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

SALAH H. ABID, UDAY J. QUAEZ
Department of Mathematics, Education College, Al-Mustansiriya University, Baghdad, Iraq
E-mail: abidsalah@gmail.com, 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

$$k(y; \mu, S) = 2g(y; \mu, S) \psi(y - \mu) , \quad y \in \mathbb{R}^d$$

where the function $\psi$ satisfies $0 \leq \psi(y) \leq 1$ and $\psi(-y) = 1 - \psi(y)$, for any $y \in \mathbb{R}^d$.

Obviously, if $\psi(y) = \frac{1}{2}$ then $Y$ has a multivariate normal distribution. As a special case if we take $\psi$ 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.


On the other hand, Shannon 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 Rényi entropy of the mixture model, therefore we consider its bounds exist. Similarity in the case of fractional relative entropy an analytic evolution of Rényi 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 $\alpha$, $\varepsilon$ and skewness parameter $\delta$ 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 $Y \in \mathbb{R}$ 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 form:

$$k(y; \mu, S, \delta_1, \delta_2) = 2g(y; \mu, S).G\left(\frac{\delta_1y}{\sqrt{1+\delta_2y^2}}\right)$$

(1)

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

$$f_d(x; \mu, S, \delta) = \frac{(\det(S))^{-\frac{1}{2}}}{(2\pi)^{d/2}}\exp\left\{-\frac{1}{2}(x - \mu)^TS^{-1}(x - \mu)\right\}$$

$$\cdot \exp\left\{(1 + \frac{2}{\pi} \arctan(\delta^TS^{-1}(x - \mu)))\right\}$$

(2)

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

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

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

(3)

$$\text{Var}(X) = S - c_\delta S \delta^TS$$

(4)

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

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

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

Definition 1. Let $X$ be a d-dimensional random vector which comes from an m-component
mixtures of multivariate skew normal-Cauchy distributions. Then the probability density function of $X \sim \text{MMSC}_d(\mu, S, \delta, \epsilon)$ is given by the following form:

$$f(x; \mu, S, \delta, \epsilon) = \sum_{i=1}^{m} \epsilon_i f(x; \mu_i, S_i, \delta_i)$$  \hspace{1cm} (5)

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

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

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

where $Z_j | Y_j \sim \text{MSN}_d(0, S_j, \delta_j Y)$, $Y_j \sim \text{HN}(0, 1)$ and $c_{\delta j} = \frac{2}{\pi \sqrt{S_j \delta_j}} m \left( \frac{1}{S_j \delta_j} \right)^{1/2}$, $i = 1, 2, \ldots, m$.

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

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

$$\text{Var}(X) = \sum_{i=1}^{m} \epsilon_i S_i - \left( \sum_{i=1}^{m} \epsilon_i (c_{\delta_i} S_i \delta_i) \right) \left( \sum_{i=1}^{m} \epsilon_i (c_{\delta_i} S_i \delta_i) \right)'$$ \hspace{1cm} (7)

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

$$H(X; \theta) = -E(\ln(f(x; \theta)))$$ \hspace{1cm} (8)

**Definition 3.** An $\alpha$th-order Rényi entropy of a continuous random vector $X \in \mathbb{R}^d$ with probability density function $f(x; \theta)$ is defined as

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

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

The second part in the right side can be written as

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

Therefore,

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

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

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

But $E_g \left( \ln \frac{g(x; \theta)^\alpha}{f(x; \theta)} \right)$ not negative value therefor, either $E_g \left( \ln \frac{g(x; \theta)^\alpha}{f(x; \theta)} \right) > 0$ then $R_\alpha$ is a decreasing with respect to $\alpha$ or $E_g \left( \ln \frac{g(x; \theta)^\alpha}{f(x; \theta)} \right) = 0$ 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 
\[
  f(x; \theta) = \begin{cases} 
  \frac{1}{\text{Vol}(B)}, & x \in B \\
  0, & x \in \overline{B}
\end{cases}
\]
for some measurable subset \( B \) in \( \mathbb{R}^d \), then 
\[
  R_\alpha(X; \theta) = \frac{1}{1-\alpha} \ln \left( \text{Vol}(B)^{\alpha-1} \right) = \ln \left( \text{Vol}(B) \right)
\]
is not dependent on the value of \( \alpha \).

