for $t > 0$, G_1 is absolutely continuous with density g_1, and g_1 is the right hand derivative of G_1 on $(0, \infty)$. Denote the class of such G_1 by $\mathcal{G}_1$. Let $\zeta_1 = \sup\{t : G_1(t) < 1\}$, and $\rho_1(t) = g_1(t)[1 - G_1(t)]^{-1}$, $0 \leq t < \zeta_1$, denote the failure rate. Under the above assumptions the process $k_1(t)$, given by (1.1), is a nonstationary Markov chain with respect to $\mathcal{F}_{1,t}$, $0 \leq t \leq \zeta_1$ (see Starr *et al.* (1976)). The random variables $V_1, ..., V_{n_2}$ satisfy the analogous assumptions. Namely, let the random variables $V_1, ..., V_{n_2}$ be independent and have a common known distribution function G_2. Suppose that $G_2(0) = 0$, $G_2(t) > 0$ for $t > 0$, G_2 is absolutely continuous with density g_2, and g_2 is the right hand derivative of G_2 on $(0, \infty)$. Denote the class of such G_2 by $\mathcal{G}_2$. Let $\zeta_2 = \sup\{t : G_2(t) < 1\}$, and $\rho_2(t) = g_2(t)[1 - G_2(t)]^{-1}$, $0 \leq t < \zeta_2$, denote the failure rate. Under the above assumptions the process $k_2(t)$, given by (1.2), is a nonstationary Markov chain with respect to $\mathcal{F}_{2,t}$, $0 \leq t \leq \zeta_2$.

The infinitesimal operator $\mathcal{A}_{1,t}$ of the processes $k_1(t)$ at $\widetilde{h}_1$ is defined by

$$(2.2)\qquad \mathcal{A}_{1,t}\widetilde{h}_1(k) := \lim_{s\to 0^+} s^{-1}E\left[\widetilde{h}_1(k_1(t+s)) - \widetilde{h}_1(k_1(t))|k_1(t) = k\right].$$

The domain $D_{\mathcal{A}_{1,t}}$ of $\mathcal{A}_{1,t}$ is the set of all bounded Borel measurable functions $\widetilde{h}_1$ on the set $\{0, 1, ..., n_1\}$ for which the limit in (2.2) exists boundedly pointwise for every $k \in \{0, 1, ..., n_1\}$. The infinitesimal operator $\mathcal{A}_{2,t}$ of the processes $k_2(t)$ is defined analogously.

To determine an optimal stopping time we use the following lemma which provides the form of the infinitesimal operator $\mathcal{A}_{1,t}$ of the process $k_1(t)$, given by (1.1).

Lemma 2.1. *Let $\widetilde{h}_1$ be a given real-valued function on the set $\{0, 1, ..., n_1\}$. The infinitesimal operator $\mathcal{A}_{1,t}$ of the process $k_1(t)$, given by (1.1), at $\widetilde{h}_1$ is of the form*

$$\mathcal{A}_{1,t}\widetilde{h}_1(k) = (n_1 - k)\left[\widetilde{h}_1(k+1) - \widetilde{h}_1(k)\right]\rho_1(t).$$

Proof: Fix $k \in \{0, 1, ..., n_1\}$. It is clear that

$$E\left[\widetilde{h}_1(k_1(t+s)) - \widetilde{h}_1(k_1(t))|k_1(t) = k\right] =$$

$$= \sum_{i=k+1}^{n_1} \left[\widetilde{h}_1(i) - \widetilde{h}_1(k)\right] P\left(k_1(t+s) = i|k_1(t) = k\right)$$

$$= \left[\widetilde{h}_1(k+1) - \widetilde{h}_1(k)\right] P\left(k_1(t+s) = k+1|k_1(t) = k\right)$$

$$+ \sum_{i=k+2}^{n_1} \left[\widetilde{h}_1(i) - \widetilde{h}_1(k)\right] P\left(k_1(t+s) = i|k_1(t) = k\right)$$

$$= \left[\widetilde{h}_1(k+1) - \widetilde{h}_1(k)\right] (n_1 - k) \frac{G_1(t+s) - G_1(t)}{1 - G_1(t)} \left[\frac{1 - G_1(t+s)}{1 - G_1(t)}\right]^{n_1 - k - 1}$$

$$+ \sum_{i=k+2}^{n_1} \left[\widetilde{h}_1(i) - \widetilde{h}_1(k)\right] P\left(k_1(t+s) = i|k_1(t) = k\right)$$

