REVSTAT – Statistical Journal
Volume 15, Number 4, October 2017, 501–522

ESTIMATING RENYI ENTROPY OF SEVERAL EXPONENTIAL DISTRIBUTIONS UNDER AN ASYMMETRIC LOSS FUNCTION

Authors: Suchandan Kayal
– Department of Mathematics, National Institute of Technology Rourkela, Rourkela-769008, India
suchandan.kayal@gmail.com
kayals@nitrkl.ac.in

Somesh Kumar
– Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur-721302, India
smsh@iitkgp.ernet.ac.in

Received: December 2014 Revised: September 2015 Accepted: January 2016

Abstract:

- The present paper takes into account the estimation of the Renyi entropy of several exponential distributions under a linex loss function. The models under study are (i) several exponential distributions with a common scale parameter and unknown but unequal location parameters and (ii) several exponential distributions with a common location parameter and unknown but unequal scale parameters. Improvements over the best affine equivariant estimator are obtained for the first model considering unrestricted and restricted parameter spaces. For the second model, sufficient conditions for improvement over affine and scale equivariant estimators are obtained and consequently, improvements over the maximum likelihood estimator and the uniformly minimum variance unbiased estimator are proposed. Sections on numerical studies have been included after each model to present comparative study of the relative risk performance of the proposed improved estimators.

Key-Words:

- *Renyi entropy; linex loss function; equivariant estimator; inadmissibility; Brewster-Zidek technique.*

AMS Subject Classification:

- 62F10, 62C15.

1. INTRODUCTION

The Shannon entropy (see Shannon (1948)) is a fundamental measure of information content and has been applied in a wide variety of fields such as statistical thermodynamics, urban and regional planning, business, economics, finance, operations research, queueing theory, spectral analysis, image reconstruction, biology and manufacturing. One may refer to Kapur (1990) and Cover and Thomas (2006) for examples of various applications. Several generalized information-theoretic measures have been proposed in the literature to measure the uncertainty of a probability distribution since the seminal work of Shannon (1948). Among these, one of the most important and applicable measures is proposed by Renyi (1961). For a random variable X with probability density function $f(x|\theta), \theta \in \Theta$, the Renyi entropy is given by

$$R_\alpha(X) = \frac{1}{1-\alpha} \ln \int_{-\infty}^{\infty} f^\alpha(x|\theta)dx, \quad \alpha\ (\neq 1) > 0. \tag{1.1}$$

Note that we are using logarithm to base e in the expression given by (1.1). Here, the unit of the information measure is nat. Golshani and Pasha (2010) provide some important properties of the measure given in (1.1):

- **(i)** The Renyi entropy can be negative,
- **(ii)** It is invariant under location transformation, but not under scale transformation, and
- **(iii)** For any $\alpha_1 < \alpha_2$, $R_{\alpha_2}(X) \leq R_{\alpha_1}(X)$ and equality holds if and only if X is a uniform random variable.

Using L'Hospital rule, it can be shown that (1.1) retrieves the Shannon entropy when α tends to 1. The Renyi entropy is more or less sensitive to the shape of the probability distributions due to the parameter α. For large values of α, the measure given in (1.1) is more sensitive to events that occur often. Likewise, for small values of α, it is more sensitive to the events which happen seldom. In many instances, the Renyi entropy is seen to be more useful than the Shannon entropy (see Nilsson (2006), Maszczyk and Duch (2008) and Pharwaha and Khehra (2009)). The measure given in (1.1) has found a lot of applications in different areas of science and technology. For example, in speech recognition, different values of α determine different concepts of noisiness. Basically, small α values tend to emphasize the noise content of signal, while large α values tend to emphasize the harmonic content of a signal (see Obin and Liuni (2012)). It is also used for ultrasonic molecular imaging (see Hughes *et al.* (2009)). For properties of Renyi entropy one may refer to Song (2001), Bercher (2008), De-Gregorio and Iacus (2009), Golshani and Pasha (2010) and Renyi (2012).

Recently, the problem of estimating a common characteristic of several independent populations has received a considerable attention. There are many situations where this problem arises. For example, this situation arises when the information from several independent studies are combined or in meta-analysis. Meta-analysis is used in clinical studies. This is also seen in many statistical designs such as balanced incomplete block designs, panel models and regression models. The present paper is concerned with the problem of estimating the Renyi entropy of several exponential populations with respect to linex loss function (see Varian (1975)). The linex loss function is given as

$$L(\triangle) = p'[\exp\{p\triangle\} - p\triangle - 1],\ \triangle = \delta - \theta,\ p \neq 0,\ p' > 0, \tag{1.2}$$

where p and p' are shape and scale parameters, respectively. Without loss of generality, we assume $p' = 1$. Note that the loss function (1.2) reduces to the squared error loss function when $|p|$ tends to 0. For more properties on linex loss function one may refer to Zellner (1986).

Let $\Pi_1, ..., \Pi_k$ be k (≥ 2) exponential populations with location and scale parameters $\underline{\mu} = (\mu_1, ..., \mu_k)$ and $\underline{\sigma} = (\sigma_1, ..., \sigma_k)$, respectively. The probability density function of the i-th population Π_i is given by

$$f_i(x|\mu_i, \sigma_i) = \begin{cases} \sigma_i^{-1} exp\{-(x - \mu_i)/\sigma_i\}, & \text{if } x > \mu_i \\ 0, & \text{otherwise,} \end{cases} \tag{1.3}$$

where $\mu_i \in \mathbb{R},\ \sigma_i > 0$ and $i = 1, ..., k$. The expression of the Renyi entropy of k exponential distributions can be obtained as $R_\alpha(\underline{\sigma}) = \sum_{i=1}^{k} \ln \sigma_i - k \ln \alpha/(1 - \alpha)$. Several authors attempted the problem of estimating entropy of various continuous probability distributions. In this direction one may refer to Misra *et al.* (2005), Kayal and Kumar (2011*a*, 2011*b*, 2013), and Kayal *et al.* (2015). Misra *et al.* (2005) showed that the best affine equivariant estimator (BAEE) of the Shannon entropy of a multivariate normal distribution is inadmissible with respect to the squared error loss function. Under linex loss function, Kayal and Kumar (2011*a*) derived an estimator improving upon the BAEE of the Shannon entropy of a shifted exponential distribution. Kayal and Kumar (2011*b*) considered the problem of estimating the Shannon entropy of several exponential distributions with respect to both squared error and linex loss functions. Generalized Bayes estimators are showed to be admissible. Kayal and Kumar (2013) obtained improved estimator upon the BAEE in estimating the Shannon entropy of several exponential distributions with a common scale but unequal location parameters with respect to the squared error loss function. Recently, Kayal *et al.* (2015) studied the problem of estimating the Renyi entropy of several exponential distributions with a common location but unequal scale parameters with respect to squared error loss function. They derived the uniformly minimum variance unbiased estimator (UMVUE) and obtained improvements over the UMVUE and the maximum likelihood estimator (MLE). In this communication we deal with the problem of estimating the Renyi entropy in similar models considered by Kayal and Kumar (2013, 2015) with respect to the linex loss function.

The rest of the paper is organized as follows. In Section 2, the problem of estimating the Renyi entropy of several exponential distributions with a common scale and unknown but unequal location parameters is considered. The BAEE is shown to be inadmissible. Further, estimators improving upon the BAEE are obtained when the parameter space is restricted. Relative risks of the proposed estimators are compared numerically. In Section 3, the problem of estimating the Renyi entropy of several exponential distributions with a common location and unequal scale parameters is considered. Inadmissibility results for the scale and affine equivariant estimators are obtained. Further, improved estimators over the MLE and the UMVUE are derived. Some concluding remarks have been added in Section 4. Finally, relative risk performance of the proposed estimators is compared numerically.