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

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

**Proof:** The probability density function \( f \) can be written as 
\[
  f(x; \mu, S) = \left( \det(S) \right)^{-\frac{1}{2}} f_{x_0} \left( S^{-\frac{1}{2}}(X - \mu), 0, 1 \right)
\]
Therefore, the Rényi entropy of \( X \) appears as 
\[
  R_\alpha(X; \mu, S) = \ln \left( \det(S) \right)^{\frac{1}{2}} + \frac{1}{1-\alpha} \ln \left( \text{E} \left( f(x_0; 0, 1) \right) \right)^{\alpha-1}
\]
\[
  = \ln \left( \det(S) \right)^{\frac{1}{2}} + R_\alpha(X; 0, 1)
\]

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

\[
  H(X; \theta) \leq \frac{1}{2} \ln \left( \det(2\pi\exp(1)R) \right)
\]
with equality if and only if \( X \sim \text{MN}_d(0, S) \).

**Proposition 4.** [Kahrari et al., (2016)] If the random vectors \( X \sim \text{MSNC}_d(0, S, \delta) \) and \( X_0 \sim \text{MN}_d(0, S) \), then 
i. For every even function \( \theta \), we have \( \theta(x) = \theta(x_0) \).
ii. \( \delta'X \sim \text{MSNC}_d(0, \delta' S \delta, \sqrt{\delta' S \delta}) \).

**Lemma 5.** [Bennett G. (1986)] Let \( a_1, a_2, \ldots, a_n \) and \( x_1, x_2, \ldots, x_n \) be two arbitrary sets of real numbers. If \( \alpha \) is a positive integer then the following equality 
\[
  \left( \sum_{i=1}^{n} a_i x_i \right)^{\alpha} = \left( \sum_{i=1}^{n} a_i x_i \right)^{\alpha} + \sum_{k=1}^{n} \left( \sum_{i=k}^{n} a_i x_i \right)^{\alpha-1} \left( \sum_{i=k}^{n} a_i x_i \right)^{\alpha} + \sum_{k \in A} \prod_{i \in [k]} \left( \prod_{i \in [k]} a_i x_i \right)^{\alpha-1}
\]
is accomplished. The set \( A \) is defined as 
\[
  A = \left\{ k \in \mathbb{N} ; 0 \leq k \leq n, \sum_{k=0}^{n} k = 0, k_{i+1} = k_{i+2} = \ldots \right\}
\]

**Lemma 6.** [Bennett G. (1986)] Let \( \epsilon_1, \epsilon_2, \ldots, \epsilon_n \geq 0 \) and \( r_1, r_2, \ldots, r_n \geq 0 \) be given. Then for any real numbers \( p \geq 0 \) and \( 0 \leq \alpha \leq p \), the following inequality is holds:

\[
  \left( \sum_{i=1}^{n} \epsilon_i r_i \right)^{p-\alpha} \geq \sum_{i=1}^{n} \left( \sum_{k=1}^{i} \epsilon_k \right)^{p-\alpha} \left( r_i^p - r_{i+1}^p \right)
\]
(12)

**Proposition 7.** Let \( X_0 \sim \text{MN}_d(\mu, S) \). Then the Rényi entropy of \( X_0 \) can be written in the following form 
\[
  R_\alpha(X_0; \mu, S) = \frac{1}{2} \ln \left( \det(2\pi S) \right)
\]
\[
  + \frac{1}{\alpha-1} \ln \left( \text{E} \left( \exp \left( -(x_0 - \mu)S^{-1}(x_0 - \mu) \right) \right) \right)^{\alpha-1}
\]
By using the transform \( Z = \sqrt{\delta} S^{-\frac{1}{2}}(X_0 - \mu) \), 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 \( \alpha \) and \( \delta \) with the values of Rényi 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 \( X \sim \text{MSNC}_{\alpha}(\mu, S, \delta) \), then the expected information in \( X \) of the random function

\[
E \left( \ln \left( 1 + \frac{2}{\pi} \arctan \left( \frac{1}{\delta} S^{-1}(x - \mu) \right) \right) \right)
\]

appears in the following form:

\[
= E \left\{ \ln \left( 1 + \frac{2}{\pi} \arctan \left( \frac{1}{\sqrt{\delta}} y \right) \right) \right\}
\]

(14)