$$\leq \left[\widetilde{h}_1(k+1) - \widetilde{h}_1(k)\right] (n_1 - k) \frac{G_1(t+s) - G_1(t)}{1 - G_1(t)} \left[\frac{1 - G_1(t+s)}{1 - G_1(t)}\right]^{n_1 - k - 1}$$

$$+ 2 \sup_{i \leq n_1} \left|\widetilde{h}_1(i)\right| P\left(k_1(t+s) \geq k+2|k_1(t) = k\right)$$

$$= \left[\widetilde{h}_1(k+1) - \widetilde{h}_1(k)\right] (n_1 - k) \frac{G_1(t+s) - G_1(t)}{1 - G_1(t)} \left[\frac{1 - G_1(t+s)}{1 - G_1(t)}\right]^{n_1 - k - 1}$$

$$+ 2 \sup_{i \leq n_1} \left|\widetilde{h}_1(i)\right| \left\{1 - \left[\frac{1 - G_1(t+s)}{1 - G_1(t)}\right]^{n_1 - k} \left[1 - (n_1 - k) \frac{G_1(t+s) - G_1(t)}{[1 - G_1(t+s)]}\right]\right\}.$$

Now it is easy to see that

$$\lim_{s \to 0^+} \frac{E\left[\widetilde{h}_1(k_1(t+s)) - \widetilde{h}_1(k_1(t))|k_1(t) = k\right]}{s} =$$

$$= (n_1 - k)\left[\widetilde{h}_1(k+1) - \widetilde{h}_1(k)\right] \rho_1(t)$$

and the lemma is proved. □

The infinitesimal operator $\mathcal{A}_{2,t}$ of the processes $k_2(t)$ is calculated analogously and we have

$$\mathcal{A}_{2,t}\widetilde{h}_2(k) = (n_2 - k)\left[\widetilde{h}_2(k+1) - \widetilde{h}_2(k)\right] \rho_2(t).$$

Let $\widetilde{h}_1(k) = h_1(k)$ for $k = 1, ..., n_1$ and $\widetilde{h}_1(0) = 0$, and $\widetilde{h}_2(k) = h_2(k)$ for $k = 1, ..., n_2$ and $\widetilde{h}_2(0) = 0$. The following theorem determines the optimal stopping time τ^* for a large class of possible h_1 and h_2.

Theorem 2.1. *Suppose that $G_1 \in \mathcal{G}_1$ has non-increasing failure rate ρ_1, $G_2 \in \mathcal{G}_2$ has non-increasing failure rate ρ_2, and the functions $h_1(k)$ and $h_2(k)$ in* *formula (1.5) are such that $h_1(k)-h_1(k+1)$ is non-increasing for $k \in \{1,...,n_1-1\}$ and $h_2(k)-h_2(k+1)$ is non-increasing for $k \in \{1,...,n_2-1\}$. Then the stopping time*

$$\begin{aligned}\tau^* &= \inf\Big\{t \geq t_1 :\ \mathcal{A}_{1,t}\widetilde{h}_1(k_1(t)) + \mathcal{A}_{2,t}\widetilde{h}_2(k_2(t)) + c_1'(t) + c_2'(t) \geq 0\Big\}\\ &= \inf\Big\{t \geq t_1 :\ (n_1-k_1(t))[h_1(k_1(t)) - h_1(k_1(t)+1)]\rho_1(t)\\ &\qquad + (n_2-k_2(t))[h_2(k_2(t)) - h_2(k_2(t)+1)]\rho_2(t) \leq c_1'(t)+c_2'(t)\Big\} \end{aligned} \tag{2.3}$$

minimizes the expected risk given by (2.1) over all stopping times $\tau \geq t_1$, $\tau \in \mathcal{T}$.