2. COMMON SCALE BUT UNEQUAL LOCATION PARAMETERS

As mentioned earlier, in this section, we consider k independent exponential populations with unknown and possibly unequal location parameters $\underline{\mu}$ and a common but unknown scale parameter σ. This model arises in reliability engineering where location parameters can be interpreted as minimum guarantee times of several equipments, whereas the common scale parameter can be considered as unknown but possibly equal failure rate of those equipments. This model is also useful in economy where one may assume the unknown location parameters as the income levels below which the tax filling is not required in different locations. However, the average income levels may be same due to overall economic policies of the country. Let $(X_{i1}, ..., X_{in_i})$ be a random sample of size n_i drawn from the i-th $(i = 1, ..., k)$ population with the probability density function

$$f_i(x|\mu_i, \sigma) = \begin{cases} \sigma^{-1}\ exp\{-(x-\mu_i)/\sigma\}, & \text{if } x > \mu_i \\ 0, & \text{otherwise}, \end{cases} \tag{2.1}$$

where $\mu_i \in \mathbb{R}$ and $\sigma > 0$. For a population with probability density function (2.1), the Renyi entropy can be obtained as $R_\alpha(\sigma) = k \ln\sigma - k\ln\alpha/(1-\alpha)$. It should be mentioned that the problem of estimating $R_\alpha(\sigma)$ with respect to the loss function of the form $L(\theta, \delta) = W(\delta - \theta)$ is equivalent to that of estimating $Q_1(\sigma) = \ln\sigma$. Here, the loss function is given by

$$L^1(Q_1(\sigma), \delta) = exp\{p(\delta - \ln\sigma)\} - p(\delta - \ln\sigma) - 1,\ p \neq 0. \tag{2.2}$$

Note that for the i-th population, $(X_{i(1)}, Y_i)$ is a complete and sufficient statistic of (μ_i, σ), where $X_{i(1)} = min_{1\le j\le n_i} X_{ij}$ and $Y_i = \sum_{j=1}^{n_i} X_{ij}$. We denote $\underline{X}_{(1)} = (X_{1(1)}, ..., X_{k(1)})$, $T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - X_{i(1)})$ and $n = \sum_{i=1}^{k} n_i$. Following the factorization criterion (see Lehmann and Casella, 1998, pp. 35), it can be showed that $(\underline{X}_{(1)}, T)$ is a complete and sufficient statistic of $(\underline{\mu}, \sigma)$. Further, $\underline{X}_{(1)}$ and T are independently distributed. It is seen that $X_{i(1)}$ follows an exponential distribution with location parameter μ_i and scale parameter σ/n_i, and $2T/\sigma$ follows a chi-square distribution with $2(n-k)$ degrees of freedom. The MLE and the UMVUE of $Q_1(\sigma)$ are given by $\delta^1_{ML} = \ln T - \ln n$ and $\delta^1_{MV} = \ln T - \psi(n-k)$, respectively, where ψ denotes digamma function and is given by $\psi(x) = \frac{d}{dx}(\ln\Gamma(x))$.

2.1. The best affine equivariant estimator and its improvement

In this section, we introduce invariance to the problem under study and obtain an improvement over the BAEE. Let $G_{a,b_i} = \{g_{a,b_i}(x_{ij}) : g_{a,b_i}(x_{ij}) = ax_{ij} + b_i, a > 0, b_i \in \mathbb{R}\},\ j = 1, ..., n_i,\ i = 1, ..., k$ be an affine group of transformations. Under the transformation $g_{a,b_i}(x_{ij}) = ax_{ij} + b_i$, the form of an affine equivariant estimator can be obtained as

$$\delta^1_c(\underline{X}_{(1)}, T) = \ln T - c, \tag{2.3}$$

where c is an arbitrary constant. In the following theorem we obtain the BAEE of $Q_1(\sigma)$. The proof is omitted as it is straightforward.

Theorem 2.1. *Under the linex loss function* (2.2), *the BAEE of* $Q_1(\sigma)$ *is* $\delta^1_{c_0}(\underline{X}_{(1)}, T)$, *where* $c_0 = -(1/p)\ln[\Gamma(n-k)/\Gamma(n-k+p)]$.

We consider a group of scale transformations $G_a = \{g_a(x_{ij}) = ax_{ij}, a > 0\},\ j = 1, ..., n_i,\ i = 1, ..., k$. Under the transformation $g_a(x_{ij}) = ax_{ij}$, the form of a scale equivariant estimator is

$$\delta_\phi(\underline{W}, T) = \ln T + \phi(\underline{W}), \tag{2.4}$$

where $\underline{W} = (W_1, ..., W_k),\ W_i = X_{i(1)}/T$ and ϕ is a real valued measurable function. In the following theorem, we prove a general inadmissibility result for the estimators of the form (2.4). First, define

$$\phi_0(\underline{w}) = \begin{cases} \ln u - \frac{1}{p}\ln\left(\frac{\Gamma(n+p)}{\Gamma(n)}\right), & \text{if } \underline{w} \in (B_1 \bigcap B_2) \bigcup (B_3 \bigcap B_2^c) \\ \phi(\underline{w}), & \text{otherwise,} \end{cases} \tag{2.5}$$

where $B_1 = \{\underline{w} : w_{(1)} > 0\},\ B_2 = \{\underline{w} : u < exp(\phi(\underline{w}) + (1/p)\ln(\Gamma(n+p)/\Gamma(n)))\}$, $B_3 = \{\underline{w} : w_{(k)} < 0\}$, $u = \sum_{i=1}^k n_i w_i + 1$, $w_{(1)} = min\{w_1, ..., w_k\}$, $w_{(k)} = max\{w_1, ..., w_k\}$ and $w_i = x_{i(1)}/t,\ i = 1, ..., k$.

Theorem 2.2. *Let* δ_ϕ *be a scale equivariant estimator of the form* (2.4) *and* $\phi_0(\underline{w})$ *be as defined in* (2.5). *If there exists* $(\underline{\mu}, \sigma)$ *such that* $P_{(\underline{\mu},\sigma)}(\phi_0(\underline{W}) \neq \phi(\underline{W})) > 0$, *then under linex loss function* (2.2), *the estimator* δ_{ϕ_0} *dominates* δ_ϕ.

Proof: The risk function of the estimators of the form (2.4) is

$$R(\underline{\mu}, \sigma, \delta_\phi) = E^{\underline{W}} R_1(\underline{\mu}, \sigma, \underline{W}, \delta_\phi),$$

where R_1 denotes the conditional risk of δ_ϕ given $\underline{W} = \underline{w}$, and is given by

$$\begin{aligned} R_1(\underline{\mu}, \sigma, \underline{w}, \delta_\phi) \;=\; & E[(exp\{p(\ln T + \phi(\underline{W}) - \ln\sigma)\} \\ & -p(\ln T + \phi(\underline{W}) - \ln\sigma) - 1)|\underline{W} = \underline{w}]. \end{aligned} \tag{2.6}$$

Note that the conditional risk function $R_1(\underline{\mu}, \sigma, \underline{w}, \delta_\phi)$ given in (2.6) is a function of the ratio $\underline{\mu}/\sigma$. Hence, without loss of generality we may assume σ to be 1. Moreover, the conditional risk is a convex function of ϕ, therefore the choice of ϕ minimizing $R_1(\underline{\mu}, \sigma, \underline{w}, \delta_\phi)$ can be obtained as

$$\hat{\phi}(\underline{\mu}, \underline{w}) = -p^{-1} \ln(E(T^p|\underline{W} = \underline{w})). \tag{2.7}$$

To get improvement over δ_ϕ, it is required to obtain the supremum and infimum of $\hat{\phi}(\underline{\mu}, \underline{w})$ given in (2.7). These can be derived along the arguments of the proof of Theorem 2 of Kayal and Kumar (2013). We omit the details here.

Case **(i):** When all μ_i's, $(i = 1, ..., k)$ are non-negative, the respective supremum and infimum of $\hat{\phi}(\underline{\mu}, \underline{w})$ can be obtained as

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = \ln u - p^{-1} \ln(\Gamma(n+p)/\Gamma(n)) \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = -\infty.$$

Case **(ii):** Assume that μ_i's are negative for $i = 1, ..., k$. Under this restriction, it is required to take into account three possibilities on w_i's: (**a**) all w_i's are non-negative, (**b**) all w_i's are negative and (**c**) some of w_i's, $(i = 1, ..., k)$ are non-negative and remaining are negative. In the following we consider these three sub-cases separately and obtain supremum and infimum of $\hat{\phi}(\underline{\mu}, \underline{w})$.

(**a**) Under the assumption that w_i's are non-negative, we obtain

$$\hat{\phi}(\underline{\mu}, \underline{w}) = \ln u - p^{-1} \ln(\Gamma(n+p)/\Gamma(n)).$$

(**b**) When w_i's are negative, note that the value of u may be positive or negative. For $u > 0$,

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = +\infty \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = \ln u - p^{-1} \ln(\Gamma(n+p)/\Gamma(n));$$

and for $u < 0$,

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = +\infty \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = -\infty.$$

(**c**) Let some of w_i's $(i = 1, ..., k)$ assume non-negative values and the remaining w_i's assume negative values. Thus u may be positive or negative. When $u > 0$, then

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = +\infty \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = \ln u - p^{-1} \ln(\Gamma(n+p)/\Gamma(n));$$

and when $u < 0$, then

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = +\infty \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = -\infty.$$

Case **(iii):** Under the constraints that some of μ_i's are non-negative and remaining are negative, we consider the following sub-cases:

(**a**) For the case when $w_1, ..., w_r, ..., w_k > 0$, we obtain

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = \ln u - p^{-1} \ln(\Gamma(n+p)/\Gamma(n)) \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = -\infty.$$

(**b**) Assume that $w_1, ..., w_r > 0$ and $w_{r+1}, ..., w_k < 0$. Then, for $u \neq 0$,

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = +\infty \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = -\infty.$$

(**c**) Let $w_1, ..., w_r > 0$ and within $(k-r)$, some w_i's be non-negative and remaining be negative. Then, for $u \neq 0$, we obtain

$$\sup_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = +\infty \text{ and } \inf_{\underline{\mu}} \hat{\phi}(\underline{\mu}, \underline{w}) = -\infty.$$

An application of the Brewster and Zidek technique (see Brewster and Zidek (1974)) on the function $R_1(\underline{\mu}, \sigma, \underline{w}, \delta_\phi)$, then completes the proof of the theorem. □

As a consequence of the Theorem 2.2, we get the following corollary which shows that the BAEE obtained in Theorem 2.1 is inadmissible.