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

**Proof:** Replacing \( S^{\frac{1}{2}} S^{-\frac{1}{2}} \) by the vector \( \delta \) and transforming \( Z = S^{\frac{1}{2}} (X - \mu) \) then we have \( Z \sim \text{MSNC}_{\alpha}(0, 1, \delta) \). From proposition 2., \( W = \delta'Z \sim \text{MSNC}_{\alpha}(0, \delta', \delta, \sqrt{\delta'}) \). Therefore,

\[
E \left( \ln \left( 1 + \frac{2}{\pi} \arctan \left( \frac{1}{\delta} S^{-1}(x - \mu) \right) \right) \right)
= f_R \left\{ \ln \left( 1 + \frac{2}{\pi} \arctan \left( \frac{1}{\sqrt{\delta}} y \right) \right) \right\} \, \phi(w; 0, \delta') \, dw
\]

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

**Proposition 9.** Let \( X_0 \sim \text{MN}_{\alpha}(\mu, S) \) and \( X \sim \text{MSNC}_{\alpha}(\mu, S, \delta) \). Then the Shannon entropy can be written as follows:

\[
H(X; \mu, S, \delta) = H(X_0; \mu, S) - \check{C}_\delta
\]

(15)

where,
From proposition 4., we obtain $W=\delta'U \sim \text{MSNC}_4\left(0, \delta', \sqrt{\delta'} \right)$. Consequently,

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

From proposition 4., we obtain $W=\delta'U \sim \text{MSNC}_4\left(0, \delta', \sqrt{\delta'} \right)$. Consequently,

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

But,

$$
E \left( 1 + \frac{2}{\pi} \arctan(w) \right)^{\alpha-1} = \frac{\alpha^{-\frac{d}{2}} (\text{det}(S))^{-\frac{1}{2}}}{(2\pi)^\frac{d}{2}} \int_R \left( 1 + \frac{2}{\pi} \arctan(w) \right)^{\alpha-1} dw
$$

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

**Example 1.**

As a simple illustrative example, explains the relationship between the parameters $\alpha$ and $\delta$ with Rényi entropy in one, two, three and four dimensions spaces. Assume that $\|\delta\| = \text{trace}(\delta \delta')$. Consider $X \sim \text{MSNC}_d(\mu, S, \delta)$ 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):**

$d=2$, $\mu = \left(\frac{3}{2}\right)$, $S = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 3 \end{pmatrix}$, $\delta = \left(\frac{1}{\sqrt{3}}\right)$, $\delta = \left(\frac{2}{\sqrt{5}}\right)$ and $\delta = \left(\frac{4}{3}\right)$.

**Case (3):**

$d=3$, $\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\delta = \left(\frac{1}{\sqrt{2}}\right)$, $\delta = \left(\frac{2}{1}\right)$ and $\delta = \left(\frac{4}{3}\right)$.

**Case (4):**

$d=4$, $\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $S = \text{eye}(4)$, $\delta = \left(\frac{1}{\sqrt{3}}\right)$, $\delta = \left(\frac{3}{2}\right)$ and $\delta = \left(\frac{3}{\sqrt{2}}\right)$.

**TABLE 1.** Rényi entropy of $\text{MSNC}_d(\mu, S, \delta)$ is computed for $\alpha = 2, 3, 4, 5, 6, 8, 10$ and $\alpha$ converges to infinite, $\|\delta\|^d = 1, 2, 3, 4$ and $\delta = 5$ in one, two, three and four dimensions.
For each case in example 1., table 1. Summarizes the values of Rényi entropy for \( \|\delta\| = 1,2,3,5 \) and \( \alpha = 1,2,3,4,5,6,8,10 \), \( \alpha \) converges to infinite value. We note that there is a relationship between the values of Rényi entropy and the values of the parameters \( \alpha, \delta \) and \( d \).

The results in this section are shown in Figure 1. For dimensional parameter \( d=1,2,3,4 \), dispersion matrix \( S \), skewness parameter \( \delta \) 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 \( \delta \). Also, we observe that the Rényi entropy converges to a finite value for the values of \( \alpha \). 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 \( \alpha \) on Rényi entropy is

![Figure 1](image-url)
slow whenever the value of $\alpha$ 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 RÉNYI 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, Hölder 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 $X \sim \text{MMSNC}_d(\mu, S, \delta, \varepsilon)$.
Then the upper and lower bounds for Shannon entropy $(\alpha = 1)$ of $X$ are obtained in the following forms:

$$C_{\text{upper}} = \frac{1}{2} \ln \left( \det(2\pi\exp(1)R) \right)$$

(20)

$$C_{\text{lower}} = \sum_{i=1}^{m} \frac{1}{2} \ln \left( \det(2\pi\exp(1)S_i) \right) - \sum_{i=1}^{m} \varepsilon_i \hat{C}_{\delta_i}$$