Proof: The proof follows Starr *et al.* (1976), Theorem 2.1. Using Dynkin's formula, we have

$$\begin{aligned}&E[\widetilde{h}_1(k_1(\tau)) + \widetilde{h}_2(k_2(\tau)) + c_1(\tau) + c_2(\tau)] \ = \\ &\qquad = E\left\{\int_0^\tau [\mathcal{A}_{1,t}\widetilde{h}_1(k_1(t)) + \mathcal{A}_{2,t}\widetilde{h}_2(k_2(t)) + c_1'(t) + c_2'(t)]dt\right\}\end{aligned}$$

for all stopping times τ. In particular for $\tau \geq t_1$ we have

$$\begin{aligned}&E[h_1(k_1(\tau^*)) + h_2(k_2(\tau^*)) + c_1(\tau^*) + c_2(\tau^*)] \ - \\ &\qquad -\ E[h_1(k_1(\tau)) + h_2(k_2(\tau)) + c_1(\tau) + c_2(\tau)] \ = \\ &=\ E[\widetilde{h}_1(k_1(\tau^*)) + \widetilde{h}_2(k_2(\tau^*)) + c_1(\tau^*) + c_2(\tau^*)] \\ &\quad -\ E[\widetilde{h}_1(k_1(\tau)) + \widetilde{h}_2(k_2(\tau)) + c_1(\tau) + c_2(\tau)] \\ &= E\left\{\int_\tau^{\tau^*} [\mathcal{A}_{1,t}\widetilde{h}_1(k_1(t)) + \mathcal{A}_{2,t}\widetilde{h}_2(k_2(t)) + c_1'(t) + c_2'(t)]dt\mathbf{1}(\tau < \tau^*)\right\} \\ &\quad - E\left\{\int_{\tau^*}^{\tau} [\mathcal{A}_{1,t}\widetilde{h}_1(k_1(t)) + \mathcal{A}_{2,t}\widetilde{h}_2(k_2(t)) + c_1'(t) + c_2'(t)]dt\right\}\mathbf{1}(\tau > \tau^*).\end{aligned} \tag{2.4}$$

Taking into account the assumptions concerning the function $h_1(k)$, $h_2(k)$, $c_1(t)$, $c_2(t)$, $\rho_1(t)$ and $\rho_2(t)$ we have that (2.4) is less or equal to zero. Thus, the stopping time τ^* is optimal. □

3. SPECIAL CASE

In this section we use the solutions of Section 2 to estimate a common location parameter of two normal distributions under the squared error loss

$$\mathcal{L}(\theta, d) = (d-\theta)^2 \tag{3.1}$$

and under a LINEX loss function

$$\mathcal{L}(\theta, d) = \exp[a(d-\theta)] - a(d-\theta) - 1, \tag{3.2}$$

where $a \neq 0$. Taking MRE estimators as optimal estimators of a location parameter of two normal distributions, we construct optimal sequential estimation procedures under the aforementioned loss functions in the model with observations which are available at random times.

Let X_i, $i = 1, ..., n_1$, be independent random variables from the normal distribution $\mathcal{N}(\theta, \sigma_1^2)$ and Y_i, $i = 1, ..., n_2$, be independent random variables from the normal distribution $\mathcal{N}(\theta, \sigma_2^2)$, where $\theta \in \mathbb{R}$ is an unknown location parameter and σ_1, $\sigma_2 > 0$ are known. We assume that the samples $(X_1, ..., X_{n_1})$ and $(Y_1, ..., Y_{n_2})$ are independent and $\sigma_1 \neq \sigma_2$.

Let

$$\overline{X}_{k_1(t)} = \frac{1}{k_1(t)} \sum_{i=1}^{k_1(t)} X_i, \quad \overline{Y}_{k_2(t)} = \frac{1}{k_2(t)} \sum_{i=1}^{k_2(t)} Y_i$$

denote the sample means based on the random size sample $\mathbf{X}_{k_1(t)} = \big(X_1, ..., X_{k_1(t)}\big)$ and $\mathbf{Y}_{k_2(t)} = \big(Y_1, ..., Y_{k_2(t)}\big)$, respectively, where $k_1(t)$ is given by (1.1) and $k_2(t)$ is given by (1.2).

The following theorem provides the MRE estimator of the parameter θ and the corresponding risk function under the loss function given by (3.1) and (3.2), respectively.