Corollary 2.1. The BAEE $\delta^1_{c_0}$ of $Q_1(\sigma)$ is dominated by the estimator

$$\delta^1_{IB} = \begin{cases} \ln(uT) - p^{-1}\ln(\Gamma(n+p)/\Gamma(n)), & \text{if } \underline{w} \in (B_1 \bigcap C_1) \bigcup (B_3 \bigcap C_1^c) \\ \ln T + p^{-1}\ln(\Gamma(n-k)/\Gamma(n-k+p)), & \text{otherwise,} \end{cases}$$

where $C_1 = \{\underline{w} : u < exp(d)\}$ and $d = p^{-1}\ln(\Gamma(n-k)\Gamma(n+p)/\Gamma(n-k+p)\Gamma(n))$.

In this part of the paper we consider the problem of estimating $Q_1(\sigma)$ in restricted parameter spaces. Here we consider the restriction on μ_i's. However, it is seen that it affects the improvement results for the estimation of $Q(\sigma)$. First assume that all μ_i's are bounded below. This arises when the minimum guarantee time of components is known to be more than a pre-specified constant. Without loss of generality, we may assume that $\mu_{(1)} \geq 0$, where $\mu_{(1)} = min\{\mu_1, ..., \mu_k\}$. In this case, δ^1_{ML} is the MLE of $Q_1(\sigma)$. Along the arguments of Case (**i**) of the proof of the Theorem 2.2, the inadmissibility of the BAEE can be established and the improved estimator is

$$\delta_{IB}^{1+} = \begin{cases} \ln(uT) - p^{-1}\ln(\Gamma(n+p)/\Gamma(n)), & \text{if } \underline{w} \in C_1, \\ \\ \ln T + p^{-1}\ln(\Gamma(n-k)/\Gamma(n-k+p)), & \text{otherwise.} \end{cases}$$

We also consider the other case when the guarantee times of the components are known to be bounded above. Without loss of generality we assume $\mu_{(k)} < 0$, where $\mu_{(k)} = max\{\mu_1, ..., \mu_k\}$. In this case, the MLE of $Q_1(\sigma)$ is $\delta_{RM}^1 = \ln T^0 - \ln n$, where $T^0 = \sum_{i=1}^{k}(Y_i - n_i X_{i(1)}^0)$, $X_{i(1)}^0 = min\{0, X_{i(1)}\}$, $i = 1, ..., k$. Along the arguments of Case (**ii**) of the Theorem 2.2, the inadmissibility of the BAEE can be established. The improved estimator is given by

$$\delta_{IB}^{1-} = \begin{cases} \ln(uT) - \frac{1}{p}\ln(\frac{\Gamma(n+p)}{\Gamma(n)}), \text{ if } \underline{w} \in B_1 \bigcup (B_3 \bigcap C_1^c) \bigcup (C_2 \bigcap C_3 \bigcap C_1^c) \\ \\ \ln T + \frac{1}{p}\ln(\frac{\Gamma(n-p)}{\Gamma(n-k+p)}), \text{ otherwise,} \end{cases}$$

where $C_2 = \{\underline{w} : w_{(r)} < 0\}$, $C_3 = \{\underline{w} : w_{(r+1)} > 0\}$.

2.2. Numerical comparisons

In this section, we present the relative risk performance of δ_{IB}^1, δ_{IB}^{1+} and δ_{IB}^{1-} over the BAEE $\delta_{c_0}^1$ through graphs for the case $k = 2$. We assume $\sigma = 1$, as the conditional risk in (2.6) is a function of $(\frac{\mu_1}{\sigma}, \frac{\mu_2}{\sigma})$. It should be mentioned that the risk values of various estimators were calculated using Monte-Carlo simulation based on $10,000$ samples of different combinations of (n_1, n_2) and different values of (μ_1, μ_2) and p. However, we present few of them considering sample sizes $(5, 5)$, $(5, 10)$, $(10, 5)$, $(10, 10)$ and $p = +0.2, -0.2$. It is worthwhile to remark that we observe similar pattern of the relative risk for other values of p and (n_1, n_2).

Based on the Fig. 1 we can conclude the following points:

- (**i**) The margin and the region of the relative risk improvement (RRI) of δ_{IB}^1 over $\delta_{c_0}^1$ becomes small when we increase sample sizes (n_1, n_2).
- (**ii**) We observe considerable RRI of δ_{IB}^1 over $\delta_{c_0}^1$ when both μ_1 and μ_2 approach to origin.
- (**iii**) For fixed (n_1, n_2), the RRI of δ_{IB}^1 over $\delta_{c_0}^1$ is marginally better for negative values of p than positive values of p. For example, the RRI of δ_{IB}^1 over $\delta_{c_0}^1$ is 8.59% at $(\mu_1 = 0.18, \mu_2 = 0)$ for $(n_1 = 5, n_2 = 5)$ and $p = -2$, whereas for the same values of (μ_1, μ_2) and (n_1, n_2), the RRI of δ_{IB}^1 over $\delta_{c_0}^1$ is 8.24% when $p = 2$.

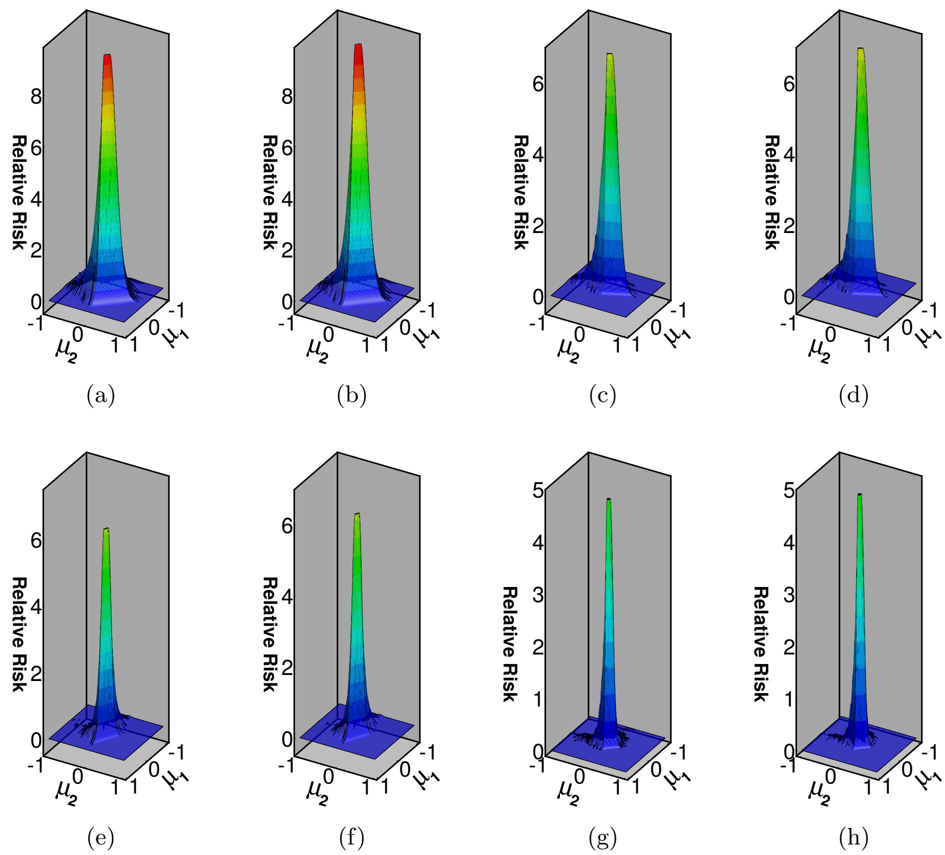


Figure 1: Fig. (a), (b), (c), (d), (e), (f), (g) and (h) represent relative percentage risk improvement plots of δ^1_{IB} over $\delta^1_{c_0}$ for (5,5,+0.2), (5,5,-0.2), (10,5,+0.2), (10,5,-0.2), (5,10,+0.2), (5,10,-0.2), (10,10,+0.2) and (10,10,-0.2), respectively when $(\mu_1, \mu_2) \in \mathbb{R}_2$. The first and second components of the triplet represent the sample sizes of the first and second population, respectively whereas the third component represents the value of p.

