Journal of Theoretical and Applied Information Technology
15th July 2019. Vol.97. No 13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RÈNYI ENTROPY FOR MIXTURE MODEL OF ULTIVARIATE SKEW NORMAL-CAUCHY DISTRIBUTIONS

[1]SALAH H. ABID, [2]UDAY J. QUAEZ

Department of Mathematics, Education College, Al-Mustansiriya University, Baghdad, Iraq
E-mail: [1]abidsalah@gmail.com, [2]uoday1977@ gmail.com

ABSTRACT

Rényi entropy is the important concept developed by Rényi in the context of entropy theory. We study in detail this measure of information in case of multivariate skew normal Cauchy distributions. Mixture model of these distributions is proposed. In addition, upper and lower bounds of entropy both types Shannon and Rényi are found on this model. Also, an asymptotic expression for Rényi entropy for a mixture of skew distributions is given in approximation by using some inequalities, multinomial theorem and properties of L^p -spaces. Finally, we give a real data examples to illustrate the behavior of Rényi entropy of the proposed mixture model.

Keywords: *Rényi Entropy, Mixture Model, Multivariate Skew Normal Cauchy Distribution, Multinomial Theorem, Approximate Entropy.*

1. INTRODUCTION

In multivariate analysis, Azzalini A., & Dalla Valle A. (1996) introduced the multivariate skew normal distribution as an alternative to multivariate normal distribution to deal with skew in the data.

Genton M., & Loperfido N. (2005) derived the generalization of multivariate skew normal distribution whose the probability density function is as follows

$$k(y;\mu,S) = 2g(y;\mu,S).\psi(y-\mu)\ , \qquad y \in R^d$$

where the function ψ satisfies $0 \le \psi(y) \le 1$ and $\psi(-y) = 1-\psi(y)$, for any $y \in R^d$.

Obviously, if $\psi(y) = \frac{1}{2}$ then Y has a multivariate normal distribution. As a special case if we take ψ is a distribution function such as normal , Logistic, Laplace or other distributions such that $\psi(y) = G(\delta' y)$ then we get on the generalized skew normal distribution. Huang W., Su N., & Gupta A. (2013) derived the explicit forms of moment generated function of these multivariate skew distributions. More recently, Kahrari et al., (2016) studied some the main probabilistic properties of multivariate skew normal Cauchy distribution.

Huang W., Su N., & Gupta A. (2013) derived the explicit forms of moment generated function of these multivariate skew distributions. More recently, Kahrari et al., (2016) studied some the main probabilistic properties of multivariate skew normal Cauchy distribution. They derived simple expression of the moments, covariance matrix and moment generated function of this distribution. Lin T., Lee J., & Wan H. (2007) proposed the development of mixture of skew normal and skew t-models. Mixture models of multivariate skew normal and skew t-distributions were studied by Pyne S., et al., (2009). Lee S., & McLachlan G. (2014) provided an overview of developments of a mixture of skew t- distributions.

On the other hand, Shanon C. E. (1948) presented measure to quantify the uncertainty of an event. Rényi A. (1961) generalized this measure for probability distribution which means the sensitive to the fine details of a density function. Javier E., & Contreras-Reyes J. (2016) discussed the Rényi entropy of flexible class of skew normal distributions. Wood R., Blythe R. & Evans M. (2017) introduced some results for the Rényi entropy of the totally asymmetric exclusion process. Also, they calculated explicitly Rényi entropy whereby the squares of configuration probabilities are

summed. It is important to refer that the authors Contreras-Reyes J., & Cortés D. (2016) showed the bounds and approximation of Rényi entropy of a class of mixture models of multivariate skew Gaussian by using the multinomial theorem and generalized HÖlder's inequality. In fact, there are no analytical expressions for the Rènyi entropy of the mixture model, therefor we consider its bounds exist. Similarity in the case of fractional relative entropy an analytic evolution of Rènyi entropy is also impossible.

In this paper, we propose a model of mixture of multivariate skew normal Cauchy distributions. The explicit expression of Rényi entropy of multivariate skew normal Cauchy distribution is derived. By using generalized HÖlder's inequality and some properties of multinomial theorem, we have derived the upper and lower bounds for Rényi entropy of mixture model. An approximate value of these entropies can be calculated. In addition, an asymptotic expression for Rényi entropy is given by the approximation and by using some inequalities and properties of L^p -spaces. Finally, we give a real data examples to illustrate the behavior of Rényi entropy with the parameters α , ε and skewness parameter δ of the proposed mixture model.

The remainder of this paper is organized as follows: In section 2. we begin with a preliminary material of mixture models and the measure of information (Rényi entropy). Section 3. provides a description of asymptotic expression for Rényi entropy of multivariate skew normal Cauchy distributions by using some methods of numerical integration such as Monte Carlo and importance sampling methods. By using some theorems which related to multinomial theorem and generalized HÖlder's inequality, in section 4. we find upper and lower bounds and also we study an approximate Rényi entropy for a mixture of multivariate skew normal Cauchy distributions. The conclusion of this paper is shown in Section 5.

2. PRELIMINARY MATERIAL

In this section, we introduce some basic definitions, notations and lemmas related to entropy theory and mixture models of skew distributions. Multivariate skew normal distribution has been proposed by Azzalini and Capitaino (2003). We say that a random variable $Y \in R$ has the skew generalized normal distribution denoted by $Y \sim SGN(\mu, S, \delta_1, \delta_2)$ if its density function is in the following the form:

$$k(y;\mu,S,\delta_1,\delta_2) = 2g(y;\mu,S).\, G\left(\frac{\delta_1 y}{(1+\delta_2\, y^2)^{\frac{1}{2}}}\right) \quad (1)$$

where, $\delta_1 \in R$, $\delta_2 \geq 0$, g and G are univariate normal density function with mean μ , variance S and the distribution function of univariate standard normal respectively. Also, a d-dimensional random vector $X \in R^d$ has a multivariate skew normal-Cauchy distribution denoted by $(X \sim MSNC_d(\mu, S, \delta))$ if its probability density function is given by

$$f_d(x;\mu,S,\delta) = \frac{(\det(S))^{-\frac{1}{2}}}{(2\pi)^{\frac{d}{2}}} \;.\; \exp\left\{\begin{matrix}\left(-\frac{1}{2}(x-\mu)'S^{-1}(x-\mu)\right) \\ .\left(1+\frac{2}{\pi}\arctan\left(\delta'\tilde{S}^{-1}(x-\mu)\right)\right)\end{matrix}\right\} \quad (2)$$

where, $\mu \in R^d$ is a location parameter , a scale matrix $S \in R^{d\times d}$ is a positive definite , a skewness vector is $\delta \in R^d$ and $\tilde{S} = \mathrm{diag}(s_{11}, s_{22}, \ldots, s_{dd})^{\frac{1}{2}}$, S=(s_{ij}) , i ,j =1,2,…,d.

The stochastic representation of $X \sim MSNC_d(\mu, S, \delta)$ can be obtained as a mixture of multivariate skew normal $X|Y \sim MSN_d(\mu, S, \delta y)$ and half standard normal distributions $Y \sim HN(0,1)$. (see, e.g., Kahrari et al., (2016)). The mean vector and covariance matrix of X are derived by Kahrari et al., (2016) and other authors in the following form:

$$E(X) = \mu + c_\delta S\,\delta \quad (3)$$

$$Var(X) = S - c_\delta^2\, S\,\delta\delta' S \quad (4)$$

where, $c_\delta = \frac{2}{\pi\sqrt{\delta' S\delta}}\, m(\frac{1}{\sqrt{\delta' S\delta}})$, $m(x) = \frac{G(-x;0,1)}{g(x;0,1)}$ is the Mill's ratio, g and G defined in (1). The moment generated function of a random vector $X \sim MSNC_d(\mu, S, \delta)$ is given by

$$M_X(r) = 2\exp\left(\frac{r' S r}{2}\right)\left(1 - K(0;0,1,\delta' S\, r, \delta' S\,\delta)\right)$$

where, K is the distribution function of the $SGN(0,1,\delta_1,\delta_2)$ distribution.