Theorem 3.1. *For any stopping time t*

(**a**) *If the loss function is given by (3.1), then the MRE estimator of the parameter θ is*

$$d_S^* \left(\mathbf{X}_{k_1(t)}, \mathbf{Y}_{k_2(t)}\right) = \omega \overline{X}_{k_1(t)} + (1-\omega) \overline{Y}_{k_2(t)}$$

with $\omega \in (0,1)$, and the risk function of the estimator d_S^* has the form

$$\mathcal{R}_t\left(\theta, d_S^*\right) = \frac{\omega^2 \sigma_1^2}{2k_1(t)} + \frac{(1-\omega)^2 \sigma_2^2}{2k_2(t)} + c_A k_1(t) + c_B k_2(t) + c_1(t) + c_2(t).$$

(**b**) *If the loss function is given by (3.2), then the MRE estimator of the parameter θ is*

$$\begin{aligned} d_L^*\big(\mathbf{X}_{k_1(t)}, \mathbf{Y}_{k_2(t)}\big) &= \omega \left(\overline{X}_{k_1(t)} - \frac{a\sigma_1^2}{2k_1(t)}\right) + (1-\omega) \left(\overline{Y}_{k_2(t)} - \frac{a\sigma_2^2}{2k_2(t)}\right) \\ &\quad + \omega(1-\omega)\, a \left(\frac{\sigma_1^2}{2k_1(t)} + \frac{\sigma_2^2}{2k_2(t)}\right) \end{aligned}$$

with $\omega \in (0,1)$, and the risk function of the estimator d_L^* has the form

$$\mathcal{R}_t(\theta, d_L^*) = \frac{\omega^2 a^2 \sigma_1^2}{2k_1(t)} + \frac{(1-\omega)^2 a^2 \sigma_2^2}{2k_2(t)} + c_A k_1(t) + c_B k_2(t) + c_1(t) + c_2(t).$$

Proof: The forms of the MRE estimators d_S^* and d_L^* are obtained from the general formula for the MRE estimators of the location parameter under the loss function (1.4) (see e.g. Shao (2003), Theorem 4.5). The formulas for the risk functions $\mathcal{R}_t(\theta, d_S^*)$ and $\mathcal{R}_t(\theta, d_L^*)$ follow from straightforward calculations. □

On the basis of Theorems 2.1 and 3.1 we construct optimal sequential estimation procedures of the form $\left(\tau^*, d^*\left(\mathbf{X}_{k_1(\tau^*)}, \mathbf{Y}_{k_2(\tau^*)}\right)\right)$, where τ^* is defined by (2.3), and d^* is the corresponding sequential MRE estimator of θ based on the random size samples $\mathbf{X}_{k_1(\tau^*)}$ and $\mathbf{Y}_{k_2(\tau^*)}$.

The next theorem determines the optimal sequential estimation procedure under the loss function $\mathcal{L}(\theta, d)$ given by (3.1) and (3.2), respectively.

Theorem 3.2. *Suppose that $G_1 \in \mathcal{G}_1$ has non-increasing failure rate ρ_1 and $G_2 \in \mathcal{G}_2$ has non-increasing failure rate ρ_2.*

(**a**) *Under the loss function $\mathcal{L}_t(\theta, d)$ given by (1.3) with $\mathcal{L}(\theta, d)$ of the form (3.1), the sequential estimation procedure $\left(\tau_S^*, d_S^*\left(\mathbf{X}_{k_1(\tau_S^*)}, \mathbf{Y}_{k_2(\tau_S^*)}\right)\right)$, where*

$$\tau_S^* = \inf\left\{t \geq t_1 :\ (n_1 - k_1(t))\left[\frac{\omega^2\sigma_1^2}{2k_1(t)} - \frac{\omega^2\sigma_1^2}{2(k_1(t)+1)} - c_A\right]\rho_1(t) \right.$$
$$\left. + (n_2 - k_2(t))\left[\frac{(1-\omega)^2\sigma_2^2}{2k_2(t)} - \frac{(1-\omega)^2\sigma_2^2}{2(k_2(t)+1)} - c_B\right]\rho_2(t) \leq c_1'(t) + c_2'(t)\right\}$$

and

$$d_S^*\left(\mathbf{X}_{k_1(\tau_S^*)}, \mathbf{Y}_{k_2(\tau_S^*)}\right) = \omega\overline{X}_{k_1(\tau_S^*)} + (1-\omega)\overline{Y}_{k_2(\tau_S^*)}$$

is optimal.