(21)

where,

$$\hat{C}_{\delta_i} = E \left\{ \left( 1 + \frac{2}{\pi} \arctan \left( \frac{\sqrt{\delta_i y}}{\hat{\delta}_i y} \right) \right) \right\}$$

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

$Y \sim \text{N}(0,1)$

**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

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

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 $X \sim \text{MMSNC}_d(\mu, S, \delta, \varepsilon)$, then for any positive integer $\alpha$ the following inequality is hold

$$R_{\alpha}(X; \mu, S, \delta, \varepsilon, \varepsilon) \leq C_{\text{Upper}}$$

(22)

where,

$$C_{\text{Upper}} = \frac{1}{1-\alpha} \ln \left( \exp \left( \frac{R_{\alpha}(X; \mu, S, \delta, \varepsilon, \varepsilon)}{\sum_{k=1}^{i} \varepsilon_k} \right) \right)$$

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

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

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

$$\left( \sum_{i=1}^{m} \varepsilon_i f(x; \mu_i, S_i, \delta_i) \right)^{\alpha} \geq f(x; \mu_m, S_m, \delta_m)^{\alpha}$$

$$+ \sum_{k=1}^{i} \varepsilon_k \left( \sum_{i=1}^{m} \varepsilon_i f(x; \mu_i, S_i, \delta_i) \right)^{\alpha}$$

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

$$\int_{R^d} (f(x; \mu, S, \delta, \varepsilon))^{\alpha} \, dx \geq \int_{R^d} f(x; \mu_m, S_m, \delta_m)^{\alpha} \, dx$$

$$+ \sum_{k=1}^{i} \varepsilon_k \int_{R^d} \left( f(x; \mu_i, S_i, \delta_i) \right)^{\alpha} \, dx$$

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

$$R_{\alpha}(X; \mu, S, \delta, \varepsilon, \varepsilon) \leq \frac{1}{1-\alpha} \ln \left( \int_{R^d} f(x; \mu, S, \delta, \varepsilon)^{\alpha} \, dx \right)$$

$$+ \sum_{k=1}^{i} \varepsilon_k \int_{R^d} \left( f(x; \mu_i, S_i, \delta_i) \right)^{\alpha} \, dx$$

**Lemma 14.** Let $X \sim \text{MMSNC}_d(\mu, S, \delta, \varepsilon)$. Then for any positive integers $k_1, k_2, \ldots, k_m$ such that $\sum_{i=1}^{m} k_i = \alpha$ the following approximation
\[ \frac{1}{\alpha} \ln \left( \frac{\alpha^n}{\prod_{k=1}^{m} \gamma_k} \right) \prod_{i=1}^{m} \left( e^{\gamma_i f(x; \mu_i, S_i, \delta_i)} \right) \leq - \sum_{i=1}^{m} y_i \ln \left( \frac{\gamma_i}{e^{\gamma_i f(x; \mu_i, S_i, \delta_i)}} \right) \]

(23)

is satisfied as \( \alpha \to \infty \), where \( y_i = \frac{k_i}{\alpha}, i = 1, 2, \ldots, m \).

**Proof:** we start the proof from the left side

\[ \frac{1}{\alpha} \ln \left( \frac{\alpha^n}{\prod_{k=1}^{m} \gamma_k} \right) \prod_{i=1}^{m} \left( e^{\gamma_i f(x; \mu_i, S_i, \delta_i)} \right) \]

= \frac{1}{\alpha} \ln(\alpha) - \frac{1}{\alpha} \sum_{i=1}^{m} k_i \ln(k_i) 
+ \frac{1}{\alpha} \sum_{i=1}^{m} k_i \ln(e^{\gamma_i f(x; \mu_i, S_i, \delta_i)})

Approximating factorial in above equality implies that

\[ \frac{1}{\alpha} \ln \left( \frac{\alpha^n}{\prod_{k=1}^{m} \gamma_k} \right) \prod_{i=1}^{m} \left( e^{\gamma_i f(x; \mu_i, S_i, \delta_i)} \right) \]

= \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} y_i \ln(e^{\gamma_i f(x; \mu_i, S_i, \delta_i)})

Assume that \( y_i = \frac{k_i}{\alpha}, i = 1, 2, \ldots, m \), then \( \sum_{i=1}^{m} y_i = 1 \).