Based on the Fig. 2 we get the following observations.

(i) The region as well as the margin of the RRI of δ^{1+}_{IB} over $\delta^1_{c_0}$ become smaller for larger values of (n_1, n_2).

(ii) When μ_i tends to the zero, the RRI of δ^{1+}_{IB} over $\delta^1_{c_0}$ first increases and then decreases, $i = 1, 2$.

(iii) For fixed sample sizes (n_1, n_2), the RRI is marginally better for negative values of p than positive values of p. The RRI of δ^{1+}_{IB} over $\delta^1_{c_0}$ is 9.79% at $(\mu_1 = 0.02, \mu_2 = 0.08)$ for $(n_1 = 5, n_2 = 5)$ and $p = -2$, whereas for the same values of (μ_1, μ_2) and (n_1, n_2), the RRI of δ^1_{IB} over $\delta^1_{c_0}$ is 9.39%, when $p = 2$.

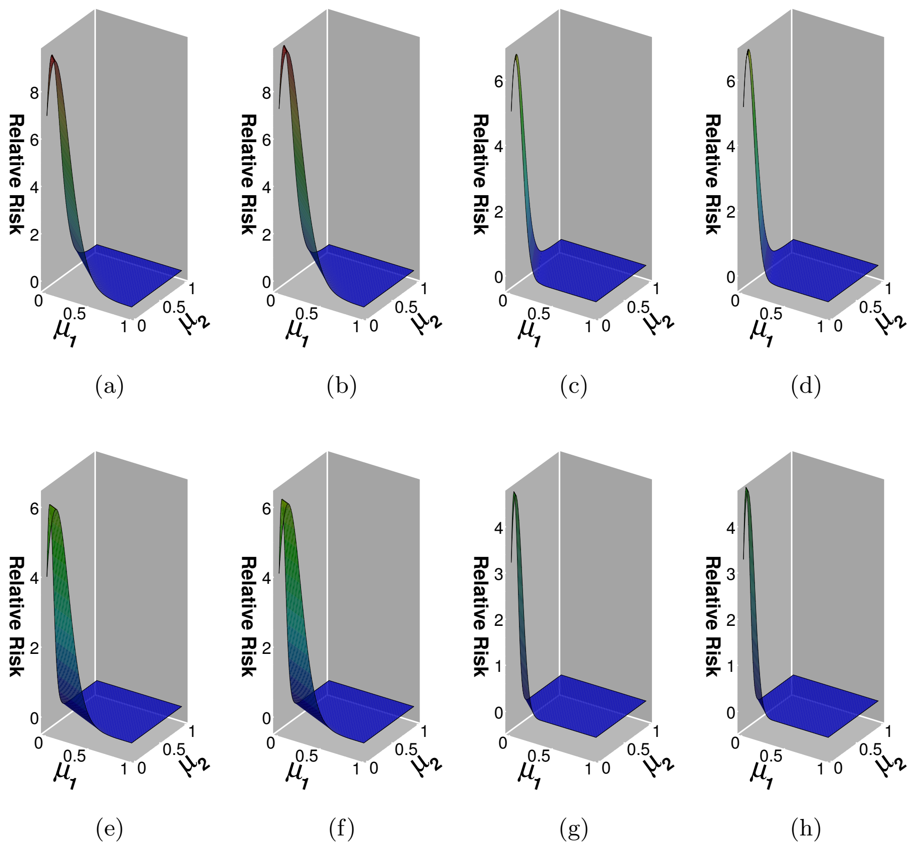


Figure 2: Fig. (a), (b), (c), (d), (e), (f), (g) and (h) represent relative percentage risk improvement plots of δ_{IB}^{1+} over $\delta_{c_0}^{1}$ for (5,5,+0.2), (5,5,-0.2), (10,5,+0.2), (10,5,-0.2), (5,10,+0.2), (5,10,-0.2), (10,10,+0.2) and (10,10,-0.2), respectively when $(\mu_1, \mu_2) \in \mathbb{R}_2^+$. The first and second components of the triplet represent the sample sizes of the first and second population, respectively whereas the third component represents the value of p.

Based on the Fig. 3, we notice the following points.

(i) The margin and the region of the RRI of δ_{IB}^{1-} over $\delta_{c_0}^{1}$ become small when we increase the values of (n_1, n_2).

(ii) When $(\mu_1, \mu_2) \to (0, 0)$, the RRI of δ_{IB}^{1-} over $\delta_{c_0}^{1}$ increases and it attains maximum at some point near origin.

(iii) For fixed (n_1, n_2), the RRI of δ_{IB}^{1-} over $\delta_{c_0}^{1}$ is marginally better for negative values of p than positive values of p. For example, the RRI of δ_{IB}^{1-} over $\delta_{c_0}^{1}$ is 18.98% at $(\mu_1 = -0.01, \mu_2 = -0.01)$ for $(n_1 = 5, n_2 = 5)$ and $p = 2$, whereas for the same values of (μ_1, μ_2) and (n_1, n_2), the RRI of δ_{IB}^{1-} over $\delta_{c_0}^{1}$ is 19.20%, when $p = -2$.

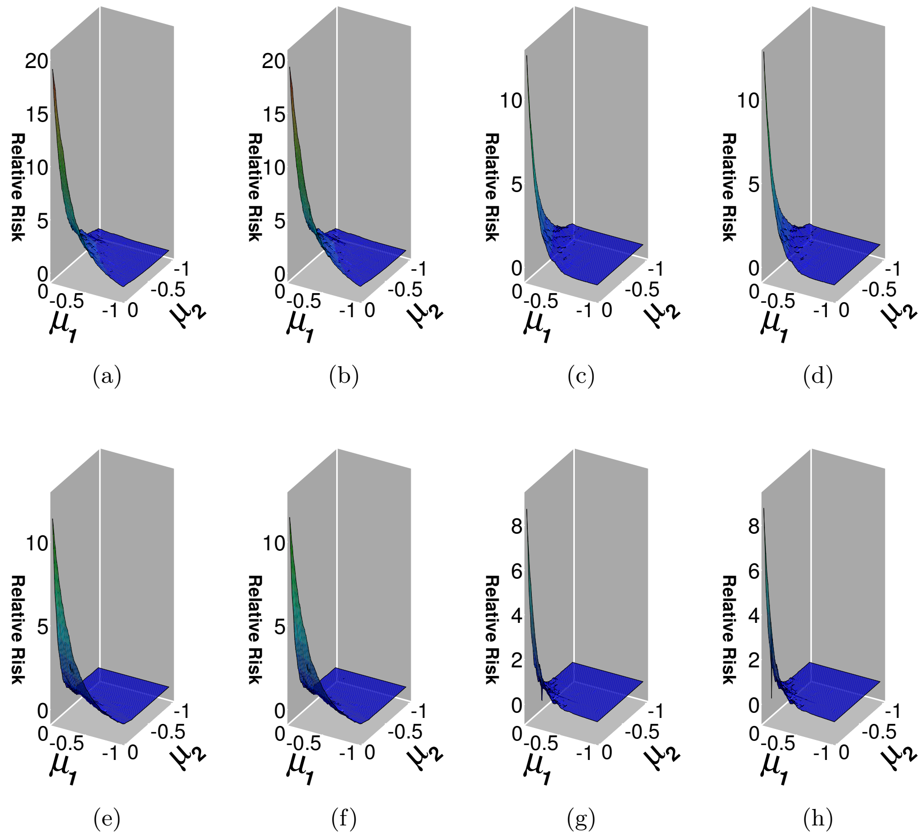


Figure 3: Fig. (a), (b), (c), (d), (e), (f), (g) and (h) represent relative percentage risk improvement plots of δ_{IB}^{1-} over $\delta_{c_0}^{1}$ for (5,5,+0.2), (5,5,-0.2), (10,5,+0.2), (10,5,-0.2), (5,10,+0.2), (5,10,-0.2), (10,10,+0.2) and (10,10,-0.2), respectively when $(\mu_1, \mu_2) \in \mathbb{R}_2^-$. The first and second components of the triplet represent the sample sizes of the first and second population, respectively whereas the third component represents the value of p.