(**b**) *Under the loss function $\mathcal{L}_t(\theta, d)$ given by (1.3) with $\mathcal{L}(\theta, d)$ of the form (3.2), the sequential estimation procedure $\left(\tau_L^*, d_L^*\left(\mathbf{X}_{k_1(\tau_L^*)}, \mathbf{Y}_{k_2(\tau_L^*)}\right)\right)$, where*

$$\tau_L^* = \inf\left\{t \geq t_1 :\ (n_1 - k_1(t))\left[\frac{\omega^2a^2\sigma_1^2}{2k_1(t)} - \frac{\omega^2a^2\sigma_1^2}{2(k_1(t)+1)} - c_A\right]\rho_1(t) \right.$$
$$\left. + (n_2 - k_2(t))\left[\frac{(1-\omega)^2a^2\sigma_2^2}{2k_2(t)} - \frac{(1-\omega)^2a^2\sigma_2^2}{2(k_2(t)+1)} - c_B\right]\rho_2(t) \leq c_1'(t) + c_2'(t)\right\}$$

and

$$\begin{aligned} d_L^*\left(\mathbf{X}_{k_1(\tau_L^*)}, \mathbf{Y}_{k_2(\tau_L^*)}\right) = {} & \omega\left(\overline{X}_{k_1(\tau_L^*)} - \frac{a\sigma_1^2}{2k_1(\tau_L^*)}\right) \\ & + (1-\omega)\left(\overline{Y}_{k_2(\tau_L^*)} - \frac{a\sigma_2^2}{2k_2(\tau_L^*)}\right) \\ & + \omega(1-\omega)\, a\left(\frac{\sigma_1^2}{2k_1(\tau_L^*)} + \frac{\sigma_2^2}{2k_2(\tau_L^*)}\right) \end{aligned}$$

is optimal.

Proof: We have to show that the assumptions of Theorem 2.1 are satisfied, i.e., the functions $h_1(k) - h_1(k+1)$ and $h_2(k) - h_2(k+1)$ are non-increasing on the set $\{1, ..., n_1 - 1\}$ and $\{1, ..., n_2 - 1\}$, respectively. Hence, we need to verify the condition $2h_1(k+1) - h_1(k) - h_1(k+2) \le 0$ and $2h_2(k+1) - h_2(k) - h_2(k+2) \le 0$, which are equivalent to $h_1(k+1) \le (h_1(k) + h_1(k+2))/2$ and $h_2(k+1) \le (h_2(k) + h_2(k+2))/2$. This can be reduced to the verification that h_1 and h_2 are convex on the interval $[1, n_1 - 1]$ and $[1, n_2 - 1]$, respectively. It is easy to see that

(**a**) $$h_1''(k) = \frac{\omega^2 \sigma_1^2}{k^3}, \quad h_2''(k) = \frac{(1-\omega)^2 \sigma_2^2}{k^3}$$

and $h_1''(k) > 0$, $h_2''(k) > 0$ for $k \ge 1$;

(**b**) $$h_1''(k) = \frac{\omega^2 a^2 \sigma_1^2}{k^3}, \quad h_2''(k) = \frac{(1-\omega)^2 a^2 \sigma_2^2}{k^3}$$

and $h_1''(k) > 0$, $h_2''(k) > 0$ for $k \ge 1$. □

4. SIMULATION RESULTS

In this section we present some results of the numerical study. The first table contains the results of the simulation study for $X_1, ..., X_{n_1} \sim \mathcal{N}(0,1)$, $n_1 = 30$ and $Y_1, ..., Y_{n_2} \sim \mathcal{N}(0,25)$, $n_2 = 50$: the means of τ_S^*, d_S^*, τ_L^* and d_L^* for $a = 2$, over the 1000 replications, when $\omega = 0.25$, $\rho_1(t) = 1\,(U_i \sim \mathcal{E}(1))$, $c_1(t) = t^2$ and $\rho_2(t) = (2 \cdot \sqrt{3t})^{-1}, (V_i \sim \mathcal{W}e(1/2,3))$, $c_2(t) = e^t - 1$.

c_A	c_B	Mean(τ_S^*)	Mean(d_S^*)	Mean(τ_L^*)	Mean(d_L^*)
0.005	0.000001	0.2388	−0.0094	0.4705	−0.8777
0.000001	0.005	0.2339	−0.0099	0.4629	−0.8949
0.005	0.005	0.2281	0.0123	0.4537	−0.8943
0.000001	0.000001	0.2455	−0.0406	0.4687	−0.9203