Definition1. Let X be a d-dimensional random vector which comes from an m-component

ISSN: www.jatit.org E-ISSN:

mixtures of multivariate skew normal-Cauchy distributions. Then the probability density function of $X \sim MMSNC_d(\mu, S, \delta, \varepsilon)$ is given by the following form:

$$f(x; \mu, S, \delta, \varepsilon) = \sum_{i=1}^{m} \varepsilon_i f(x; \mu_i, S_i, \delta_i) \quad (5)$$

where, ε_i denotes the mixing probability with $\varepsilon_i \geq 0, \sum_{i=1}^{m} \varepsilon_i = 1$, $f(x; \mu_i, S_i, \delta_i)$ and $f(x; \mu, S, \delta, \varepsilon)$ represent respectively the probability density functions of an i-th component and mixture model with parameter vectors set $(\mu, S, \delta, \varepsilon)$; $\mu = \{\mu_1, \mu_2, \dots, \mu_m\}$ a set of vectors represent location parameters, $S = \{S_1, S_2, \dots, S_m\}$ a set of dispersion matrices and the shape vector parameter is $\delta = \{\delta_1, \delta_2, \dots, \delta_m\}$.

If $\kappa = (\kappa_1, \kappa_2, \dots, \kappa_m)$ represent a set of m latent allocations for densities of observations x then $f(x; \mu, S, \delta, \varepsilon) = \prod_{j=1}^{m} f(x; \mu, S, \delta, \kappa_j)$, where $\Pr(\kappa_j = i | \varepsilon) = \varepsilon_i$ then for any j-th component density in (4) is obtained as

$$X_j | (\kappa_j = i) \overset{d}{=} \mu_i + Z_j | Y_j \quad, j = 1, 2, \dots, m,$$

where $Z_j | Y_j \sim MSN_d(0, S_j, \delta_j Y_j)$, $Y_j \sim HN(0,1)$ and $c_{\delta_j} = \frac{2}{\pi \sqrt{\delta_j' S_j \delta_j}} m(\frac{1}{\sqrt{\delta_j' S_j \delta_j}})$, $i = 1, 2, \dots, m$. Equations (3-4) gives the first and second moments for each i-th component of X respectively. Therefore, the mean and covariance of X can be obtained as follow

$$E(X) = \sum_{i=1}^{m} \varepsilon_i (\mu_i + c_{\delta_i} S_i \delta_i) \quad (6)$$

$$Var(X) = \sum_{i=1}^{m} \varepsilon_i S_i - \left(\sum_{i=1}^{m} \varepsilon_i (c_{\delta_i} S_i \delta_i)\right)\left(\sum_{i=1}^{m} \varepsilon_i (c_{\delta_i} S_i \delta_i)\right)' \quad (7)$$

Definition 2. The Shannon entropy of a continuous random vector $X \in R^d$ with probability density function $f(x; \theta)$ is given by

$$H(X; \theta) = -E(\ln(f(x; \theta))) \quad (8)$$

Definition3. An αth-order Rényi entropy of a continuous random vector $X \in R^d$ with probability density function $f(x; \theta)$ is defined as

$$R_\alpha(X; \theta) = \begin{cases} \frac{1}{1-\alpha} \ln(E(f(x; \theta))^{\alpha-1}) & , 0 < \alpha < \infty, \alpha \neq 1 \\ -E\left(\ln\left(f(x; \theta)\right)\right) & \alpha = 1 \end{cases} \quad (9)$$

The relationship between Shannon and Rényi entropies is obtained by the limit $H(X; \theta) = \lim_{\alpha \to 1} R_\alpha(X; \theta)$. Translation does not change the entropy $H(X + C) = H(X) + C$ where, C is constant.

Proposition 1. If $\alpha_1 < \alpha_2$, then $R_{\alpha_1}(X; \theta) \geq R_{\alpha_2}(X; \theta)$ with equality if and only if the distribution of X is uniformly.

Proof: Assume that $\alpha \neq 1$, then the partial derivative of $R_\alpha(X; \theta)$ with respect to α is given as

$$\frac{\partial}{\partial \alpha} R_\alpha(X; \theta) = \frac{E\left((f(x;\theta))^{\alpha-1} \ln(f(x;\theta))\right)}{(1-\alpha)E(f(x;\theta))^{\alpha-1}} + \frac{\ln\left(E(f(x;\theta))^{\alpha-1}\right)}{(1-\alpha)^2}$$

The second part in the right side can be written as

$$\frac{\ln\left(E(f(x;\theta))^{\alpha-1}\right)}{(1-\alpha)^2} = \frac{E\left((f(x;\theta))^{\alpha-1} \ln\left(E(f(x;\theta))^{\alpha-1}\right)\right)}{(1-\alpha)^2 E(f(x;\theta))^{\alpha-1}}$$

Therefore,

$$\frac{\partial}{\partial \alpha} R_\alpha(X; \theta) = \frac{1}{(1-\alpha)^2} \left\{ \frac{E\left((f(x;\theta))^{\alpha-1} \ln(f(x;\theta))^{\alpha-1}\right)}{E(f(x;\theta))^{\alpha-1}} + \frac{E\left((f(x;\theta))^{\alpha-1} \ln\left(E(f(x;\theta))^{\alpha-1}\right)\right)}{E(f(x;\theta))^{\alpha-1}} \right\}$$

$$= \frac{-1}{(1-\alpha)^2} \left\{ \frac{E\left((f(x;\theta))^{\alpha-1} \ln\left(\frac{(f(x;\theta))^{\alpha-1}}{E(f(x;\theta))^{\alpha-1}}\right)\right)}{E(f(x;\theta))^{\alpha-1}} \right\}$$

Define the probability density function g as $g(x; \theta) = \frac{(f(x;\theta))^\alpha}{E(f(x;\theta))^{\alpha-1}}$ then

$$\frac{\partial}{\partial \alpha} R_\alpha(X; \theta) = \frac{-1}{(1-\alpha)^2} \left\{ E_g\left(\ln\left(\frac{g(x;\theta)}{f(x;\theta)}\right) \right) \right\}$$

But $E_g\left(\ln\left(\frac{g(x;\theta)}{f(x;\theta)}\right)\right)$ not negative value therefor, either $E_g\left(\ln\left(\frac{g(x;\theta)}{f(x;\theta)}\right)\right) > 0$ then R_α is a decreasing with respect to α or $E_g\left(\ln\left(\frac{g(x;\theta)}{f(x;\theta)}\right)\right) = 0$ this implies that $f = g$

almost everywhere, hence by assumption $\alpha \neq 1$ then f is only uniformly distributed. Conversely, if f is uniform distribution then it can be written as $f(x;\theta) = \begin{cases} \frac{1}{Vol(B)} & , x \in B \\ 0 & , x \notin B \end{cases}$,for some measurable subset B in R^n . then $R_\alpha(X;\theta) = \frac{1}{1-\alpha}\ln(Vol(B)^{\alpha-1}) = \ln(Vol(B))$ is not depend on the value of α.

The following proposition shows that for any location, scale model the Rènyi entropy does not depend upon location parameter.

Proposition 2. Let $f_x(x;\mu,S)$ be a location scale probability density function with location vector $\mu \in R^d$, $S \in R^{d\times d}$ is dispersion matrix and let $X_0 = S^{-\frac{1}{2}}(X-\mu)$ be a standardized version of X. Then the Rènyi entropy does not depend on μ.

Proof: The probability density function f can be written as

$$f_x(x;\mu,S) = \left(\det(S)\right)^{-\frac{1}{2}} f_{x_0}\left(S^{-\frac{1}{2}}(X-\mu),0,I\right)$$

Therefore, the Rènyi entropy of X appears as

$$R_\alpha(X;\mu,S) = \ln\left(\det(S)\right)^{\frac{1}{2}} + \frac{1}{1-\alpha}\ln\left(E\left(f(x_0;0,I)\right)\right)^{\alpha-1}$$
$$= \ln\left(\det(S)\right)^{\frac{1}{2}} + R_\alpha(X_0;0,I)$$

Lemma 3. **[**Cover T. M. (2006)**]** If a random vector $X \in R^d$ (not necessary normal) has zero mean and covariance matrix $R = E(XX')$, then the following inequality is accomplished

$$H(X;\theta) \leq \frac{1}{2}\ln\left(\det(2\pi\exp(1)R)\right) \quad (10)$$

with equality if and only if $X \sim MN_d(0,S)$.

Proposition 4. **[**Kahrari et al., (2016)**]** If the random vectors $X \sim MSNC_d(0,S,\delta)$ and $X_0 \sim N_d(0,S)$, then

i. For every even function ϑ , we have $\vartheta(x) \overset{d}{=} \vartheta(x_0)$.
ii. $\delta' X \sim MSNC_1(0,\delta' S\,\delta,\sqrt{\delta' S\,\delta})$.