3. COMMON LOCATION BUT UNEQUAL SCALE PARAMETERS

In this section, we consider k exponential populations with a common location parameter μ and unknown but unequal scale parameters $\underline{\sigma}$. This model arises in life testing and reliability, where the common location parameter can be considered as minimum guarantee time of operation of several components and scale parameters are interpreted as unknown and possibly unequal failure rates of these components. Let the probability density function of the i-th population be

$$f_i(x|\mu,\sigma_i) = \begin{cases} \sigma_i^{-1} exp\{-(x-\mu)/\sigma_i\}, & \text{if } x > \mu \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

where $\mu \in \mathbb{R},\ \sigma_i > 0,\ i = 1, ..., k$. Let $(X_{i1}, ..., X_{in_i})$ be a random sample of size n_i drawn from the i-th population $(i = 1, ..., k)$ with probability density function given in (3.1). The expression of the Renyi entropy is $R_\alpha(\underline{\sigma}) = \sum_{i=1}^{k} \ln \sigma_i - k \ln \alpha/(1-\alpha)$. It is worthwhile to mention that the problem of estimating $R_\alpha(\underline{\sigma})$ with respect to the loss function of the form $L(\theta, \delta) = W(\delta - \theta)$ is equivalent to the problem of estimating $Q_2(\underline{\sigma}) = \sum_{i=1}^{k} \ln \sigma_i$. We consider the loss function as

$$L^2(Q_2(\underline{\sigma}), \delta) = \exp\Big\{p\Big(\delta - \sum_{i=1}^{k} \ln \sigma_i\Big)\Big\} - p\Big(\delta - \sum_{i=1}^{k} \ln \sigma_i\Big) - 1,\ p \neq 0. \tag{3.2}$$

Denote $Z_i = Y_i - n_i X_{i(1)},\ i = 1, ..., k$. For the i-th population, $(X_{i(1)}, Z_i)$ is a complete and sufficient statistic for (μ, σ_i). Moreover, Z_i and $X_{i(1)}$ are independently distributed, where $2\sigma_i^{-1} Z_i$ follows chi-square distribution with $2(n_i - 1)$ degrees of freedom and $X_{i(1)}$ follows an exponential distribution with location parameter μ and scale parameter σ_i/n_i. Further, define $X = min\{X_{1(1)}, ..., X_{k(1)}\}$ and $T_i = Y_i - n_i X$. It is easy to show that $(X, \underline{T})$ is a joint complete and sufficient statistic for $(\mu, \underline{\sigma})$, where $\underline{T} = (T_1, ..., T_k)$. The MLE of $Q_2(\underline{\sigma})$ is $\delta^2_{ML} = \sum_{i=1}^{k} \ln T_i - \ln(\prod_{i=1}^{k} n_i)$. Also, X and $\underline{T}$ are independently distributed with respective probability density function

$$f_X(x) = \tau \exp\{-\tau(x - \mu)\}, \qquad x > \mu \tag{3.3}$$

and

$$f_{\underline{T}}(\underline{t}) = Nq\eta l \tau^{-1} \Big(\prod_{i=1}^{k} t_i^{n_i} - 1\Big)\ \exp\Big\{ - \Big(\sum_{i=1}^{k} t_i \sigma_i^{-1}\Big)\Big\}, \quad t_i > 0, \tag{3.4}$$

where $\eta = \Big(\prod_{i=1}^{k} \sigma_i\Big)^{-n_i},\ \tau = \sum_{i=1}^{k} n_i \sigma_i^{-1},\ N = \sum_{i=1}^{k} n_i(n_i - 1),\ l = \Big(\prod_{i=1}^{k} \Gamma(n_i)\Big)^{-1},$ $q = \sum_{i=1}^{k} t_i^{-1}$ and $\underline{t} = (t_1, ..., t_k)$. Following steps analogous to Kayal *et al.* (2015), the UMVUE of $Q_2(\underline{\sigma})$ can be obtained as

$$\delta^2_{MV} = \sum_{i=1}^{k} \ln T_i - \sum_{i=1}^{k} \frac{1 - (JT_i)^{-1}}{n_i - 1} - \sum_{i=1}^{k} \psi(n_i - 1), \tag{3.5}$$

where $J = \sum_{i=1}^{k} T_i^{-1}$.

3.1. Affine equivariant estimator

The estimation problem under study is invariant under $G_{a,b}$, a group of affine transformations, where $G_{a,b} = \{g_{a,b} : g_{a,b}(x) = ax + b,\ a > 0,\ b \in \mathbb{R}\}$. The form of an affine equivariant estimator can be obtained as (see Kayal *et al.* (2015))

$$\begin{aligned} \delta_\eta(X,\underline{T}) &= k\ln T_1+\eta(W_1,...,W_{k-1}) \\ &= k\ln T_1+\eta(\underline{W}), \end{aligned} \tag{3.6}$$

where $\underline{W}=(W_1,...,W_{k-1})$, $W_i=(T_{i+1}/T_1),\ i=1,...,k-1$ and η is a real valued measurable function. The following theorem provides a general inadmissibility result for an affine equivariant estimator of the form (3.6).

Theorem 3.1. Let δ_η be the form of an affine equivariant estimator given in (3.6). Further, define the estimator δ_η^* by

$$\delta_\eta^*=\begin{cases}\delta_\eta, & \text{if} \quad \eta(\underline{w})\geq\eta_0(\underline{w}) \\ \\ \delta_{\eta_0}, & \text{if} \quad \eta(\underline{w})<\eta_0(\underline{w}),\end{cases}$$

where $\underline{w}=(w_1,...,w_{k-1})$ and $\eta_0(\underline{w})=\ln[k^k(\prod_{i=1}^{k-1}w_i)]-\frac{1}{p}\ln\left(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}\right)$. Then under the linex loss function (3.2), δ_η^* improves δ_η if there exists $(\mu,\underline{\sigma})$ such that $P_{\mu,\underline{\sigma}}(\eta(\underline{W})<\eta_0(\underline{W}))>0$.

Proof: The risk function of δ_η can be written as

$$R(\mu,\underline{\sigma},\delta_\eta)=E^{\underline{W}}R_1(\mu,\underline{\sigma},\underline{W},\delta_\eta),$$

where $R_1(\mu,\underline{\sigma},\underline{w},\delta_\eta)$ denotes the conditional risk of δ_η given $\underline{W}=\underline{w}$, and is given by

$$\begin{aligned} R_1(\mu,\underline{\sigma},\underline{w},\delta_\eta) &= E\Big[\Big(\exp\Big\{p\Big(k\ln T_1+\eta(\underline{W})-\sum_{i=1}^{k}\ln\sigma_i\Big)\Big\} \\ &\qquad -p\Big(k\ln T_1+\eta(\underline{W})-\sum_{i=1}^{k}\ln\sigma_i\Big)-1\Big)\Big|\underline{W}=\underline{w}\Big]. \end{aligned}$$

Note that R_1 is a convex function in η and minimized at

$$\hat{\eta}(\underline{\sigma},\underline{w})=\frac{1}{p}\ln\left(\frac{\Big(\prod_{i=1}^{k}\sigma_i\Big)^p}{E(T_1^{kp}|\underline{W}=\underline{w})}\right). \tag{3.7}$$

To evaluate $\hat{\eta}(\underline{\sigma},\underline{w})$ in (3.7), we need to derive the conditional distribution of T_1 given $\underline{W}=\underline{w}$ which is given by

$$f_{T_1|\underline{W}}(t_1|\underline{w})=\Gamma^{-1}(n-1)s^{n-1}\ t_1^{n-2}\ e^{-st_1},\ t_1>0,\ w_i>0, \tag{3.8}$$

where $s=\sigma_1^{-1}+\sum_{i=1}^{k-1}w_i\sigma_{i+1}^{-1}$ and $n=\sum_{i=1}^{k}n_i$. Using (3.8) we obtain

$$E(T_1^{kp}|\underline{W}=\underline{w})=\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}\frac{1}{s^{kp}}.$$

Putting $E(T_1^{kp}|\underline{W}=\underline{w})$ in (3.7), we get

$$\hat{\eta}(\underline{\sigma},\underline{w}) = \ln\left(s^k \prod_{i=1}^{k} \sigma_i\right) - \frac{1}{p}\ln\left(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}\right). \tag{3.9}$$

For fixed values of $\underline{w}$, the supremum and infimum of $\hat{\eta}(\underline{\sigma},\underline{w})$ over $\underline{\sigma}$ can be obtained as

$$\begin{aligned} \sup_{\underline{\sigma}} \hat{\eta}(\underline{\sigma},\underline{w}) &= +\infty, \\ \inf_{\underline{\sigma}} \hat{\eta}(\underline{\sigma},\underline{w}) &= \ln\left[k^k\Big(\prod_{i=1}^{k-1} w_i\Big)\right] - \frac{1}{p}\ln\left(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}\right) \\ &= \phi_0, \text{ say.} \end{aligned} \tag{3.10}$$