The second table contains the results of the simulation study for $X_1, ..., X_{n_1} \sim \mathcal{N}(0,1)$, $n_1 = 30$ and $Y_1, ..., Y_{n_2} \sim \mathcal{N}(0,25)$, $n_2 = 50$: the means of τ_S^*, d_S^*, τ_L^* and d_L^* for $a = 2$, over the 1000 replications, when $\omega = 0.5$, $\rho_1(t) = (2 \cdot \sqrt{3t})^{-1}\,(U_i \sim \mathcal{W}e(1/2,3))$, $c_1(t) = e^t - 1$ and $\rho_2(t) = 1\,(V_i \sim \mathcal{E}(1))$, $c_2(t) = t^2$.

c_A	c_B	Mean(τ_S^*)	Mean(d_S^*)	Mean(τ_L^*)	Mean(d_L^*)
0.005	0.000001	0.3634	0.0036	0.6073	−0.4517
0.000001	0.005	0.3657	−0.0088	0.6144	−0.4831
0.005	0.005	0.3610	−0.0097	0.6102	−0.4881
0.000001	0.000001	0.3665	0.0134	0.6165	−0.4453

The third table contains the results of the simulation study for $X_1, ..., X_{n_1} \sim \mathcal{N}(0,1)$, $n_1 = 30$ and $Y_1, ..., Y_{n_2} \sim \mathcal{N}(0,25)$, $n_2 = 50$: the means of τ_S^*, d_S^*, τ_L^* and d_L^* for $a = 2$, over the 1000 replications, when $\omega = 0.75$, $\rho_1(t) = (2 \cdot \sqrt{3t})^{-1}$ ($U_i \sim \mathcal{W}e(1/2, 3)$), $c_1(t) = t^2$ and $\rho_2(t) = 1$ ($V_i \sim \mathcal{E}(1)$), $c_2(t) = e^t - 1$.

c_A	c_B	Mean(τ_S^*)	Mean(d_S^*)	Mean(τ_L^*)	Mean(d_L^*)
0.005	0.000001	0.2975	−0.0107	0.5129	−0.1816
0.000001	0.005	0.2956	0.0284	0.5076	−0.1504
0.005	0.005	0.2908	0.0014	0.5010	−0.1417
0.000001	0.000001	0.3021	0.0106	0.5110	−0.1710

Simulation results above are consistent with expectations. The both procedures are working properly. In case of Linex loss function, decision function is biased, however it is MRE estimator because ω is fixed. It could be applicable especially in a case when one sample is more preferable than second one.

REFERENCES

[1] Rao, M.S. and Reddy, M.K. (1988). Estimation of a common location parameter of two distributions, *Aligarh J. Statist.*, **8**, 6–13.

[2] Baklizi, A. (2004). Shrinkage estimation of the common location parameter of several exponentials, *Comm. Statist. Simulation Comput.*, **33**(2), 321–339.

[3] Farsipour, N.S. and Asgharzadeh, A. (2002). On the admissibility of estimators of the common mean of two normal populations under symmetric and asymmetric loss functions, *South African Statist. J.*, **36**(1), 39–54.

[4] Mitra, P.K. and Sinha, B.K. (2007). On some aspects of estimation of a common mean of two independent normal populations, *J. Statist. Plann. Inference*, **137**(1), 184–193.

[5] Chang, Y.; Oono, Y. and Shinozaki, N. (2012). Improved estimators for the common mean and ordered means of two normal distributions with ordered variances, *J. Statist. Plann. Inference*, **142**(9), 2619–2628.

[6] Nair, K.A. (1982). An estimator of the common mean of two normal populations, *J. Statist. Plann. Inference*, **6**(2), 119–122.

[7] Elfessi, A. and Pal, N. (1992). A note on the common mean of two normal populations with order restricted variances, *Comm. Statist. Theory Methods*, **21**(11), 3177–3184.

[8] Graybill, F.A. and Deal, R.B. (1959). Combining unbiased estimators, *Biometrics*, **15**, 543–550.

[9] Starr, N.; Wardrop, R. and Woodroofe, M. (1976). Estimating a mean from delayed observations, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **35**(2), 103–113.

[10] Magiera, R. (1996). On a class of sequential estimation problems for one-parameter exponential families, *Sankhyā Ser. A*, **58**(1), 160–170.

[11] Jokiel-Rokita, A. and Stępień, A. (2009). Sequential estimation of a location parameter from delayed observations, *Statist. Papers*, **50**(2), 363–372.

[12] Lehmann, E.L. and Casella, G. (1998). *Theory Of Point Estimation*, Springer-Verlag, New York.

[13] Shao, J. (2003). *Mathematical Statistics*, Springer-Verlag, New York.