Consequently,

\[ \frac{1}{\alpha} \ln \left( \frac{\alpha^n}{\prod_{k=1}^{m} \gamma_k} \right) \prod_{i=1}^{m} \left( e^{\gamma_i f(x; \mu_i, S_i, \delta_i)} \right) \]

= \frac{1}{2\alpha} \ln \left( \frac{(2\pi \alpha)^{-n}}{|\prod_{i=1}^{m} e^{\gamma_i f(x; \mu_i, S_i, \delta_i)}|} \right) 
- \sum_{i=1}^{m} y_i \ln(e^{\gamma_i f(x; \mu_i, S_i, \delta_i)})

But \( \lim_{\alpha \to \infty} \sum_{i=1}^{m} y_i \ln(e^{\gamma_i f(x; \mu_i, S_i, \delta_i)}) = 0 \)

This gives the right side.

**Lemma 15.** The approximation

\[ R_\alpha(X; \mu, S, \delta, \epsilon) \]

\[ \approx \frac{1}{1 - \alpha} \ln \left( \sum_{k_1 \cdots k_m \in B} \prod_{i=1}^{m} (e^{\alpha f(x; \mu_i, S_i, \delta_i)})^{-k_i} \right) 
\prod_{i=1}^{m} e^{\epsilon_i} \exp \left( \left( 1 - \alpha \right) R_\alpha(X; \mu_i, S_i, \delta_i) \right) \]

(24)

is accomplished as \( \alpha \to \infty \).

where, \( \sum_{k_1 \cdots k_m \in B} \frac{\alpha^m}{\prod_{i=1}^{m} k_i} = m^\alpha \), \( B = \{ k_i \in \mathbb{N}, k_i > 0, \sum_{i=1}^{m} k_i = \alpha, i = 1, 2, \ldots, m \} \).

**Proof:** Applying the multinomial theorem in the equality

\[ \int_{R_\alpha} (f(x; \mu, S, \delta, \epsilon))^\alpha \, dx = 
\int_{R_\alpha} (\sum_{i=1}^{m} e^{\epsilon_i} f(x; \mu_i, S_i, \delta_i))^\alpha \, dx \]

To obtain

\[ \int_{R_\alpha} (f(x; \mu, S, \delta, \epsilon))^\alpha \, dx = 
\int_{R_\alpha} (\sum_{k_1 \cdots k_m \in B} \prod_{i=1}^{m} (e^{\alpha f(x; \mu_i, S_i, \delta_i)})^{-k_i}) \, dx \]

where, \( \sum_{k_1 \cdots k_m \in B} \frac{\alpha^m}{\prod_{i=1}^{m} k_i} = m^\alpha \), \( B = \{ k_i \in \mathbb{N}, k_i > 0, \sum_{i=1}^{m} k_i = \alpha, i = 1, 2, \ldots, m \} \).

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

\[ \int_{R_\alpha} (f(x; \mu, S, \delta, \epsilon))^\alpha \, dx \approx 
\prod_{i=1}^{m} \gamma_i \exp \left( \left( 1 - \alpha \right) \sum_{i=1}^{m} y_i \ln \left( \frac{\gamma_i}{e^{\gamma_i f(x; \mu_i, S_i, \delta_i)}} \right) \right) \]

The last approximation can be written as

\[ \int_{R_\alpha} (f(x; \mu, S, \delta, \epsilon))^\alpha = 
\sum_{k_1 \cdots k_m \in B} \prod_{i=1}^{m} (\epsilon_i)^{k_i} \int_{R_\alpha} (f(x; \mu_i, S_i, \delta_i))^\alpha \, dx \]

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 \( X \sim \text{MMSNC}_\alpha(\mu, S, \delta, \epsilon) \) appears as follow

\[ R_\alpha(X; \mu, S, \delta, \epsilon) \geq \text{C\textsubscript{Lower}} \]

(25)

where,

\[ \text{C\textsubscript{Lower}} = \frac{1}{1 - \alpha} \ln \left( \sum_{k_1 \cdots k_m \in B} \frac{\alpha^m}{\prod_{i=1}^{m} k_i} \prod_{i=1}^{m} (\epsilon_i)^{k_i} \right) 
\exp \left( \left( 1 - \alpha \right) \sum_{i=1}^{m} y_i \ln \left( \frac{\gamma_i}{e^{\gamma_i f(x; \mu_i, S_i, \delta_i)}} \right) \right) \]