An application of the Brewster-Zidek technique on R_1, then completes the proof the theorem. $\square$

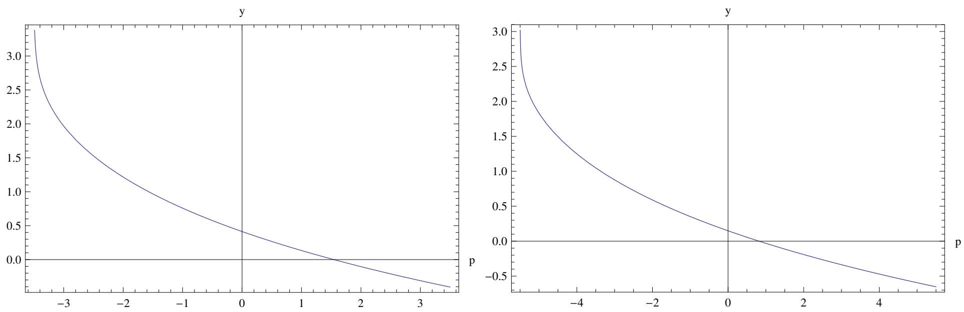
Note that δ^2_{ML} and δ^2_{MV} belong to the class of affine equivariant estimators of the form (3.6) when $\eta(\underline{w})$ is equal to $\ln\left(\frac{T_2}{T_1}\cdots\frac{T_k}{T_1}\right) - k\ln(\prod_{i=1}^k n_i)$ and $\ln\left(\frac{T_2}{T_1}\cdots\frac{T_k}{T_1}\right) - \sum_{i=1}^k \frac{1-(JT_i)^{-1}}{n_i-1} - \sum_{i=1}^k \psi(n_i-1)$, respectively. The Theorem 3.1 then leads to the following corollaries.

Corollary 3.1. The MLE δ^2_{ML} is inadmissible and dominated by

$$\delta^2_{IML} = \begin{cases} \ln[(kT_1)^k(\prod_{i=1}^{k-1} w_i)] - \frac{1}{p}\ln(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}), \\ \qquad\qquad \text{if } \ln(k^k\prod_{i=1}^k n_i) - \frac{1}{p}\ln(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}) > 0 \\ \sum_{i=1}^k \ln T_i - \ln(\prod_{i=1}^k n_i), \qquad \text{otherwise.} \end{cases}$$

Corollary 3.2. The UMVUE δ^2_{MV} is inadmissible and dominated by

$$\delta^2_{IMV} = \begin{cases} \ln[(kT_1)^k(\prod_{i=1}^{k-1} W_i)] - \frac{1}{p}\ln(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}), \\ \qquad\qquad \text{if } \ln(k^k) - \frac{1}{p}\ln(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)}) \\ \qquad\qquad\qquad + \sum_{i=1}^k \psi(n_i-1) + \sum_{i=1}^k \frac{1-(JT_i)^{-1}}{n_i-1} > 0 \\ \sum_{i=1}^k \ln T_i - \sum_{i=1}^k \frac{1-(JT_i)^{-1}}{n_i-1} - \sum_{i=1}^k \psi(n_i-1), \qquad \text{otherwise.} \end{cases}$$



(a) (b)

Figure 4: Fig. (a) represents the plot of $\ln(k^k \prod_{i=1}^{k} n_i) - \frac{1}{p}\ln(\frac{\Gamma(n+kp-1)}{\Gamma(n-1)})$ for $(n_1 = 4, n_2 = 4)$ when $k = 2$; and Fig. (b) represents the plot of that for $(n_1 = 4, n_2 = 8)$, when $k = 2$;

3.2. Scale equivariant estimator

In this section, we introduce invariance to the problem under the group of scale transformations $G_a = \{g_a(x) = ax,\ a > 0\}$. The form of a scale equivariant estimator is obtained as

$$\begin{aligned}\delta_\xi(X, \underline{T}) &= k\ln T_1 + \xi\Big(\frac{X}{T_1}, \frac{T_2}{T_1}, \frac{T_3}{T_1}, ..., \frac{T_k}{T_1}\Big) \\ &= k\ln T_1 + \xi(\underline{V}), \text{(say)},\end{aligned} \tag{3.11}$$

where $\underline{V} = (V_1, V_2, ..., V_k)$, $V_1 = X/T_1$, and $V_i = T_i/T_1$, $i = 2, 3, ..., k$. The risk function of δ_ξ given in (3.11) is

$$R(\mu, \underline{\sigma}, \delta_\xi) = E^{\underline{V}} R_1(\mu, \underline{\sigma}, \underline{V}, \delta_\xi),$$

where $R_1(\mu, \underline{\sigma}, \underline{v}, \delta_\xi)$ denotes the conditional risk of δ_ξ given $\underline{V} = \underline{v}$, is given by

$$\begin{aligned}R_1(\mu, \underline{\sigma}, \underline{v}, \delta_\xi) &= E\Big[\Big(\exp\Big\{p\Big(k\ln T_1 + \xi(\underline{V}) - \sum_{i=1}^{k}\ln\sigma_i\Big)\Big\} \\ &\quad -p\Big(k\ln T_1 + \xi(\underline{V}) - \sum_{i=1}^{k}\ln\sigma_i\Big) - 1\Big)\Big|\underline{V} = \underline{v}\Big]\end{aligned}$$

which is minimized at

$$\hat{\xi}(\mu, \underline{\sigma}, \underline{v}) = \frac{1}{p}\ln\left(\frac{\Big(\prod_{i=1}^{k}\sigma_i\Big)^p}{E(T_1^{kp}|\underline{V} = \underline{v})}\right). \tag{3.12}$$

The joint probability density function of T_1 and $\underline{V}$ is

$$f(t_1,\underline{v}) \;=\; C\exp\{\mu\tau\}\Big(\prod_{i=2}^{k} v_i^{n_i-1}\Big)\Big(1+\sum_{i=2}^{k} v_i^{-1}\Big) \times\; \exp\Big\{-\Big(v_1\tau+\Big(\sigma_1^{-1}+\sum_{i=2}^{k} v_i\sigma_i^{-1}\Big)\Big)t_1\Big\}t_1^{n\,-\,1}, \tag{3.13}$$

where $C = N\eta l,\ t_1 > 0, t_1v_1 > \mu, v_2 > 0, v_3 > 0, ..., v_k > 0$. Note that to obtain the supremum and infimum of $\hat{\xi}(\mu,\underline{\sigma},\underline{v})$, it is required to derive the conditional distribution of $T_1|\underline{V}=\underline{v}$, which can be obtained through the arguments of the cases considered in Section 3.2 of the paper by Kayal *et al.* (2015). Hence we omit the details.

Case **(i):** Under the assumptions that $\mu > 0$ and $v_1 > 0$, we obtain

$$\hat{\xi}(\mu,\underline{\sigma},\underline{v}) = \ln\Big(\prod_{i=1}^{k}\sigma_i\Big) - \frac{1}{p}\ln(I_1^*/I_1), \tag{3.14}$$

where $I_1^* = \int_{\mu/v_1}^{\infty} exp\{-At_1\}\ t_1^{kp+n-1}dt_1,\ I_1 = \int_{\mu/v_1}^{\infty} exp\{-At_1\}\ t_1^{n-1}dt_1$ and $A = v_1\tau + (\sigma_1^{-1} + \sum_{i=2}^{k} v_i\sigma_i^{-1})$. Using the monotone likelihood ratio property it is not hard to show that

$$\sup_{\mu,\underline{\sigma}}\hat{\xi}(\mu,\underline{\sigma},\underline{v}) = +\infty \quad\text{and}\quad \inf_{\mu,\underline{\sigma}}\hat{\xi}(\mu,\underline{\sigma},\underline{v}) = -\infty.$$

Case **(ii):** When $\mu < 0$ and $v_1 > 0$, we get

$$\hat{\xi}(\mu,\underline{\sigma},\underline{v}) = \ln\Big(A^k\prod_{i=1}^{k}\sigma_i\Big) - \frac{1}{p}\ln\Big(\frac{\Gamma(n+kp)}{\Gamma(n)}\Big). \tag{3.15}$$

It is easy to see that supremum of $\widehat{\xi}(\mu,\underline{\sigma},\underline{v})$ is $+\infty$. The infimum of $\widehat{\xi}(\mu,\underline{\sigma},\underline{v})$ can be obtained by applying geometric mean-harmonic mean inequality to the variables $(\sigma_1/n_1v_1+1), (\sigma_2/n_2v_1+v_2), ..., (\sigma_k/n_kv_1+v_k)$ and is given by

$$\inf_{\mu,\underline{\sigma}}\widehat{\xi}(\mu,\underline{\sigma},\underline{v}) = \ln\Big(k^k(n_1v_1+1)\prod_{i=2}^{k}(n_iv_1+v_i)\Big) - \frac{1}{p}\ln\Big(\frac{\Gamma(n+kp)}{\Gamma(n)}\Big).$$

Case **(iii):** Let $\mu < 0$ and $v_1 < 0$. Under these assumptions, we have

$$\hat{\xi}(\mu,\underline{\sigma},\underline{v}) = \ln\Big(\prod_{i=1}^{k}\sigma_i\Big) - \frac{1}{p}\ln(I_2^*/I_2), \tag{3.16}$$

where $I_2^* = \int_0^{\mu/v_1} exp\{-At_1\}t_1^{kp+n-1}dt_1$ and $I_2 = \int_0^{\mu/v_1} exp\{-At_1\}t_1^{n-1}dt_1$. Note that the value of A may be positive or negative. For both $A > 0$ and $A < 0$, it can be shown that the supremum and infimum of $\hat{\xi}(\mu,\underline{\sigma},\underline{v})$ are $+\infty$ and $-\infty$, respectively. As in the full sample space the supremum and infimum choices of $\hat{\xi}(\mu,\underline{\sigma},\underline{v})$ are $+\infty$ and $-\infty$, respectively, therefore we do not get any improvement over the BAEE. But if we restrict the parameter space to $\mu<0$, then an improvement over the BAEE exists, which is shown in the next theorem. Define

$$\xi_0(\underline{v}) = \begin{cases} \ln(v^*) - \frac{1}{p}\ln(\frac{\Gamma(n+kp)}{\Gamma(n)}), & \text{if } v_1 > 0 \text{ and} \\ \qquad v^* > exp\{\xi(\underline{v}) + \frac{1}{p}\ln(\frac{\Gamma(n+kp)}{\Gamma(n)})\} & \\ \xi(\underline{v}), & \text{otherwise,} \end{cases} \tag{3.17}$$

where $v^* = k^k(n_1v_1+1)\prod_{i=2}^{k}(n_iv_1+v_i)$.

Theorem 3.2. Let δ_ξ be a scale equivariant estimator of the form (3.12) and $\xi_0(\underline{v})$ be as defined in (3.17). If there exists a $(\mu,\underline{\sigma})$ such that $P_{(\mu,\underline{\sigma})}(\xi_0(\underline{V}) \neq \xi(\underline{V})) > 0$, then the estimator δ_{ξ_0} dominates δ_ξ, with respect to the linex loss function, when $\mu<0$.

As a consequence of the Theorem 3.2, the following corollary immediately follows.

Corollary 3.3. When $\mu<0$, the MLE and the UMVUE are inadmissible and dominated by

$$\delta_{IML}^{2-} = \begin{cases} \ln(T_1^kV^*) - \frac{1}{p}\ln(\frac{\Gamma(n+kp)}{\Gamma(n)}), & \text{if } V_1>0 \text{ and} \\ \qquad V^* > exp\{\sum_{i=2}^k \ln V_i - \ln(\prod_{i=1}^k n_i) + \frac{1}{p}\ln(\frac{\Gamma(n+kp)}{\Gamma(n)})\} & \\ \sum_{i=1}^k \ln T_i - \ln(\prod_{i=1}^k n_i), & \text{otherwise} \end{cases}$$

and

$$\delta_{IMV}^{2-} = \begin{cases} \ln(T_1^kV^*) - \frac{1}{p}\ln(\frac{\Gamma(n+kp)}{\Gamma(n)}), & \text{if } V_1>0 \\ \qquad \text{and } V^* > exp\{\sum_{i=2}^k \ln V_i - \sum_{i=1}^k \frac{1-(JT_i)^{-1}}{n_i-1} & \\ \qquad\qquad - \sum_{i=1}^k \psi(n_i-1) + \frac{1}{p}\ln(\frac{\Gamma(n+kp)}{\Gamma(n)})\} & \\ \sum_{i=1}^k \ln T_i - \sum_{i=1}^k \frac{1-(JT_i)^{-1}}{n_i-1} - \sum_{i=1}^k \psi(n_i-1), & \text{otherwise,} \end{cases}$$

where $V^* = k^k(n_1V_1+1)\prod_{i=2}^{k}(n_iV_1+V_i)$ and $J=\sum_{i=1}^k T_i^{-1}$.

3.3. Numerical comparisons

Here we present risk and relative risk of various estimators derived in Section 3. As in Section 2.2, the risk values were calculated using Monte-Carlo simulation based on $10,000$ samples of different combinations of sample sizes (n_1, n_2) and different values of $(\mu, \sigma_1, \sigma_2)$ and p. We present few of them in tabular form below. Table 1 is for the risk values of δ^2_{MV} and δ^2_{IMV} when $k = 2$. The RRI of the estimators δ^{2-}_{IMV}, δ^{2-}_{IML} and δ^2_{MV} over δ^2_{ML} is presented in Table 2 and Table 3 for $k = 2$. Different combinations of (n_1, n_2) and different values of p, σ_1, σ_2 have been chosen. We have considered sample sizes $(5, 5), (5, 10), (10, 5)$ and $(10, 10)$. The values of p have been chosen as $-0.5, -1, +0.5$ and $+1$. Here, we have presented very few values, however, similar observations are made for various other values of (n_1, n_2), p and μ.

Table 1: The risk values of δ^2_{MV} and δ^2_{IMV} for $k = 2$.

p	μ	(n_1, n_2)	(σ_1, σ_2)	δ^2_{MV}	δ^2_{IMV}	p	μ	(n_1, n_2)	(σ_1, σ_2)	δ^2_{MV}	δ^2_{IMV}
-0.75	0.2	(5,5)	(0.5,0.5)	0.164190	0.157732	-2	-0.2	(10,5)	(0.5,0.5)	1.157259	1.068251
			(0.5,1.0)	0.159419	0.153242				(0.5,1.0)	1.135408	1.009288
			(1.0,0.5)	0.169174	0.162430				(1.0,0.5)	1.227956	1.157454
			(1.0,1.0)	0.164190	0.157732				(1.0,1.0)	1.157259	1.068251
-1.5	0.5	(5,10)	(0.5,0.5)	0.561950	0.560949	-2.5	-0.5	(10,10)	(0.5,0.5)	0.974965	0.669685
			(0.5,1.0)	0.657281	0.653271				(0.5,1.0)	1.007035	0.680643
			(1.0,0.5)	0.622936	0.621872				(1.0,0.5)	0.931061	0.654255
			(1.0,1.0)	0.561950	0.552783				(1.0,1.0)	0.974965	0.669685

Table 2: The relative percentage risk improvement over δ^2_{ML} by δ^{2-}_{IMV}, δ^{2-}_{IML} and δ^2_{MV} for $k = 2$.

p	μ	(n_1, n_2)	(σ_1, σ_2)	δ^{2-}_{IMV}	δ^{2-}_{IML}	δ^2_{MV}	p	μ	(n_1, n_2)	(σ_1, σ_2)	δ^{2-}_{IMV}	δ^{2-}_{IML}	δ^2_{MV}
-0.5	-0.2	(5,5)	(0.5,0.5)	1.18	1.50	41.46	0.5	-0.2	(5,5)	(0.5,0.5)	0.38	0.69	17.72
			(0.5,1.0)	1.55	2.59	41.68				(0.5,1.0)	0.59	1.31	17.81
			(1.0,0.5)	3.38	3.75	41.29				(1.0,0.5)	2.23	2.41	17.85
			(1.0,1.0)	5.19	7.03	41.47				(1.0,1.0)	3.21	4.05	17.72
-0.5	-0.2	(5,10)	(0.5,0.5)	0.15	0.23	35.57	0.5	-0.2	(5,10)	(0.5,0.5)	0.12	0.16	14.02
			(0.5,1.0)	0.41	0.82	37.06				(0.5,1.0)	0.23	0.47	15.64
			(1.0,0.5)	0.08	0.16	33.86				(1.0,0.5)	0.02	0.07	13.02
			(1.0,1.0)	0.39	1.25	35.57				(1.0,1.0)	0.42	0.53	14.02
-0.5	-0.2	(10,5)	(0.5,0.5)	0.12	0.25	36.64	0.5	-0.2	(10,5)	(0.5,0.5)	0.04	0.09	15.87
			(0.5,1.0)	0.23	0.55	35.37				(0.5,1.0)	0.08	0.23	15.63
			(1.0,0.5)	0.24	1.31	37.83				(1.0,0.5)	0.10	0.45	16.71
			(1.0,1.0)	2.66	3.68	36.65				(1.0,1.0)	0.90	1.66	15.87
-0.5	-0.2	(10,10)	(0.5,0.5)	0.01	0.02	27.68	0.5	-0.2	(10,10)	(0.5,0.5)	0.02	0.05	13.95
			(0.5,1.0)	0.23	0.34	28.20				(0.5,1.0)	0.04	0.22	14.72
			(1.0,0.5)	0.42	0.48	26.89				(1.0,0.5)	0.07	0.31	14.23
			(1.0,1.0)	0.77	1.42	27.68				(1.0,1.0)	0.43	0.82	13.95