Lemma 5. **[**Bennett G. (1986)**]** Let $a_1,a_2,\dots,a_n$ and $x_1,x_2,\dots,x_n$ be two arbitrary sets of real numbers. If α is a positive integer then the following equality

$$\left(\sum_{i=1}^n a_i x_i\right)^\alpha = \left(\sum_{i=1}^n a_i\right)^\alpha x_n^\alpha + \sum_{i=1}^{n-1}\left(\sum_{k=1}^i a_k\right)^\alpha (x_i^\alpha - x_{i+1}^\alpha) + \sum_{k_t\in A}\frac{\alpha!}{\prod_{j=1}^n k_j!}\left(\prod_{t=1}^i(\varepsilon_t)^{k_t}\right)\left[\left(\prod_{t=1}^{i-1} x_t^{k_t}\right) - x_i^{\alpha-k_i}\right] \quad (11)$$

is accomplished. The set A is defined as $A = \{k_t \in N; 0 < k_t < \alpha, \sum_{i=1}^n k_i = \alpha, k_{t+1} = k_{t+2} = \cdots = k_n = 0\}$

Lemma 6. **[**Bennett G. (1986)**]** Let $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n \geq 0$ and $r_1,r_2,\dots,r_n \geq 0$ be given. Then for any real numbers $p \geq 0$ and $0 \leq \alpha \leq p$, the following inequality is holds:

$$\left(\sum_{i=1}^n \varepsilon_i r_i\right)^\alpha \|r\|_p^{p-\alpha} \geq \sum_{i=1}^{n-1}(i)^{1-\frac{\alpha}{p}}\left(\sum_{k=1}^i \varepsilon_k\right)^\alpha (r_i^p - r_{i+1}^p) + n^{1-\frac{\alpha}{p}}\left(\sum_{k=1}^n \varepsilon_k\right)^\alpha r_n^p \quad (12)$$

where, $\|r\|_p = \left(\sum_{k=1}^n {r_k}^p\right)^{\frac{1}{p}}$

Proposition 7. Let $X_0 \sim MN_d(\mu,S)$. Then the Rényi entropy of X_0 can be written in the following form

$$R_\alpha(X_0;\mu,S) = \begin{cases} \frac{1}{2}\ln\left(\det(2\pi\exp(1)\,S)\right) & , \alpha = 1 \\ \frac{1}{2}\ln\left(\det(2\pi S)\right) - \frac{d}{2(1-\alpha)}\ln(\alpha) & ,0<\alpha<\infty, \alpha \neq 1 \end{cases} \quad (13)$$

Proof : firstly, assume that $\alpha = 1$, then we use the change of variables $Z = S^{-\frac{1}{2}}(X_0-\mu)$ to have directly that $H(X_0;\theta) = -E\left(-\frac{1}{2}\ln\left(\det(2\pi S)\right) - \frac{1}{2}z'z\right)$ but $z'z \sim N(0,1)$. Therefore, $H(X_0;\theta) = \left(\frac{1}{2}\ln\left(\det(2\pi S)\right) - \frac{1}{2}\right)$. If $\alpha \neq 1$ then the Rènyi entropy can be written as

$$R_\alpha(X_0;\mu,S) = \frac{1}{2}\ln\left(\det(2\pi S)\right) + \frac{1}{1-\alpha}\ln\left(E\left(\exp(-(x_0-\mu)S^{-1}(x_0-\mu))\right)^{\alpha-1}\right)$$

By using the transform $Z = \sqrt{\alpha}S^{-\frac{1}{2}}(X_0-\mu)$, we get on the result of this proposition.

3. RÉNYI ENTROPY OF MULTIVARIATE SKEW NORMAL-CAUCHY DISTRIBUTIONS

In this section we derive simple expressions for Rényi entropy of multivariate skew normal-Cauchy distributions by using some properties of transformation and integration. Also, we give an illustrative example explains the relationship between the parameters α and δ with the values of Rènyi entropy by using some methods of numerical integrations such as Monte Carlo and importance sampling methods.

The following lemma and proposition present the simple expression of Rényi entropy when $\alpha = 1$.

Lemma 8. If $X \sim MSNC_d(\mu, S, \delta)$, then the expected information in X of the random function $\ln\left(1+\frac{2}{\pi}\arctan\left(\delta'\tilde{S}^{-1}(x-\mu)\right)\right)$ appears in the following form:

$$E\left(\ln\left(1+\frac{2}{\pi}\arctan\left(\delta'\tilde{S}^{-1}(x-\mu)\right)\right)\right) = E\left\{\begin{matrix}\ln\left(1+\frac{2}{\pi}\arctan\left(\sqrt{\tilde{\delta}'\tilde{\delta}}\,y\right)\right)\\ .\left(1+\frac{2}{\pi}\arctan(\tilde{\delta}'\tilde{\delta}\,y)\right)\end{matrix}\right\} \quad (14)$$

where, $Y \sim N(0,1)$ and $\tilde{\delta} = S^{\frac{1}{2}}\tilde{S}^{-1}\delta$

Proof : Replacing $S^{\frac{1}{2}}\tilde{S}^{-1}\delta$ by the vector $\tilde{\delta}$ and transforming $Z = S^{-\frac{1}{2}}(X-\mu)$ then we have $Z \sim MSNC_d(0, I, \tilde{\delta})$. From proposition 2., $W = \tilde{\delta}'Z \sim MSNC_1(0, \tilde{\delta}'\tilde{\delta}, \sqrt{\tilde{\delta}'\tilde{\delta}})$. Therefore,

$$E\left(\ln\left(1+\frac{2}{\pi}\arctan\left(\delta'\tilde{S}^{-1}(x-\mu)\right)\right)\right) = \int_R \left\{\begin{matrix}\ln\left(1+\frac{2}{\pi}\arctan(w)\right)g(w;0,\tilde{\delta}'\tilde{\delta})\\ .\left(1+\frac{2}{\pi}\arctan(w)\right)\end{matrix}\right\} dw$$

The transform $Y = \left(\tilde{\delta}'\tilde{\delta}\right)^{-\frac{1}{2}}W$ finishes the proof.

Proposition 9. Let $X_0 \sim MN_d(\mu, S)$ and $X \sim MSNC_d(\mu, S, \delta)$. Then the Shannon entropy can be written as follows:

$$H(X;\mu,S,\delta) = H(X_0;\mu,S) - \hat{C}_{\tilde{\delta}} \quad (15)$$

where,

$$\hat{C}_{\tilde{\delta}} = E\left\{\begin{matrix}\ln\left(\left(1+\frac{2}{\pi}\arctan\left(\sqrt{\tilde{\delta}'\tilde{\delta}}\,y\right)\right)\right)\\ .\left(\left(1+\frac{2}{\pi}\arctan(\tilde{\delta}'\tilde{\delta}\,y)\right)\right)\end{matrix}\right\} \quad (16)$$

, $Y \sim N(0,1)$ and $\tilde{\delta} = S^{\frac{1}{2}}\tilde{S}^{-1}\delta$

Proof: By taking the natural logarithm and expectation for both sides of equation (2), we get

$$E\left(\ln(f_d(x;\mu,S,\delta))\right) = \ln\frac{(\det(S))^{-\frac{1}{2}}}{(2\pi)^{\frac{d}{2}}} - E\left(\left(\frac{1}{2}(x-\mu)'S^{-1}(x-\mu)\right)\right) + E\left(\ln\left(1+\frac{2}{\pi}\arctan\left(\delta'\tilde{S}^{-1}(x-\mu)\right)\right)\right)$$

Using proposition 2., we obtain

$$H(X;\mu,S,\delta) = H(X_0;\mu,S) - E\left(\ln\left(1+\frac{2}{\pi}\arctan\left(\delta'\tilde{S}^{-1}(x-\mu)\right)\right)\right)$$

Lemma 8. gives us the required result of this Proposition.

Next,, we drive the expression of Rènyi entropy of $X \sim MSNC_d(\mu, S, \delta))$ when $\alpha \neq 1$.

Lemma 10. Suppose that $X \sim MSNC_d(\mu, S, \delta)$. Then for any $\alpha > 0, \alpha \neq 1$ we have

$$\int_{R^d}\left(f_d(x;\mu,S,\delta)\right)^{\alpha}dx = \frac{(\det(2\pi S))^{\frac{1}{2}(1-\alpha)}}{\alpha^{\frac{d}{2}}} E\left(1+\frac{2}{\pi}\arctan\left(\sqrt{\hat{\delta}'\hat{\delta}}y\right)\right)^{\alpha} \quad (17)$$

where, $\hat{\delta} = \frac{1}{\sqrt{\alpha}}S^{\frac{1}{2}}\tilde{S}^{-1}\delta$ and $Y \sim N(0,1)$

proof: Replacing $S^{\frac{1}{2}}\tilde{S}^{-1}\delta$ by $\tilde{\delta}$ and transforming the variable $Z = S^{-\frac{1}{2}}(X-\mu)$ associated with Jacobian matrix $S^{\frac{1}{2}}$ of equation (2), we get

$$\int_{R^d}\left(f_d(x;\mu,S,\delta)\right)^{\alpha}dx = \frac{(\det(2\pi S))^{\frac{1}{2}(1-\alpha)}}{(2\pi)^{\frac{d}{2}}}\int_{R^d}\exp\left(\frac{-\alpha z'z}{2}\right)\left(1+\frac{2}{\pi}\arctan(\tilde{\delta}'z)\right)^{\alpha}dz$$

Again replacing $\frac{1}{\sqrt{\alpha}}\tilde{\delta}$ by $\hat{\delta}$, and we change the variable $\sqrt{\alpha}\,Z$ by U, to get

ISSN: www.jatit.org E-ISSN:

$$\int_{R^d}\left(f_d(x;\mu,S,\delta)\right)^{\alpha}dx=\frac{\alpha^{-\frac{d}{2}}\left(\det(S)\right)^{\frac{1}{2}(1-\alpha)}}{(2\pi)^{\frac{d}{2}}}\cdot\int_{R^d}\exp\left(\frac{-u'u}{2}\right)\left(1+\frac{2}{\pi}\arctan(\hat{\delta}'u)\right)^{\alpha}du$$

From proposition 4., we obtain $W=\hat{\delta}'U\sim MSNC_1\left(0,\hat{\delta}'\hat{\delta},\sqrt{\hat{\delta}'\hat{\delta}}\right)$. Consequently,

$$\int_{R^d}\left(f_d(x;\mu,S,\delta)\right)^{\alpha}dx=\frac{\alpha^{-\frac{d}{2}}\left(\det(S)\right)^{\frac{1}{2}(1-\alpha)}}{(2\pi)^{\frac{d}{2}}}E\left(1+\frac{2}{\pi}\arctan(w)\right)^{\alpha-1} \quad (18)$$

But,

$$E\left(\left(1+\frac{2}{\pi}\arctan(w)\right)^{\alpha-1}\right)=\frac{(\hat{\delta}'\hat{\delta})^{-\frac{1}{2}}}{(2\pi)^{\frac{d}{2}}}\int_R\exp\left(\frac{-w^2}{2\hat{\delta}'\hat{\delta}}\right)\left(1+\frac{2}{\pi}\arctan(w)\right)^{\alpha}dw$$

If we change $\frac{W}{\sqrt{\hat{\delta}'\hat{\delta}}}$ by the variable Y then the proof is completed.

Corollary 11.
Let $X\sim MSNC_d(\mu,S,\delta)$, then

$$R_{\alpha}(X;\mu,S,\delta)=\begin{cases}\frac{1}{2}\ln\left(\det(2\pi\exp(1)\,S)\right)-\hat{C}_{\tilde{\delta}}\ , & \alpha=1\\ \frac{1}{2}\ln\left(\det(2\pi S)\right)-\frac{d}{2(1-\alpha)}\ln(\alpha)+C_{\hat{\delta},\alpha} & ,\alpha\neq 1\end{cases} \quad (19)$$

where,

$$\hat{C}_{\tilde{\delta}}=E\left\{\begin{array}{l}\ln\left(\left(1+\frac{2}{\pi}\arctan\left(\sqrt{\tilde{\delta}'\,\tilde{\delta}}\,y\right)\right)\right)\\ \left(\left(1+\frac{2}{\pi}\arctan(\tilde{\delta}'\,\tilde{\delta}\,y)\right)\right)\end{array}\right\}$$

$$C_{\hat{\delta},\alpha}=\frac{1}{(1-\alpha)}\ln\left(E\left(1+\frac{2}{\pi}\arctan\left(\sqrt{\hat{\delta}'\hat{\delta}}\,y\right)\right)^{\alpha}\right),$$

$\hat{\delta}=\frac{1}{\sqrt{\alpha}}S^{\frac{1}{2}}\tilde{S}^{-1}\delta$ and $Y\sim N(0,1)$

Proof: The proof is immediate from lemma 10. and proposition 9.

Example 1.
As a simple illustrative example, explains the relationship between the parameters α and δ with Rényi entropy in one, two, three and four dimensions spaces. Assume that $\|\delta\|=\operatorname{trace}(\delta\delta')$.
Consider $X\sim MSNC_d(\mu,S,\delta)$ with the following parameters:

Case (1):
d=1, $\mu=0.3$, $S=1.5$, $\delta=1$, $\delta=2$, $\delta=3$ and $\delta=5$

Case (2):
d=2, $\mu=\begin{pmatrix}3\\2\end{pmatrix}$, $S=\begin{pmatrix}0.7 & 0.3\\0.3 & 3\end{pmatrix}$, $\delta=\begin{pmatrix}0\\1\end{pmatrix}$, $\delta=\begin{pmatrix}1\\\sqrt{3}\end{pmatrix}$, $\delta=\begin{pmatrix}2\\\sqrt{5}\end{pmatrix}$ and $\delta=\begin{pmatrix}4\\3\end{pmatrix}$

Case (3):
d=3, $\mu=\begin{pmatrix}0\\0\\0\end{pmatrix}$, $S=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$, $\delta=\begin{pmatrix}0\\1\\0\end{pmatrix}$, $\delta=\begin{pmatrix}1\\1\\\sqrt{2}\end{pmatrix}$, $\delta=\begin{pmatrix}2\\2\\1\end{pmatrix}$ and $\delta=\begin{pmatrix}0\\4\\3\end{pmatrix}$

Case (4):
d=4, $\mu=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$, $S=\operatorname{eye}(4)$, $\delta=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$, $\delta=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$, $\delta=\begin{pmatrix}2\\0\\1\\2\end{pmatrix}$ and $\delta=\begin{pmatrix}3\\3\\2\\\sqrt{3}\end{pmatrix}$

TABLE1. Rényi entropy of $MSNC_d(\mu,S,\delta)$ is computed for α = 2,3,4,5,6,8,10 and α converges to infinite, $\|\delta\|$=1,2,3 and 5 in one , two, three and four dimensions.

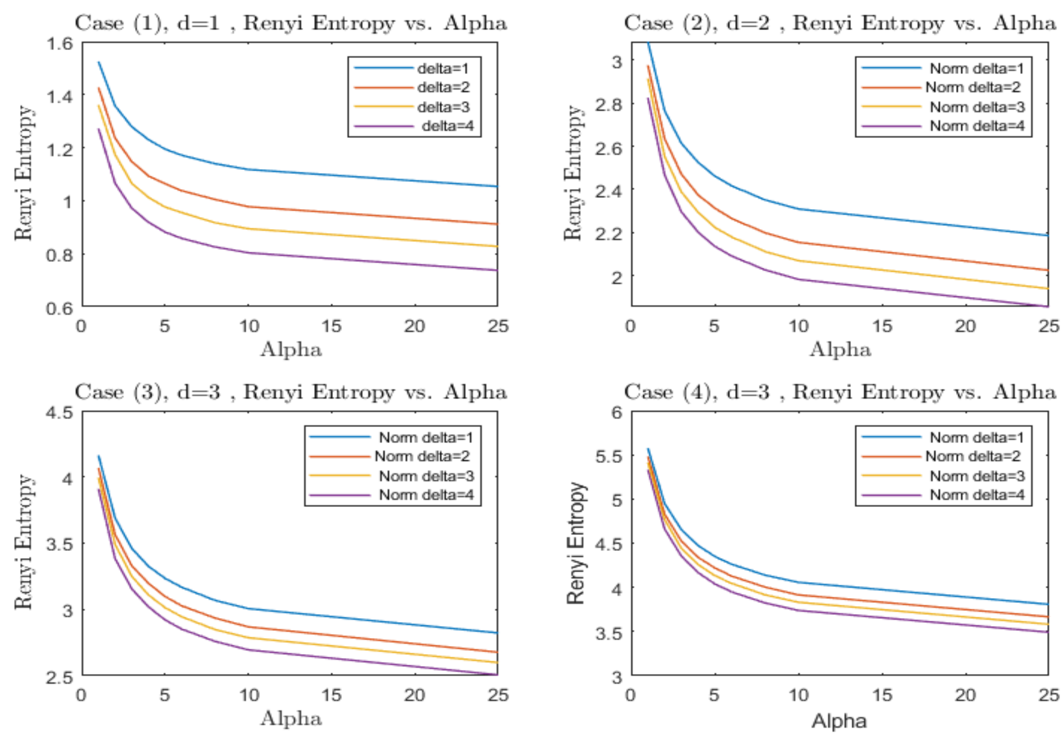
			Shannon entropy				*$R_{\alpha}(x;\mu,S,\delta)$*				
ase	d	$\|\delta$	$H(x;\mu,S,\delta$	$\alpha=2$	$\alpha=3$	$\alpha=4$	$\alpha=5$	$\alpha=6$	$\alpha=8$	$\alpha=10$	$\alpha\to\infty$
1	1	1	1.5256	1.3576	1.2793	1.2305	1.1948	1.1716	1.1398	1.1179	1.0538
	1	2	1.4280	1.2358	1.1474	1.0939	1.0640	1.0378	1.0049	0.9775	0.9117
	1	3	1.3610	1.1739	1.0656	1.0124	0.9764	0.9554	0.9171	0.8944	0.8277
	1	5	1.2727	1.0657	0.9718	0.9190	0.8817	0.8580	0.8258	0.8036	0.7374
2	2	1	3.0864	2.7662	2.6151	2.5265	2.4627	2.4174	2.3522	2.3097	2.1859
	2	2	2.9772	2.6361	2.4723	2.3751	2.3141	2.2659	2.1991	2.1547	2.0250
	2	3	2.9151	2.5541	2.3891	2.2948	2.2245	2.1803	2.1119	2.0697	1.9397

ISSN: www.jatit.org E-ISSN:

	2	5	2.8261	2.4679	2.2966	2.2028	2.1364	2.0916	2.0273	1.9829	1.8556
	3	1	4.1626	3.6896	3.4613	3.3267	3.2349	3.1680	3.0691	3.0069	2.8234
3	3	2	4.0682	3.5648	3.3296	3.1979	3.0982	3.0278	2.9356	2.8691	2.6777
	3	3	3.9939	3.4911	3.2510	3.1101	3.0141	2.9454	2.8487	2.7874	2.5998
	3	5	3.9103	3.3838	3.1558	3.0206	2.9217	2.8522	2.7598	2.6956	2.5068
4	4	1	5.5763	4.9496	4.6550	4.4745	4.3516	4.2633	4.1399	4.0577	3.8088
	4	2	5.4821	4.8296	4.5255	4.3411	4.2216	4.1294	4.0045	3.9153	3.6670
	4	3	5.4151	4.7677	4.4437	4.2596	4.1340	4.0470	3.9168	3.8322	3.5830
	4	5	5.3309	4.6613	4.3560	4.1684	4.0405	3.9499	3.8251	3.7400	3.4930

For each case in example 1., table 1. Summarizes the values of Rényi entropy for $\|\delta\| = 1,2,3,5$ and $\alpha = 1,2,3,4,5,6,8,10$, α converges to infinite value . we note that there is relationship between the values of Rényi entropy and the values of the parameters α, δ and d.

FIGURE 1. The horizontal line represents the values of parameter α and the vertical line is the Rényi entropy of $X \sim MSNC_d(\mu, S, \delta)$ with parameters in example 1.



The results in this section are shown in Figure 1. For dimensional parameter d=1,2,3,4, dispersion matrix S, skewness parameter δ with $\|\delta\| = 1,2,3,5$ and $\alpha = 1,2,3,4,5,6,8,10,25$. We can see that Rényi entropy is minimized and decreasing to increasing the values of norm of δ. Also, we observe that the Rényi entropy converges to a finite value for the values of α. It can be seen that the Rényi entropy is increasing with increasing dimensional parameter d.
The determine of dispersion matrices play important role in determining the value of Rényi entropy where it increases with its increasing. Also, we note that the effectiveness of α on Rényi entropy is

slow whenever the value of α is large and vice versa when it is small, Rényi entropy tends to a constant value when $\alpha \to \infty$, this is obvious when $\alpha > 25$.

4. UPPER, LOWER BOUNDS AND APPROXIMATION OF RÈNYI ENTROPY OF MIXTURE MODELS

In this section, we study the entropy of a proposal mixture model of skew Cauchy normal distributions. In other word, we find the upper and lower bounds of Rényi entropy by using the multinomial theorem , Holder inequality and some properties of L^p –spaces for this model. Also, an asymptotic expression for the Rényi entropy of mixture model is obtained.

Lemma 12. Let $X \sim MMSNC_d(\mu, S, \delta, \varepsilon)$. Then the upper and lower bounds for Shannon entropy $(\alpha = 1)$ of X are obtained in the following forms:

$$C_{upper} = \frac{1}{2}\ln\big(\det(2\pi\exp(1)R)\big) \quad (20)$$

$$C_{lower} = \sum_{i=1}^{m} \frac{\varepsilon_i}{2}\ln\big(\det(2\pi\exp(1)S_i)\big) - \sum_{i=1}^{m} \varepsilon_i \hat{C}_{\tilde{\delta}_i} \quad (21)$$

where,

$$\hat{C}_{\tilde{\delta}_i} = E\left\{\begin{array}{l} \ln\left(1 + \frac{2}{\pi}\arctan\left(\sqrt{\tilde{\delta}_i{}' \, \tilde{\delta}_i}\, y\right)\right) \\ .\left(1 + \frac{2}{\pi}\arctan\left(\tilde{\delta}_i{}' \, \tilde{\delta}_i\, y\right)\right) \end{array}\right\}$$

$$R = \sum_{i=1}^{m} \varepsilon_i S_i - \left(\sum_{i=1}^{m} \varepsilon_i \left(c_{\delta_i} S_i \, \delta_i \right)\right)\left(\sum_{i=1}^{m} \varepsilon_i \left(c_{\delta_i} S_i \, \delta_i \right)\right)'$$

$Y \sim N(0,1)$

Proof: The upper bound is obtained directly from lemma 1. and equation (7). The Shannon entropy of mixture model in equation (4) can be written as

$$H(X; \mu, S, \delta, \varepsilon) = -E\left(\ln\left(\sum_{i=1}^{m} \varepsilon_i f(x; \mu_i, S_i, \delta_i)\right)\right)$$

But the function $-\ln(x)$ is a concave, then by applying Jensen's inequality, we get

$$H(X; \mu, S, \delta, \varepsilon) \geq \sum_{i=1}^{m} \varepsilon_i H(X; \mu_i, S_i, \delta_i)$$

the Shannon entropy of each component is obtained by proposition 7. This complete the proof.

Lemma 13. If $X \sim MMSNC_d(\mu, S, \delta, \varepsilon)$, then for any positive integer α the following inequality is hold

$$R_\alpha(X; \mu, S, \delta, \varepsilon) \leq \mathfrak{C}_{Upper} \quad (22)$$

where,

$$\mathfrak{C}_{Upper} = \frac{1}{1-\alpha}\ln\Big\{\exp\big((1-\alpha)R_\alpha(X; \mu_m, S_m, \delta_m)\big) + \sum_{i=1}^{m-1}\Big(\sum_{k=1}^{i}\varepsilon_k\Big)^\alpha \left(\begin{array}{l} \exp\big((1-\alpha)R_\alpha(X; \mu_i, S_i, \delta_i)\big) \\ -\exp\big((1-\alpha)R_\alpha(X; \mu_{i+1}, S_{i+1}, \delta_{i+1})\big) \end{array}\right)\Big\}$$

Proof: The mixture probability density function in (5) implies that

$$\big(f(x; \mu, S, \delta, \varepsilon)\big)^\alpha = \Big(\sum_{i=1}^{m} \varepsilon_i f(x; \mu_i, S_i, \delta_i)\Big)^\alpha$$

Applying the result in lemma 4. when $p = \alpha$, we get

$$\Big(\sum_{i=1}^{m} \varepsilon_i f(x; \mu_i, S_i, \delta_i)\Big)^\alpha \geq f(x; \mu_m, S_m, \delta_m)^\alpha + \sum_{i=1}^{m-1}\Big(\sum_{k=1}^{i}\varepsilon_k\Big)^\alpha \left\{\begin{array}{l} f(x; \mu_i, S_i, \delta_i)^\alpha \\ -f(x; \mu_{i+1}, S_{i+1}, \delta_{i+1})^\alpha \end{array}\right\}$$

If we take the integral for both sides of above inequality over R^d , then

$$\int_{R^d}\big(f(x; \mu, S, \delta, \varepsilon)\big)^\alpha dx \geq \int_{R^d} f(x; \mu_m, S_m, \delta_m,)^\alpha dx + \sum_{i=1}^{m-1}\Big(\sum_{k=1}^{i}\varepsilon_k\Big)^\alpha \int_{R^d}\left[\begin{array}{l} f(x; \mu_i, S_i, \delta_i)^\alpha \\ -f(x; \mu_{i+1}, S_{i+1}, \delta_{i+1})^\alpha \end{array}\right] dx$$

Again, taking natural logarithm and multiplying by $\frac{1}{1-\alpha}$ for both sides of the last inequality , we have

$$R_\alpha(X; \mu, S, \delta, \varepsilon) \leq \frac{1}{1-\alpha}\ln\Big\{\int_{R^d} f(x; \mu_m, S_m, \delta_m)^\alpha dx + \sum_{i=1}^{m-1}\Big(\sum_{k=1}^{i}\varepsilon_k\Big)^\alpha \int_{R^d}\left[\begin{array}{l} \big(f(x; \mu_i, S_i, \delta_i)\big)^\alpha \\ -f(x; \mu_{i+1}, S_{i+1}, \delta_{i+1})^\alpha \end{array}\right] dx\Big\}$$

Lemma 14. Let $X \sim MMSNC_d(\mu, S, \delta, \varepsilon)$. Then for any positive integers $k_1, k_2, \ldots, k_m$ such that $\sum_{i=1}^{m} k_i = \alpha$ the following approximation

$$\frac{1}{\alpha}\ln\left\{\left(\frac{\alpha!}{k_1!\ k_2!\dots k_m!}\right)\prod_{i=1}^{m}\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^{k_i}\right\} \cong -\sum_{i=1}^{m}\gamma_i\ln\left(\frac{\gamma_i}{\varepsilon_i f(x;\mu_i,S_i,\delta_i)}\right) \quad (23)$$

is satisfied as $\alpha \to \infty$, where $\gamma_i = \frac{k_i}{\alpha}, i = 1,2,\dots,m$

Proof: we start the proof from the left side

$$\frac{1}{\alpha}\ln\left\{\left(\frac{\alpha!}{k_1!\ k_2!\dots k_m!}\right)\prod_{i=1}^{m}\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^{k_i}\right\} = \frac{1}{\alpha}\ln(\alpha!) - \frac{1}{\alpha}\sum_{i=1}^{m}\ln(k_i!) + \frac{1}{\alpha}\sum_{i=1}^{m}k_i\ln\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)$$

Approximating factorial in above equality implies that

$$\frac{1}{\alpha}\ln\left\{\left(\frac{\alpha!}{k_1!\ k_2!\dots k_m!}\right)\prod_{i=1}^{m}\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^{k_i}\right\} = \ln(\alpha) - 1 + \frac{1}{2\alpha}\ln(2\pi\alpha) - \frac{1}{\alpha}\sum_{i=1}^{m}k_I\ln(k_i) + \frac{1}{\alpha}\sum_{i=1}^{m}k_i - \frac{1}{2\alpha}\sum_{i=1}^{m}\ln(2\pi k_i) + \sum_{i=1}^{m}\gamma_i\ln\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)$$

Assume that $\gamma_i = \frac{k_i}{\alpha}$, $i = 1,2,\dots,m$, then $\sum_{i=1}^{m}\gamma_i = 1$.

Consequently,

$$\frac{1}{\alpha}\ln\left\{\left(\frac{\alpha!}{k_1!\ k_2!\dots k_m!}\right)\prod_{i=1}^{m}\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^{k_i}\right\} = \frac{1}{2\alpha}\left[\ln\left(\frac{(2\pi\alpha)^{1-n}}{\prod_{i=1}^{n}\gamma_i}\right)\right] - \sum_{i=1}^{m}\gamma_i\ln\left(\frac{\gamma_i}{\varepsilon_i f(x;\mu_i,S_i,\delta_i)}\right)$$

But $\lim_{\alpha\to\infty}\sum_{i=1}^{m}\gamma_i\ln\left(\frac{\gamma_i}{\varepsilon_i f(x;\mu_i,S_i,\delta_i)}\right) = 0$

This gives the right side.

Lemma 15. The approximation

$$R_\alpha(X;\mu,S,\delta,\varepsilon) \cong \frac{1}{1-\alpha}\ln\left(\sum_{k_i\in B}\left(\prod_{i=1}^{m}(\gamma_i)^{-k_i}\right)\prod_{i=1}^{m}\varepsilon_i^{k_i}\exp\left((1-\alpha)R_{k_i}(X;\mu_i,S_i,\delta_i)\right)\right) \quad (24)$$

is accomplished as $\alpha \to \infty$.

where, $\sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!} = m^\alpha$, $B = \{k_i \in N, k_i > 0, \sum_{i=1}^{m}k_i = \alpha, i = 1,2,\dots,m\}$

Proof: Applying the multinomial theorem in the equality

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^\alpha dx = \int_{R^d}\left(\sum_{i=1}^{m}\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^\alpha$$

To obtain

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^\alpha dx = \int_{R^d}\sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^{k_i}dx$$

where, $\sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!} = m^\alpha$, $B = \{k_i \in N, k_i > 0, \sum_{i=1}^{m}k_i = \alpha, i = 1,2,\dots,m\}$

By replacing the right side of equation (38) in equation (36), we get

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^\alpha dx \cong \int_{R^d}\sum_{k_i\in B}\exp\left\{-\alpha\sum_{i=1}^{m}\gamma_i\ln\left(\frac{\gamma_i}{\varepsilon_i f(x;\mu_i,S_i,\delta_i)}\right)\right\}dx$$

The last approximation can be written as

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^\alpha = \sum_{k_i\in B}\left[\prod_{i=1}^{m}(\gamma_i)^{-k_i}\right]\left[\prod_{i=1}^{m}\int_{R^d}\left(\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^{k_i}dx\right]$$

If we take the natural logarithm and multiplying by $\frac{1}{1-\alpha}$ for both sides of last approximation, then the proof is completed.

Lemma 16. For any positive integer $\alpha \geq 0$, $\alpha \neq 1$ the lower bound of Rényi entropy of $X \sim MMSNC_d(\mu,S,\delta,\varepsilon)$ appears as follow

$$R_\alpha(X;\mu,S,\delta,\varepsilon) \geq \mathfrak{C}_{Lower} \quad (25)$$

where,

$$\mathfrak{C}_{Lower} = \frac{1}{1-\alpha}\ln\left(\sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}(\varepsilon_i)^{k_i}\exp\left\{\frac{(1-\alpha)}{\alpha}\sum_{i=1}^{m}k_i R_\alpha(X;\mu_i,S_i,\delta_i)\right\}\right)$$

Proof By assumption $\alpha \neq 1$, then

$$R_\alpha(X;\mu,S,\delta,\varepsilon) = \frac{1}{1-\alpha}\ln\left(E\left(f(x;\mu,S,\delta,\varepsilon)\right)^{\alpha-1}\right) = \frac{1}{1-\alpha}\ln\left(\int_{R^d}\left(\sum_{i=1}^{m}\varepsilon_i f(x;\mu_i,S_i,\delta_i)\right)^\alpha dx\right)$$

By using the multinomial theorem, we have

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^\alpha dx = \sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}(\varepsilon_i)^{k_i}\int_{R^d}\prod_{i=1}^{m}\left(f(x;\mu_i,S_i,\delta_i)\right)^{k_i}dx$$

Applying generalized HÖlder's inequality, we get

$$\int_{R^d}f(x;\mu,S,\delta,\varepsilon)^\alpha dx \leq \sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}(\varepsilon_i)^{k_i}\prod_{i=1}^{m}\left(\int_{R^d}f(x;\mu_i,S_i,\delta_i)^{p_i k_i}dx\right)^{\frac{1}{p_i}}$$

where, $p_1, p_2, \dots, p_m > 0$ and $\sum_{i=1}^{m}\frac{1}{p_i} = 1$.

Therefore, the last equation can be written as

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^{\alpha}dx \le \sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}(\varepsilon_i)^{k_i} .\exp\left\{\sum_{i=1}^{m}\left(\frac{(1-p_ik_i)}{p_i}R_{p_ik_i}(X;\mu_i,S_i,\delta_i)\right)\right\}$$

If we choose $p_i=\frac{\alpha}{k_i}$, $i=1,2,\dots,m$ such that $\sum_{i=1}^{m}\frac{1}{p_i}=\sum_{i=1}^{m}\frac{k_i}{\alpha}=1$ and $1\le\frac{\alpha}{k_i}\le\alpha$, then

$$\int_{R^d}\left(f(x;\mu,S,\delta,\varepsilon)\right)^{\alpha}dx \le \sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}(\varepsilon_i)^{k_i}\exp\left\{\frac{(1-\alpha)}{\alpha}\sum_{i=1}^{m}k_iR_\alpha(X;\mu_i,S_i,\delta_i,)\right\}$$

We complete the proof by taking the natural logarithm and multiplying by $\frac{1}{1-\alpha}$ for both sides of last inequality.