**Proof** By assumption \( \alpha \neq 1 \), then

\[ R_\alpha(X; \mu, S, \delta, \epsilon) \]

\[ = \frac{1}{1 - \alpha} \ln \left( e^{\sum_{i=1}^{m} \epsilon_i f(x; \mu_i, S_i, \delta_i)} \right) \]

By using the multinomial theorem, we have

\[ \int_{R_\alpha} (f(x; \mu, S, \delta, \epsilon))^\alpha \, dx \approx 
\sum_{k_1 \cdots k_m \in B} \prod_{i=1}^{m} (\epsilon_i)^{k_i} \int_{R_\alpha} (f(x; \mu_i, S_i, \delta_i))^\alpha \, dx \]

Applying generalized Hölder’s inequality, we get

\[ \int_{R_\alpha} (f(x; \mu, S, \delta, \epsilon))^\alpha \, dx \leq 
\sum_{k_1 \cdots k_m \in B} \prod_{i=1}^{m} (\epsilon_i)^{k_i} \prod_{i=1}^{m} \left( \int_{R_\alpha} (f(x; \mu_i, S_i, \delta_i))^\alpha \, dx \right)^{\frac{1}{\alpha}} \]

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

Therefore, the last equation can be written as
\[ \int_{R_a} \left( f(x; \mu, S, \delta, \varepsilon) \right)^\alpha \, dx \leq \sum_{k_i \in B} \frac{a^{l_i}}{l_i!} \prod_{i=1}^l (\varepsilon_i)^k_i \exp \left\{ \sum_{i=1}^l \left( -\alpha \right) k_i R_a \left( X; \mu_i, S_i, \delta_i \right) \right\} \]

If we choose \( p_i = \frac{\alpha}{k_i} \), \( i = 1, 2, \ldots, m \) such that \( \sum_{i=1}^l \frac{1}{p_i} = \sum_{i=1}^l \frac{k_i}{\alpha} = 1 \) and \( 1 \leq \frac{\alpha}{k_i} \leq \alpha \), then

\[ \int_{R_a} \left( f(x; \mu, S, \delta, \varepsilon) \right)^\alpha \, dx \leq \sum_{k_i \in B} \frac{a^{l_i}}{l_i!} \prod_{i=1}^l (\varepsilon_i)^k_i \exp \left\{ -\alpha \sum_{i=1}^l k_i R_a \left( X; \mu_i, S_i, \delta_i \right) \right\} \exp \left\{ \left( 1 - \alpha \right) \sum_{i=1}^l R_a \left( X; \mu_i, S_i, \delta_i \right) \right\} \]

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 \( X \sim \text{MMSNC}_d(\mu, S, \delta, \varepsilon) \). Then the approximation of Renyi entropy of \( X \) can be written as

\[ R_a(X; \mu, S, \delta, \varepsilon) = \frac{1}{2(1-\alpha)} \ln \left( \sum_{k_i \in B} \frac{a^{l_i}}{l_i!} \prod_{i=1}^l (\varepsilon_i)^k_i \exp \left\{ -\alpha \sum_{i=1}^l k_i R_a \left( X; \mu_i, S_i, \delta_i \right) \right\} \right) \]

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 Renyi 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 \( X \sim \text{MMSNC}_d(0, S, \delta, \varepsilon) \) with the following parameters:

**Case (1):** \( d=1 \)

\begin{align*}
\text{m}=2, & \varepsilon = (0.2, 0.8), \quad S = (1.5, 5), \quad \delta = (0.3, 0.4) \\vspace{1mm} \\
\text{m}=3, & \varepsilon = (0.2, 0.3, 0.5), \quad S = (1.5, 5, 3), \quad \delta = (0.3, 0.4, 2.2) \\vspace{1mm} \\
\text{m}=4, & \varepsilon = (0.1, 0.2, 0.2, 0.5), \quad S = (1.5, 5, 3, 2), \quad \delta = (0.3, 0.4, 2.2, 1) \\vspace{1mm} \\
\text{m}=5, & \varepsilon = (0.2, 0.2, 0.2, 0.2, 0.2), \quad S = (1.5, 5, 3, 2, 5), \quad \delta = (0.3, 0.4, 2.2, 1, 2.1)
\end{align*}