Table 3: The relative percentage risk improvement over δ^2_{ML} by δ^{2-}_{IMV}, δ^{2-}_{IML} and δ^2_{MV} for $k = 2$.

p	μ	(n_1, n_2)	(σ_1, σ_2)	δ^{2-}_{IMV}	δ^{2-}_{IML}	δ^2_{MV}	p	μ	(n_1, n_2)	(σ_1, σ_2)	δ^{2-}_{IMV}	δ^{2-}_{IML}	δ^2_{MV}
-1	-0.2	(5,5)	(0.5,0.5)	1.60	2.08	50.16	1	-0.2	(5,5)	(0.5,0.5)	0.27	0.44	0.43
			(0.5,1.0)	2.25	3.21	50.37				(0.5,1.0)	0.34	0.78	0.42
			(1.0,0.5)	4.09	4.76	49.88				(1.0,0.5)	1.01	1.58	0.77
			(1.0,1.0)	6.67	8.56	50.16				(1.0,1.0)	2.55	2.98	0.44
-1	-0.2	(5,10)	(0.5,0.5)	0.15	0.23	43.88	1	-0.2	(5,10)	(0.5,0.5)	0.10	0.11	0.98
			(0.5,1.0)	0.40	1.07	45.22				(0.5,1.0)	0.21	0.37	1.03
			(1.0,0.5)	0.06	0.35	42.00				(1.0,0.5)	0.02	0.05	0.76
			(1.0,1.0)	0.51	1.81	43.88				(1.0,1.0)	0.29	0.42	1.20
-1	-0.2	(10,5)	(0.5,0.5)	0.03	0.45	44.71	1	-0.2	(10,5)	(0.5,0.5)	0.02	0.13	2.01
			(0.5,1.0)	0.59	1.12	43.14				(0.5,1.0)	0.08	0.21	2.66
			(1.0,0.5)	0.75	1.98	45.95				(1.0,0.5)	0.09	0.29	2.47
			(1.0,1.0)	4.06	5.56	44.71				(1.0,1.0)	0.45	1.15	2.36
-1	-0.2	(10,10)	(0.5,0.5)	0.12	0.16	33.43	1	-0.2	(10,10)	(0.5,0.5)	0.04	0.06	5.59
			(0.5,1.0)	0.18	0.27	33.84				(0.5,1.0)	0.05	0.09	6.49
			(1.0,0.5)	0.16	0.30	32.81				(1.0,0.5)	0.07	0.33	4.14
			(1.0,1.0)	1.24	1.73	33.43				(1.0,1.0)	0.34	0.56	5.59

The following conclusions are evident from Table 2 and Table 3:

(i) We observe marginal RRI over δ^2_{ML} by δ^{2-}_{IML}, δ^{2-}_{IMV}, but substantial improvement by δ^2_{MV}.

(ii) For fixed (n_1, n_2) and μ, the RRIs for the estimators δ^{2-}_{IMV} and δ^{2-}_{IML} over δ^2_{ML} are marginally better, whereas we observe substantial RRIs for the estimator δ^2_{MV} over δ^2_{ML} for negative values of p than positive values of p.

(iii) For fixed μ, p and (n_1, n_2) the RRI of δ^{2-}_{IML} and δ^{2-}_{IMV} over δ_{ML} approximately increases with (σ_1, σ_2), but we do not observe such behaviour for the RRI of δ^2_{MV} over δ^2_{ML}.

4. CONCLUDING REMARKS

In this paper, the problem of estimating the Renyi entropy of several exponential distributions has been investigated with respect to a linex loss function. The concept of invariance has been used to derive improved estimators over the standard ones such as MLE and UMVUE. We have considered two distinct models here. Both these models have various applications in real life experiments. In the first model, the location parameters are distinct but the scale parameter is assumed to be common. Improved estimators over the BAEE have been obtained when parameters space is restricted as well as unrestricted. In the second model, the scale parameters are distinct but the location parameter is assumed to be common. Affine and scale equivariant estimators improving over the UMVUE and MLE are obtained under restrictions on the parameter space. Finally, margins of relative risk improvements by new estimators are determined numerically using simulations.

ACKNOWLEDGMENTS

The authors wish to thank the Associate Editor and three anonymous referees for their comments which have greatly improved the initial version of this manuscript.

REFERENCES

[1] Bercher, J.F. (2008). On some entropy functionals derived from Renyi information divergence, *Information Sciences*, **178**, 2489–2506.

[2] Brewster, J.F. and Zidek, J.V. (1974). Improving on equivariant estimators, *Annals of Statistics*, **2**, 21–38.

[3] Cover, T.M. and Thomas, J.A. (2006). *Elements of Information Theory*, Wiley, New York.

[4] De-Gregorio, A. and Iacus, S.M. (2009). On Renyi information for ergodic diffusion processes, *Information Sciences*, **179**, 279–291.

[5] Golshani, L. and Pasha, E. (2010). Renyi entropy rate for Gaussian processes, *Information Sciences*, **180**, 1486–1491.

[6] Hughes, M.S.; Marsh, J.N.; Arbeit, J.M.; Neumann, R.G.; Fuhrhop, R.W.; Wallace, K.D.; Thomas, L.; Smith, J.; Agyem, K.; Lanza, G.M.; Wickline, S.A. and McCarthy, J.E. (2009). Application of Renyi entropy for ultrasonic molecular imaging, *The Journal of the Acoustical Society of America*, **125**, 3141–3145.

[7] Kapur, J.N. (1990). *Maximum Entropy Models in Science and Engineering*, Wiley, New York.

[8] Kayal, S. and Kumar, S. (2011a). Estimating the Shannon entropy of several exponential populations, *International Journal of Statistics and Economics*, **7**, 42–52.

[9] Kayal, S. and Kumar, S. (2011b). Estimating entropy of an exponential population under linex loss function, *Journal of Indian Statistical Association*, **49**, 91–112.

[10] Kayal, S. and Kumar, S. (2013). Estimation of the Shannon's entropy of several shifted exponential populations, *Statistics and Probability Letters*, **83**, 1127–1135.

[11] Kayal, S.; Kumar, S. and Vellaisamy, P. (2015). Estimating the Renyi entropy of several exponential populations, *Brazilian Journal of Probability and Statistics*, **29**, 94–111.

[12] Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*, Second Ed., Springer, New York.

[13] Maszczyk, T. and Duch, W. (2008). Comparison of Shannon, Renyi and Tsallis entropy used in decision trees, *In Artificial Intelligence and Soft Computing-ICAISC, Proceedings*, 5097, 643–651.

[14] Misra, N.; Singh, H. and Demchuk, E. (2005). Estimation of the entropy of a multivariate normal distribution, *Journal of Multivariate Analysis*, **92**, 324–342.

[15] Nilsson, M. (2006). *Entropy and Speech*, Ph.D. Dissertation, Royal Institute of Technology, Stockholm, KTH.

[16] Obin, N. and Liuni, M. (2012). On the generalization of Shannon entropy for speech recognition, *IEEE Workshop on Spoken Language Technology, Miami, USA*, December 2–5, 97–102.

[17] Pharwaha, A.P.S. and Khehra, B.S. (2009). Shannon and non-Shannon measures of entropy for statistical texture feature extraction in digitized mammograms, *Proceedings of the World Congress on Engineering and Computer Science, San Francisco, USA*, October 20–22, 2, 1286–1291.

[18] Renyi, A. (1961). On measures of entropy and information, *Proceedings of Fourth Berkeley Symposium on Mathematics, Statistics and Probability*, University of California Press, Berkeley, 1, 1286–1291.

[19] Renyi, A. (2012). *Probability Theory*, North-Holland, Amsterdam.

[20] Shannon, C. (1948). The mathematical theory of communication, *Bell System Technical Journal*, **27**, 379–423.

[21] Song, K.S. (2001). Renyi information, loglikelihood and an intrinsic distribution measure, *Journal of Statistical Planning and Inference*, **93**, 51–69.

[22] Varian, H.R. (1975). *A Bayesian approach to real estate assessment.* In "Studies in Bayesian Econometrics and Statistics in Honour of L.J. Savage" (S.E. Feinberg and A. Zellner, Eds.), Amsterdam North Holland, 195–208.

[23] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions, *Journal of American Statistical Association*, **81**, 446–451.