Theorem 17. Let $X\sim MMSNC_d(\mu,S,\delta,\varepsilon)$. Then the approximation of Renyi entropy of X can be written as

$$R_\alpha(X;\mu,S,\delta,\varepsilon)= \frac{1}{2(1-\alpha)}\Big\{\ln\Big(\sum_{k_i\in B}\frac{\alpha!}{\prod_{i=1}^{m}k_i!}\prod_{i=1}^{m}(\varepsilon_i)^{k_i} \exp\left\{\frac{(1-\alpha)}{\alpha}\sum_{i=1}^{m}k_iR_\alpha(X;\mu_i,S_i,\delta_i)\right\} + (1-\alpha)R_\alpha(X;\mu_m,S_m,\delta_m) +\ln\Big(\sum_{i=1}^{m-1}\big(\sum_{k=1}^{i}\varepsilon_k\big)^{\alpha} \big(\exp((1-\alpha)R_\alpha(X;\mu_i,S_i,\delta_i)) -\exp((1-\alpha)R_\alpha(X;\mu_{i+1},S_{i+1},\delta_{i+1}))\big)\Big)\Big\} \quad (26)$$

Proof: The proof is directed from lemmas 13. and 16., by taking the mean of upper and lower bounds.

Example 2. To study the behavior of approximate Rènyi entropy in theorem 17. and its bounds in lemma 12., equations (22) and (25), some cases in this example are simulated for one, two and three dimension spaces. Consider $X\sim MMSNC_d(0,S,\delta,\varepsilon)$ with the following parameters:

Case (1) : d=1

m=2, $\varepsilon=(0.2,0.8)$, $S=(1.5,5)$, $\delta=(0.3,4)$

m=3 , $\varepsilon=(0.2,0.3,0.5)$, $S=(1.5,5,3)$, $\delta=(0.3,4,2.2)$

m=4, $\varepsilon=(0.1,0.2,0.2,0.5,$ $S=(1.5,5,3,2)$, $\delta=(0.3,4,2.2,1)$

m=5 , $\varepsilon=(0.2,0.2,0.2,0.2,0.2)$, $S=(1.5,5,3,2,5)$, $\delta=(0.3,4,2.2,1,2.1)$

Case (2): d=2

m=2, $\varepsilon=(0.2,0.8)$, $S=\left(\begin{pmatrix}0.7&0.3\\0.3&3\end{pmatrix},\begin{pmatrix}0.12&0.13\\0.13&3\end{pmatrix}\right)$, $\delta=\left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix}\right)$

m=3 , $\varepsilon=(0.2,0.3,0.5)$, , $S=\left(\begin{pmatrix}0.7&0.3\\0.3&3\end{pmatrix},\begin{pmatrix}0.12&0.13\\0.13&3\end{pmatrix},\begin{pmatrix}0.18&0.6\\0.6&4\end{pmatrix}\right)$

, $\delta=\left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix},\begin{pmatrix}2.6\\1\end{pmatrix}\right)$

m=4 , $\varepsilon=(0.1,0.2,0.2,0.5)$, $S=\left(\begin{pmatrix}0.7&0.3\\0.3&3\end{pmatrix},\begin{pmatrix}0.12&0.13\\0.13&3\end{pmatrix},\begin{pmatrix}0.18&0.6\\0.6&4\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix}\right)$

, $\delta=\left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix},\begin{pmatrix}2.6\\1\end{pmatrix},\begin{pmatrix}0.6\\1\end{pmatrix}\right)$

m=5 , $\varepsilon=(0.2,0.2,0.2,0.2,0.2)$, $S=\left(\begin{pmatrix}0.7&0.3\\0.3&3\end{pmatrix},\begin{pmatrix}0.12&0.13\\0.13&3\end{pmatrix},\begin{pmatrix}0.18&0.6\\0.6&4\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix}\right)$

, $\delta=\left(\begin{pmatrix}0.16\\0.59\end{pmatrix},\begin{pmatrix}2.3\\3.1\end{pmatrix},\begin{pmatrix}2.6\\1\end{pmatrix},\begin{pmatrix}0.6\\1\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right)$

Case (3): d=3

m=2, $\varepsilon=(0.2,0.8)$, $S=\left(\begin{pmatrix}0.7&0.3&0.5\\0.3&3&0.3\\0.5&0.3&1\end{pmatrix},\begin{pmatrix}5&0.3&2\\0.3&5&1\\2&1&3\end{pmatrix}\right)$, $\delta=\left(\begin{pmatrix}0.16\\0.59\\0.1\end{pmatrix},\begin{pmatrix}2.3\\3.1\\1.5\end{pmatrix}\right)$

ISSN: www.jatit.org E-ISSN:

m=3 , $\varepsilon = (0.2,0.3,0.5)$, $S = \left(\begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.3 \\ 0.5 & 0.3 & 1 \end{pmatrix}, \begin{pmatrix} 5 & 0.3 & 2 \\ 0.3 & 5 & 1 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$

, $\delta = \left(\begin{pmatrix} 0.16 \\ 0.59 \\ 0.1 \end{pmatrix}, \begin{pmatrix} 2.3 \\ 3.1 \\ 1.5 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}\right)$

m=4 , $\varepsilon = (0.1,0.2,0.2,0.5)$, $S = \left(\begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.3 \\ 0.5 & 0.3 & 1 \end{pmatrix}, \begin{pmatrix} 5 & 0.3 & 2 \\ 0.3 & 5 & 1 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$

,$\delta = \left(\begin{pmatrix} 0.16 \\ 0.59 \\ 0.1 \end{pmatrix}, \begin{pmatrix} 2.3 \\ 3.1 \\ 1.5 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right)$

m=5 , $\varepsilon = (0.2,0.2,0.2,0.2,0.2)$,

, $S = \left(\begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.3 \\ 0.5 & 0.3 & 1 \end{pmatrix}, \begin{pmatrix} 5 & 0.3 & 2 \\ 0.3 & 5 & 1 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$

, $\delta = \left(\begin{pmatrix} 0.16 \\ 0.59 \\ 0.1 \end{pmatrix}, \begin{pmatrix} 2.3 \\ 3.1 \\ 1.5 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}\right)$

For each cases in example 2., table 2. summarize the values of approximate Rènyi (Shannon) entropy and its bounds of $X \sim MMSNC_d(0, S, \delta, \varepsilon)$ with above parameters for $\alpha = 1$. We see that the values of approximation localized between the lower and upper bounds. Also, we observe that the Rényi entropy converges to finite value for each vector ε. it can be seen that the value of m changes the values of entropy.

TABLE 2. Shannon entropy of $X \sim MMSNC_d(0, S, \delta, \varepsilon)$ is computed for m = 2,3,4 and 5 in one , two and three dimensions.

Approximate Shannon entropy						
Case	d	m	C_{lower}	C_{upper}	H	Error
1	1	2	1.8558	1.8671	1.8614	0.0057
	1	3	1.7741	1.7826	1.7783	0.0042
	1	4	1.7290	1.7374	1.7332	0.0042
	1	5	1.7972	1.8136	1.8054	0.0082
2	2	2	2.2154	2.3827	2.2990	0.0836
	2	3	2.2427	2.4911	2.3669	0.1242
	2	4	2.4786	2.8261	2.6523	0.1738
	2	5	2.5190	2.8414	2.6802	0.1612
3	3	2	5.5641	5.6966	5.6304	0.0663
	3	3	4.6287	5.0595	4.8441	0.2154
	3	4	4.4237	4.8133	4.6185	0.1948
	3	5	4.4452	4.8519	4.6486	0.2033

ISSN: www.jatit.org E-ISSN:

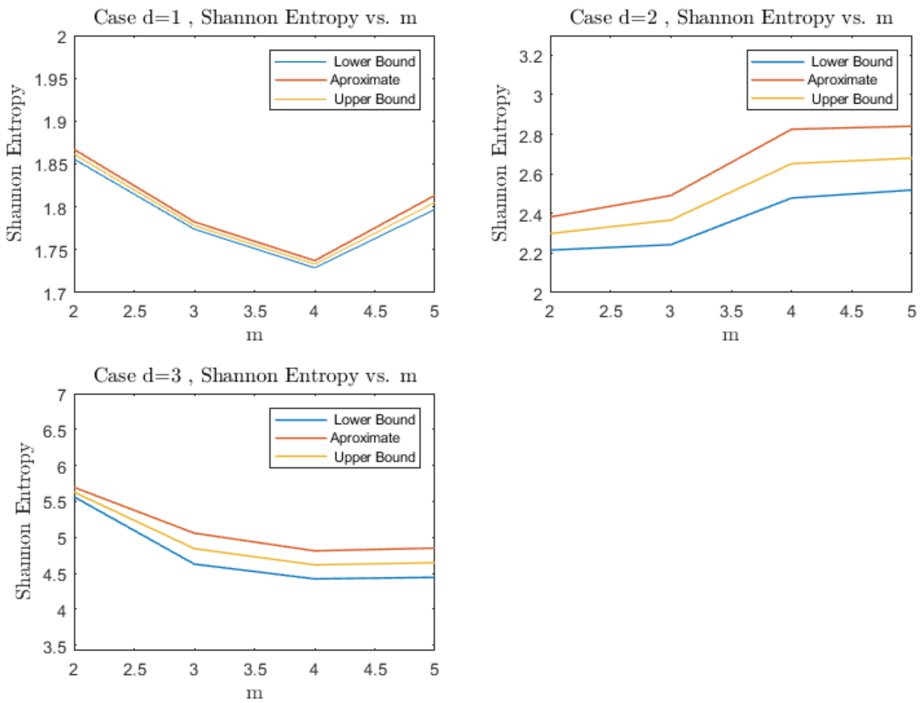


FIGURE 2. The horizontal line represents the values of parameter m and the vertical line is a Shannon entropy of $X \sim MMSNC_d(0, S, \delta, \varepsilon)$ with parameters in example 2.

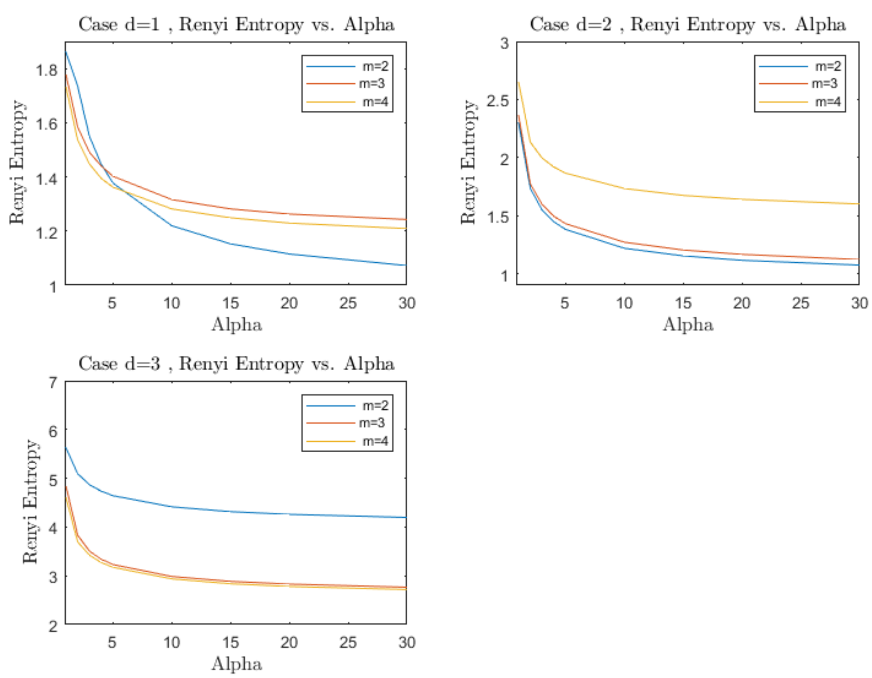


Figure 3. The Horizontal Line Represents The Values Of Parameter α And The Vertical Line Is A Renyi Entropy Of $X \sim MmsCn_D(0, S, \Delta, E)$ With Parameters In Example 2.

The results in this section are shown in Figures 2. and 3. For dimensional parameter d=1,2,3, dispersion matrix S, skewness parameter δ and $\alpha = 1,2,3,4,5,10,\ 15,\ 20,30$. We can see that Rényi entropy is minimized and decreasing for increasing the values of α and tends to a constant value when $\alpha \rightarrow \infty$, this is obviously when $\alpha > 30$. Also, we note that the Rényi entropy converges to finite value for each cases. it can be seen that the Rényi entropy is increasing with increasing dimensional parameter d.

5. CONCLUSION

We have derived the upper and lower bounds on the Rènyi entropy of $X{\sim}MMSCN_d(0, S, \delta, \varepsilon)$. Using the mean of these bounds, the approximate values of entropy can be calculated. These values are localized between bounds of Rènyi entropy. The entropy both types (Shannon and Rényi) converges to a finite value for any values of α , m and d. It has been established that the Rènyi entropy of $X{\sim}MMSCN_d(0, S, \delta, \varepsilon)$ depends proportionally upon the parameters α , S and d but there is not clear relation between this entropy and the parameter m. In fact, there is not an analytical method to find the exact value of the Rényi entropy of mixture model of distributions therefore, our approximation is effective and more accurate. We have seen through the examples given in this article that the error in the values of Rényi entropy by approximation was almost acceptable.

REFRENCES:

[1]. Arellano-Vall R., Gomez H., & Quintana F. (2003). A New Class of Skew Normal Distributions. Journal Communication in Statistic-Theory and Methods, Vol. 33, Issue 7.
[2]. Azzalini A., & Capitanio A. (2003). Distribution Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution. Journal of the Royal Statistical Society: series B, 65 (2), pp. 367-389.
[3]. Azzalini A., & Dalla Valle A. (1996). The Multivariate Skew-normal distribution. Biometrika, 83(4), pp. 715-726.
[4]. Bennett G. (1986). Lower Bounds for Matrices. Elsevier Science Publishing Co. Inc., 52 Vanderbilt Ave., NewYork, NY 10017.
[5]. Contreras-Reyes J., & Cortés D. (2016). Bounds on Rényi and Shannon Entropies for Finite Mixtures of Multivariate Skew-Normal Distributions: Applications to Swordfish (Xiphias gladius Linnaeus). Entropy 11 (382); doi:10.3390/e18110382.
[6]. Cover T. M. (2006). Elements of Information Theory. Wily and Son, 2nd ed., New York. NY, USA.
[7]. Genton M., & Loperfido N. (2005). Generalized skew elliptical distributions and their quadratic forms. Annals of the Institute of Statistical Mathematics, 57 (2), pp. 389–401.
[8]. Huang W., Su N., & Gupta A. (2013). A study of Generalized Skew-symmetric Models. Journal of Multivariate Analysis. 101 (6), pp. 1434-1444.
[9]. Javier E. and Contreras-Reyes J. (2016). Rényi Entropy and Complexity Measure for Skew-Gaussian Distributions and Related Families. arXiv:1406.0111v2[physics. Data-an] 8 May.
[10]. Javier E., & Contreras-Reyes J. (2016). Rényi Entropy and Complexity Measure for Skew-Gaussian Distributions and Related Families. Retrieved from arXiv.org website: https://arxiv.org/abs/ arXiv:1406.0111v21
[11]. José C. (2010). Principe Information Theoretic Learning Renyi's Entropy and Kernal Perspectives. Springer, New York, NY, USA.
[12]. Kahrari F., Lazard, Rezael M., Yousefzadeh F., & Arellano-Vall R. (2016). On Multivariate Skew-normal-Cauchy Distribution. Elsevier, Statistic and Probability Letters 117.
[13]. Lee S., & McLachlan G. (2014). Finite Mixtures of Multivariate Skew t-distributions. Statistics and Computing, 24(2).
[14]. Lin T., Lee J., & Wan H. (2007). Robust Mixture Modeling Using the Skew t-Distribution. Statistics and Computing, 17(2), pp.81-92.
[15]. Pyne S., Hu X, Wang K., Rossin E., Lin T., Maier L., Baecher-Allan C., McLachlan G., Tamayo P., Halfer D., De J.,& Mesirow J. (2009). Automated High-dimensional Flow Cytometric data Analysis. Proceedings of the National Academy of Sciences USA 106:8519-8524.
[16]. Rényi A. (1961). On Measures of Information and entropy. Proceeding of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability pp. 547-561.

[17]. Shanon C. E. (1948). A mathematical Theory of Communication. Bell systems technology, j., 27, pp. 379-423.
[18]. Wood R., Blythe R. & Evans M. (2017). Rényi Entropy of the Totally Asymmetric Exclusion Process. Retrieved from arXiv.org website: https://arxiv.org/abs/arXiv:1708.003031.