**Case (2):** \( d=2 \)

\begin{align*}
\text{m}=2, & \varepsilon = (0.2, 0.8), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.5 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix} \\vspace{1mm} \\
\text{m}=3, & \varepsilon = (0.2, 0.3, 0.5), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.5 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix} \\vspace{1mm} \\
\text{m}=4, & \varepsilon = (0.1, 0.2, 0.2, 0.5), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.5 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix} \\vspace{1mm} \\
\text{m}=5, & \varepsilon = (0.2, 0.2, 0.2, 0.2, 0.2), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.5 \\ 0.3 & 3 & 0.5 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix}
\end{align*}

**Case (3):** \( d=3 \)

\begin{align*}
\text{m}=2, & \varepsilon = (0.2, 0.8), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.3 \\ 0.3 & 3 & 0.3 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix} \\vspace{1mm} \\
\text{m}=3, & \varepsilon = (0.2, 0.3, 0.5), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.3 \\ 0.3 & 3 & 0.3 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix} \\vspace{1mm} \\
\text{m}=4, & \varepsilon = (0.1, 0.2, 0.2, 0.5), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.3 \\ 0.3 & 3 & 0.3 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix} \\vspace{1mm} \\
\text{m}=5, & \varepsilon = (0.2, 0.2, 0.2, 0.2, 0.2), \quad S = \begin{pmatrix} 0.7 & 0.3 & 0.3 \\ 0.3 & 3 & 0.3 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0.16 & 0.23 \\ 0.59 & 3.1 \end{pmatrix}
\end{align*}
For each case in example 2., table 2. summarizes the values of approximate Rényi (Shannon) entropy and its bounds of \( X \sim \text{MMSNC}_d(0, S, \delta, \varepsilon) \) 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 \( \varepsilon \). It can be seen that the value of \( m \) changes the values of entropy.

**TABLE 2.** Shannon entropy of \( X \sim \text{MMSNC}_d(0, S, \delta, \varepsilon) \) is computed for \( m = 2, 3, 4 \) and 5 in one, two and three dimensions.

<table>
<thead>
<tr>
<th>Case</th>
<th>( d )</th>
<th>( m )</th>
<th>( C_{\text{lower}} )</th>
<th>( C_{\text{upper}} )</th>
<th>( H )</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1.8558</td>
<td>1.8671</td>
<td>1.8614</td>
<td>0.0057</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>3</td>
<td>1.7741</td>
<td>1.7826</td>
<td>1.7783</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4</td>
<td>1.7290</td>
<td>1.7374</td>
<td>1.7332</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>1.7972</td>
<td>1.8136</td>
<td>1.8054</td>
<td>0.0082</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.2154</td>
<td>2.3827</td>
<td>2.2990</td>
<td>0.0836</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>2.2427</td>
<td>2.4911</td>
<td>2.3669</td>
<td>0.1242</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>2.4786</td>
<td>2.8261</td>
<td>2.6523</td>
<td>0.1738</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
<td>2.5190</td>
<td>2.8414</td>
<td>2.6802</td>
<td>0.1612</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5.5641</td>
<td>5.6966</td>
<td>5.6304</td>
<td>0.0663</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>4.6287</td>
<td>5.0595</td>
<td>4.8441</td>
<td>0.2154</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>4.4237</td>
<td>4.8133</td>
<td>4.6185</td>
<td>0.1948</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5</td>
<td>4.4452</td>
<td>4.8519</td>
<td>4.6486</td>
<td>0.2033</td>
</tr>
</tbody>
</table>
FIGURE 2. The horizontal line represents the values of parameter $m$ and the vertical line is a Shannon entropy of $X \sim \text{MMSNC}_d(0, S, \delta, \epsilon)$ with parameters in example 2.

Figure 3. The horizontal line represents the values of parameter $\alpha$ and the vertical line is a Renyi entropy of $X \sim \text{MmsCr}_d(0, S, \Delta, E)$ 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 α = 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 α → ∞, this is obviously when α > 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 Rényi entropy of $X \sim \text{MMSCN}_d(0, S, \delta, \epsilon)$. Using the mean of these bounds, the approximate values of entropy can be calculated. These values are localized between bounds of Rényi 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 Rényi entropy of $X \sim \text{MMSCN}_d(0, S, \delta, \epsilon)$ 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.

REFERENCES:


